Effects of landscape spatial structure and composition on models of pine warbler (*Dendroica pinus*) habitat selection

Amy Gilboy
School of Natural Resources and Environment
University of Michigan
agilboy@umich.edu

Abstract

An emerging application of GIS within ecology is the construction of fundamental niche models to predict possible geographic distributions of species. Through the construction of these models, geographic information science (GIS) can be used to map possible species' habitat. These maps can then aide in wildlife conservation by predicting new areas of possible habitat use for threatened and endangered species, as well as assessing existing protected areas.

One quantitative approach that has shown to be robust is genetic algorithms, which have been implemented in the Genetic Algorithm for Rule-Set Production (GARP) modeling program. The aim of this study is to test the value of forest structural information (both horizontal and vertical), over vegetation type alone, in predicting bird species occurrence through the use of the Desktop GARP modeling program.

Field-based point observations of pine warbler occurrences were used as input to GARP to create a series of models.   The pine warbler is a habitat generalist and usually occurs in mature upland coniferous forests.

Models utilizing vegetation type were created and compared to models made with the addition of landscape spatial structure. The predictive abilities of the models were evaluated by comparing predictions with a subset of observations that was set aside and not used in model creation. Also, the output maps were compared to a map interpreted from the Northwoods Database, a database that lists habitat for species in the Upper Great Lakes Region. Models run with horizontal and vertical structure as input produced greater variability among prediction as well as finer niche maps than the model run with vegetation type alone. The results of the study can be used in the creation of a new generation of habitat models, which incorporate more detailed information about vegetation structure.

Introduction

Prior to 1960, ecological studies generally assumed spatial homogeneity. Spatial variation was beyond the capacity of most simulation tools and space was simply ignored (Wiens 1995). However, since the development of the theory of island biogeography (MacArthur and Wilson 1967), the landscape has quickly become a main focus of ecological research. The term *landscape* refers to a heterogeneous land area composed of interacting ecosystems of different land-cover/land-use (lc/lu) types. The field of landscape ecology is based on the idea that if the composition and spatial arrangement of the landscape were different, ecological processes, both biotic and abiotic, would also be different (Wiens 1995). For example, various spatial patterns of vegetation in the landscape have shown to influence distribution patterns among animals (Trzcinski et al. 1999, With and King 2001).

The development of geographic information science (GIS) and the availability of satellite imagery have escalated studies of landscapes patterns and their effects on ecological processes. An emerging application of GIS within ecology is the construction of fundamental niche models to predict possible geographic distributions of species. A fundamental niche describes the range of environmental conditions that a species requires (Hutchinson 1957). A species may be limited by interaction with other species due to competition, predation, or parasitism which can decrease its niche into a smaller subset known as the realized niche (Hutchinson 1957). The fundamental niche may be affected by landscape heterogeneity and patterns, such as fragmentation, patch size, and edge effects (Blake and Karr 1987, Mazerolle and Villard 1999).

Increasingly, both GIS and remote sensing are being employed with bird location data (Lyon 1983, Herr and Queen 1993, Austin et al. 1996, Hepinstall and Sader 1997, Debinski et al. 1999, Osborne et al. 2001). These advancements have enhanced both the quality and quantity of information that can be incorporated into ecological niche modeling. More detailed maps of geographic distributions provided by quantitative spatial models can lead to better assessment of established nature areas, the improvement of ground surveys and the discovery of more precise ranges for endangered species (Stockwell and Peters 1999). Also, once suitable habitat has been identified, spatial models can aid in prediction of the likely affect of future land-use change (Austin et al. 1996).

One quantitative approach that has shown to be robust is the use of genetic algorithms. Genetic algorithms are based on the theory of evolution through natural selection, and the idea that solutions to problems evolve the same way organisms evolve. A set of possible solutions to a problem are formed and, through a series of iterations, the solutions are modified and tested until the best solution is found (Stockwell and Noble 1992). The possible solutions are different types of rule sets. One program in which genetic algorithms have been implemented is the Genetic Algorithm for Rule-Set Production (GARP) modeling program. GARP inputs point locations of observations of species and environmental layers from raster grids and produces maps of possible habitat. When compared with other modeling methods, published results suggest that GARP has greater predictive capability (Stockwell and Peterson 2002).

