A Spatial Analysis of the Economic and Ecological Efficacy of Land Retirement

Elizabeth P. Marshall
Department of Agricultural Economics and Rural Sociology
The Pennsylvania State University

Frances R. Homans
Department of Applied Economics, University of Minnesota

May 15, 2001

Paper to be presented at the 2001 Annual Meeting of the American Agricultural Economics Association, Chicago IL, August 5-8

Abstract

Most land management policies, such as land retirement, have multiple objectives. This study uses a cellular automata simulation model to explore how various spatial characteristics of land parcels on a hypothetical landscape contribute to the efficacy of land retirement in the presence of multiple retirement objectives– hydrological improvement, habitat improvement, and cost. Statistical analysis of the simulation results is used to tie particular spatial characteristics back to achievement of the three distinct objectives. In order to combine the three objectives into a measure that allows decision-makers to rank the desirability of different retirement strategies, linear and nonlinear goal programming frameworks are introduced. These frameworks are explored to determine what each implies about the tradeoffs that must be made among objectives and among the spatial land parcel characteristics that contribute to those objectives.

A Spatial Analysis of the Economic and Ecological Efficacy of Land Retirement

1. INTRODUCTION

Although references to location and the importance of spatial relationships are sparingly scattered throughout the early history of economics, in the past 20 years the exploration of these themes has noticeably broadened. Economists are joining ecologists and engineers in their consideration of the world as a patchy place, complete with scattered resource endowments, spatial flows (of material or information), and asymmetric interactions.

One particular branch of exploration that has seized on the importance of an explicit analysis of location has been, not surprisingly, issues of land use. The explosion of a theoretical interest in the importance of location and land use reflects a phenomenon that is anything but theoretical. As the pace and pattern of economic activity across the globe have changed over the last few decades, land use conflicts have become a contentious and visible symptom of how large the externalities generated by certain land uses can be. The externalities associated with land use range from nuisance effects, where neighbors are bothered by the odors of adjacent livestock facilities or by the lights of an industrial development, to the flow effects associated with point- and non-point-source pollution. Positive externalities exist as well, as seen when adjacent open space increases residential housing values or when one natural area increases the viability of a second natural area simply by being within species dispersal distance. Certain land uses generate a combination of types of externalities– a single land use may generate positive externalities for some neighbors and negative externalities for others.

This study uses simulation techniques and a hypothetical scenario to explore the relative strengths of different external land use effects that can exist for a single land parcel and how these effects must be traded off when land managers consider questions of appropriate land use. Consider a land manager assigned the task of acquiring for the purposes of retirement a parcel of agricultural land. There is a reason for this retirement; in fact, there are several. The region has been grappling with water quality issues associated with intensive agriculture and is beginning to suffer the effects of that decline in water quality in terms of decreased agricultural productivity. In addition, the natural habitat of that area has been seriously diminished and fragmented by the conversion of natural tracts to agriculture. One species, considered an umbrella species, is of particular concern in this region. Land retirement has therefore been suggested as a way both to ameliorate the water quality issues in the region and to supplement the existing system of natural areas on behalf of the species of concern.

A simulation model is used to explore how important location characteristics, both absolute and relative, are in determining how successful a land parcel is at achieving each of these objectives. Certain measures of location may be relevant to both of the manager's objectives, but not necessarily in a consistent way. Goal programming is presented as a decision-making framework that accommodates multiple objectives and explicitly determines how spatial characteristics should be traded off in cases where different, and perhaps opposing, characteristics are important for achieving the distinct management objectives.

2. MODEL FORMULATION

The simulation model used in this study is a cellular automata programmed in Matlab (MathWorks, 1999). The model generates a virtual landscape comprising 12 x 12 cells, each of which is assigned one of three possible land uses– urban, agriculture, or natural area. The model then tracks what happens to landscape water quality, habitat quality, and land values over time as agricultural activity progresses. Each agricultural cell on the landscape represents a candidate strategy for a land manager interested in retiring agricultural land, which the model translates as converting agricultural land to natural area. The model is used to explore how effective each of these agricultural land parcels (or retirement "strategies") would be at achieving the goals of a land manager who has three considerations in retiring land– landscape-level hydrological improvement, landscape habitat enhancement, and expenditure minimization. In order to measure the effectiveness of each parcel, the simulation model enumerates all retirement possibilities– it converts each agricultural cell to a habitat cell, simulates the behavior of the landscape over a period of time called the "management horizon," and at the end of the management horizon measures the change in the objective levels that have resulted from retirement of that cell. The model therefore produces for each candidate parcel three performance scores, one for each of the retirement objectives. The components of the model are described in more detail below.

2.1 Probability of Survival

The ultimate goal in wildlife management is species survival, but the uncertainty inherent in most environmental processes makes it more useful to express the management goal in terms of a certain probability of species survival (Haight, 1995, Marshall *et al.*, 1998). This component of the model uses a production function for the probability of species survival that includes total area in habitat and a measure of landscape connectivity (or fragmentation of that habitat) as production inputs. For a given generated landscape, the model measures landscape connectivity before and after each agricultural cell has been converted to habitat, and uses this information to calculate the marginal change in probability of survival associated with retiring each nonhabitat cell. The model also calculates the effect of the increase in area on survival probability, but this effect is constant landscape-wide, as all of the agricultural parcels are assumed to be the same size.

