AN ABSTRACT OF THE THESIS OF

Maribeth Todd for the degree of Master of Science in Agricultural and Resource Economics presented on June 11, 2008.
Title: Urban Development in Coastal Oregon: Discrete-Choice Estimation with Spatial Autocorrelation.

Abstract approved:

_____

Andrew J. Plantinga

Development along the Oregon coast is continuing in areas that are known to be hazardous, in spite of strict zoning and other laws. The coast commonly sees erosion that can wash away cliffs and undermine structures as well as accretion that can bury houses in sand. However the much more acute risk is the potential for a tsunami triggered by an earthquake in the ocean. This paper examines the factors influencing development along the Oregon coast using satellite imagery and other spatial data. An econometric model for land use change at the pixel level is developed. The scale at which the data (satellite imagery) are collected is different from the scale at which development occurs, leading to spatial correlation among pixels that are geographically close to each other. Estimating a standard probit model in this case leads to parameter estimates that are inconsistent. Incorporating a spatial lag of the dependent variable will account for the spatial autocorrelation but makes maximum likelihood estimation nearly impossible so a Bayesian approach is used instead. A Gibb's sampling algorithm is implemented to estimate the conditional distribution of each parameter in the model, from which parameter estimates can be derived. The Bayesian spatial probit estimation is very computationally intensive however the results indicate that the spatial lag is a crucial part of the model. This approach also allows projections of the spatial

pattern of future development, unlike estimation methods that sample the data to remove autocorrelation among observations. The spatial lag model is then used to project future development patterns in several regions along the Oregon coast and explore potential applications of this projection model. In particular, predicted development patterns in the tsunami hazard zone are examined in the Waldport area, on the central coast of Oregon.

Urban Development in Coastal Oregon:
Discrete-Choice Estimation with Spatial Autocorrelation

by
Maribeth Todd

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 11, 2008
Commencement June 2009

Master of Science thesis of Maribeth Todd presented on June 11, 2008.

APPROVED:

Major Professor, representing Agricultural and Resource Economics

---

Head of the Department of Agricultural and Resource Economics

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Maribeth Todd, Author

ACKNOWLEDGEMENTS

I would like to thank…

Scott, my soon to be husband, for his patience and encouragement the past few years

Andrew and Julia for their advice on this project as well as their support through some tough times and transitions

My friends, roommates and classmates for keeping life fun and interesting in spite of all the hard work

My family, for loving and supporting me in everything that I do

Support for this research was provided by an NSF IGERT graduate fellowship (NSF award 0333257) in the Ecosystem Informatics IGERT program at Oregon State University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

Urban Development in Coastal Oregon:
Discrete-Choice Estimation with Spatial Autocorrelation

1. Introduction

The Oregon coast presents unique challenges to urban development. The land that is available for development lies in a narrow strip between the Pacific Ocean and the mountains of the Oregon Coast Range. The coastline is characterized by low lying estuaries and beaches in between rocky headlands. Oceanfront properties are prone to long term erosion and accretion that vary with the seasons and weather patterns, but also face the risk of more acute hazards such as landslides, floods and rare but devastating tsunamis.

Oregon has a long history of protecting its coastal resources, beginning with Governor Oswald West. In 1913, West and the state legislature designated Oregon's beaches a state highway in order to guarantee public access to the area between the low and high tide lines. The State Highway Commission was also established that year, and in 1925 the commission was given the authority to start acquiring land near the roadway for parks and campgrounds (ODOT 1998). That authorization allowed the state to establish 36 state parks along the coast by 1950 (NOAA 1998). Today there are 84 parks on the Oregon coast according to the Oregon Parks and Recreation Department. When full public access to Oregon's beaches was called into question in 1966, Governor Tom McCall was quick to react. The Beach Bill was signed into law in 1967 and extended public access to include all beach sand up to the vegetation line regardless of ownership (NOAA 1998).

A more comprehensive land use planning system was established in 1973 and laid out 19 statewide planning goals. There are three goals that pertain specifically to protecting the coast. Goals 17, 18 and 19 provide guidelines for managing coastal shorelands, beaches and dunes, and ocean resources respectively (DLCD 1997). The state planning system also established strict zoning laws to protect important resource lands such as agriculture and timber.

In spite of the protections provided by these laws, development is still occurring in environmentally sensitive or hazardous areas, particularly in the tsunami hazard zone. The tsunami hazard zone is defined by the tsunami inundation line that was mapped by the Oregon Department of Geology and Mineral Industries (DOGAMI). Few restrictions are placed on development within the tsunami hazard zone and the regulations that are in place are enforced through the state building code. These regulations place limits only on construction of public facilities that would be difficult to evacuate in case of emergency, such as jails or schools, and emergency facilities like fire stations (WSSPC 1997). This leaves the tsunami hazard zone open to the construction of homes, condos, hotels and other structures. Tsunamis have historically occurred every 300 to 600 years in Oregon, with the last tsunami striking the coast in 1700 (WSSPC 1997). The next major earthquake to cause a tsunami in Oregon could happen tomorrow or it may be another 200 years. This is a large window of uncertainty compared to the timescale of human experience. The loss will be devastating when it occurs, but because of this uncertainty, developers and homebuyers rarely take this risk into consideration when choosing a location (personal communication with Lisa Phipps, resource planner with Tillamook County).

Chronic risks associated with erosion and accretion are also serious problems for oceanfront property on the Oregon coast. From one year to the next, the coast may sees erosion that can wash away cliffs and undermine structures as well as accretion that can bury houses in sand. This constant movement of sand leads

many property owners to put riprap (piles of large boulders) or other protective structures in front of their homes to prevent further erosion. However, hardening the shoreline in this manner can cause a variety of unintended negative externalities, such as increasing erosion rates in other areas or destroying public beaches. Oregon has strong laws to protect public access to beaches and the development that is going on along the coast is not necessarily consistent with maintaining the rights of the public. There are negative externalities of oceanfront development associated with the tsunami risk as well. Development removes natural vegetation and replaces it with homes and other buildings and, as mentioned, can destroy beaches if homeowners try to protect their property from erosion. Vegetation and beaches both provide buffers against an incoming tsunami. Without them, as some studies of the Indian Ocean tsunami in 2004 have suggested, a giant wave can reach much further inland causing more extensive damage (Danielson et al. 2005).

Beyond simply a lack of restrictions on development in hazardous coastal areas, government policies may even be encouraging more development in these areas. By providing disaster relief programs like the Federal Emergency Management Agency (FEMA), the federal government has reduced the financial risk to individual property owners in the event of a disaster like a major storm or tsunami by spreading the financial burden across all taxpayers. This leads to a moral hazard problem that causes inefficiencies in the level of development in hazardous coastal areas. Individuals no longer have to bear the full burden of their risky decision to build in a high risk area, which results in a higher level of development in those areas than would be seen under the efficient outcome. In addition to the financial risks associated with this development, as more people choose to live by the water, more lives may also be at risk in the event of a tsunami. Unnecessary costs are being placed on society in the form of disaster relief aid and even loss of life because too much development is continuing in areas that are known to be high risk.

Land-use planners on the coast face the additional challenge of complying with goal 14 of the statewide planning system. Goal 14 relates to urbanization and requires the establishment of urban growth boundaries that can "accommodate long range urban population, consistent with a 20-year population forecast coordinated with affected local governments" (DLCD 1997). Coastal land-use planners have indicated that they are lacking a sound method to determine just how much land they will need over a 20 year time horizon because development on the coast is not driven by population change alone (personal communication with Lisa Phipps). Tourism is a major driver and much of the development in coastal communities consists of condos, vacation homes and other tourist destinations. Having a model to predict development patterns along the coast would be helpful to resource planners, emergency managers and others.

To address any of these issues, it would be useful to understand how development is progressing on the coast, particularly in areas that are prone to erosion, accretion and tsunamis. The objective of this paper is to develop a model to predict development on the Oregon coast based on satellite imagery and other pixel level data. A land-use change model based on pixel level data has its shortcomings. Decisions about land use are made by landowners at the parcel level, not at the scale of 30 meter pixels (Irwin 2001). However satellite imagery and other raster type data is fairly cheap and easy to come by compared to parcel data for large study areas. A predictive model based on remotely sensed data and other attributes that can be derived through GIS calculations would be an economical alternative to a model based on parcel data.

We have one observation of land-use change based on satellite imagery acquired by the National Oceanic and Atmospheric Administration (NOAA) in 1996 and 2001. We observe whether each previously undeveloped pixel gets developed or not over the five year period. The 30 meter pixel resolution of the satellite

imagery results in clusters of pixels that get developed at the same time, because an undeveloped parcel of land is generally larger that 900 square meters. This type of spatial autocorrelation in a binary choice setting is difficult to model and many previous studies have ignored autocorrelation or sampled the data to reduce or remove the problem. Sampling creates a dataset of nonnearest neighbors by choosing points or parcels with a specified minimum distance between them. This approach does reduce or remove the spatial dependence among observations, and therefore should lead to unbiased parameter estimates (Carrión-Flores and Irwin 2004). However, sampling treats the spatial autocorrelation as a nuisance rather than an integral part of the process of land-use change and obscures the spatial relationships among observations. In this case, it would be desirable to preserve the clustered pattern of development for the purpose of projecting future land use changes, so we employ a binary choice (probit) model with a spatial lag of the dependent variable. This model is estimated using a Bayesian approach, a Gibb's sampling algorithm that has been developed by James P. LeSage and others (LeSage 2000, Holloway, Shankar and Rahman 2002, Thomas 2006).

This Bayesian estimation method is very computationally intensive, and cannot be implemented on the full dataset at once. Therefore, the study area has been divided along natural breaks in the geography and 15 regional models have been estimated. Once regional parameter estimates were obtained for the observed data from 1996 to 2001, three regions were chosen to produce projections of future development over the five year period from 2001 to 2006. In order to explore some potential applications for these projection models, several different types of analysis were performed. Two different growth scenarios are compared using the spatial probit model and then differences are examined between the spatial probit results and the development that would be projected by a standard probit model. Some of the relationships between projected development and other characteristics of the land are also examined, including zoning, slope and tsunami risk.

2. Literature Review

There are numerous examples in the literature of land-use change models based on satellite imagery and other GIS data. These studies have focused mainly on deforestation in developing countries. The models typically involve landowners choosing the profit maximizing land use from a set of options including different forms of agriculture, forest and other uses. These decisions are driven by spatial characteristics that generally include factors like rainfall, soil quality, and proximity to cities and highways as well as other explanatory variables. See, for example, Chomitz and Gray (1996) and Nelson and Hellerstein (1997).

The difficulties associated with discrete choice models with spatial effects are well documented. By incorporating a spatial lag or another assumption for the pattern of spatial autocorrelation into the probit, we create two problems. The implied error structure is heteroscedastic and the resulting likelihood function involves a multidimensional integral that cannot be estimated directly (McMillen 1992, Fleming 2004). McMillen proposes an estimation method for the spatial autoregressive probit model that is based on an estimation and maximization (EM) algorithm. This method estimates the expectation of the underlying latent variable and substitutes this value into the model in place of the discrete dependent variable. The log-likelihood function of this modified model can then be maximized to obtain parameter estimates. These two steps are repeated until the estimates converge. One shortcoming of the EM algorithm arises with the covariance matrix of the estimates. McMillen admits that the results of the proposed method for estimating the covariance matrix for this model will be biased. The EM algorithm produces consistent parameter estimates but misleading confidence intervals.

A Bayesian approach is preferable to the EM algorithm for several reasons. First, this method allows us to derive valid measures of dispersion for the parameter

estimates unlike the biased estimates of the EM approach. The Bayesian spatial probit relies on Gibb's sampling, which uses a large number of parameter draws to produce a posterior conditional distribution for each parameter in the model. These draws can also be used to construct measures of dispersion for each parameter. Another shortcoming of the EM algorithm is the necessity to define a functional form for the heteroscedasticity present in the error terms. Heteroscedasticity is handled in the Bayesian approach by sampling for the variance terms from a chi-squared prior distribution with r degrees of freedom. The chi-squared distribution can be characterized by this one parameter, the degrees of freedom (r), which is chosen based on the degree of heteroscedasticity that is believed to occur in the data. Thus the terms in the covariance matrix can be estimated by adding one assumption about r into the model, which is less restrictive than specifying a particular functional form for the error terms (LeSage 1998).

Many previous examples of probit models using spatial data have ignored the spatial autocorrelation problem, see for example Nelson and Hellerstein (1997). Bockstael (1997) estimates a discrete choice model of land-use change in Maryland, by first estimating a hedonic model of land values in different uses and then estimating conversion probabilities. The author acknowledges the likely presence of spatial autocorrelation among the error terms but makes no attempt to model the spatial relationship. In other studies the dataset is sampled to try and remove the autocorrelation problem. Carrión-Flores and Irwin (2004) estimate a standard probit model of residential development in the urban-rural fringe. The authors test for spatial autocorrelation in the error term but treat autocorrelation as a nuisance rather than an integral part of the model. To correct for the spatially correlated errors, the dataset is sampled to ensure that no two parcels in the study are within a certain distance of each other. The authors dismiss computationally intensive estimation methods like the Bayesian spatial probit as being too difficult to implement on their dataset of 9,760 observations.

Few examples seem to exist of econometric applications of the Bayesian spatial probit despite the availability of several primers on the theory and application of the model as well as tools to implement it. LeSage provides the most thorough description of the Gibb's sampling approach to discrete choice models with spatial autocorrelation in the form of a web book titled "Spatial Econometrics" (1998). Holloway, Shankar and Rahmann (2002) provide a primer on Bayesian estimation of the probit model with a spatial lag of the dependent variable as well as an example of an application. They begin by describing how the Gibb's sampler works in the most basic normal means estimation problem, then expand to the spatial autoregressive probit and then apply the estimation technique to high-yielding variety rice adoption, a research question first examined by Case (1992).

Timothy Thomas seems to be one of few researchers to implement the Bayesian spatial probit method on a regular basis. His primer on this technique includes one example of his research, an application of the model to deforestation in Madagascar. The results of this study point to the difficulties associated with choosing the best model specification, between a spatial autoregressive model, a model with spatially dependent error terms or a combination of the two. In this example, computing log-likelihoods using the simulated latent variable and comparing the values for these different specifications produced misleading results. Thomas has also developed Matlab code to implement the estimation technique, based on tools originally created by LeSage for his econometrics toolbox for Matlab. Both researchers provide these programs for download from their websites, making this estimation method more accessible to researchers.

The September 2007 issue of the Journal of Agricultural Economics contained a special feature titled "Spatial Issues in Agricultural Economics" underscoring the continued relevance of this problem. In their contribution to the issue, De Pinto and Nelson lament the current lack of a theory based approach to discrete choice

models with spatial effects. Instead, they say, the land-use literature is still employing “ad hoc” estimation methods to minimize the impact of spatial effects on model results.

3. Materials and Methods

3.1 Study Area

The coast of Oregon is relatively undeveloped compared to other coastal states. This is due in part to the lack of large deep water ports on the coast and the cities that usually grow up around them. The coast is also isolated from the major cities in the Willamette Valley by the Coast Range, although road improvements are in the works to make the coast more accessible from the valley. For example, US Highway 20 between Corvallis and Newport is currently under construction to straighten and shorten the route. There is the potential to see more development pressure on coastal communities as the commute between the valley and the coast becomes easier and population grows statewide.

A large portion of the coast is publicly owned, primarily in the form of oceanfront state parks and state and nationally owned forest land in the Coast Range. Privately owned land is governed by strict zoning laws which can be broadly categorized into zones that allow for development and those that do not. This results in urban and developed areas that are separated from each other by public lands and areas that are zoned for resource uses like agriculture and forestry. In addition, mountains and headlands also form natural barriers between communities along the coast.

Under current laws, we would not expect to see significant development on land that is not zoned for development. Therefore, this study considers only lands that are currently zoned for development. This includes five general zoning categories on the coast: urban, rural residential, rural commercial, rural industrial and rural service center. In addition, the study area has been restricted to land that is located within five miles of the ocean in order to focus on areas that are truly coastal.

3.2 Data

The primary dataset is satellite imagery of the Pacific Coast available from NOAA. Landsat Thematic Mapper and Landsat Enhanced Thematic Mapper satellite imagery of 30 meter resolution were classified by NOAA into 22 land cover types, including high intensity and low intensity development. The coastal land-cover data provides two snapshots of the Oregon coast, in 1996 and 2001, in the form of raster datasets with 30 x 30 meter pixels. While the datasets provide many different land cover classifications, we are concerned here with whether land is in a developed state or a non-developed state and if the area is in a land cover type that is suitable for development. Categories that are not suitable for development, such as water, were excluded from the dataset.

A land ownership layer was used to exclude land that is publicly owned and a state zoning map was used to reduce the dataset to only land that is currently zoned for development. Once a dataset of all privately owned land that is suitable for development was created, additional characteristics of the pixels were derived from other GIS layers. A digital elevation map was used to determine slope and aspect of each pixel. A map of highways was used to determine the distance from each pixel to the nearest highway and the state zoning map was used to determine the zoning of each pixel.