Recently, management of avian species across broad heterogeneous scales has become a focus of research (Collins 1983, Blake and Karr 1987, Bolger et al. 1997, Azevedo et al. 2000). In the context of ecological niche modeling of birds, it is important to consider the landscape structure in addition to other environmental factors as a principal component of habitat preference. Spatial composition and variability of vegetation contribute to the horizontal structure (McGarigal and Marks 1995), while height and biomass describe the volumetric aspect of vegetation. When these components of vegetation are combined, a realistic multi-dimensional aspect of habitat can be illustrated.

In this paper, point locations of the pine warbler (who typically exploit mature pine forests) were used as input to GARP along with different environmental layers to create a series of models. Habitat models were run in an iterative fashion within GARP, adding with each model more input concerning forest structural information (both horizontal and vertical) to aid in niche modeling of the pine warbler. It is the intention of this paper to refine niche modeling at a regional scale by utilizing GIS and Desktop GARP, as well as examining the predictive capabilities of GARP at a landscape level.

To date, GARP has been run only at state, national, or continental scales (Godown and Peterson 2000, Peterson et al. 2000, Peterson 2001, Anderson et al. 2002, Feria and Peterson

2002, Peterson et al. 2002). This study is the first usage of GARP known at a landscape scale. Factors such as precipitation and climate have been used as input in previous studies, but the small area of this study assumes those factors to be constant. For this reason, only layers based on vegetation were implemented within GARP.

Study site

The study area is approximately a 20X20 kilometer swath located mostly within the Hiawatha National Forest, Eastern Division, in the Upper Peninsula of Michigan. Intensive vegetation surveys were conducted in four-hectare homogenous stands (Bergen et al. 1995) which included the following vegetational composition and structure: red pine (*Pinus resinosa*) – seedling, sapling, pole, mature; jack pine (*Pinus banksiana*) – seedling, sapling, pole, mature; eastern white pine (*Pinus strobus*) – mature; black spruce (*Picea mariana*) – mature; northern white cedar (*Thuja occidentalis*) – mature; aspen species – sapling, pole, mature; and northern hardwoods – pole, mature. Northern hardwood species include old-growth maple (*Acer*), beech (*Fagus*), and birch (*Betula*).

The study site contains a diversity of physiographic regions. The Raco Plains, found in the northeast quadrant of the site, is an area of excessively drained glacial outwash sands bordered by agricultural lake plain. A large poorly-drained wetland dominates the southeastern area of the site. The western half is composed of well-drained moraine features and low-lying poorly-drained areas. Forested moraine till continues west outside the site and the northern edge is bordered by Lake Superior (Bergen et al. 1995).

The mean annual temperature for the region is 5° C (41° F). The average temperature in July is 24.5^o C (76.1^o F) and in January is 14° C (6.8^o F). The mean annual precipitation is 79 centimeters (31.1 inches) (Bergen et al. 1995).

Methods
Bird Data Collection and Vegetation Surveys

Bird community surveys were conducted within the previously surveyed test stands from July 11 through August 8, 2002 from sunrise until approximately 11:00 a.m. Circular plot surveys (Reynolds et al. 1980) were performed within the four-hectare stands, as near to the center as possible. Once the center was located, a GPS coordinate was taken, and the azimuth and north direction were noted. The surveys lasted five minutes to remain consistent with the protocol of other bird surveys being conducted in Hiawatha. When a bird was detected, the species was noted along with an approximate distance and direction. After completion of the surveys, GPS points were differentially corrected with data from the base station located in Newberry, Michigan.