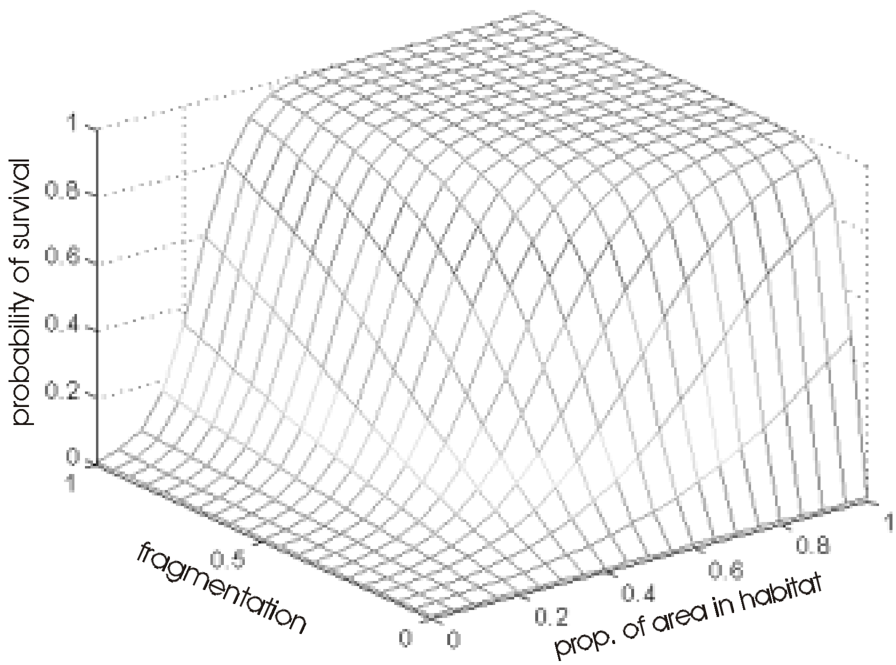
The production function used to represent survival probability was selected to meet a number of criteria. The relationship between area and species survival probability (for a fixed level of habitat fragmentation) is assumed to resemble an s-curve. As the level of landscape connectivity increases (and fragmentation decreases) one would expect the s-curve to shift to the left, so that survival probability increases more quickly, and begins its ascent at a lower total habitat amount, when habitat is clustered rather than widely dispersed. The function should range from 0, which is the probability of survival if no habitat is available, to 1 (in the limit) as the amount of habitat increases and its dispersal decreases. As the habitat becomes highly fragmented, one might also expect the maximum survival probability achievable to drop below 1.

A production function based on the cumulative logistic growth function satisfies these requirements. One example of this function, adjusted to reflect the interaction between the two input variables, is:

$$Z = \frac{1}{1 + 100e^{-20Y^{.5}X}} \qquad (1)$$

In this function , Z represents the probability of species survival (over some specified time horizon, such as 200 years), X represents the total area in habitat (measured as a proportion of total habitat available), and Y is a measure of habitat connectivity. A graph of this function appears in figure 1.

Figure 1: Production of probability of survival.



A number of measures of landscape connectivity appear in the literature, including correlation length, number of habitat patches, length of habitat edge, and the ratio of habitat edge to area, and largest patch size (Andren, 1994; Gustafson, 1998; With et al, 1997). This study uses largest habitat patch (as a proportion of total habitat available) as a measure of landscape connectivity. This metric behaves nicely when landscapes are in the critical transition phase from completely disconnected to predominantly connected. When additional habitat cells behave as stepping stones, for instance, connecting existing patches of habitat, this measure captures not only the addition of a habitat cluster to the largest previously known cluster, but the size of the cluster added on as well.

2.2 Hydrological Flow and Contaminant Accumulation

The groundwater component of the model roughly simulates a groundwater gradient that moves from the left column of the grid (representing the top of the watershed) to the right column of the grid (the watershed's drainage plain). Surface agricultural activities result in a constant accumulation in the groundwater of some generic pollutant; this pollutant can be considered either an applied substance that reaches the water table, such as applied nitrogen from fertilizer, or a naturally occurring soil substance that is concentrated by the leaching effect of applied irrigation water, such as salt or selenium. In either case the rate of accretion depends on the surface land use: natural habitat and urban areas add nothing to the pollutant pool in each cell, while agriculture increments the pollutant level by a constant amount in each time period for as long as the cell remains an agricultural cell.

In each time period, a percentage of the pollutant stock in each cell flows to the right. Therefore, in each period, the pollutant stock of each cell is augmented by surface accretion and by flow from upstream cells, and decreased by flow toward cells below. There is no drainage from cells in the final column, so contamination simply accumulates there. A certain amount of lateral flow, or smoothing, occurs as well; in each time step a percentage of the difference in pollutant concentration moves from cells of higher concentration to lateral cells of lower concentration.

A saturation threshold is also incorporated into the groundwater model. When cells in reach a critical threshold value of contamination, they no longer "accept" contamination from upstream cells. The contaminant level in a saturated cell stays fixed at the threshold level, and all excess pollutants from upstream backtrack up a single row until an unsaturated cell is reached. At this point, lateral smoothing also occurs among cells in all columns; a percentage of the difference in cell concentrations moves from cells of higher concentration in a column to adjacent cells of lower concentration.

The result of this flow structure is that upstream agricultural activity creates a plume of increased contamination downstream. Similarly, retiring an upstream parcel of land results in a downstream plume of improved hydrological quality relative to a scenario in which no retirement occurs.