3.3 Model

The econometric framework is based on a random utility model. The owner of a piece of land derives utility from the land from either developing it or using it for some other purpose. The utility of developing a piece of land is $U_d = \overline{U}_d + \varepsilon_d$ and the utility of using the land for any other use is $U_o = \overline{U}_o + \varepsilon_o$. These utilities

depend on the returns the owner can realize from the land in a developed or undeveloped state. The land will be devoted to the highest value use. The probability of development is

$$\begin{aligned}\Pr[Y=1] &= \Pr[U_d > U_o] \\ &= \Pr\left[\overline{U}_d + \varepsilon_d > \overline{U}_o + \varepsilon_o\right] \\ &= \Pr\left[\overline{U}_d - \overline{U}_o > \varepsilon_o - \varepsilon_d\right]\end{aligned}$$

In the probit model, we assume that the errors are normally distributed with mean zero, variances σ_d^2 and σ_o^2 and covariance σ_{do}. So the probability of development is given by

$$\Pr[Y=1] = \int_{-\infty}^{\overline{U}_d - \overline{U}_o} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} dt$$

where $\sigma^2 = \sigma_d^2 + \sigma_0^2 - 2\sigma_{do}$ is the variance of $\varepsilon_o - \varepsilon_d$. The deterministic portion of the random utility model is $\overline{U}_d - \overline{U}_o$, equal to the difference in the utilities associated with developing and not developing a piece of land. We do not directly observe these utilities, we only observe the outcome of the decision process. When $U_d > U_o$, the pixel is developed and we observe $Y = 1$, otherwise the pixel is not developed and $Y = 0$. Utility is assumed to reflect the returns, or the benefits net of the costs, that could be realized from using the land for a particular purpose. In the case of, for example, a housing development, returns would depend on the total sale price of all homes in the development minus the costs of developing the land. These returns to the land could be estimated with a hedonic model based on the characteristics of each pixel. Bockstael (1996) estimates a model of land-use change in Maryland in two steps, first estimating the value of land in the alternative uses with a hedonic and second estimating the probability of conversion. The hedonic model used in that study includes variables such as distance to major roads, distance to amenities like the waterfront, and variables that describe the surrounding landscape. Similar variables are used here, however

these characteristics are used directly in the probability model rather than estimating returns separately.

Location is an important determinant of land value, particularly on the coast. Several characteristics were included to measure the desirability of the location. Construction that is closer to the ocean and potentially has an ocean view is worth more than inland development. So OCEAN measures the distance, in hundreds of feet, from each pixel to the ocean and WEST is dummy variable for western facing slopes. Pixels that are located in an urban area are generally closer to business districts and are developed at higher densities, so URBAN is a dummy variable for urban zoning. Otherwise, the land is located in one of the rural zoning designations. Zoning is taken as fixed and exogenous in this model, because the time horizon considered here is five to ten years. Major changes to urban growth boundaries and other zoning designations occur infrequently so this assumption is reasonable for this study. Over longer time periods, however, changes in the urban growth boundary would also need to be modeled.

Several factors are also included to measure the costs of development. Costs are influenced by accessibility, measured here by HWY. This variable is the distance from the nearest existing highway measured in hundreds of feet. The costs of extending infrastructure such as water and sewer lines will increase as the distance from existing development increases. DEVPROX is a measure of proximity to current development, calculated by finding the average distance to the nearest eight pixels that have already been developed. This means that if a pixel is completely surrounded by development, it will have a DEVPROX value of 30 meters, or about 98 feet. This measurement will be much higher in areas that are sparsely developed. SLOPE is measured in degrees and is a proxy for the increased engineering and construction costs of building on steeper slopes.

Initially, several soil characteristics were included as explanatory variables, such as limitations on construction due to shrink-swell, propensity for flooding, and depth to hard bedrock. These factors could significantly impact returns to the land in developed or other uses. However, we decided that in the final model parsimony was more valuable than the additional information from these explanatory variables. The data have also been mean centered at the suggestion of a statistician (personal communication with Lisa Madsen) to help promote convergence of the Gibb's sampling algorithm, which is necessary to obtain useful parameter estimates. The variables included in the model are summarized in table 1.

Table 1. Variables included in the final model.

Variable Definitions	
WEST	Dummy variable if the pixel is a western facing slope
HIGHWAY	Distance from nearest highway in hundreds of feet
OCEAN	Distance from ocean, including large estuaries, in hundreds of feet
URBAN	Dummy variable for urban zoning vs. rural zoning
SLOPE	Slope of the pixel in degrees
DEVPROX	Measure of proximity to existing development

The scale at which the NOAA land cover and other data are collected is different from the scale at which development occurs, leading to spatial correlation among pixels that are geographically close to each other. A probit model is chosen to ensure that the predicted outcomes are in the range between zero and one, while a spatial lag of the dependent variable accounts for the correlation among pixels with the same outcome. The clustered pattern of development is captured by a spatial autoregressive probit model of the form $z = \rho Wz + x\mathrm{B} + \varepsilon$, where:

- $z_{(\mathrm{Nx1})}$ denotes the latent variable that underlies the discrete observations of 1 if a parcel was developed during the observed timeframe and 0 if not

- $\rho_{(1x1)}$ is the spatial parameter and conveys the level of spatial correlation among observations
- $W_{(NxN)}$ is the spatial weights matrix defining the spatial relationship between observations
- $x_{(NxK)}$ denotes the matrix of observations on the regressors
- $B_{(Kx1)}$ denotes the vector of coefficients to be estimated
- $\varepsilon_{(Nx1)}$ is a random error term with mean zero and variance $\sigma^2 V$ where V has non-constant terms on the diagonal to reflect heteroscedasticity

The observed discrete choice can be thought of as a reflection of an underlying regression involving the continuous latent variable, z, where we observe $y_i = 1$ when $z_i > 0$ and $y_i = 0$ when $z_i \leq 0$ for pixel i. In the random utility framework, $\overline{U}_d - \overline{U}_o = \rho W z + x\mathrm{B}$. So if the utility of developing pixel i is greater than the utility of not developing, then $\rho W z + x\mathrm{B} > 0$ and there is a high probability of observing development, or $y_i = 1$.

By introducing the spatially lagged dependent variable into the probit model, we make direct maximum likelihood estimation nearly impossible because the resulting likelihood function is a multidimensional integral that cannot be simplified. This also creates errors that are heteroscedastic and correlated. Rearranging the model equation, $y = (I - \rho W)^{-1} x\mathrm{B} + (I - \rho W)^{-1} \varepsilon$, results in the error term $u = (I - \rho W)^{-1} \varepsilon$ with heteroscedastic and spatially correlated variance $E(uu') = (I - \rho W)^{-1} E(\varepsilon\varepsilon')\left[(I - \rho W)^{-1}\right]'$. Discrete choice models with spatial autocorrelation and heteroscedasticity pose difficult estimation problems with few solutions. In addition to the implications of the model specification, this approach also allows us to model heteroscedasticity that may arise in the error term ε.

LeSage and others have developed a body of work describing the use of Bayesian methods to estimate this type of model. All of the parameters in this model, B, ρ, σ and V, share a multivariate posterior distribution that can be derived from combining the likelihood function and the prior distribution for the parameters. Denote the likelihood for the entire data as $\ell(\mathrm{B}, \rho, \sigma, V | y)$ and the prior probability density function (pdf) as $\pi(\mathrm{B}, \rho, \sigma, V)$ where the prior pdf reflects any knowledge or assumptions we have about the distribution of these parameters. Then the posterior distribution is found by combining our prior beliefs with the likelihood function of the data via Bayes' rule, $\pi(\mathrm{B}, \rho, \sigma, V | y) \propto \ell(\mathrm{B}, \rho, \sigma, V | y)\pi(\mathrm{B}, \rho, \sigma, V)$. Theoretically, we would like to estimate the full posterior distribution from which these parameters are derived however the function obtained by this method is very complicated and difficult to work with (Holloway 2002).

The Gibb's sampler allows us to produce parameter estimates for this model using conditional distributions instead of the full posterior distribution. Gibb's sampling takes a draw from the conditional distribution for each parameter in the model in sequence until a large sample of draws is collected. Gelfand and Smith (1990) show that the parameter estimates obtained by sampling the full set of conditional distributions converge to the true parameter values of the full joint posterior distribution. The algorithm for the spatial probit simulates the latent variable, z, underlying the observed binary response, y. Then the other parameters in the linear regression $z = \rho W z + x\mathrm{B} + \varepsilon$ are estimated by taking draws from the conditional distribution of each parameter contingent on all other parameters in the model. These simulated draws are saved and used to update the conditional distribution of each parameter. A "burn in" period, during which the draws are not saved, is implemented to allow the Gibb's sampler to establish a steady state.

Matlab code to implement this method has been graciously provided by Dr. Timothy Thomas of the World Bank. This program calls on many routines that

were created by LeSage and are available in his econometrics toolbox for Matlab. Thomas's algorithm is as follows:

- Simulate the latent variable z, which is drawn from a truncated normal distribution. If the observed y value is 0, z is drawn from a left truncated normal distribution whereas if y is observed to be 1, z is drawn from a right truncated normal distribution.
- B vector is estimated using Gibb's sampling from a multivariate normal distribution. Since we do not have strong prior beliefs about the distribution of B, the prior mean is zero and the prior covariance matrix consists of large variances along the diagonal.
- Estimate the variance parameters σ^2 and V. σ^2 is drawn from a gamma distribution with mean zero and zero degrees of freedom. This assumption, again, reflects that we do not have strong prior beliefs about the distribution of this parameter.
- V has nonzero diagonal elements (not all 1) and zeroes off the diagonal due to heteroscedasticity. The diagonal elements of V are drawn from a chi-squared distribution with r degrees of freedom, where r represents the degree of hesteroscedasticity that we expect.
- The spatial parameter, ρ, must be estimated using a Metropolis-Hastings random walk algorithm since it cannot be drawn from a standard distribution.

The sampling approach for the spatial parameter, ρ, is slightly more complicated than for the other parameters in the model. We do not have a conditional distribution for ρ in closed form, so the Metropolis-Hastings algorithm provides a way to estimate this parameter using a random walk procedure. This algorithm draws from a known distribution and then accepts or rejects the draw based on a probability rule that takes into account the newly updated target distribution at each step. The accepted sequence of draws should then converge to the desired distribution.

Specification of the spatial weights matrix, W, is another important consideration in this estimation method. The weights matrix determines which neighboring pixels will influence the outcome on a given pixel of land. In this case, W is based on inverse distances to the neighboring pixels. The nonzero elements are set equal to the inverse of the distance between two pixels. Beyond 130 meters, the neighbors are assumed to have no influence and the values in the weights matrix are set to zero. This threshold value was chosen to capture the majority of the clusters of development seen in the observed dataset however this is not the only option and may not be the best choice. The specification of W can have significant impacts on the parameter estimates and this is one area that could benefit from additional research.

The satellite imagery and other raster data used in this study form a regular grid of 30 meter pixels. However, because of the ownership and zoning restrictions placed on the dataset, some of the pixels in the grid will not be a part of the dataset of developable land. So some pixels will have more neighbors than others. With a threshold of 130 meters, each pixel could have up to 68 neighbors influencing it through the weights matrix. Figure 1 shows the spatial pattern of neighbors induced by this specification. Numbers represent the distance in meters from the center pixel, while all 68 pixels shown in the grid fall within the threshold of 130 meters from the center pixel. Of course, not all pixels in the dataset will have a full 68 neighbors due to the patchiness of land in the dataset that is suitable for development.

Figure 1. Spatial pattern of influence induced by the chosen specification of the weights matrix.

		*	*	120	*	*		
	*	*	*	90	*	*	*	
*	*	*	*	60	*	*	*	*
*	*	*	*	30	*	*	*	*
120	90	60	30	0	30	60	90	120
*	*	*	*	30	*	*	*	*
*	*	*	*	60	*	*	*	*
	*	*	*	90	*	*	*	
		*	*	120	*	*		

Row standardization of the weights matrix is generally desirable, as it alleviates some computation problems and assures that the spatial parameter, ρ, is between -1 and 1. The process of row standardization normalizes each row so that it sums to 1. However, this can be problematic with distance based weights. Consider a pixel that has two neighbors that are both 120 meters away. Row standardization would give each of these pixels a weight of 0.5. Whereas, for a pixel with four neighbors that are 30 meters away, each neighbor would have a weight of 0.25. Presumably, a pixel with adjacent neighbors should be more influenced by its neighbors than a pixel with neighbors that are far away, regardless of how many total neighbors are within the threshold distance. Adjacent pixels have a higher likelihood of being part of the same parcel than those that are farther apart, and development occurs at the parcel level. Row standardization maintains the distance relationship among neighbors of a particular pixel, but not the distance relationship among neighbors of different pixels.

We use matrix standardization, a solution suggested by Thomas. This method, rather than dividing the elements of each row by their own row total, finds the

largest row sum in the matrix and divides the elements of each row by this number. This maintains all of the distance relationships in the matrix and produces a spatial parameter that can be roughly interpreted as a correlation measure.

Estimating the spatial probit is very computationally intensive and the full dataset for the Oregon coast contains measurements on around 600,000 pixels. Anything more than about 35,000 pixels is too much for a standard personal computer to handle, so implementing this estimation method requires breaking the full dataset into smaller regional models. Given the size limitations on the datasets and the natural breaks that occur between developed areas due to public lands, mountains, headlands and resource zoning, this process is fairly straightforward. The coast was divided into 15 different regions, which are shown in figures 2, 3 and 4.

Each region on the northern portion of the Oregon coast consists of several cities along with areas that are zoned for rural development. The Astoria region consists of Hammond, Warrenton and Astoria and lies at the northernmost point in Oregon on the mouth of the Columbia River. Just south of Astoria is the Seaside region, containing the beachside towns of Gearhart and Seaside. Tillamook Head borders Seaside on the south side and separates the Seaside region from the Cannon Beach region. This region consists of a mix of tourist towns, like Cannon Beach, Rockaway and Manzanita (not shown on the map) and fishing communities like Wheeler. The Tillamook region contains Tillamook, Bay City and Garibaldi, and some smaller oceanfront communities.

Moving to the central coast, again the regions generally consist of multiple cities and rural areas. The Lincoln City region can be seen at the top of figure 3. Depoe Bay, Toledo and the portion of Newport that lies north of Yaquina Bay make up the Newport region. The Waldport study area consists of the portion of Newport to the south of Yaquina Bay, Waldport and Yachats. Florence and Dunes City

comprise the Florence region. Reedsport is the smallest region in this study and is the only developed area on the coast in Douglas County.

On the southern coast, shown in figure 4, cities are generally further apart and most of the regions on this section of the coast contain only one urban area and some rural land. The cities of Lakeside, Bandon, Port Orford, Gold Beach and Brookings each comprise their own regions, along with some land zoned for rural development. The dataset for the region consisting of Coos Bay and North Bend turned out to be too large to analyze, with no natural way to divide it any further, so it has been omitted from this study.

Figure 2. Areas zoned for development for study regions on the north coast.

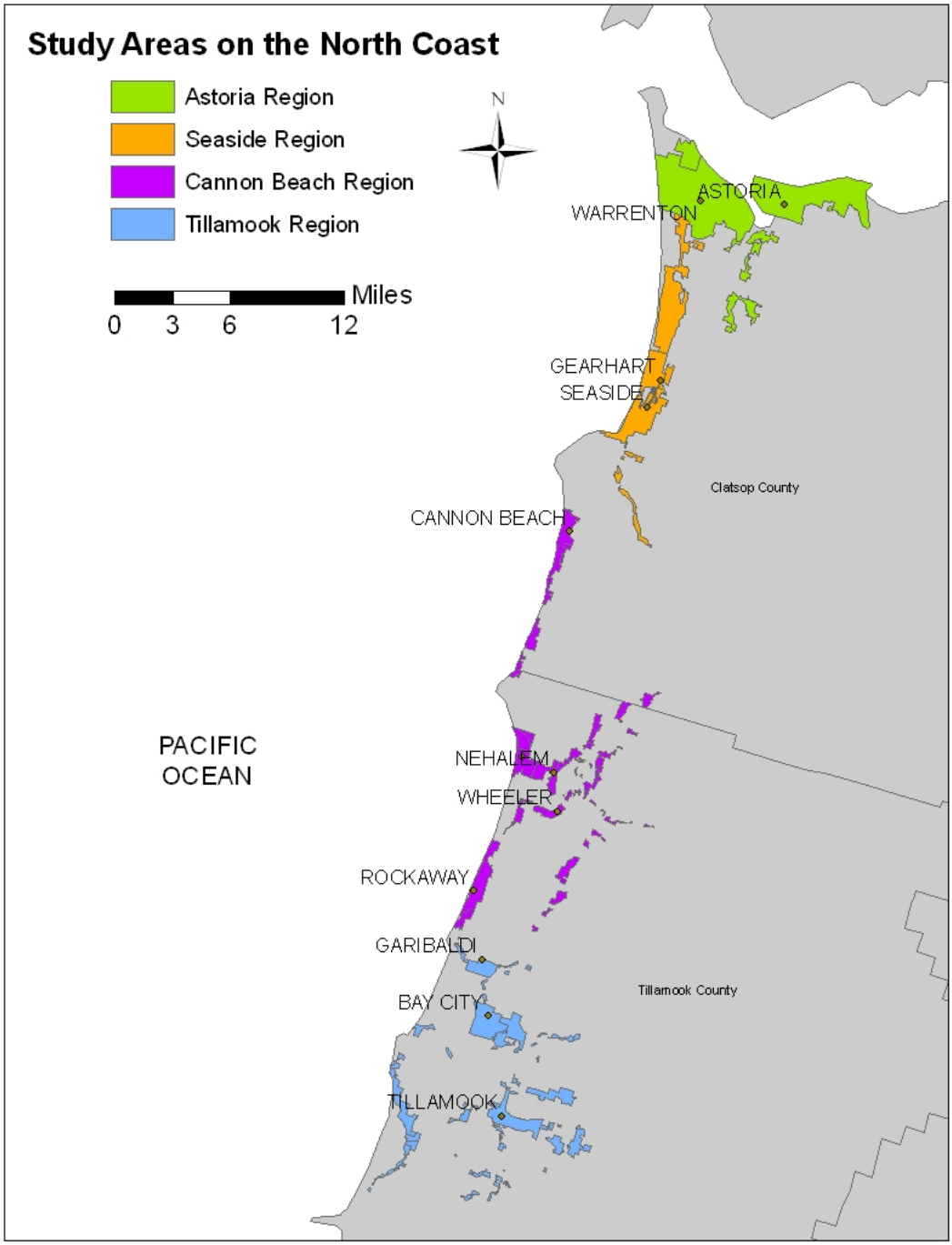


Figure 3. Areas zoned for development for study regions on the central coast.