Between June 11 to July 28, 2002, surveys were also being conducted by the Forest Service within the study site and data from these surveys was used for analysis (Langstaff 2002). These surveys were generally located within a 100-meter radius of homogenous vegetation, mostly jack and red pine stands. The location of the survey was recorded using a GPS receiver and the surveys were five minutes in length. Any birds detected were identified along with proximity to surveyor. The combined surveys detected 60 individual pine warblers which were used as input into GARP.

Minimal vegetation surveys were conducted to verify stand vegetation and assess for growth and disturbance since the original vegetation surveys (Bergen et al. 1995). Upon arriving at the survey site, a description of the stand was first noted (i.e., young jack pine) as well as disturbance (i.e., thinning of stand). Within a specified radius around the survey point, diameter-at-breast height (dbh), height, and crown height were noted for all non-herbaceous plants.

The Northwoods Database

The Northwoods Database (Northwoods) (Benyus et al. 1992) was used to aide in the validation of maps produced by GARP. Northwoods lists habitat associations of wildlife that inhabit the Upper Great Lakes region of Minnesota, Wisconsin, and Michigan. Northwoods was compiled from wildlife habitat inventories from seven National Forests in the region. Experts from each Forest assembled a database containing information about the wildlife community such as species' names and habitats. These habitat relationships were based on literature reviews and field experience. A majority scoring system was used to determine whether or not a species would be considered present for each habitat listed. If 50% or more of the Forests listed a species as present in a certain habitat, then the species was listed as present in Northwoods. The pine warbler is listed in Northwoods as having main habitat – mature upland coniferous forests.

Spatial Analysis Methods
Desktop GARP

Desktop GARP was implemented in the modeling effort for the pine warbler. The GARP modeling system works through a set of eight programs: RASTERIZE, PRESAMPLE, INITIAL, EXPLAIN, VERIFY, PREDICT, IMAGE, and TRANSLATE (Payne and Stockwell 1996).

The first two steps, RASTERIZE and PRESAMPLE, prepare the input data for use in GARP. RASTERIZE converts species point data into contiguous raster layers. This step compresses information by clearing the data of duplicates caused by localized intensive sampling.

PRESAMPLE takes the newly created raster layers and creates training and testing data sets by randomly sampling the data set prepared in RASTERIZE. The training set is necessary to construct a model while the testing set allows for the assessment of the model's accuracy. PRESAMPLE outputs a set of 2500 points, 1250 of which are re-sampled from actual location points to create a large amount of data representing presence. The other 1250 are re-sampled from the total geographic space to replicate absence data, termed *background*.

After the training set is generated, it is input into the next program INITIAL. This creates an initial model which is the starting point for the GARP algorithm. The initial model is a set of rules that influence the development of the subsequent models.

There are four types of inductive rules that are the basis for modeling: atomic, BIOCLIM, range and logit rules. The simplest form of the rules is the atomic rule. In this case, a model is developed when only a single variable within the precondition of the rule is used. An example **atomic rule** would be: *if the average snowfall is 12 inches then the species is present*. The second format of rules is the **BIOCLIM rules**. This type of rule originates from the BIOCLIM biogeography model (Nix 1986) and is based on the idea that species have ecological tolerances beyond which they cannot survive. A model is developed by enclosing the range of the environmental values in an envelope where species may occur. If a point is outside the range of tolerance, then the species is predicted to be absent. The third type of rule, a **range rule**, is a

generalization of the BIOCLIM rule which allows for negation. The final class of rules, **logit rules,** are based on logistic multiple regression models where there is a positive dependence between events.

The fourth module EXPLAIN applies the genetic algorithm to improve the primary models and then outputs the best of these models. The GARP genetic algorithm behaves differently than other genetic algorithms because it creates a rule archive and does not converge on a single rule. This allows GARP to utilize different rules and select the best to create each model output. After the best rules are placed within the archive, the program checks the archive for any considerable changes. If the archive has not changed, the program terminates. If the rule archive has changed significantly, the program will continue to create a new population by modifying archived rules with genetic recombination, known as heuristic operators. Three heuristic operators may be utilized in the EXPLAIN module: *join*, *crossover*, and *mutate*. *Join* is simply the joining of two rules to produce a longer rule. The *crossover* operation occurs when two rules exchange a part of their binary code. In this way, two new rules are created. The *mutation* operator can change a rule by randomly changing a single value. After new rules are made by genetic recombination, GARP measures the fitness of the new rules and the more successful an operator is, the more it will be used in future generations.