Landscape hydrological quality is measured as the total amount of contamination on agricultural cells exceeding threshold 1. To measure the impact of a retirement on landscape hydrological quality, the model removes the cell from agriculture, runs the flow model for the remaining agricultural cells for a period of time called the "management horizon" and then measures the difference in hydrological quality between the scenarios with and without retirement. The impact of retirement is sensitive to location because although the total amount of contamination entering the landscape will decrease by the same amount regardless of where a cell is retired, the extent to which the decreased contamination benefits agricultural productivity will vary depending on the surrounding and downstream land uses.

2.3 Land Value and Cost

The model uses a gravitational model to estimate agricultural land values in the shadow of an urban area (Shi *et al.*, 1997). This model posits that the effect of an urban area on an agricultural parcel's market value is inversely related to the distance between the two sites and directly proportional to the size of the urban area. This model is generalizable to the case where land values are influenced by several different urban areas, however in this study only a single urban area is considered, and it's urban influence on surrounding agricultural areas is measured as, where R_i is the urban influence index for agrigultural cell i, N_j is the population of the urban area, and D_{ij} is the distance between agricultural cell i and the urban area j.

$$R_i = \frac{N_j}{D_{ij}^2} \tag{2}$$

The market price of each agricultural cell on the landscape is then estimated as:

$$V_i = P_i + 2.5R_i \tag{3}$$

where P_i is a measure of the expected net returns from agricultural use of the land and R_i is as described above. P_i is calculated as a discounted stream of returns, where each period's return, p_i, is sensitive to the level of the hydrological contamination in the cell. For low levels of contamination, p_i is equal to a landscape-wide maximum possible return for each period, $\hat{p}$. As a cell becomes more contaminated, however, two threshold contamination levels are reached. At the first, t_1, agricultural productivity begins to decline, and at the second, t_2, contamination makes agriculture untenable, and the land is removed from agriculture altogether. The relationship that exists between contamination and productivity levels when the contamination level is between these two thresholds is:

$$p_i = \hat{p} - \frac{\hat{p}}{(t_2 - t_1)^2}(h_i - t_1)^2 \,. \tag{4}$$

Total expected net returns are then calculated using the discount rate *r* as shown below:

$$P_i = \sum_{t=0}^{\infty} e^{-rt} p_i \,. \tag{5}$$

When applied to a landscape in which no contamination levels exceed t_1, the gravitational model results in concentric circles of increasing land values that radiate out from the central urban area. Over time, however, as increasing contamination erodes the productivity of downstream cells, the pattern of land values reflects this effect.

3 METHODOLOGY

The purpose of this research, however, is to explore the particular role that location plays in determining how desirable a land parcel is as a candidate for retirement, given a suite of retirement objectives. Each land parcel is characterized by two different types of location measurements, those arising from its absolute location on the grid, and those arising from its relative location. Absolute location refers to a cell's geographical position, as measured by variables such as latitude and longitude, x and y, etc.; for a given cell, these measures are invariant, regardless of how land use patterns change over the landscape. In contrast, relative location refers to a cell's location with respect to other cells, or more specifically, with respect to some relevant characteristics of other cells. On the virtual landscapes explored here, for instance, the dynamics of relative location are generated by differences in surface land use patterns; relevant descriptions of relative location would therefore measure position relative to the land use of other cells, such as whether a cell is adjacent to habitat or surrounded by agriculture. These measures are different for each cell and for a fixed cell vary by landscape.

Use of a simulation model allows one to isolate the effects of location on a cell's desirability by fixing all the other input parameters, such as flow and productivity variables, and running the model on a number of different landscapes, so that the only variable input is the pattern of land uses over a landscape of fixed dimensions. Each model run provides three performance measures for each agricultural cell on the landscape, indicating its potential as a retirement candidate. When the pattern of surface land use changes during the next model run, so do the performance scores for each cell. Within a given landscape, differences in performance scores may be due to differences in absolute location. However, when results for a cell with a fixed location differ between landscapes, these differences can only be due to differences in relative location, as a cell in a fixed position goes from being surrounded by habitat to being surrounded by agriculture, for instance. The structure, or land use pattern, of each individual landscape can therefore influence, to varying degrees, how individual cells perform toward each objective. Can the results of running the model on several different landscapes be used to identify what it is about those landscape configurations that is so influential?

To explore this question, 30 different landscapes were generated that differed only in the distribution of habitat across their surfaces. The model produced performance results for each of the agricultural cells on each of these landscapes. The following sections examine a number of different measures of relative location for a particular cell in an attempt to identify those that capture the relevant dynamics of landscape pattern. Two types of measures are examined– macro measures and micro measures. Macro measures are measures taken over the entire landscape; for each landscape there is a single value. Micro measures are measures made on each cell; these measures can simply be measures of location relative to other surface activities, or they can be weighted by the intensity of activity at other locations. Given the distinct nature of the three different objectives in this management problem, the measures relevant to the dynamics of each objective are likely to be different.

Ultimately, it would be useful to incorporate the measures identified for the distinct objectives into an overall framework for thinking about how points can be awarded to candidate parcels on the basis of these measures, and for evaluating what sorts of tradeoffs exist among measures of relative location. For instance, how can the importance of cell characteristic A, which contributes to the hydrology objective, be compared to the importance of characteristic B, which influences the cost objective? Can the effects of characteristics A and B be standardized somehow so that their contributions to some composite retirement objective, which encompasses both the hydrology and the cost objective, can be directly compared?