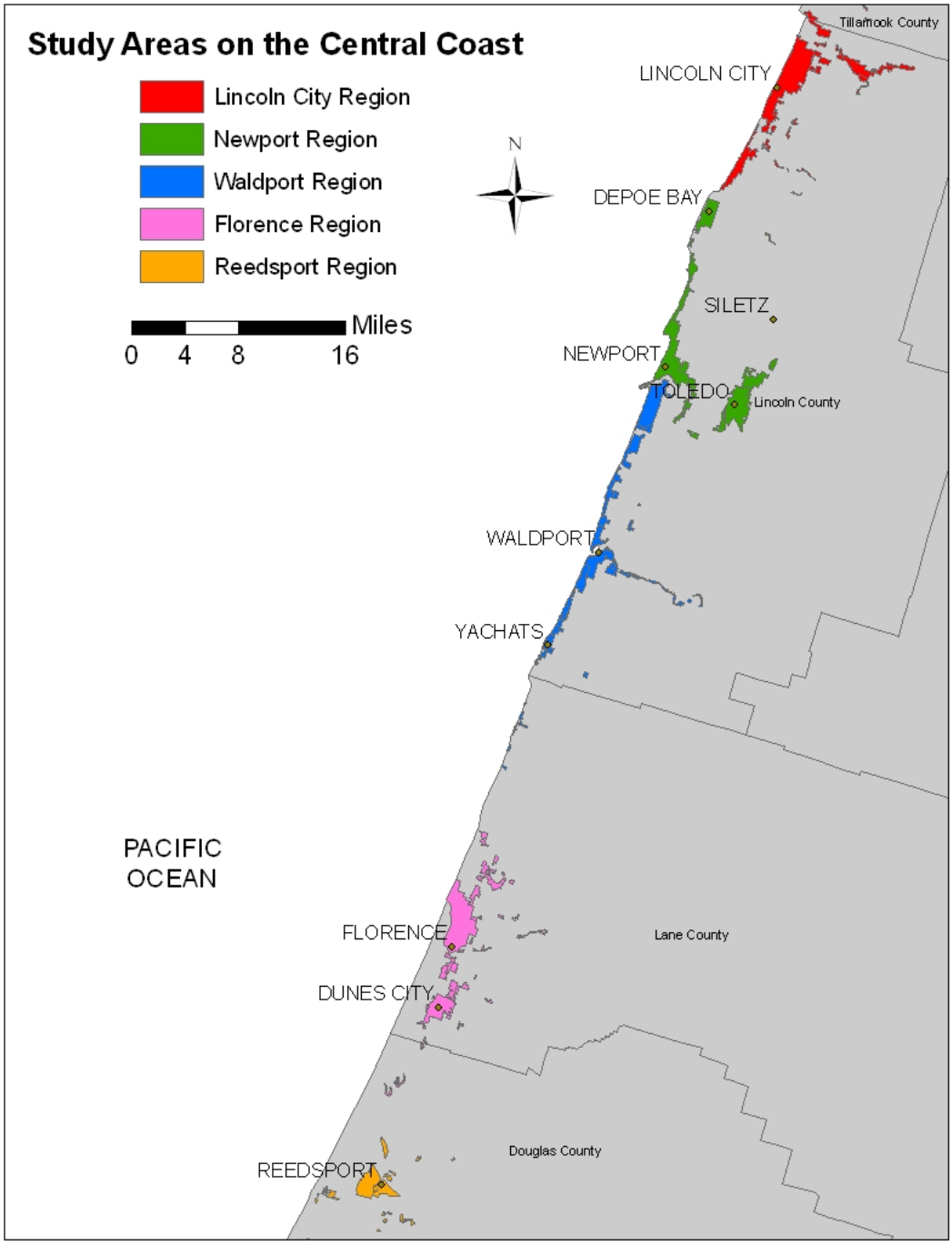
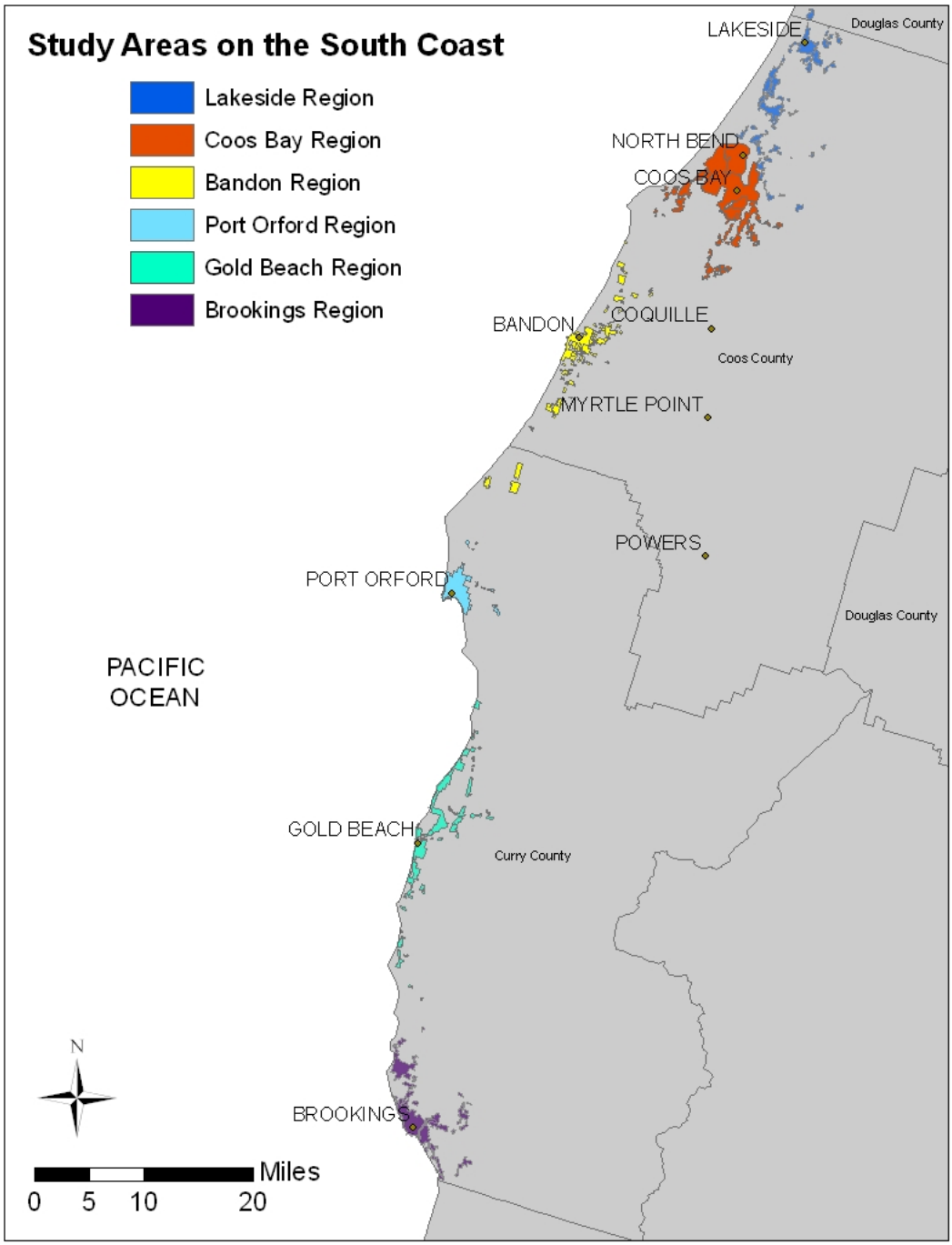


Figure 4. Areas zoned for development for study regions on the south coast.



Projections of future development were produced using the parameter estimates from the Bayesian spatial probit for three different regions, Astoria, Waldport and Newport. The NOAA dataset provides the actual land-use situation in 2001, at the end of the observed five year period. We use the land that is still undeveloped in 2001 as the starting point for the five year period from 2001 to 2006. The measurement of proximity to existing development must be recalculated given the new land-use pattern and a new spatial weights matrix must also be created for the new dataset. The other physical characteristics of the landscape, like slope, do not change over time.

The Bayesian spatial probit algorithm provides parameter estimates for the model $z = \rho Wz + x\mathrm{B} + \varepsilon$ where z is the estimated latent variable. These parameter estimates can be used to find transition probabilities in order to predict land-use changes in the future. Rearranging the model equation gives us $z = (I - \rho W)^{-1} x\mathrm{B} + (I - \rho W)^{-1}\varepsilon$ where ε is distributed normally with mean 0 and variance $\sigma^2 V$. Then z is distributed normally with mean $E(z) = (I - \rho W)^{-1} x\mathrm{B}$ and variance $E(zz') = (I - \rho W)^{-1}\sigma^2 V\left[(I - \rho W)^{-1}\right]'$. V is a diagonal matrix with non-constant variances, v_{ii}^2, on the diagonal and zeroes elsewhere. If we define P such that $P'P = V^{-1}$, then P is an n x n matrix with $\frac{1}{v_{ii}}$ on the diagonal for each observation i and zeroes off the diagonal. The standard error of z is $\sigma(I - \rho W)^{-1} P^{-1}$.

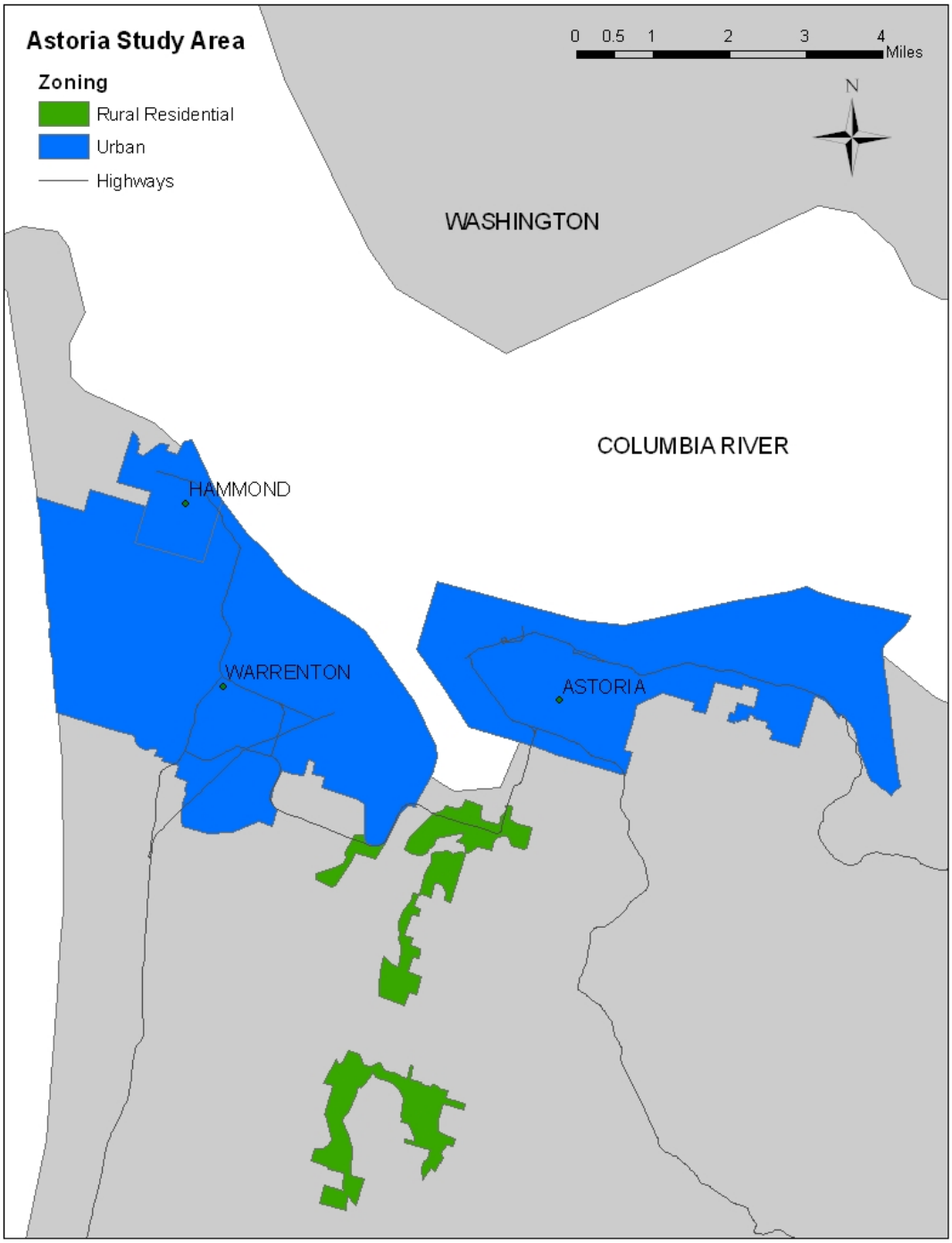
Estimates for the variance terms, σ^2 and the diagonal elements of V, are computed within the Gibb's sampling algorithm, so we can estimate the standard error of z using s in place of σ and $\hat{v}_{ii}$ to estimate P. The estimated standard error $s(I - \hat{\rho}W)^{-1}\hat{P}^{-1}$ can be used to transform the fitted value $\hat{z} = (I - \hat{\rho}W)^{-1} x\hat{\mathrm{B}}$ to a standard normal random variable, denoted $\hat{z}_*$. Then the expected probability of

development is $p = F\left(\hat{z}_*\right)$ where F is the standard normal cumulative distribution function.

The expected probability of development was calculated for each pixel and those pixels with the highest transition probabilities were assumed to convert to a developed state. Two different scenarios were considered here. In the low growth scenario, the level of development in the projection period, from 2001 to 2006, is assumed to be the same as the observed period, 1996 to 2001. Thus, roughly the same number of pixels will change in the projected period as in the observed dataset. In the high growth scenario, the rate of development is assumed to increase by 50 percent, so 50 percent more pixels will change in the projected period than in the observed dataset.

The Astoria region consists of three urban areas, Astoria, Warrenton and Hammond, and several rural residential areas, as shown in figure 5. In the observed dataset, 215 pixels were developed during the five year period from 1996 to 2001. In the low growth projection scenario, 215 pixels will be developed from 2001 to 2006 while in the high growth scenario, roughly 50 percent more for a total of 322 pixels will change.

Figure 5. Land zoned for development in the Astoria region.



The Waldport study area contains the cities of Waldport and Yachats and the portion of Newport that lies south of Newport Bay, as well as various areas zoned for rural development. A map of the region showing areas that are zoned for development can be found in figure 6. Over the observed time period from 1996 to 2001, 298 pixels were developed in the region. So for the low growth scenario, the same number of pixels, 298, is projected to convert into a developed use from 2001 to 2006. In the high growth scenario, 444 pixels are projected to change.

The Newport study area contains three urban areas, Depoe Bay, Toledo, and the portion of Newport that lies north of Yaquina Bay, as well as land zoned for various types of rural development. A map of the region showing areas that are zoned for development is shown in figure 7. From 1996 to 2001, 189 pixels were converted to a developed use in the Newport region. For the low growth scenario, 189 pixels will change from 2001 to 2006 while in the high growth scenario, 270 pixels will change.

Figure 6. Land zoned for development in the Waldport region.