The fifth module in GARP is VERIFY. This program tests the predictive accuracy of the training set on the test data set that was created in PRESAMPLE. In this way, accuracy is independent of the data used to formulate the rules and thus a more reliable estimate of how well the rules worked.

The next module PREDICT takes the newly created model and forms a prediction for each cell within the raster data set. A probability prediction exists for each rule whose precondition pertains to a particular cell. Types of output from PREDICT include predictions and uncertainty, areas where rules conflict, and the probability of occurrence.

The seventh module IMAGE takes the calculations produced in PREDICT and converts them into image formats for visualization. Finally, the TRANSLATE function screens the rule sets and eliminates rules which were not used to make predictions.

Data Pre-processing

Using ArcView 3.3 software, each survey point was plotted on a map of the study site using UTM Zone 16 coordinates, NAD 27 datum. The pine warbler was then plotted in its own shapefile according to the estimated distance and direction from the surveyor. Once the shapefiles were completed, they were converted to the geographic projection for compatibility with Desktop GARP.

The base environmental data (raster grids) used to create the models included a lc/lu layer (representing horizontal structure), a biomass layer (representing vertical structure) and two different layers incorporating surrounding vegetation (representing horizontal structure).

The lc/lu classification was obtained from the Gap Analysis program and represents vegetation. It was derived from Landsat TM scenes between 1992 and 1994 and its resolution is 30x30 meters. The original raster grid was first cut to fit the study area and smoothed in ERDAS Imagine software to rid the image of excessive patchiness. Next, the image was reclassified to include the following nine lc/lu classifications: upland conifer (jack, red, and eastern white pine), lowland conifer (cedar, spruce, fir, tamarack), northern hardwoods (beech, maple, birch, mixed deciduous/coniferous), aspen and lowland deciduous (aspen, lowland deciduous, lowland aspen,

lowland shrub, mixed lowland deciduous), grassland, agriculture, wetland, open water, and urban/barren (Figure 1).
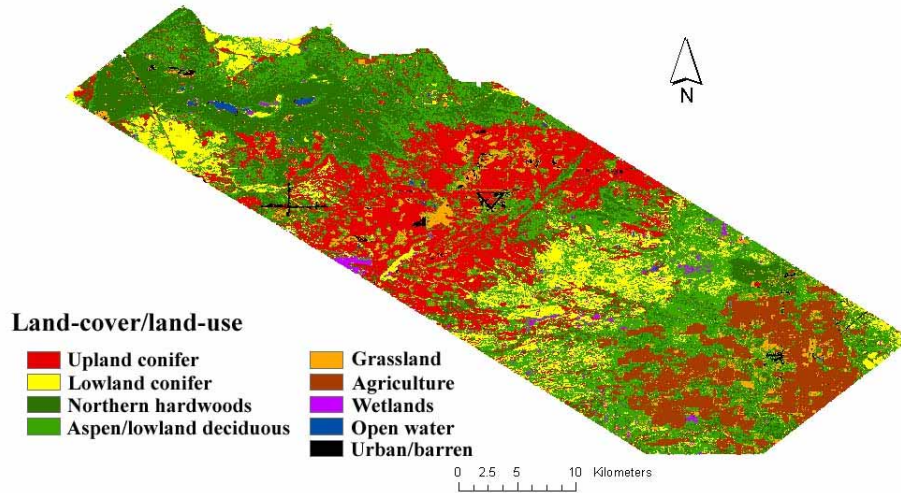
Figure 1: *Land-cover/land-use classification of study site.*

A biomass image was available from a previously completed biomass project (Dobson et al. 1995, Bergen et al 1998) and utilized to represent age and vertical structure. The pixel size was originally 20x20 meters and was resampled to 30x30 meters. The biomass pixel values range from 0-30 kg/m^2 and were transformed to the range 0-255, with a lower number indicating lower biomass. High biomass values corresponded with northern hardwoods while the lowest values corresponded with agriculture and grassland areas.