The first step in the standardization process requires formulating the composite retirement objective. This task requires use of a multi-criteria decision-making framework that translates the three-dimensional performance measures for each cell (hydrology, habitat, and cost) into a one-dimensional measure (composite score). A single score allows every cell on a landscape to be compared to every other cell on that and on every other landscapes in terms of relative desirability in achieving some overall objective. The multi-criteria decision-making (MCDM) framework used in this study is goal programming.

3.1 Goal Programming

Under the goal programming framework, a management objective function is created that minimizes a weighted aggregation of the distance of each objective from some pre-specified target level for that objective. In order to standardize the measures of distance across objectives that are measured in different units, this distance is measured as relative deviation from the target level. Given a management strategy x, the performance of that strategy toward an objective target level T_g can be measured as:

$$d_g = \frac{(T_g - z_g(x))}{(T_g - z_g^{**}(x))} \tag{6}$$

where $z_g(x)$ measures the actual, absolute level of performance achieved, and d_g measures its relative deviation from the target level. This measure of relative deviation also requires specification of an "anti-target" or "anti-ideal" level of performance, denoted $z_g^{**}(x)$. The anti-ideal value adopted for each objective is often defined to be the lowest level that objective can take among the candidate management strategies.

Given a measure of the deviation of each objective from their target values, the goal programming objective function used by the land manager in this study is:

$$\min_x s_x = \sum_{g=1}^{n} w_g f(d_g) \tag{7}$$

To complete the overall objective function formulation, the decision-maker must assign a weight w_g to each of n objectives, and select a distance function, $f(d_g)$. The distance function translates each objective's deviation, d_g, into some measure of how important a deviation of that magnitude should be in the decision-making process.[1] This function can be linear or nonlinear. A linear distance function would yield an overall objective function of:

$$\min_{x} s_x = \sum_{g=1}^{n} w_g d_g \tag{8}$$

and would imply that the individual objectives do not become more important in the decision-making process as they get farther from their ideal value. Improving performance toward objective k by moving from d_k=.9 to d_k=.8 is "worth" as much to the decision-maker, in terms of what he would have to sacrifice toward the other objectives, as moving from d_k=.2 to d_k=.1.

If instead the decision-maker would like the flexibility to allow objectives to have a relatively larger contribution to a cell's score when they have large deviations from the ideal than when they have small deviations from the ideal, then a nonlinear distance function can be used. In addition to selecting the form of the distance function itself, decision-makers often must also select the magnitude of some parameter that modulates how sensitive the distance function actually is to distance from the target value. In order to accommodate objectives whose sensitivity to distance operate independently of one another, this study uses the following distance function:

$$s_i = \sum w_g d_g e^{\lambda_g d_g} \tag{9}$$

This formulation permits specification of independent λ_g for each objective. Any objective that is to be considered linear in importance is assigned a λ_g of 0; if λ_g=0 for all objectives, then this formulation simplifies to the linear goal programming formulation described earlier.

4. RESULTS

Use of a MCDM framework such as goal programming solves the problem of standardizing the contribution each characteristic makes toward the overall objective by standardizing the units in which the objective performances are measured. When measured as deviations from some target level, performance measures for each individual objective therefore fall in the range [0,1]. For each objective, a series of regressions was run to explore the contribution of measures of absolute and relative location to each cell's performance toward the individual objectives. In order to make the magnitude of the coefficients assigned to each independent variable comparable across objectives, the dependent variable used in all these regressions is the deviation of each cell's performance from the target level established for each objective. The result is that the estimated coefficients for each of the variables will have comparable units across individual objectives.

To explore the effects of absolute location of a parcel on performance results for each individual objective, those results were regressed on sets of dummy variables representing increasingly broad

---

[1]The function is called a distance function because it measures how far a strategy's set of performance indicators (x_1, x_2, x_3) is from the overall management objective (T_1, T_2, T_3).

definitions of absolute location. Given results for multiple landscapes, it is possible to define dummies representing individual cells, columns, or aggregates of columns, for instance. The broadest definition of absolute location that could be included without sacrificing explanatory power was the scale used to delineate regions within which absolute location effects are constant.

A large number of possible measures might be effective at capturing what is relevant about a parcel's relative location in determining its performance toward the distinct objectives. Using OLS and regressing the performance measures on both the regional dummy variables and an additional measure (or measures) of relative location, a series of possible measures were examined to see how much they were able to contribute to the explanatory power of the absolute location model. Those variables included a measure of how many of the immediately adjacent cells are in habitat, a distance-weighted measure of the amount of habitat downstream, a distance-weighted measure of the amount of surrounding agriculture, and a distance-weighted measure of the amount of habitat upstream of the cell.

Distance-weighted measures are measures in which the contribution of a neighboring agricultural cell or habitat cell drops off with distance to the cell whose measurement is being taken (the base cell). There are a number of factors that will affect how effective a measurement is at capturing a particular landscape dynamic. The size of the neighborhood over which the measurement is taken is quite important. When looking at the influence of "surrounding" habitat, for instance, how large should the neighborhood be that constitutes "surrounding"? How many cells in each direction should be considered? Also, the structure of the decline in importance with distance is important–is the decline linear in distance, or nonlinear, for instance? And finally, the rate at which the decline in measurement value occurs must match the actual rate at which influence declines on the landscape.