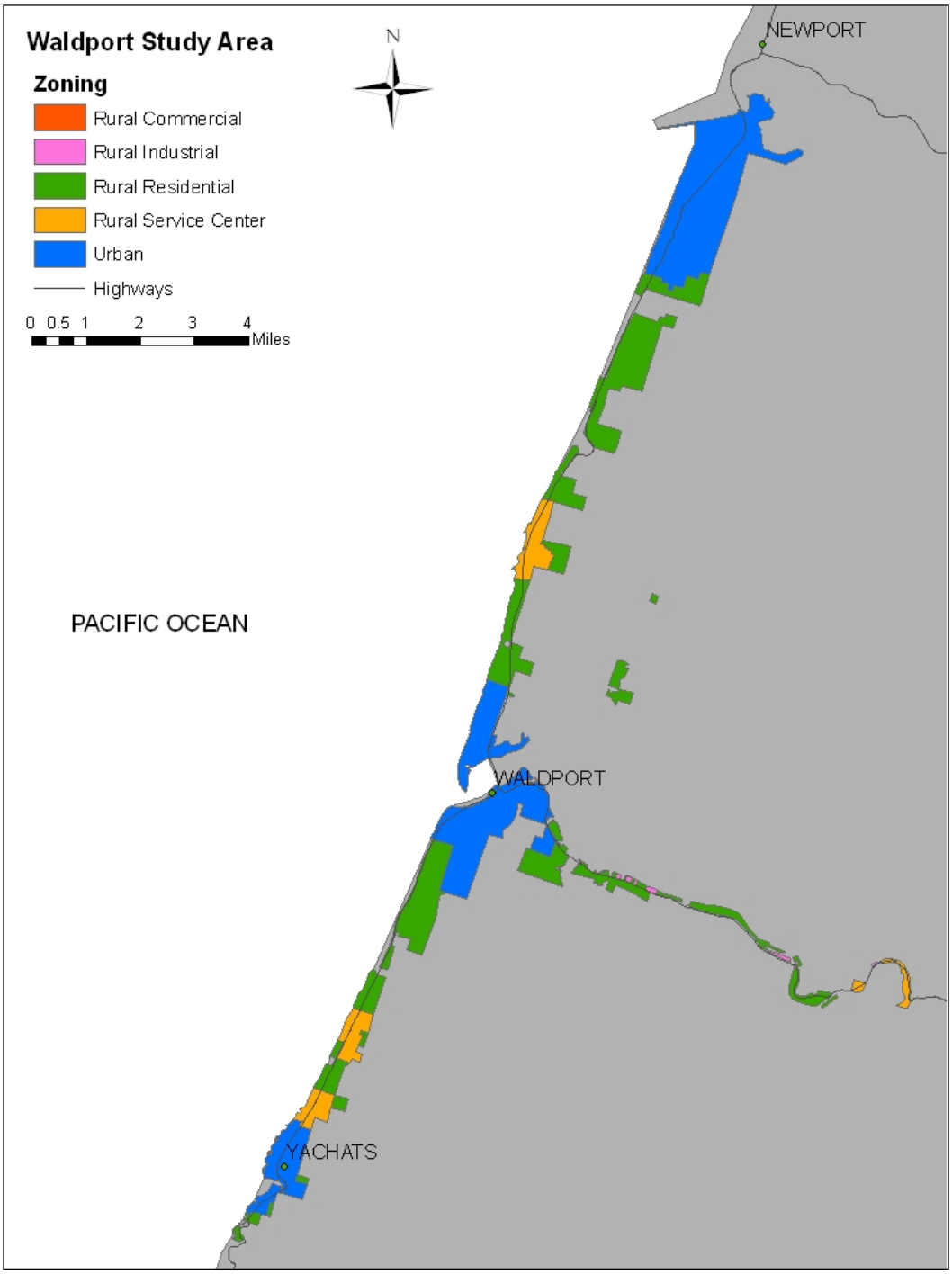
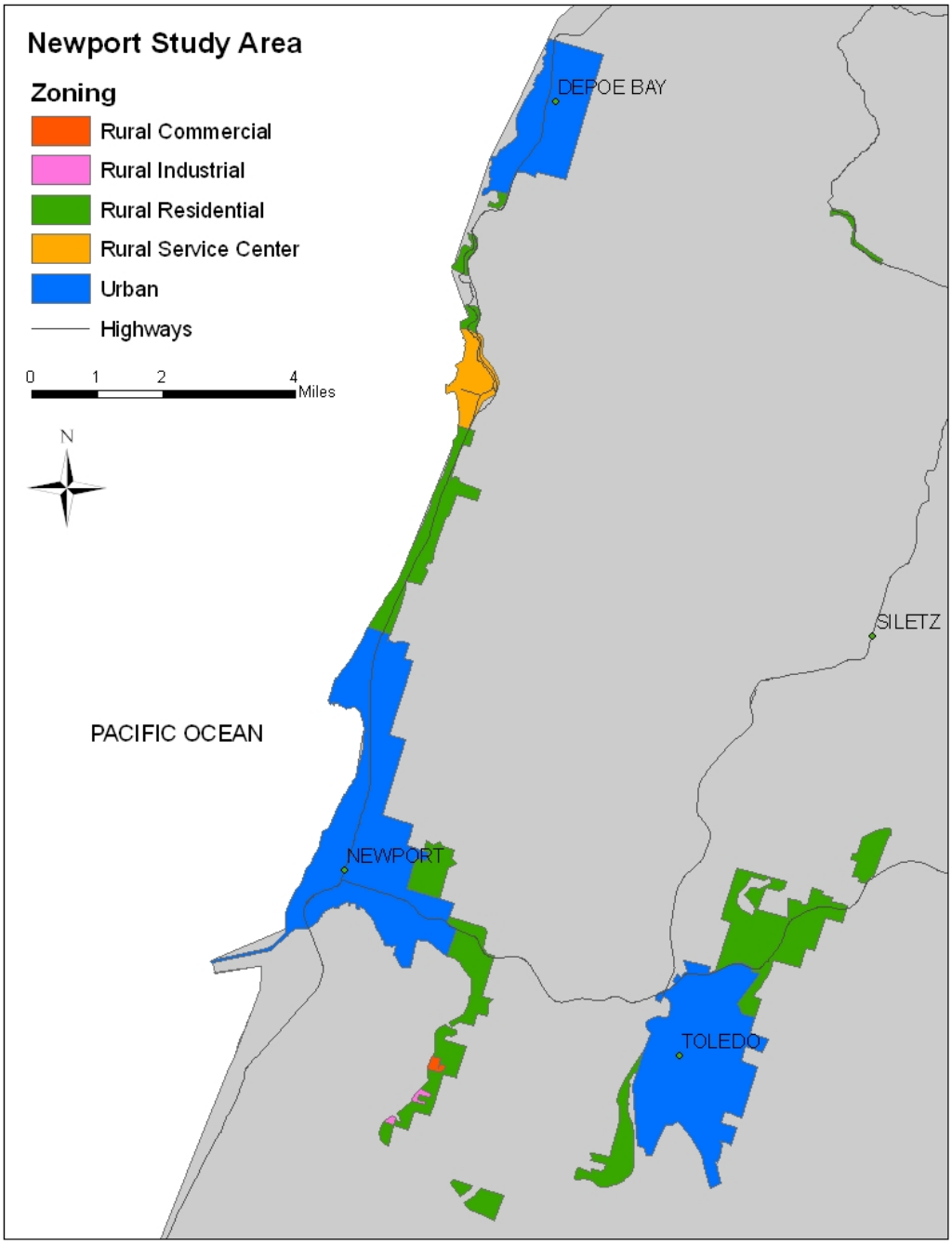


Figure 7. Land zoned for development in the Newport region.



4. Results

4.1 Regression

Table 2 includes some summary statistics of the datasets for each region. There are between 24,000 and 35,000 observations for most regions, however, Reedsport and Port Orford are significantly smaller, with 8,587 and 15,595 observations respectively. Each observation is a 900 square meter pixel, which is about 0.22 acres of land. The Brookings region saw the most development over the observed period, with 753 pixels changing, compared to an average of 183 pixels per region. Reedsport, Port Orford and Gold Beach saw little change from 1996 to 2001, with between 15 and 31 pixels becoming developed in each region. This adds up to less than 7 acres of new development in each of these three regions.

On average, about 44 percent of the land available for development in each region falls under urban zoning but this varies widely across regions. A little over 12 percent of available land in the Lakeside region is zoned for urban development, while more than 82 percent of available land in the Astoria region is urban. Lakeside also has the largest average distance to existing development, at 2,850 feet and Astoria has the smallest average measure of proximity to existing development, coming in at 1,050 feet. At first glance this may suggest an inverse relationship between urban zoning and proximity to development but this relationship does not hold in the rest of the regions.

Astoria is by far the furthest, on average, from the ocean, because most of the region is situated slightly inland along the Columbia River. The average distance to the ocean for pixels in the Astoria region is 19,400 feet. The available land in Waldport is closest to the ocean, with pixels located an average of 2,690 feet from the water. The average across all regions is 6,700 feet.

Table 2. Summary statistics for regional datasets (before centering).

Region	Categorical Variables				Continuous Variables			
	Number of Pixels				Means of Pixel Characteristics			
	Total Pixels	Developed from 1996 to 2001	WEST = 1	URBAN = 1	HIGHWAY (hundreds of feet)	OCEAN (hundreds of feet)	SLOPE (degrees)	DEVPROX (hundreds of feet)
Astoria	31,454	215	6,365	25,893	10.7	194.4	3.7	10.5
Bandon	27,982	141	10,694	4,749	9.3	73.2	3.4	16.3
Brookings	25,080	753	12,030	4,717	17.0	70.6	13.3	23.3
Cannon Beach	31,283	185	14,481	14,190	9.8	48.9	8.7	24.0
Florence	30,792	195	10,322	19,170	10.5	84.1	5.8	21.9
Gold Beach	24,787	25	11,703	5,210	14.2	33.1	11.8	24.8
Lakeside	28,491	145	11,453	3,467	16.2	96.5	10.9	28.5
Lincoln City	34,439	209	14,969	16,368	7.7	78.2	8.9	17.6
Newport	31,922	189	15,547	18,829	9.5	57.8	11.2	17.5
Port Orford	15,595	31	6,002	12,968	8.9	52.4	7.4	21.3
Reedsport	8,587	15	2,402	6,335	8.7	41.7	12.7	25.4
Seaside	26,067	214	5,181	8,935	5.0	47.1	2.1	11.4
Tillamook	28,934	128	11,081	10,643	12.6	100.0	8.6	22.4
Waldport	30,292	296	16,007	12,957	4.9	26.9	5.8	12.0
Average	25,047	183	9,882	10,962	9.7	67.0	7.6	18.5

Regression results for the spatial probit and the standard non-spatial probit are shown in table 3. Standard probit results were included for comparison to the spatial probit specification.

The WEST parameter was significant in only six of the regional models. In Port Orford, Seaside and Waldport, western facing slopes had a significant and positive effect on development, however in Lakeside, Newport and Reedsport, this dummy variable had a significant and negative impact on development.

The coefficient on HIGHWAY was significant and positive in eight of the regions indicating that more development occurs as distance from the highway increases. However three regions have coefficients that are significant and negative. In these regions, the increased costs of accessing a piece of land may outweigh the benefits of building further from the highway. Given the shape of the study areas, ease of access may not be a serious consideration for development. Areas that are zoned for development generally fall along the major highways, which can be seen in the detailed study area maps in figures 5, 6 and 7. So the land that is available for development is rarely very far from the existing highway.

Distance from the ocean was insignificant in all but two regions. The parameter estimate for OCEAN was significant and positive in Cannon Beach and negative in Brookings.

Urban zoning had a positive and significant impact on development in most regions. However, the coefficient on URBAN for Gold Beach was negative and significant, indicating that during the period 1995 to 2001, more development was occurring in rural areas than urban. This parameter also had an insignificant coefficient in two models.

The coefficient on slope was significant and negative in seven models, indicating development potential decreases as slope increases, as expected. SLOPE is positive and significant in Port Orford and insignificant elsewhere.

The measure of proximity to existing development was generally negative across all of the regions, but is significantly different from zero in only five models. The coefficient on DEVPROX is significant and negative in Astoria, Bandon, Cannon Beach, Port Orford and Waldport, indicating that more development occurs in close proximity to existing development. In other regions, this parameter is insignificant.

Finally, the spatial parameter RHOLAG is positive and significant in all regions. This coefficient is roughly a correlation measure and ranges in magnitude from 0.22 in Lakeside to 0.68 in Port Orford.

The standard probit results follow the same general pattern as the spatial probit estimates, however the magnitudes of the effects are very different. For parameter estimates that are considered significant in both models, the signs always match with one exception. The effect of HIGHWAY is significant and positive in the spatial model for Port Orford and significant and negative in the standard probit. The magnitudes of the standard probit estimates are consistently larger across all regions and parameters.

The marginal effects and variances shown in table 3 were calculated at each data point using the delta method and then averaged. Though this is not the best way to find marginal effects for dummy variables (WEST and URBAN), Greene (2003) indicates that this method often provides an accurate approximation of the marginal effect of a binary variable.

Table 3. Regression results for spatial probit and standard probit models.

Region: Astoria	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0696	(-0.0722, -0.0655)			-3.3055	(-3.4677, -3.1433)		
WEST	-0.0001	(-0.0015, 0.0038)	-1.80E-10	4.27E-06	-0.4857	(-0.7468, -0.2247)	-0.00781	4.28E-03
HIGHWAY	0.0004	(0.0004, 0.0005)	5.74E-10	9.43E-07	-0.0068	(-0.0212, 0.0075)	-0.00011	2.12E-04
OCEAN	0.0000	(0, 0)	1.40E-11	4.01E-08	0.0019	(0.0014, 0.0024)	0.00003	1.02E-05
URBAN	0.0150	(0.0136, 0.0180)	2.02E-08	3.30E-05	0.8053	(0.4198, 1.1907)	0.01295	6.44E-03
SLOPE	-0.0008	(-0.0009, -0.0005)	-1.04E-09	1.71E-06	-0.0999	(-0.1241, -0.0757)	-0.00160	5.19E-04
DEVPROX	-0.0008	(-0.0011, -0.0002)	-1.13E-09	1.98E-06	-0.1072	(-0.1265, -0.0878)	-0.00172	4.97E-04
RHOLAG	0.2780	(0.2756, 0.2855)						

Region: Bandon	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0583	(-0.0628, -0.0500)			-4.2630	(-4.6269, -3.8991)		
WEST	0.0007	(-0.0013, 0.0043)	6.87E-09	5.18E-05	-0.1435	(-0.3213, 0.0343)	-0.00179	1.66E-04
HIGHWAY	0.0002	(0.0002, 0.0003)	2.15E-09	4.56E-06	0.0198	(0.0083, 0.0314)	0.00025	2.94E-05
OCEAN	0.0001	(0, 0.0001)	6.16E-10	1.41E-06	0.0018	(-0.0005, 0.0041)	0.00002	2.28E-03
URBAN	0.0055	(0.0027, 0.0100)	5.64E-08	1.33E-04	-0.0208	(-0.2069, 0.1652)	-0.00026	5.01E-04
SLOPE	-0.0003	(-0.0004, 0.0002)	-2.62E-09	7.40E-06	-0.0583	(-0.0936, -0.0231)	-0.00073	8.06E-04
DEVPROX	-0.0002	(-0.0002, -0.0001)	-1.84E-09	3.95E-06	-0.1659	(-0.197, -0.1347)	-0.00207	0.00E+00
RHOLAG	0.4123	(0.4006, 0.4349)						

Table 3. Continued.

Region: Brookings	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.1262	(-0.1283, -0.1204)			-4.4293	(-4.8795, -3.9791)		
WEST	-0.0027	(-0.0050, 0.0054)	-2.24E-06	5.28E-03	-0.0521	(-0.1375, 0.0334)	-0.00246	4.77E-04
HIGHWAY	0.0006	(0.0004, 0.0008)	5.04E-07	8.18E-04	-0.0008	(-0.0100, 0.0084)	-0.00004	1.60E-04
OCEAN	-0.0002	(-0.0002, -0.0001)	-1.51E-07	2.45E-04	-0.0030	(-0.0061, 0.0001)	-0.00014	9.69E-03
URBAN	0.0776	(0.0728, 0.0854)	6.40E-05	9.96E-02	1.3095	(1.1900, 1.4290)	0.06178	3.39E-04
SLOPE	-0.0021	(-0.0022, -0.0015)	-1.69E-06	2.65E-03	-0.0269	(-0.0328, -0.0210)	-0.00127	1.36E-03
DEVPROX	-0.0003	(-0.0004, 0.0001)	-2.59E-07	4.54E-04	-0.1015	(-0.1257, -0.0773)	-0.00479	0.00E+00
RHOLAG	0.3227	(0.2984, 0.3627)						

Region: Cannon Beach	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0628	(-0.0663, -0.0566)			-8.6076	(-3902.34, 3885.12)		
WEST	0.0000	(-0.0014, 0.0043)	-1.40E-09	1.58E-04	-0.0708	(-0.2020, 0.0604)	-0.00100	0.00E+00
HIGHWAY	-0.0002	(-0.0003, 0)	-6.74E-09	1.94E-05	0.0306	(0.0161, 0.0451)	0.00043	0.00E+00
OCEAN	0.0001	(0.0001, 0.0002)	3.70E-09	1.02E-05	0.0044	(-0.0017, 0.0104)	0.00006	1.24E+01
URBAN	0.0070	(0.0046, 0.0100)	2.12E-07	5.77E-04	6.3122	(-7204.30, 7216.93)	0.10090	0.00E+00
SLOPE	-0.0002	(-0.0003, -0.0001)	-6.70E-09	1.82E-05	-0.0302	(-0.0444, -0.0160)	-0.00043	0.00E+00
DEVPROX	-0.0002	(-0.0002, -0.0001)	-5.52E-09	1.51E-05	-0.1569	(-0.1948, -0.1189)	-0.00222	0.00E+00
RHOLAG	0.3757	(0.3671, 0.3928)						

Table 3. Continued.

Region: Florence	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0677	(-0.0703, -0.0597)			-7.3570	(-3399.69, 3384.98)		
WEST	0.0002	(-0.0018, 0.0055)	3.95E-08	7.92E-04	0.1222	(-0.0252, 0.2696)	0.00184	2.50E-04
HIGHWAY	-0.0003	(-0.0004, -0.0002)	-5.58E-08	1.13E-04	-0.0391	(-0.0512, -0.0270)	-0.00059	4.27E-05
OCEAN	-0.0001	(-0.0001, 0)	-8.84E-09	1.78E-05	-0.0045	(-0.0070, -0.0021)	-0.00007	8.36E-01
URBAN	0.0097	(0.0075, 0.0126)	1.63E-06	3.28E-03	5.4094	(-9163.06, 9173.883)	3.33410	2.79E-04
SLOPE	-0.0001	(-0.0002, 0.0001)	-2.10E-08	7.28E-05	0.0091	(-0.0091, 0.0272)	0.00014	8.92E-04
DEVPROX	-0.0001	(-0.0001, 0)	-8.67E-09	1.77E-05	-0.1598	(-0.1982, -0.1214)	-0.00240	0.00E+00
RHOLAG	0.3999	(0.3917, 0.4166)						

Region: Gold Beach	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0324	(-0.0337, -0.0302)			-4.7564	(-5.8592, -3.6536)		
WEST	-0.0003	(-0.0009, 0.0017)	-6.88E-11	6.15E-07	0.1380	(-0.1367, 0.4127)	0.00041	4.71E-05
HIGHWAY	0.0001	(0.0001, 0.0002)	2.08E-11	7.83E-08	0.0127	(0.0048, 0.0205)	0.00004	3.64E-05
OCEAN	-0.0001	(-0.0001, 0)	-1.44E-11	5.69E-08	-0.0063	(-0.0156, 0.0029)	-0.00002	1.16E-03
URBAN	-0.0041	(-0.0048, -0.0020)	-8.78E-10	3.11E-06	-0.0432	(-0.3834, 0.2970)	-0.00013	8.26E-05
SLOPE	-0.0001	(-0.0002, -0.0001)	-3.11E-11	1.12E-07	-0.0133	(-0.0350, 0.0083)	-0.00004	3.23E-04
DEVPROX	0.0000	(0, 0)	-5.25E-12	2.69E-08	-0.0882	(-0.1437, -0.0327)	-0.00026	0.00E+00
RHOLAG	0.3215	(0.3145, 0.3404)						

Table 3. Continued.