To represent horizontal structure (vegetation surrounding the bird points), two different raster grids were created in Arc Workstation using neighborhood statistics. Two neighborhood layers were created for this analysis – variety and majority. A variety statistic determines the number of unique values within a neighborhood. The majority statistic determines the value that occurs most often in the neighborhood. In this study, the neighborhood was defined by a circle with a 6-cell radius, approximately 180 meters surrounding the point location in all directions. The variety and majority statistics were determined using the original, non-smoothed vegetation grid.

Before any of the environmental layers could be employed, they were modified to be compatible with Desktop GARP. GARP uses latitudinal and longitudinal ASCII grids, so first the grids were re-projected from the UTM coordinate system to the geographic coordinate system in Arc Toolbox. Next, the raster grids were converted into ASCII grid format in Arc Workstation for input into GARP.

The Desktop GARP interface has a number of default settings that were used in this analysis. The individual species data points were divided in half by GARP for training and testing data. Twenty runs using a maximum of 1000 iterations were completed, each of which output 20 separate maps ranging in variability. All four rule types were included in the models.

GARP was run for the pine warbler utilizing three different combinations of environmental grids. First, GARP was run to model the occurrence of the pine warbler with vegetation as the only environmental input. Next, GARP was run with both the vegetation and the biomass layer to assess the added value of vertical structure on model prediction. GARP was

then run a third time combining vegetation, biomass, and neighborhood statistics to test the further influence of horizontal spatial structure on model prediction.

Each time the model was run, 20 separate maps were produced as ASCII grids. These grids were imported into ArcMap for visualization purposes and then summarized into one composite map used for the final display results. Model accuracy was assessed in three ways: 1) a chi-square analysis was calculated for the raw data; 2) training, testing, and *p-values* were observed and compared to the pine warbler maps modeled with only half the total location points; and 3) the GARP maps were analyzed by tabulating areas of lc/lu chosen as presence compared with areas interpreted as presence in Northwoods.
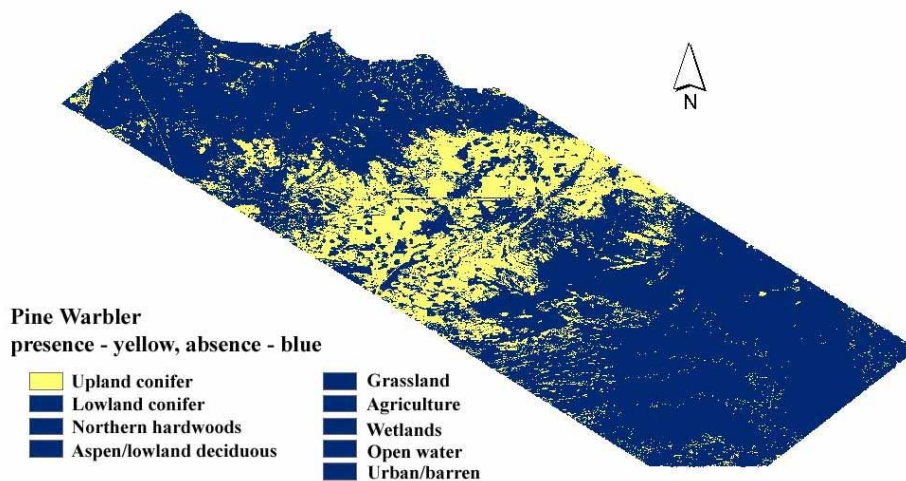
Results
Chi-square of raw data

Chi-square tests were performed on the author's bird survey data to determine if species and forest type were related. Two chi-square tests were performed - one on total number of individuals and one on the total number of sites where the species was noted, regardless of how many individuals were present.