To determine an appropriate structure and rate for decline, variables based on several different structures were explored to determine which of these was most effective at explaining a cell's capacity for landscape hydrological improvement. The structures evaluated are shown below, where each component $x_{i,h}$ represents a neighbor cell h's contribution to the base cell i's measurement:

$$x_{i,h}=\eta^{D_{i,h}} \tag{10}$$

$$x_{i,h}=\frac{1}{e^{D_{i,h}}} \tag{11}$$

$$x_{i,h}=\frac{1}{e^{D_{i,h}^2}} \tag{12}$$

$$x_{i,h}=\max(1-\gamma\bullet D_{i,h},0) \tag{13}$$

$$D_{i,h}=\sqrt{v^2+w^2}\ , \tag{14}$$

where v is distance between the neighbor cell and the base cell along the y axis (in rows), w is the distance between them along the x axis (in columns), and η and γ are parameters from 0 to 1 representing how quickly the contribution of a cell decreases with distance from the base cell. As

η increases, the contribution of distant cells to a base cell's measurement increases, and the total measurement value achievable by a cell increases. As the parameter γ increases, on the other hand, the rate of decline of influence of neighboring cells increases, and the measurement value of a given cell decreases.

In exploring which of the measurements was most effective at capturing variability in agricultural productivity, cell measurements were calculated for each of the structures for several different values of η and γ.[2] Testing also included several different neighborhood configurations. The regression results for each of the objectives are described below.

4.1 Hydrological objective

The hydrology analysis determines which of the cells on the landscape can be retired and have the maximum effect on the hydrology of the rest of the landscape. The analysis of absolute location effects identified discrete regions within which cells had comparable effects when retired. For the performance of cells retired in time period 1 whose effects are measured in time period 4, for instance, a regression on dummy variables representing individual cells has an $\overline{R}^2$ =0.275, while a regression on columns has an $\overline{R}^2$ =0.288. The coefficients resulting from the regression on columns suggested regions of similar overall desirability for retirement. On average, both the upstream region and the downstream region are less desirable than the middle region in achieving the hydrological objective. The results of this column regression divide the landscape into seven regions–the most upstream and downstream three columns each have their own regions, and the middle six columns represent the midstream region. A final regression on dummy variables representing these seven regions yields an $\overline{R}^2$ =0.289. The regions defined by regressing on the column dummy variable will therefore be used as a proxy for location along the gradient and as an efficient measure to capture the effect of absolute location. The matrix of regional dummy variables generated therefore has seven columns, one for each region, and will be denoted $I_{2,1}$.

Results of the exploration of relative location effects indicated that the most effective single measure of relative location was a distance-weighted measure of surrounding habitat made on a neighborhood of four cells in every direction and using the neighborhood function:

[2] In the regressions described in this and the following sections, possible transformations of both the response variable and the predictors were also explored. The Box and Cox approach was used to select possible transformations of the response variables, and a method suggested by Weisberg (1985) was used to explore possible power transformations of the predictors. In general, no transformations of the predictors were called for. Although in a few cases these approaches suggested transforming the response variables, the transformations generally yielded only slight improvements in the fit of the model. Because such transformations, especially of the response variable, complicate the derivation of the point systems described in section 4.5 and yield only a very small increase in fit, they were not applied.

$$\hat{x}_{i,h} = \max(1 - 0.27 \cdot D_{i,h}, 0)\,. \tag{15}$$

Regressions and predictions using various measures of relative location revealed that forcing the effect of a relative location variable to be landscape-wide restricted the explanatory power of the variable. In fact, the effect of surrounding habitat on a downstream cell's hydrological performance may be considerably different from the effect of surrounding habitat on an upstream cell's hydrological performance. Modifying the model in order to allow the effect of the relative location to vary by absolute location, i.e. to vary by region, improved the explanatory power of the model. The final model selected to explain and predict hydrological performance was

$$d_{2,1} = \mathbf{I}_{2,1} \cdot \boldsymbol{\alpha}_{2,1} + \sum_r (\beta^r_{3,1} \cdot i^r_{2,1} \cdot X_3) + e_{2,1} \tag{16}$$

where

$$X_3 = \sum_h \hat{x}_{i,h}\ .$$

Estimating the model yields the results shown in table 1.

A positive coefficient on the region-specific X_3 variables ($\beta^r_{3,1}$) indicates that surrounding habitat tends to increase the deviation from the ideal (relative to the benchmark value given by the coefficient on the region-specific constant $\alpha^r_{2,1}$), making cells surrounded by habitat less desirable as a retirement option when hydrological improvement is being considered. This result arises because when a cell has a great deal of surrounding habitat, the water quality improvements derived from the retirement are spread over habitat cells rather than over agricultural cells. The water quality of habitat cells, however, is not included in the measure of landscape water quality used in this study.

4.2 Habitat Objective

Having formulated the production function that calculates species survival as a function of habitat quantity and configuration, it is a straightforward task to identify what types of cells will contribute to survival probability when retired. Characteristics of a cell that are important, regardless of the landscape, are: whether or not the cell forms a corridor between two existing patches, whether or not the cell is adjacent to an existing patch, and whether or not the cell is adjacent to the existing largest patch. Exactly how important each of these characteristics is, however, will depend on the configuration of a particular landscape.

This section explores the relative effectiveness of each of these measures in explaining the suitability of an agricultural cell for conversion for the purposes of optimizing habitat quality. As in the sections above, estimating this effectiveness required regressing a measurement of how well each agricultural cell performed toward the habitat objective on a number of cell characteristics.