Region: Lakeside	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0605	(-0.0609, -0.0594)			-6.9201	(-7.9445, -5.8956)		
WEST	-0.0052	(-0.0067, -0.0024)	-5.24E-08	1.11E-04	-0.1536	(-0.3757, 0.0684)	-0.00162	2.86E-04
HIGHWAY	0.0001	(0.0001, 0.0002)	1.35E-09	3.33E-06	0.0095	(-0.0068, 0.0257)	0.00010	7.43E-05
OCEAN	0.0000	(0, 0)	3.20E-10	6.56E-07	0.0050	(0.0010, 0.0091)	0.00005	4.03E-03
URBAN	0.0243	(0.0238, 0.0279)	2.44E-07	4.98E-04	0.4008	(0.1994, 0.6022)	0.00422	6.83E-04
SLOPE	-0.0006	(-0.0008, -0.0002)	-6.12E-09	1.28E-05	-0.0860	(-0.1173, -0.0547)	-0.00090	1.05E-03
DEVPROX	0.0000	(-0.0001, 0)	-3.41E-10	8.79E-07	-0.1586	(-0.1989, -0.1182)	-0.00167	0.00E+00
RHOLAG	0.2315	(0.2240, 0.2458)						

Region: Lincoln City	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0600	(-0.0621, -0.0557)			-4.0432	(-4.3898, -3.6967)		
WEST	-0.0034	(-0.0069, 0.0011)	-7.17E-08	2.19E-04	-0.0790	(-0.1971, 0.0391)	-0.00117	2.12E-04
HIGHWAY	-0.0002	(-0.0003, 0)	-4.64E-09	1.11E-05	-0.0172	(-0.0316, -0.0028)	-0.00025	5.60E-05
OCEAN	0.0000	(0, 0)	-5.97E-10	1.33E-06	-0.0065	(-0.0101, -0.0030)	-0.00010	4.13E-03
URBAN	0.0064	(0.0056, 0.0088)	1.36E-07	2.94E-04	0.7630	(0.5571, 0.9690)	0.01127	1.50E-04
SLOPE	0.0001	(0, 0.0003)	1.95E-09	5.62E-06	-0.0055	(-0.0162, 0.0051)	-0.00008	4.82E-04
DEVPROX	-0.0001	(-0.0002, 0.0001)	-1.58E-09	5.49E-06	-0.0813	(-0.1077, -0.0550)	-0.00120	0.00E+00
RHOLAG	0.3436	(0.3401, 0.3509)						

Table 3. Continued.

Region: Newport	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0599	(-0.0651, -0.0505)			-7.1504	(-3941.12, 3926.82)		
WEST	-0.0023	(-0.0036, -0.0002)	-1.99E-07	8.19E-04	0.0991	(-0.0233, 0.2216)	0.00138	0.00E+00
HIGHWAY	-0.0002	(-0.0002, -0.0001)	-1.41E-08	5.30E-05	0.0043	(-0.0027, 0.0114)	0.00006	0.00E+00
OCEAN	0.0000	(0, 0.0001)	2.29E-09	9.42E-06	-0.0003	(-0.0038, 0.0033)	0.00000	7.31E+00
URBAN	0.0024	(0.0011, 0.0057)	2.11E-07	9.95E-04	5.0287	(-9590.01, 9600.07)	0.07531	0.00E+00
SLOPE	-0.0001	(-0.0003, 0.0002)	-1.14E-08	6.75E-05	-0.0390	(-0.0520, -0.0260)	-0.00054	0.00E+00
DEVPROX	-0.0001	(-0.0001, 0)	-8.04E-09	3.39E-05	-0.1925	(-0.2348, -0.1502)	-0.00268	0.00E+00
RHOLAG	0.4875	(0.4851, 0.4953)						

Region: Port Orford	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0305	(-0.0368, -0.0205)			-5.6608	(-7.0292, -4.2924)		
WEST	0.0032	(0.0025, 0.0064)	2.44E-06	5.47E-03	0.2074	(-0.1095, 0.5243)	0.00094	2.54E-03
HIGHWAY	0.0003	(0.0002, 0.0005)	1.91E-07	4.41E-04	-0.3233	(-0.4954, -0.1512)	-0.00147	8.22E-05
OCEAN	0.0000	(0, 0.0001)	5.55E-09	4.01E-05	0.0056	(-0.0022, 0.0134)	0.00003	7.80E-03
URBAN	-0.0029	(-0.0041, 0.0008)	-2.24E-06	5.20E-03	-0.9599	(-1.4754, -0.4445)	-0.00437	3.07E-04
SLOPE	0.0002	(0.0001, 0.0004)	1.25E-07	3.07E-04	0.0433	(0.0263, 0.0602)	0.00020	3.47E-04
DEVPROX	-0.0002	(-0.0002, -0.0001)	-1.43E-07	2.93E-04	-0.0357	(-0.0632, -0.0082)	-0.00016	0.00E+00
RHOLAG	0.6819	(0.6542, 0.7160)						

Table 3. Continued.

Region: Reedsport	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0435	(-0.0460, -0.0389)			-32.0732	(-46478, 46413.8)		
WEST	-0.0063	(-0.0073, -0.0038)	-2.34E-09	4.66E-06	-6.9703	(-99402, 99388.08)	-0.04647	6.65E-03
HIGHWAY	0.0002	(0.0001, 0.0004)	5.92E-11	1.48E-07	0.7976	(0.3450, 1.2501)	0.00096	2.88E-03
OCEAN	-0.0001	(-0.0001, 0)	-3.25E-11	6.78E-08	-0.1667	(-0.4563, 0.1229)	-0.00020	4.72E-01
URBAN	-0.0060	(-0.0073, 0.0001)	-2.23E-09	4.63E-06	-9.4664	(-51938.2, 51919.23)	-0.07105	2.25E-03
SLOPE	0.0000	(-0.0001, 0)	-9.83E-12	3.83E-08	-0.2405	(-0.4297, -0.0512)	-0.00029	9.13E-03
DEVPROX	-0.0001	(-0.0001, 0)	-3.08E-11	6.29E-08	-0.7084	(-1.5628, 0.1460)	-0.00085	0.00E+00
RHOLAG	0.2194	(0.2027, 0.2494)						

Region: Seaside	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0747	(-0.0757, -0.0723)			-5.4722	(-5.983, -4.9614)		
WEST	0.0096	(0.0075, 0.0143)	3.44E-08	5.93E-05	0.2273	(0.0571, 0.3974)	0.00386	4.97E-04
HIGHWAY	-0.0006	(-0.0009, -0.0001)	-2.20E-09	3.98E-06	-0.0410	(-0.0619, -0.02)	-0.00070	1.04E-04
OCEAN	0.0000	(0, 0)	-4.39E-11	1.46E-07	0.0105	(0.0065, 0.0146)	0.00018	1.42E-02
URBAN	0.0197	(0.0148, 0.0263)	7.05E-08	1.19E-04	1.7992	(1.2986, 2.2998)	0.03053	4.87E-04
SLOPE	-0.0001	(-0.0005, 0.0008)	-5.26E-10	2.58E-06	-0.0070	(-0.03, 0.0159)	-0.00012	1.95E-03
DEVPROX	-0.0001	(-0.0002, 0)	-3.86E-10	7.58E-07	-0.3112	(-0.3612, -0.2611)	-0.00528	0.00E+00
RHOLAG	0.2630	(0.2471, 0.3124)						

Table 3. Continued.

Region: Tillamook	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0597	(-0.0608, -0.0575)			-4.0596	(-4.4368, -3.6824)		
WEST	-0.0080	(-0.0108, 0.0001)	-5.90E-07	2.33E-03	-0.4437	(-0.6605, -0.2269)	-0.00492	1.45E-04
HIGHWAY	0.0000	(-0.0001, 0.0002)	-1.89E-09	3.11E-05	-0.0089	(-0.0218, 0.0041)	-0.00010	1.27E-05
OCEAN	0.0000	(0, 0)	-1.16E-09	4.66E-06	0.0007	(-0.0005, 0.0018)	0.00001	1.88E-03
URBAN	0.0101	(0.0070, 0.0158)	7.47E-07	2.87E-03	-0.0116	(-0.1854, 0.1622)	-0.00013	2.08E-04
SLOPE	-0.0003	(-0.0004, -0.0002)	-2.21E-08	7.94E-05	-0.0283	(-0.0449, -0.0118)	-0.00031	3.83E-04
DEVPROX	0.0000	(-0.0001, 0.0001)	-2.44E-09	1.73E-05	-0.0815	(-0.1037, -0.0593)	-0.00090	0.00E+00
RHOLAG	0.2518	(0.2446, 0.2742)						

Region: Waldport	Spatial Probit				Non-spatial Probit			
Variable	Parameter Estimate	95% Highest Posterior Density	Marginal Effect (ME)	ME Standard Error	Parameter Estimate	95% Confidence Interval	Marginal Effect (ME)	ME Standard Error
INTERCEPT	-0.0710	(-0.0772, -0.0589)			-5.1951	(-5.6194, -4.7708)		
WEST	0.0029	(0.0012, 0.0071)	5.31E-07	1.23E-03	0.1040	(-0.0459, 0.2538)	0.00190	7.25E-04
HIGHWAY	0.0012	(0.0011, 0.0015)	2.26E-07	4.12E-04	0.0877	(0.0652, 0.1102)	0.00160	1.92E-04
OCEAN	-0.0001	(-0.0001, 0)	-9.65E-09	1.83E-05	-0.0222	(-0.0283, -0.0161)	-0.00041	1.54E-02
URBAN	0.0143	(0.0094, 0.0205)	2.64E-06	5.00E-03	1.7720	(1.2760, 2.2680)	0.03235	1.33E-03
SLOPE	-0.0010	(-0.0012, -0.0003)	-1.76E-07	3.37E-04	-0.1771	(-0.2161, -0.1382)	-0.00323	1.27E-03
DEVPROX	-0.0003	(-0.0004, -0.0002)	-5.70E-08	1.04E-04	-0.1948	(-0.2277, -0.1620)	-0.00356	0.00E+00
RHOLAG	0.5350	(0.5298, 0.5456)						

4.2 Projections

Projection results are presented for three regions, Astoria, Waldport and Newport. In order to demonstrate some examples of how this projection model might be used, the analysis shown here is slightly different for each region. For all three, we look at the relationship between zoning and projected development as well as a comparison of the low and high growth scenarios. For the Astoria region, we focus on the differences between the spatial probit projection and a standard probit projection. In the Waldport region, we again compare spatial and standard probit results and look at projected development in areas prone to erosion and tsunami hazards. For Newport, we relate the development projections back to the regression model by looking at a visual comparison of development to slope.

Figure 8 is a map of the Astoria study area showing the general zoning categories that were included in the model (urban and rural) as well as projected development for the entire region under the high growth scenario. It is clear from this map that most of the projected development is expected to occur in urban areas. This map also indicates that a large cluster of land at the northern edge of Hammond is projected to develop all at once.

The map in figure 9 shows the historical and projected changes in the urban areas of the region in more detail. The cluster of development in the north seems to follow the pattern of existing development, which has grown up in large clusters along the Columbia River in Hammond. Smaller clumps of land are expected to convert to a developed use in Astoria and Warrenton, with only a few pixels changing in the rural areas. Figure 10 compares the low growth projection based on the spatial probit to a projection of the same level of development from the non-spatial probit in the Astoria region. Both projections show a clustered pattern of development, however the clusters are in different locations. The spatial probit

projects more development occurring in Hammond and Warrenton than in Astoria while the non-spatial probit predicts the opposite, with more development in occurring in Astoria than elsewhere in the region.

Figure 8. Projected development in the Astoria region under the high growth scenario.

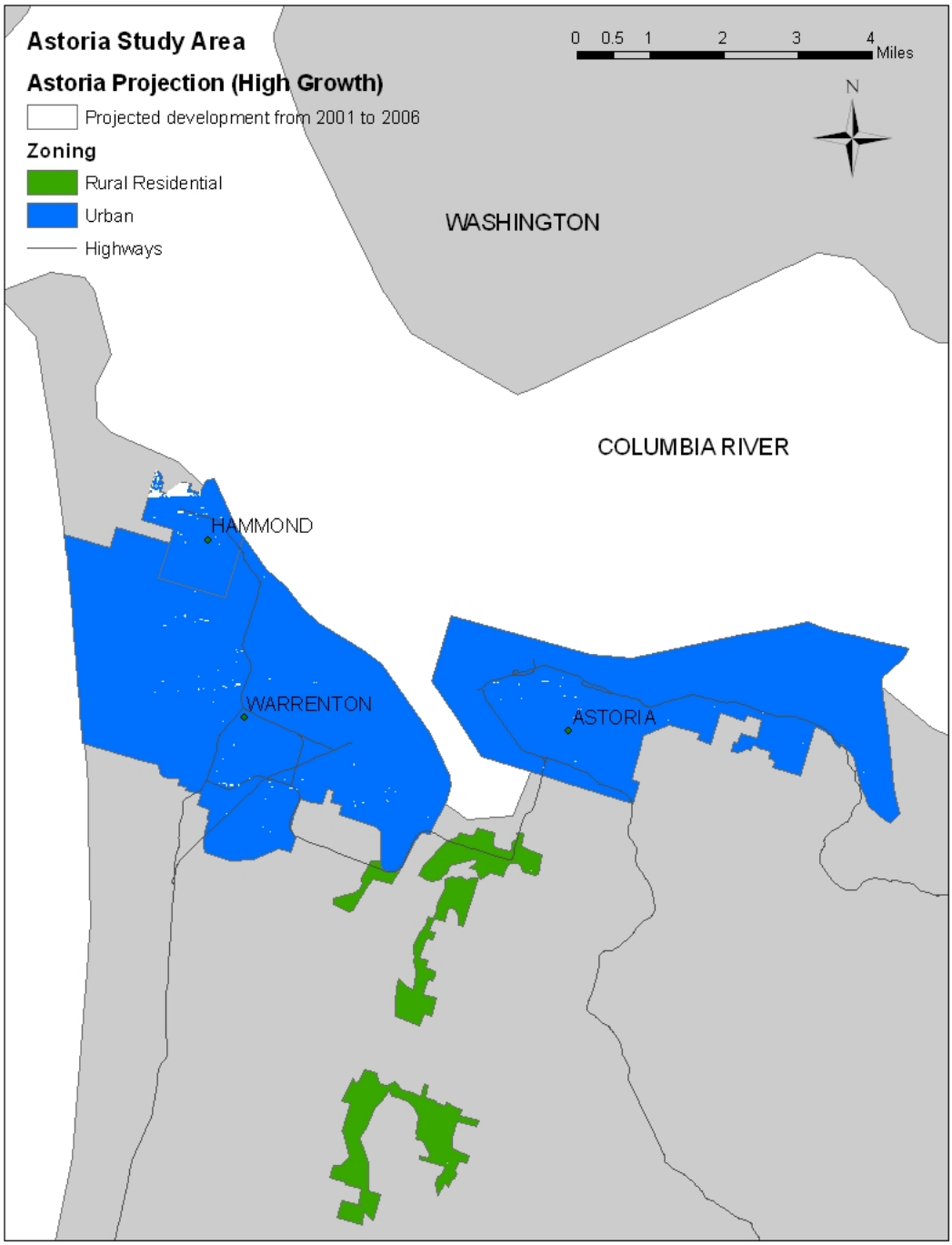


Figure 9. Projected and existing development in urban areas of Astoria region under low and high growth scenarios.

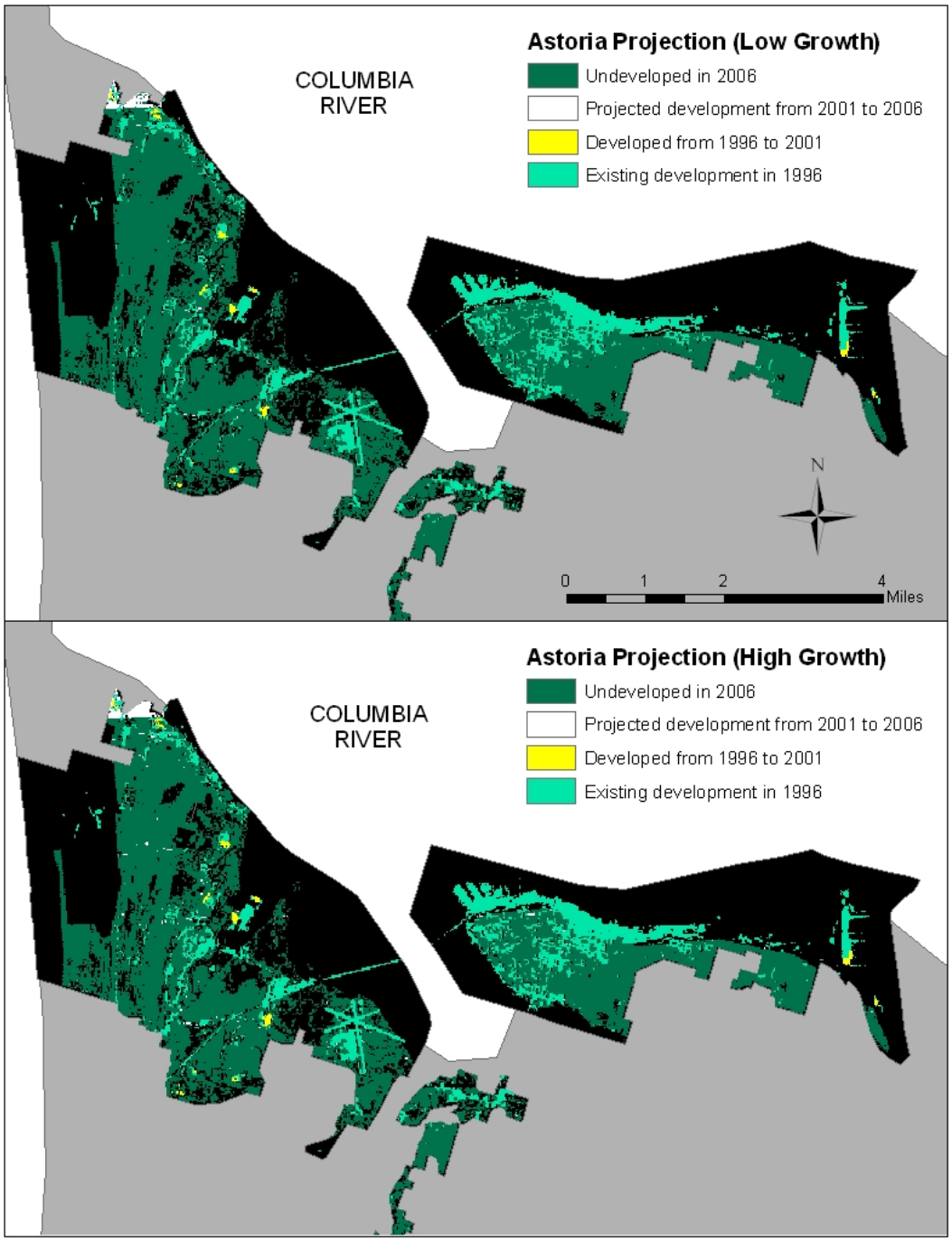


Figure 10. Comparison of spatial probit and standard probit projection results under low growth scenario in the Astoria region.