The vegetation types surveyed were grouped into three categories for the purpose of the chi-square analysis – mature pine, young pine, and mature deciduous. The author's surveys detected 24 individual pine warblers at 14 different sites. Fifteen individuals were found in mature pine, nine in young pine, and none in deciduous. A significance value of 0.05 was selected and values less than this indicated a relationship between bird and forest type. Chi-square analysis for individuals and sites produced *p-values* approximately equal to 0.0.

### *Map constructed from Northwoods*

According to Northwoods, *only mature upland coniferous forest* is listed as habitat for the pine warbler. A predicted presence/absence map was made from this information and only upland conifer was selected as *presence*, represented by yellow on the map (Figure 2). Blue areas designate *absence*.
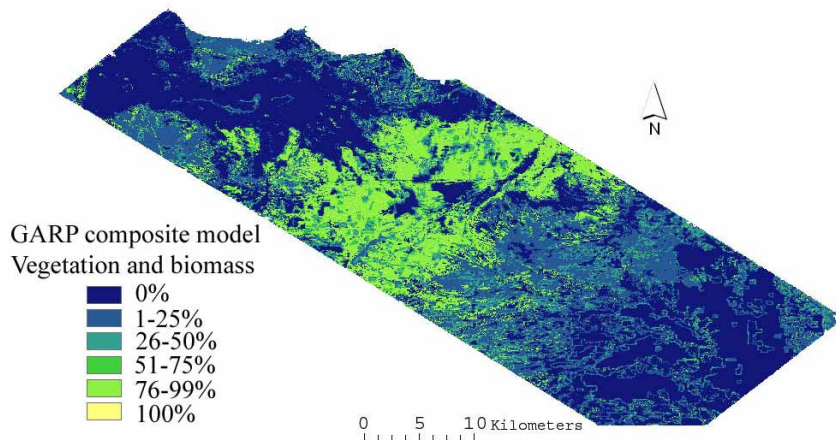

Pine Warbler
presence - yellow, absence - blue

- Upland conifer
- Lowland conifer
- Northern hardwoods
- Aspen/lowland deciduous
- Grassland
- Agriculture
- Wetlands
- Open water
- Urban/barren

Figure 2: *Map of pine warbler habitat interpreted from the Northwoods Database.*
Desktop GARP output

Each run of GARP produced 20 separate maps and these maps were combined to form a composite. These composite maps are presented below in Figures 3-5. The maps display six categories representing presence – 0%, 1-25%, 26-50%, 51-75%, 76-99%, and 100%. These categories correspond to how many times each pixel in the twenty separate maps appeared as presence upon completion of the composite. For example, if a pixel falls into the 26-50% category, that pixel was classified as *presence* in 26-50% of the twenty maps.