Table1: Regression results for the hydrology objective.

variable	estimated coefficient	t-statistic
region 1	0.5848	50.17
region 2	0.3174	23.25
region 3	0.1151	7.28
region 4	0.1562	25.72
region 5	0.4204	32.57
region 6	0.5753	48.65
region 7	0.6701	64.81
X_x•region 1	0.550	10.02
X_3•region 2	0.0942	18,97
X_3•region 3	0.1136	21.91
X_3•region 4	0.0889	47.02
X_3•region 5	0.0221	5.15
X_3•region 6	-0.0085	-1.93
X_3•region 7	-0.0191	-4.02

Dep variable: hydrological effects derived from retiring a cell in the first time period, with effects measured in the fourth time period as deviations from the ideal ($d_{2,1}$)

$\overline{R}^2 = 0.644$

As above, the first step involved exploring the importance of absolute location in determining a cell's suitability for the habitat objective. Dummy variables by cell or region explain almost none of the variation in habitat suitability; the $\overline{R}^2$ for the regression on only regional dummies is 0.01. Because the existing habitat is randomly placed on a landscape, this result is not unexpected. The coefficients on the dummy variables are all quite close to 0.81 and are highly significant; these coefficients represent the mean deviation from the survival ideal achievable by retiring a parcel of land in that region. A deviation of 0.81 represents a survival probability of 0.19 (since the ideal survival probability is 1). This would be the case for any landscape with a total of 36 cells in habitat, including a single large cluster of 14 cells.

It is intuitive that in the case of the habitat objective relative location– in particular position with respect to existing habitat– would be much more important than absolute location in determining a cell's habitat suitability. There are a number of possible relative location measures that might be efficient predictors of a cell's habitat suitability, including whether a cell is adjacent to the an existing patch of habitat on the grid, whether the cell is adjacent to the largest patch of habitat on the grid, whether a cell forms a habitat bridge and if so how many cells it connects to an existing patch, and how many habitat cells are immediately adjacent to the base cell.

The production function that determines survival probability provides a great deal of information about which measures would be most efficient at explaining variability in survival probability. The only inputs into the production function are total amount of habitat and the connectivity of that habitat, but when only a single cell is retired the total amount of habitat will be constant regardless of which cell it is. Therefore, all of the variation in survival probability arises from differential impacts on the connectivity of habitat in the grid. Clearly, a linear regression of d_3 on the change in connectivity is not a perfectly accurate representation of the true relationship between the habitat objective and landscape connectivity. However, the OLS model is nevertheless able to explain quite a bit of the variability in habitat suitability. The OLS model

$$d_3 = \alpha_3 + \beta_4 X_4 + \beta_5 X_4 X_5 + e_3 \tag{17}$$

where

X_4= a dummy variable indicating if the cell is adjacent to the largest cluster

and

X_5=a variable indicating whether the cell forms a habitat corridor and how many cells it connects to an existing patch.

provides the results shown in table 2.

Table 2: Results of regression on the untransformed habitat variable d_3.

variable	estimated coefficient	t-statistic
constant	0.8156	11009
adjacent	-0.028	-145.98
adjacent • no. of cells connected	-0.028	-329.51
dep. variable: deviation from the habitat ideal resulting from retirement of each cell (d_3)		
$\overline{R}^2 = .99$		

According to these results, retirement of any cell on the landscape will result in a maximum deviation of 0.8156, or a minimum survival probability of 0.1844. Each additional unit of habitat connectivity (i.e. each additional unit attached to the landscape's largest cluster), raises survival probability by 0.028, and therefore decreases the deviation from the habitat ideal (d_3) by 0.028. Although these coefficients represent average, rather than marginal, effects of additional landscape connectivity, they appear to be reasonable predictors of habitat suitability $(\overline{R}^2 = .99)$. An efficient measure of relative location with respect to the habitat objective is therefore the total number of additional units attached to the largest habitat cluster by retirement of a given cell, which is captured by X_4 in combination with X_4X_5.

4.3 Cost Objective

An examination of the cost objective blurs the distinction between measures of absolute and relative location. When land is being retired only in the first time period, there is no hydrological contamination to complicate the calculation of land values. The only variable on which cost depends in the first time period is $\frac{1}{D_{i,j}^2}$, where $D_{i,j}$ represents the distance between cell i and the urban cell j. Because the urban cell occupies a fixed location on each landscape, this measure is constant for any cell i, regardless of landscape. Therefore regressing cost performance on dummy measures for cell yields an $\overline{R}^2 = 1.0$. This would suggest that measures of absolute location are important, whereas in fact what is important is distance from urban center, which is more appropriately considered a measure of relative location.

It is possible to recreate the formula for cost of retirement in the first period simply by regressing the cost of retiring each cell on a constant and a variable X_1 which represents the value

$\frac{1}{D_{i,j}^2}$:

$$d_{1,1} = \alpha_{1,1} + \beta_{1,1} X_1 + e_{1,1} \tag{18}$$

The results of this regression appear in table 3.