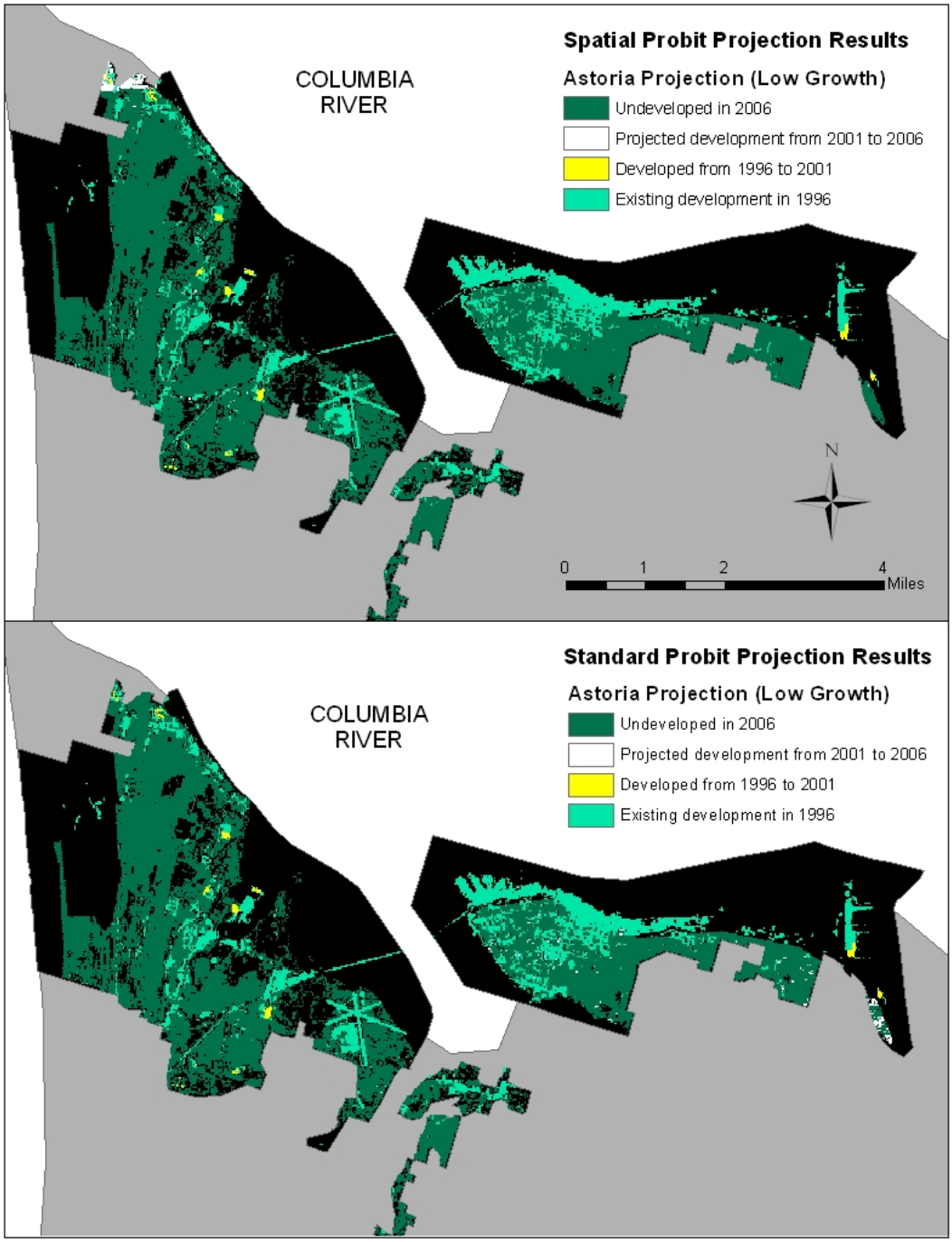


Figure 11 shows the projected development for the Waldport region under the high growth scenario. This map clearly depicts the strong relationship between zoning and development, with most future development occurring in areas that are zoned for urban use. The detail maps of the City of Waldport in figure 12 depict the connection between projected development and historical development in the city. These maps show that most of the development that was observed from 1996 to 2001 occurred in large patches adjacent to existing development. Future development is projected to follow a similar pattern under both the low and high growth scenarios, filling in gaps in the existing development as well as forming along the outer edges.

Figure 13 compares projections from the spatial model to the standard non-spatial probit. The standard probit projections seem more clustered than the spatial probit projections. Although it is not immediately clear from the map in figure 13, the standard probit model has projected far more development within the City of Waldport and less throughout the rest of the study area, compared with the spatial model projection.

Figure 11. Projected development under the high growth scenario for the Waldport region.

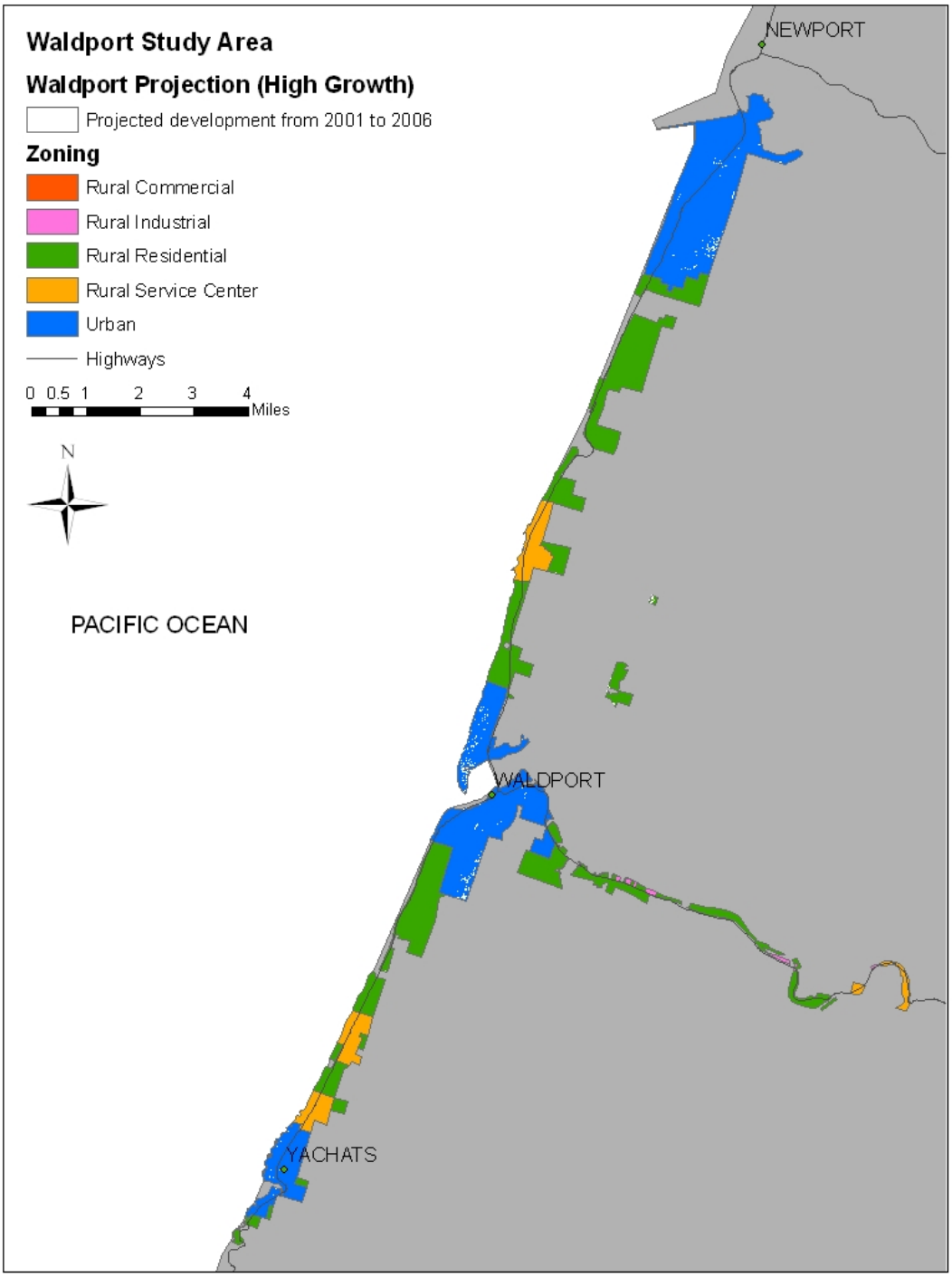


Figure 12. Projected and existing development in the City of Waldport.

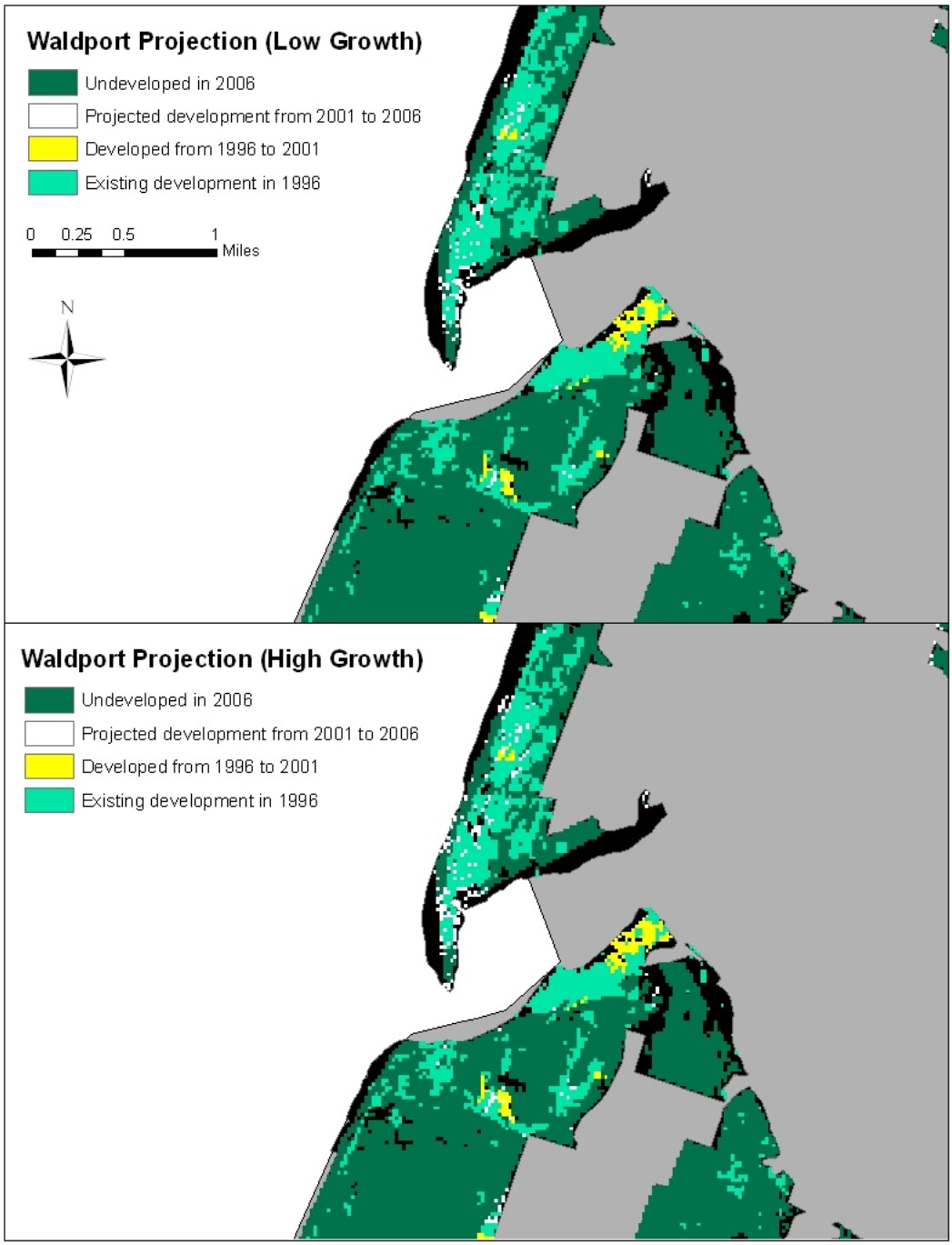


Figure 13. Comparison of spatial probit and standard probit projection results under low growth scenario in Waldport.

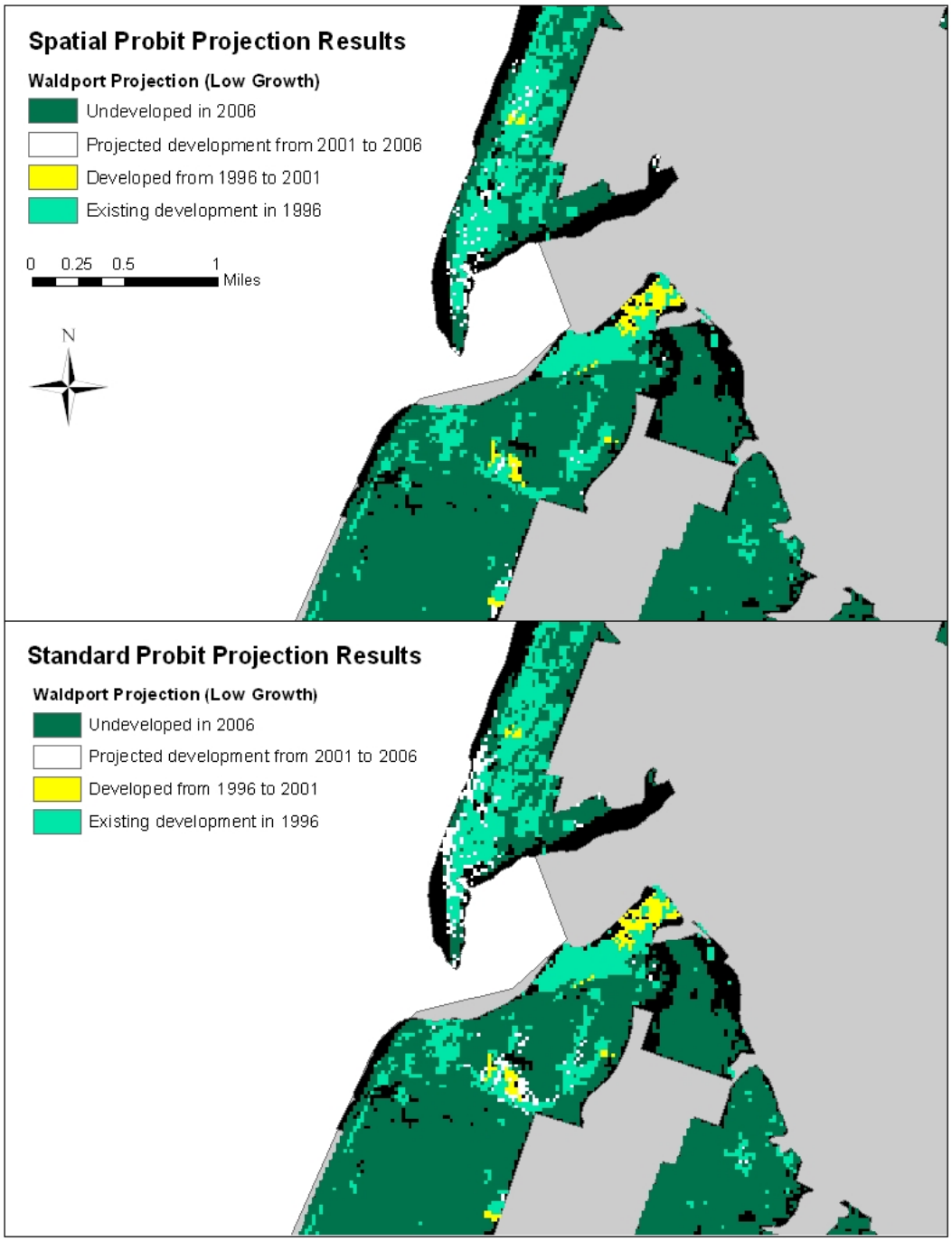


Figure 14 highlights observed and projected development within the urban growth boundary for the City of Waldport. This map also includes the tsunami inundation line mapped by DOGAMI. Areas that are located between the Pacific Ocean and the tsunami inundation line are at risk of being inundated in the case of a tsunami. Over the observed period from 1996 to 2001, 200 pixels were developed within the Waldport urban growth boundary. Of those 200 pixels, 105 were located within the tsunami hazard zone. With each pixel representing about 0.22 acres of land, this means that about 23 acres of land in the tsunami hazard zone were developed over the observed period. During the projected period, from 2001 to 2006, 269 pixels within the Waldport urban growth boundary are expected to be developed under the high growth scenario, with 120 of these pixels in the tsunami hazard zone.