Table 3: Regression results for the cost objective in the first time period.

variable	estimated coefficient	t-statistic
constant	0.4167	6.7×10^{14}
X_1	0.4167	4.0×10^{14}
dep. variable: cost of retiring a cell in the first time period measured as deviations from the ideal.		
R^2=1		

For the cost objective, the target, or "ideal," cost established is zero, and the anti-ideal is 3000. The coefficient of 0.4167 on the constant suggests that retiring any cell will cause the cost to deviate from the objective by at least 0.4167– or $1250. This value is equal to the net present value of expected agricultural returns on an uncontaminated cell. As distance to the cell from the urban area increases, the inverted measure will decrease, and the positive coefficient on the inverted distance measure suggests that deviations from the objective will decrease as well; in other words, cells become less costly with distance, as one would expect. The fact that the coefficients are equal means that at a distance of 1 (i.e. immediately adjacent to the urban area), agricultural production value and urban influence contribute equally to a cell's price. The coefficient on the distance variable captures the magnitude of the other parameters included in the price equation–the constant appearing in land value equation XX, which has been set equal to 2.5, and the population of the urban cell, N_j.[3]

4.4 Point Schemes

The results of the regressions above can be combined into a point scheme that would enable decision-makers to rank cells on an out-of-sample landscape. The measurements made on the new

---

[3]The results of exploring relevant measurements for the cost objective become more interesting when land is retired in later periods. In later periods, the cost of land reflects the effects of increasing contamination on agricultural productivity, as well as the urban influence on land prices. Therefore, in order to accurately explain the cost of land retired in later time periods, the regression must include measures of location that are able to capture which cells are likely to experience increased contamination over time. This research does not provide those results.

landscape must be the same as those used in the regression above, and the points assigned to those measured characteristics will be equal to some function of the estimated coefficients. The coefficients indicate contributions toward deviations from the ideal, so the units are comparable across objectives. Weights for the different objectives and a decision-making framework for combining objectives must still be selected, however, if the point scheme is to accurately reflect the decision-maker's preferences and the composite objective.

Weights and a decision-making framework will be used to manipulate the regression equations above, and points will then be established for the different measurements, or predictors, that will result in a ranking of cells that mimics as closely as possible how the cells would actually be ranked if the full simulation model were run on the random landscape. For simplicity, suppose first that the decision-maker has selected a linear compromise programming framework with equal weights. According to this framework, a decision-maker's objective is to minimize the weighted sum of the deviations from the ideal. The deviations from the ideal are the dependent variables in the regressions above, so the decision-maker's objective, given equal weights, is simply to find the cell whose characteristics minimize $\frac{(d_1 + d_2 + d_3)}{3}$. The point scheme is therefore straightforward-- simply award each characteristic a weight-adjusted amount of the estimated coefficient, total its score, and find the cell with the smallest overall deviation.

All of the landscape scores can be calculated using the following equation:

$$S = \frac{d_1}{3} + \frac{d_2}{3} + \frac{d_3}{3} = \frac{I_1 \cdot \alpha_1}{3} + \frac{I_2 \cdot \alpha_2}{3} + \frac{\alpha_3}{3} + \frac{\beta_1 X_1}{3} + \frac{\sum_r (\beta_2^r \cdot i_1^r \cdot X_2)}{3} + \frac{\sum_r (\beta_3^r \cdot i_2^r \cdot X_3)}{3} + \frac{\beta_4 X_4}{3} + \frac{\beta_5 X_4 X_5}{3}. \qquad (19)$$

The matrix notation in this equation obscures its underlying simplicity. Consider the equation for the score of a particular cell i in region r:

$$s_i = \frac{\alpha_1^r}{3} + \frac{\alpha_2^r}{3} + \frac{\alpha_3}{3} + \frac{\beta_1}{3} x_1 + \frac{\beta_2^r}{3} x_2 + \frac{\beta_3^r}{3} x_3 + \frac{\beta_4}{3} x_4 + \frac{\beta_5}{3} x_4 x_5 \qquad (20)$$

where the values of the αs and βs are given by the individual regressions, and the lower case x values refer to cell i's values for the measures of relative location. The "points" assigned to each of the k characteristics of relative location are set equal to $\frac{ds_i}{dx_k}$; in the linear case with equal weights they are all simply $\frac{\beta_k}{3}$. These points represent the value of characteristics of relative location, while $\frac{ds_i}{d\alpha_g}\cdot\alpha_g$ represents the value of absolute location toward each of the g objectives. The cells are ranked in desirability from smallest to largest S.

How does the point scheme change if the decision-maker chooses instead the nonlinear compromise programming approach? The principles all remain the same, though the algebra becomes more complicated. The score prediction equation now appears below:

$$
\begin{aligned}
\mathbf{s}_i &= \mathbf{w}_1(\alpha_1^r + \beta_1\mathbf{x}_1 + \beta_2^r\mathbf{x}_2)\mathbf{e}^{\lambda_1(\alpha_1^r + \beta_1\mathbf{x}_1 + \beta_2^r\mathbf{x}_2)} \\
&+ \mathbf{w}_2(\alpha_2^r + \beta_3^r\mathbf{x}_3)\mathbf{e}^{\lambda_2(\alpha_2^r + \beta_3^r\mathbf{x}_3)} \\
&+ \mathbf{w}_3(\alpha_3 + \beta_4\mathbf{x}_4 + \beta_5\mathbf{x}_4\mathbf{x}_5)\mathbf{e}^{\lambda_3(\alpha_3 + \beta_4\mathbf{x}_4 + \beta_5\mathbf{x}_4\mathbf{x}_5)}
\end{aligned}
\tag{21}
$$

When a nonlinear decision-making framework is selected, the points assigned to each cell's characteristics are no longer constant landscape-wide. Instead the points assigned depend on the level of performance that cell achieves toward each objective. Whereas with the linear case $\frac{dS}{dX_k}$ is a constant that is some weight-adjusted function of the estimated coefficients, and $\frac{dS}{dX_k}$ and $\frac{dS}{d\alpha_g}\cdot\alpha_g$ are independent of one another, this simplicity disappears with the nonlinear case. As an example, consider the points awarded to the variable representing distance from the urban center (X_1):

$$
\frac{ds_i}{dx_1} = \beta_1 w_1 e^{\lambda_1(\alpha_1^r + \beta_1 x_1)}(1 + \lambda_1(\alpha_1^r + \beta_1 x_1)). \tag{22}
$$

The points are again a weight-adjusted function of the estimated coefficient ($\beta_1 w_1$), this time with an additional adjustment factor that is a function of the overall objective level, and therefore on the current level of X_1. Just as in the linear case, the sign of this function will be determined by the sign of β_1. It would not be possible, therefore, to derive a landscape-wide scoring system that is independent of a particular cell's characteristics X_1.