The spit of land that juts into the Alsea Bay from the north is a particularly hazardous area. It is clearly located well within the tsunami hazard zone, but this area is also prone to long term erosion and accretion as weather patterns vary over the years. The winter of 1982 to 1983 produced an El Nino that severely eroded the Alsea River spit and placed dozens of homes in peril. Revetments were built to protect those homes and over the years the sand returned and buried the protective structures. During the years of sand accretion and calmer weather, additional homes were built to the west of the old revetments. Obviously these homes were in even more danger of being washed away during the next El Nino year, and when it hit, new revetments had to be put in place to protect the new homes (OSU 2006). The projection model developed here indicates that development will continue on the Alsea River spit, placing more and more homes on a sandy spit that is continually moving. This area is low lying and exposed to the open ocean, which also places it in great danger should a tsunami strike the Oregon coast.

Figure 14. Observed and projected development in the tsunami hazard zone within the Waldport urban growth boundary.

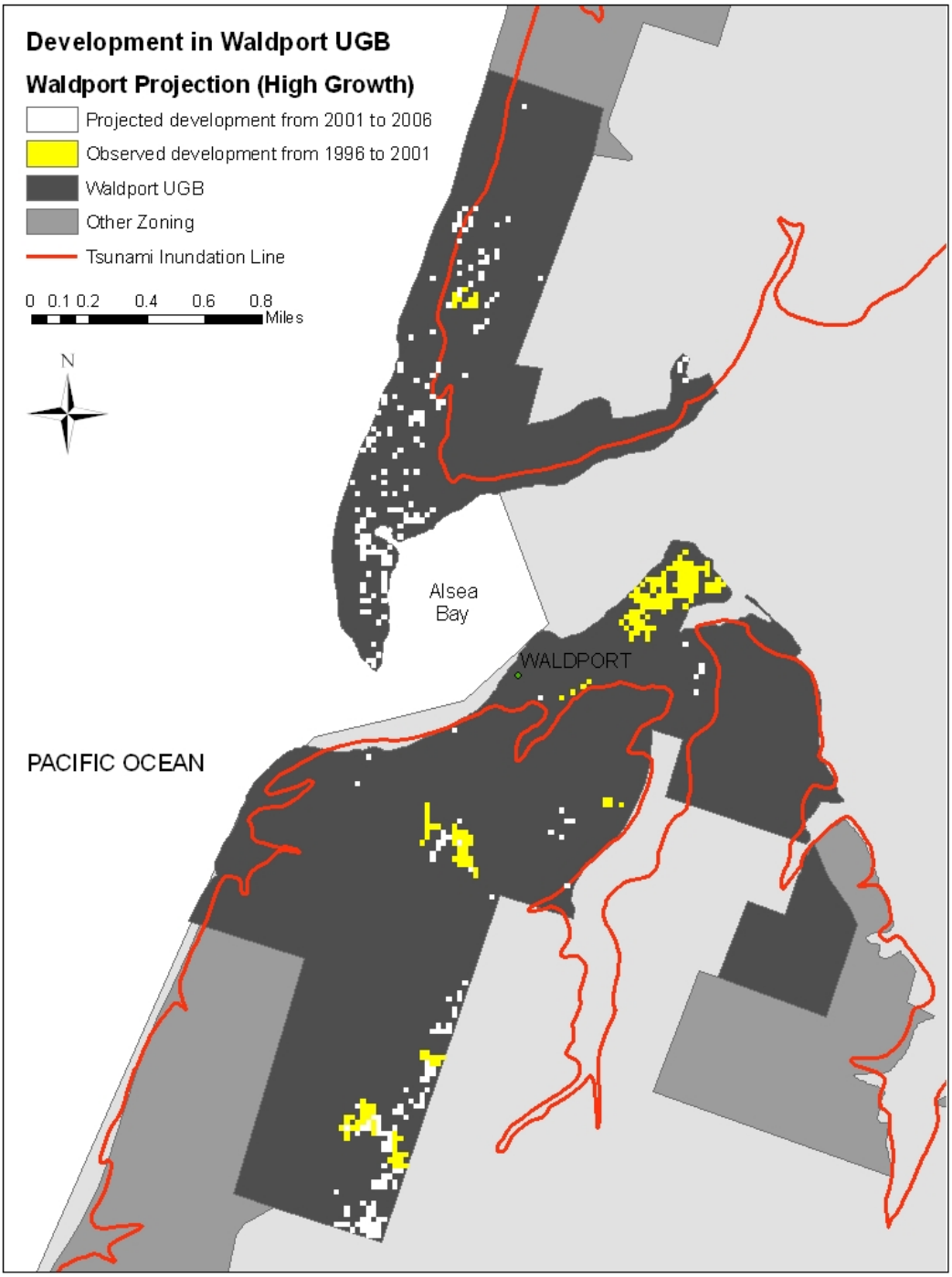


Figure 15 shows projected development in the Newport region under the high growth scenario. In this region, as in the others, the connection between projected development and zoning is clear. The majority of development projected from 2001 to 2006 is found in areas that fall under urban zoning. A detail map of land-use change in the City of Newport is included in figure 16. The development that was observed from 1996 to 2001 occurred mostly in large clusters along the eastern edge of existing development. The projections for 2001 to 2006 under both growth scenarios show new development filling into the gaps in existing development in smaller clusters.

The relationship between existing and projected developed land and slope in the Newport study area can be seen in figure 17. Development has historically happened in areas that are relatively flat and this projection indicates that trend will continue in the future. The average slope of land that was developed previous to 1996 is 4.8 degrees. The average slope among pixels that were observed to change from 1996 to 2001 is also 4.8 degrees however the average slope of the land that is still available for new development is 11.2 degrees. As development expands into the eastern portion of the Newport urban growth boundary, steeper slopes will be encountered

Figure 15. Projected development under the high growth scenario for the Newport region.

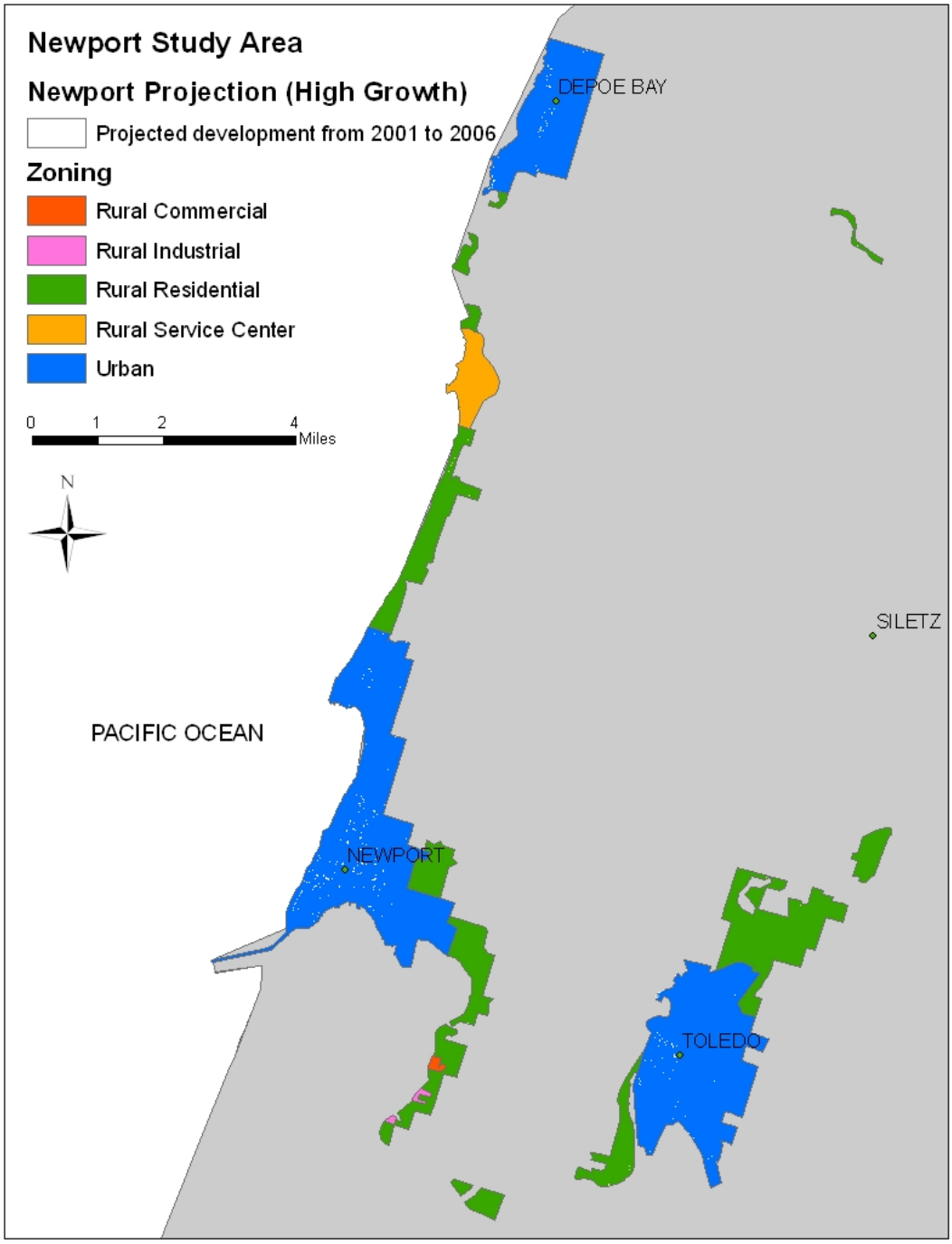


Figure 16. Projected and existing development in the City of Newport.

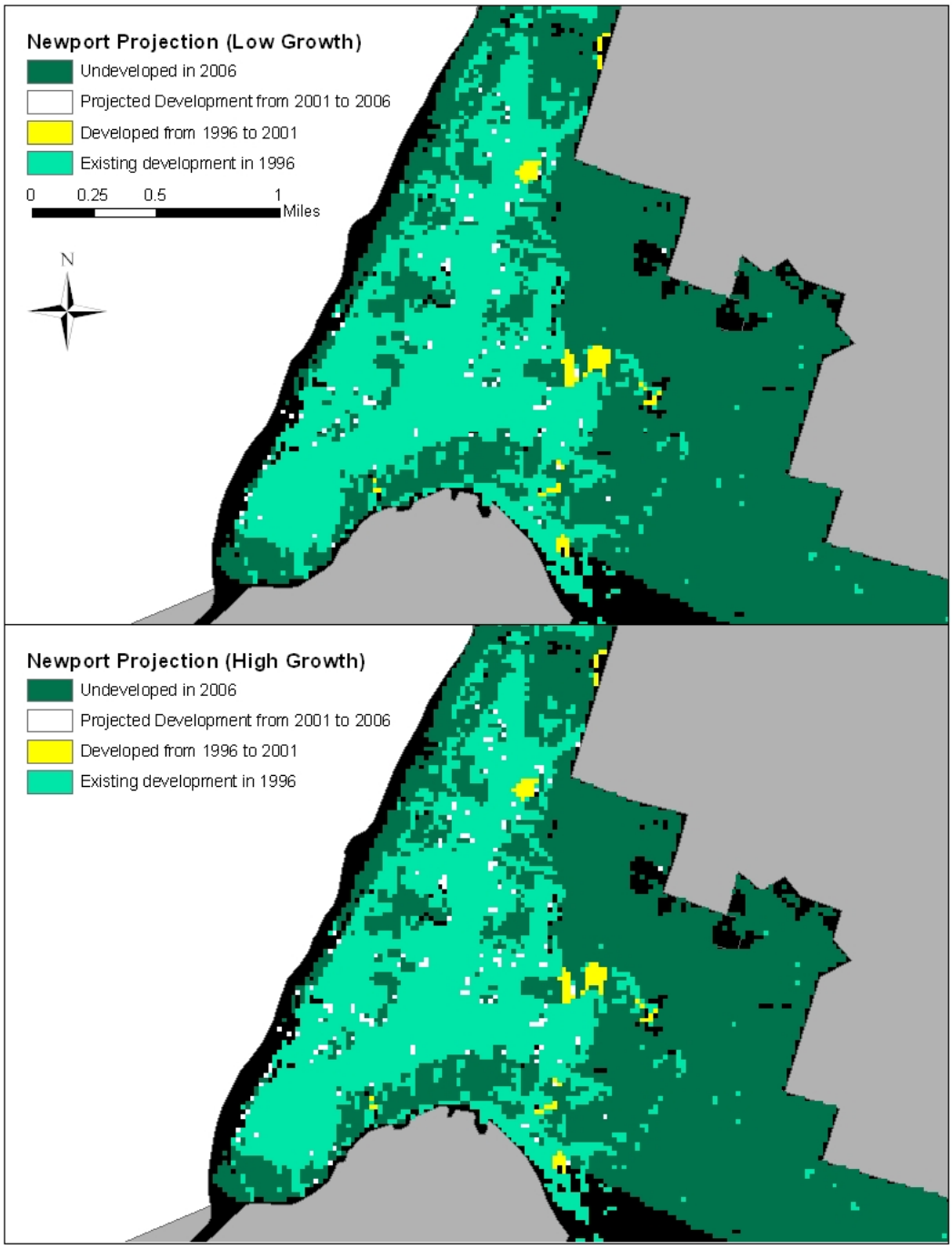
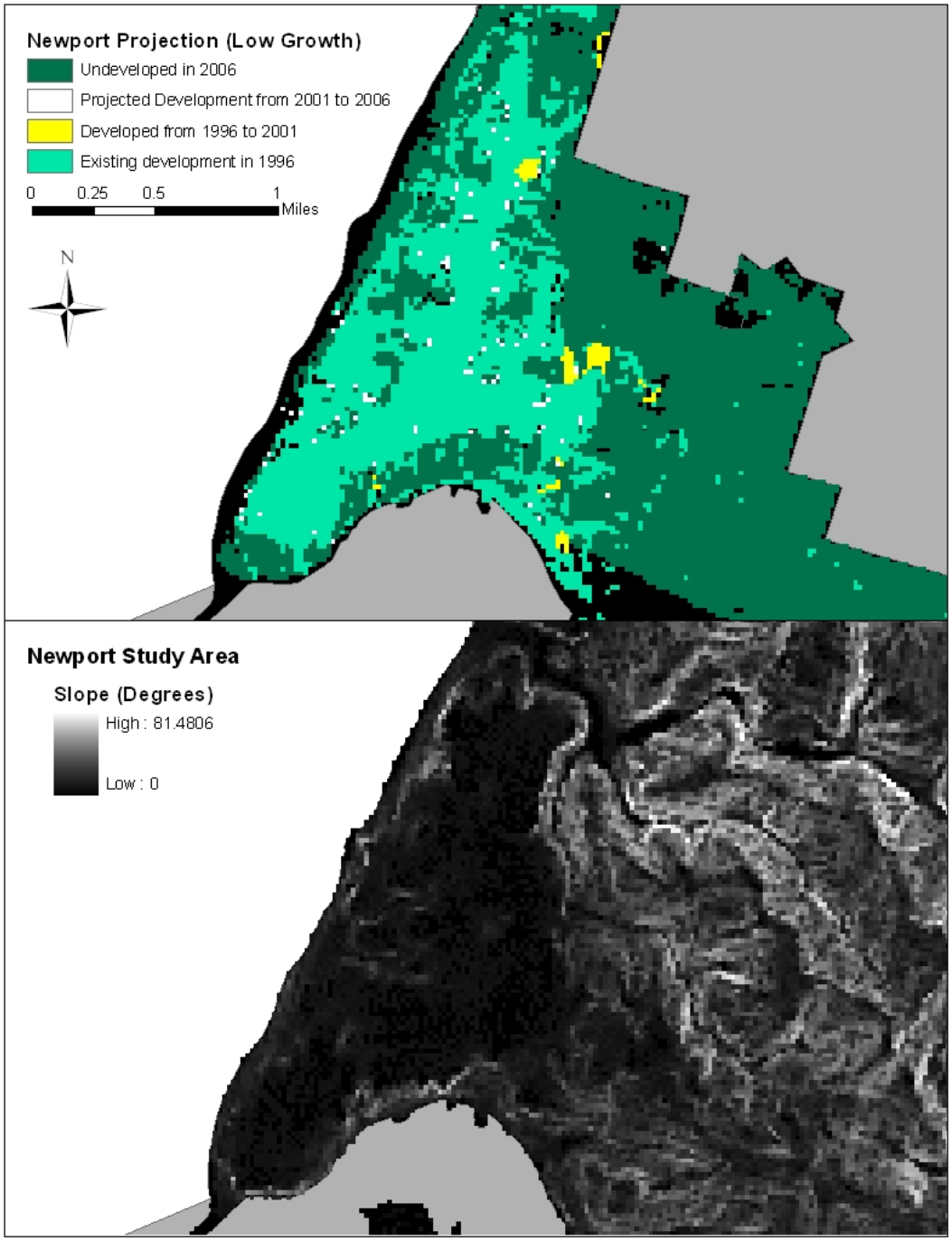


Figure 17. Comparison of existing development and low growth projection to slope in the Newport region.



5. Discussion

5.1 Regression

The differences between the standard probit estimates and the spatial probit estimates are notable. The distinction between the standard and spatial probit models is that the spatial probit model includes a lag of the latent variable in the form ρWz. The magnitude of the spatial parameter, ρ, in each region indicates that the value of the latent variable for surrounding pixels has a significant impact on the value of the latent variable for a given pixel. In the standard probit, the effects of the individual pixel characteristics are overstated because they are compensating for the impact of surrounding pixels. The impact of the spatial lag is also apparent in the marginal effects. The standard probit marginal effects are several orders of magnitude larger than the spatial probit marginal effects, indicating that the marginal effects for the standard model may be exaggerated because the spatial effects are not accounted for.

The dummy variable for western facing slopes was included as a proxy for a view of the ocean, since ocean views are highly desirable to those living or vacationing on the coast. Presumably, land facing the ocean would have a higher likelihood of being developed than land that does not have an ocean view. The negative coefficient in Newport, Reedsport and Lakeside might be explained by geography. The cities of Reedsport and Lakeside are located slightly inland from the ocean and therefore western facing slopes do not necessarily imply potential ocean views. All three regions, Reedsport, Lakeside and Newport, contain inland waterways that may be a bigger draw to development than ocean views, potentially accounting for the negative coefficient on WEST.

Distance from the nearest highway could be capturing two different, and opposite, effects. The highways used for this measurement are the primary roads on the

coast, including Highway 101 and the major roads that connect the coast to the valley. Being close to the highway reduces the cost of accessing the land for development and therefore may have a positive impact on development, but building a house near the highway can also be undesirable due to traffic noise, safety concerns and other issues. Having both positive and negative coefficients on HWY indicates that the benefit of developing further away from major roads outweighs the increased access costs in some areas, while the opposite is true in others. Positive coefficients outnumber negative ones eight to three, indicating that accessibility may not be a serious barrier to development on the coast. This is not surprising since most of the land zoned for development on the coast is located along the major highways. Therefore, the majority of available pixels of land examined here are a short distance from existing roads.

The coefficient on OCEAN is not significant in most models and this may be because there is not enough variation in the observed dataset. It is also possible that the measure of distance to the ocean is interacting with the measure of distance from the highway. Highway 101 runs parallel to the coastline for much of its length in Oregon, so these measures may capture very similar effects for land to the east of the highway and opposite effects for pixels to the west of the highway.

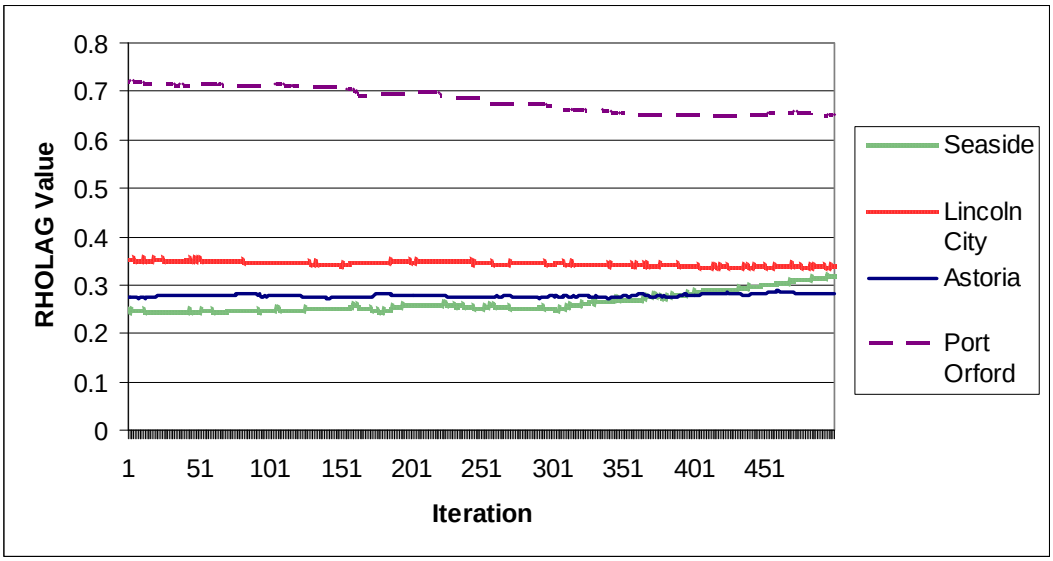
SLOPE is an indication of development costs, and as the slope of a pixel increases, development costs will usually increase. This is generally reflected in the parameter estimates, with negative and significant coefficients in seven models.

The estimate for the spatial parameter RHOLAG is positive and significant in every region, indicating that the clustered pattern of development is important to this model. The differences in the spatial parameter estimates across regions could indicate several things. Smaller spatial parameters could indicate that the region generally has smaller parcels or only partial parcels are being developed. A larger

spatial parameter could indicate larger parcels or that multiple adjacent parcels are being subdivided and developed at the same time.

In addition to the highest posterior density intervals reported in table 3, convergence of the sampling draws is an important measure of the stability of these parameter estimates. Figure 18 shows the saved Gibb's sampler draws for the RHOLAG parameter in four of the regional models. In this example, the RHOLAG parameters for Lincoln City and Astoria look fairly stable, while the draws for Seaside and Port Orford each have a definite trend. The saved draws for Seaside show an upward trend after roughly 300 iterations while the draws for Port Orford show a consistent downward trend.

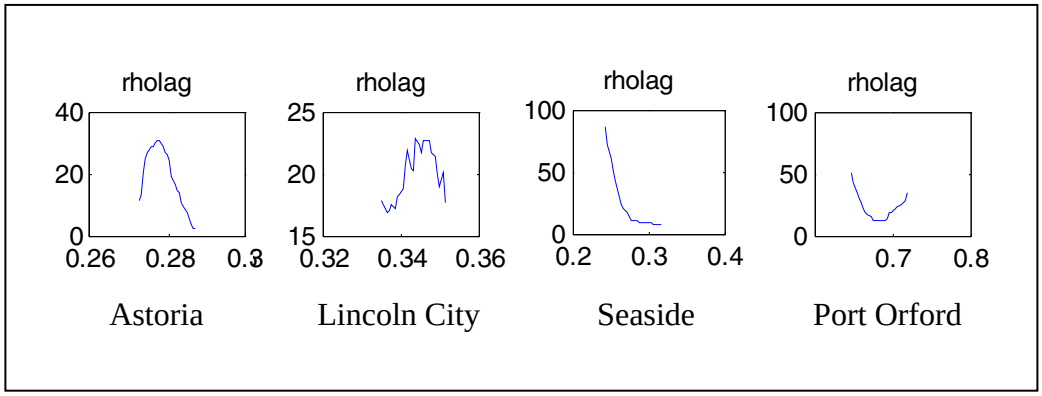
Figure 18. Saved Gibb's sampler draws for RHOLAG in four regional models.



This comparison can also be seen in the graphs of the posterior distributions for the parameter in figure 19. The graphs on the left are for the RHOLAG parameters in the Astoria and Lincoln City models. They are fairly bell shaped

and centered at each of the parameter estimates. However, the posterior graphs for Seaside and Port Orford reflect the trends that can be seen in figure 18. This analysis indicates that the RHOLAG parameter distributions for Seaside and Port Orford have not converged to the posterior distributions in 500 iterations. The RHOLAG parameter estimate was found to be statistically significant in every region however it is clear from examining the estimated posterior distributions that some of these estimates are more meaningful than others.

Figure 19. Posterior distributions of RHOLAG in four models.



When the algorithm was allowed to run for 5000 iterations for the Seaside model, the draws for the parameter estimates diverged even more indicating that the parameter posterior distributions would likely never converge. Therefore the means of the distributions that do not seem to have converged should not necessarily be used as parameter estimates. The convergence issue is something that should be examined more thoroughly. Because the estimated posterior distributions for each parameter are conditioned on all other parameters in the model, if the draws for one parameter diverge, they all tend to diverge. This is the case for some of the regions in this study, particularly Seaside, Port Orford and Brookings, and the reasons for this lack of convergence are not obvious.

The variation in all of the parameter estimates across different regions underscores the differences between the regions modeled in this study. The communities on the Oregon coast vary widely, ranging from fishing and logging towns struggling to maintain viable economies to bustling vacation destinations. Naturally these areas will have different drivers for development and different development patterns will emerge. No clear relationships among the different regions emerged that would allow us to group similar regions based on parameter estimates. Each model seems to have its own distribution of significant positive and negative parameter estimates.

In more permanent communities with stable year round populations, development may be driven primarily by population change, but in many coastal towns this is not the case. Some cities have more empty vacation homes than permanent residents, according to population and housing data from the U.S. Census Bureau. For example, Manzanita, on the northern coast, was home to 564 people as of 2000 and yet the city contained 1,078 housing units of which 771 were vacant at the time of the 2000 census. The majority of these vacant housing units are vacation homes. In Manzanita and other similar tourist towns, where more than half of all housing units are vacant, development is largely driven by tourism. In fact, the population of Gearhart declined between the 1990 and 2000 censuses and yet 345 new housing units were added to the city, most of them additional vacant homes. Contrast this with cities like Tillamook, where tourism is not the primary industry. Vacant homes account for 7.4 percent of housing units in Tillamook and the economy is based primarily on agriculture, timber and other resource industries (OSA 2008). The Oregon coast is a patchwork of tourist towns and logging, fishing and agricultural communities. These facts emphasize the difficulty faced by land-use planners on the coast when trying to predict how much developable land they will need over a given time period.

5.2 Projections

One of the primary motivations for this study is that estimation of a standard probit model based on spatial data is inappropriate. Thus, the comparisons of the spatial probit and non-spatial probit results for the Astoria and Waldport regions presented in section 4.2 are an important assessment of the value of the more complicated spatial model. Both projections show a clustered pattern of development, but these patterns arise for different reasons. In the spatial probit, the spatial weights matrix (W) and the spatial parameter (ρ) are included to capture the spatial pattern of the observed development. This pattern occurs in the dataset because parcels of land are larger than the 30 meter pixels on which measurements are collected. The spatial probit is attempting to replicate this pattern of parcels and development. In the standard probit, the clumpy development pattern emerges because the variables included in the regression model are also clumpy. For instance, measurements like slope and distance to the nearest highway do not vary much between any two neighboring 30 meter pixels. So if one pixel has a high probability of being developed based on these characteristics, then its neighboring pixels probably also have a high probability of development. The spatial probit and standard probit projections show significantly different patterns of development in both regions and the spatial probit projections should reflect the true process of land-use change better than the standard probit.

A visual inspection of figure 16 would suggest that slope has a significant impact on development in the City of Newport. The majority of existing and projected development seems to be located on land that is relatively flat. However, in the spatial probit regression for the region, the coefficient on SLOPE was not significant. This may be due to a lack of variation in the slope of pixels that were observed to change from 1996 to 2001, since the parameter estimates used for this projection are based on that one observation of change. Among the pixels that developed over the observed time period, the slope ranges from zero to 19 degrees.

However the median is 3.9 degrees and only 21 pixels out of 189 have a slope greater than 10 degrees, so slope does not vary much in the observed dataset.

6. Conclusion

This study highlights the tradeoff between using a cheap and widely available data source and the resulting need for more complicated estimation techniques. From the perspective of economic theory, a model that better describes the decision making process would be preferable. However, that requires data on individual parcels and landowners and collecting that data for a large study area is expensive and time consuming. The data source for this model was readily available, but the Bayesian estimation method was very computationally intensive.

Bayesian methods are not frequently used in econometric models and it is not difficult to see why. There is a fairly steep learning for those that have been trained in frequentist statistics to understand the Bayesian approach and the tools that are available. The computation time is also a limitation, since each model took about a day to estimate in this case. This makes a thorough examination of different model specifications very time consuming and impossible in this case. However, this approach is one of very few options available to estimate a model where the spatial pattern is considered to be an important part of the data generating process. Many other studies have treated the spatial autocorrelation as a nuisance and simply sampled it away. While these studies may produce unbiased parameter estimates, they cannot be used to project future land-use patterns because the spatial relationships between pixels or parcels have not been modeled.

The results of this study indicate the importance of modeling the spatial relationship in the dependent variable in this dataset. The spatial parameter is significantly greater in magnitude than the other parameter estimates in the model, indicating that much of the variation in the latent variable z is captured by the spatial lag, ρWz. This particular dataset is fairly clumpy, so the importance of ρ in this model is to be expected. But development does often occur in clumps in

reality, as do other types of land-use change, like deforestation. Estimating a standard probit in this case overstated the effects of the individual pixel characteristics on development, so it is clearly important to consider the spatial effects in any analysis using this type of data.

Implementing this estimation technique has been a challenge and there are several aspects of this process that could benefit from additional attention. First is specification of the weights matrix. The threshold distance beyond which spatial correlation is assumed to have no impact should be chosen carefully, as the choice of W can have a significant impact on the results of the model. W will become less sparse as the threshold distance increases, so this choice has to balance capturing the true extent of the spatial relationships with the additional computation required as more nonzero elements are added to the weights matrix. In this case, W was chosen based upon inspection of the clusters of development in the dataset and there may be a more rigorous process to compare models with different W specifications. The time required to estimate each model has precluded a thorough comparison of various model specifications for this study. Another unresolved issue is the lack of convergence of the posterior distributions in some models. This is a serious problem, as the resulting means of the distributions cannot be interpreted as meaningful parameter estimates. The causes of this non-convergence have not been explored in this study and this may be an area for future work.

Land-use planners could use a model like this to look at where development is most likely to go within areas that are already zoned for development or to compare how development might occur under different urban growth boundary expansion scenarios. However one key element that is missing in terms of practical application is a measure of development pressure. This study does not provide real world solutions to the problem of maintaining a 20 year supply of developable land within the urban growth boundary of each city. The methods

used to determine how much land will be needed over this time period are still lacking in locations where development is not necessarily driven by population change, particularly on the coast. This is another area where research is sorely needed.

Bibliography

Bockstael, N. E. 'Modeling Economics and ecology: The importance of a spatial perspective', *American Journal of Agricultural Economics*, 78 (December 1996), pp. 1168-1180.

Carrión-Flores, C. and Irwin, E. G. 'Determinants of residential land-use conversion and sprawl at the rural–urban fringe', *American Journal of Agricultural Economics*, Vol. 86, Issue 4 (November 2004), pp. 889-904.

Case, A. 'Neighborhood influence and technological change', *Regional Science and Urban Economics*, Vol. 22, Issue 3 (September 1992), pp. 491-508.

Chomitz, K. M. and Gray, D. A. 'Roads, land use, and deforestation: A spatial model applied to Belize', *The World Bank Economic Review*, Vol. 10, No. 3 (1996), pp. 487-512.

Danielsen, F., et al. 'The Asian Tsunami: A Protective Role for Coastal Vegetation', *Science*, Vol. 310 (28 October 2005), p. 643.

De Pinto, A. and Nelson, G. C. 'Modelling deforestation and land-use change: Sparse data environments', *Journal of Agricultural Economics*, Vol. 58, No. 3, (2007), pp. 502-516.

Fleming, M. 'Techniques for estimating spatially dependent discrete choice models', in Anselin L., Florax R. and Rey S. (eds.), Advances in Spatial Econometrics, pp. 145-167. Springer, Berlin, 2004.

Gelfand, A. E. and Smith, A. F. M. 'Sampling-based approaches to calculating marginal densities' *Journal of the American Statistical Association*, Vol. 85, No. 410 (June 1990), pp. 398-409.

Greene, W. H. Econometric Analysis. Prentice Hall, New Jersey, 2003.

Holloway, G., Shankar, B. and Rahman, S. 'Bayesian spatial probit estimation: a primer and an application to HYV rice adoption', *Agricultural Economics*, 27 (2002), pp. 383-402.

Irwin, E. G. and Geoghegan, J. 'Theory, data, methods: Developing spatially explicit economic models of land use change', *Agriculture, Ecosystems and Environment*, Vol. 85 (2001), pp. 7-23.

'OSU Researcher: Winter Storms and Wave Heights Escalating off Northwest Coast', Oregon State University Media Release, 7 December 2006. Accessed May 20, 2008 from http://oregonstate.edu/dept/ncs/newsarch/2006/Dec06/waveheights.html.

LeSage, J. P. Spatial Econometrics. Last updated June, 2000. Accessed Feb 11, 2008, from http://www.spatial-econometrics.com.

National Oceanic and Atmospheric Administration (NOAA). 'Which way to the beach? Oregon's beaches belong to the public', *Coastal Services Magazine*, April 1998. Accessed April 10, 2008 from http://www.csc.noaa.gov/magazine/back_issues/apr98/sec4c.html.

Nelson, G. C. and Hellerstein, D. 'Do roads cause deforestation? Using satellite images in econometric analysis of land use', *American Journal of Agricultural Economics*, 79 (February 1997), pp. 80-88.

Oregon Department of Land Conservation and Development (DLCD), 'Statewide Planning Goals', Accessed April 10, 2008 from http://www.oregon.gov/LCD/goals.shtml

Oregon Department of Transportation (ODOT). 'Agency History', Written 1998. Accessed Feb 11, 2008 from http://arcweb.sos.state.or.us/state/odot/hist/history1914.htm.

Oregon State Archives. 'Oregon Blue Book – Almanac and Fact Book', Accessed April 28, 2008 from http://bluebook.state.or.us/local/counties/counties29.htm.

Thomas, T. S. 'Lost in Space: A Primer for Probits When Geography Matters (as it Almost Always Should)' Written December, 2006. Accessed Feb 11, 2008 from http://www.timthomas.net.

Western States Seismic Policy Council (WSSPC), 'Tsunami Legislation and Risk Reduction Efforts in Oregon', Written 1997. Accessed April 28, 2008 from http://www.wsspc.org/tsunami/tsunamilaw.html.

Data Sources

Spatial Data	**Source**
Pacific Coast Land Cover	NOAA Coastal Services Administration
Available from NOAA: http://www.csc.noaa.gov/crs/lca/pacificcoast.html	
Digital Elevation Map (DEM)	U.S. Geological Survey (USGS)
Available from: http://buccaneer.geo.orst.edu/dem/	
Tsunami Inundation Line	Oregon Department of Geology and Mineral Industries (DOGAMI)
State Zoning	Department of Land Conservation and Development (DLCD)
Highways	Oregon Department of Transportation (ODOT)
Land, Public Ownership	Oregon Department of Forestry (ODF)
Available from the Oregon Geospatial Enterprise Office (GEO): http://gis.oregon.gov/DAS/EISPD/GEO/sdlibrary.shtml	

Non-Spatial Data	**Source**
Census 2000 population and housing data	US Census Bureau
Available from the Census Bureau: http://www.census.gov/main/www/cen2000.html	