The second derivative of the score function indicates how the points assigned to a cell characteristic change with the level of that particular characteristic:

$$\frac{d^2 s_i}{dx_1^2} = w_1 \beta_1^2 e^{\lambda_1(\alpha_1^r + \beta_1 x_1)} \lambda_1 (2 + \lambda_1 (\alpha_1^r + \beta_1 x_1)). \tag{23}$$

This function is easy to sign, as all of its components must be positive.[4] The interpretation of a positive sign is a bit roundabout, however. If a characteristic X_k is a constructive characteristic in the sense that increasing it increases a cell's ability to realize a particular objective, then the coefficient on that variable (and the points assigned the variable) will be negative– that characteristic decreases the distance from the ideal. According to the second derivative, if the constructive X_k increases, then the points assigned that characteristic should also increase– i.e. become less negative. Assigning fewer negative points to a constructive variable de-emphasizes the importance of that variable relative to the other variables, and of that objective relative to the other objectives in calculating the composite score. However, if X_k is a destructive characteristic, then it is assigned positive points, for increasing it inhibits achievement of an objective and increases d_3– the deviation from the ideal. Therefore, according to equation 23 (above), as the level of the destructive characteristic X_1 increases, more points are assigned to it, increasing the relative contribution of both that characteristic and that objective to the composite score.

5. CONCLUSION

Goal programming provides a flexible way for decision-makers to incorporate multiple objectives into a single composite objective. This decision-making framework translates a number of individual objectives into a single overall composite objective and measures how far the results of each strategy fall from the composite objective. In this study, this framework was applied to a management scenario with multiple objectives in order to explore how important different characteristics of absolute and relative location are at contributing to individual objectives and how that information can be used to create a ranking system with respect to an overall objective.

The importance of absolute location of parcels varies by objective. In the case of the hydrology objective, information about where on the landscape a parcel falls (without information about surrounding land uses) can be very useful in predicting the success of that parcel's retirement. Information about absolute location is not helpful at all in predicting a parcel's performance toward the habitat objective, however. Performance toward the habitat objective is entirely explained by a cell's position relative to other land uses.

Among the measures of relative location explored, the most effective single measure at

---

[4]Either or both of the β may be negative, but $\alpha_1^r + \beta_1 x_1$ must (theoretically) lie between 0 and 1. Failure to do so is due to error in the estimation of the coefficients.

explaining performance toward the hydrology objective is a distance-weighted measure of surrounding habitat. This characteristic is a destructive characteristic, in that the higher the measure of surrounding habitat, the poorer the cell's ability to improve measures of landscape hydrological quality once retired. The cost objective, when cells are retired in the first time period and their land value is not yet affected by hydrological contamination, requires only a measure of cell's distance relative to the urban center. The habitat objective depends almost entirely on a cell's ability to form a habitat stepping stone.

The individual regressions provide comparable coefficients across objectives, so that one could rank units of each characteristic by importance if in fact the value of a one-unit improvement in objective a were equivalent to the value of a one-unit improvement in objective b. Departures from this assumption of equal objective value are captured in the form of the MCDM framework–the weights and distance function used. The MCDM framework selected must therefore be used in conjunction with the regression results in determining the relative importance of different characteristics and the points that should be assigned in establishing a parcel ranking scheme.

As diverse environmental objectives become more important in the evaluation of public policy alternatives, it is imperative that they be systematically and explicitly incorporated into the decision-making process. Establishing a framework such as goal programming requires policy makers to address each objective individually and to establish for those objectives target values and structures of relative importance with respect to the other objectives. In return, the framework provides policy makers with the mechanism necessary to evaluate the acceptable tradeoffs implied by the prioritization decisions that they make and with an unambiguous means of ranking management strategies and justifying policy decisions in terms of overall management objectives.

6. REFERENCES

Andren, Henrik. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: A review. OIKOS 71: 655-66.

Gustafson, Eric J. 1998. Quantifying landscape spatial pattern: What is the state of the art? Ecosystems 1: 143-56.

Haight, Robert G. 1995. Comparing extinction risk and economic cost in wildlife conservation planning. Ecological Applications 58(3): 767-75.

Marshall, Elizabeth P. 1998. Incorporating environmental uncertainty into species management decisions: Kirtland's warbler habitat management as a case study. Conservation Biology 12(5):975-85.

Mathworks. 1999. Matlab programming software, version 5. www.mathworks.com.

Shi,Yue Jin, Timothy T. Phipps, and Dale Colyer. 1997. Agricultural land values under urbanizing influence. Land Economics 73(1): 90-100.

Weisberg, Sanford. 1985. Applied Linear Regression, 2nd Edition. New York: John Wiley & Sons.

With, Kimberly A., Robert H. Gardner, and Monica G. Turner. 1997. Landscape connectivity and population distributions in heterogeneous environments. OIKOS 78: 151-69.