The three maps presented below are the pine warbler and the different output produced with three different combinations of environmental layers – vegetation only (Figure 3), vegetation and biomass (Figure 4), and vegetation, biomass, and neighborhood values (Figure 5).
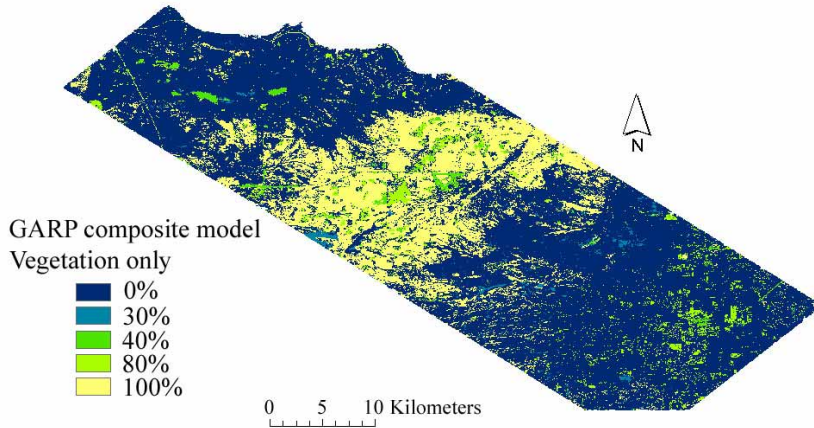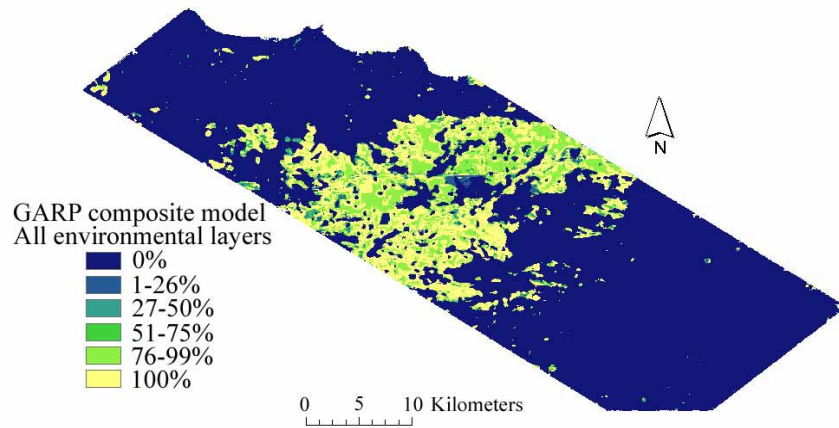

Figure 3: *GARP composite model of pine warbler locations modeled with vegetation only.*

Figure 4: *GARP composite model of pine warbler locations modeled with vegetation and biomass.*

Figure 5: *GARP composite model of pine warbler locations modeled with vegetation, biomass, and neighborhood layers.*

Desktop GARP results output

During execution, Desktop GARP creates two main output objects: the prediction maps and the results table (Scachetti-Pereira 2002). The results table contains information regarding statistical tests and training and testing accuracies. The training accuracy is calculated using the training data points. Accuracy is determined using the percent correctly classified. Testing accuracy is calculated using the data points which were set aside by GARP for accuracy assessment of the training model. In this way, points used to make the model are not being used to test the model.

A *p-value* is also calculated by GARP and is the result of a chi-square test which uses test points and the area predicted present by the output model. These values represent the probability of getting a model at random that would perform as well as the model generated by GARP. Values less than 0.05 would indicate that GARP performs better than a random model. Training and testing accuracies for all runs of GARP ranged from 69 – 83% for training and 70 – 80% for testing. The *p-values* for all three runs were less than 0.05.

When half the locations were withheld from the training set by the author (runs modeled with 30 points), values ranged from 69 - 84% for training and 61 – 80% for testing. However, *p-values* were greater than 0.05 for two out of three runs. The vegetation model had a *p-value* of 0.11, while vegetation and biomass produced a *p-value* of .14.

Vegetation percent area tabulations

To compare GARP output to the map constructed from Northwoods, the GARP maps were first overlayed with the lc/lu layer in ArcMap. Through *raster calculator*, the maps were broken down into categories to determine how often certain vegetation was predicted in the composite map as absence, presence, or categories in between. The percentages in the following tables represent the breakdown of lc/lu type and how often the different vegetation types were predicted as habitat (Tables 1-3). These figures were determined by dividing the number of cells of each lc/lu within each category (i.e., 5-25%) by the total number of cells for each lc/lu within the entire image and multiplying by 100 to arrive at a percent. If a value is in the "0%" category,

9

the percent value listed is how often that lc/lu was never chosen as *presence* (or chosen as *absence*). Conversely, if a value is in the "100%" category, that percent value corresponds to how often that lc/lu was chosen as *presence* in all of the 20 separate maps.

Pine Warbler: vegetation layer only	*Breakdown of presence in composite map (%)*				
Land-cover/land-use	0%	30%	40%	80%	100%
Upland conifer	-	-	-	-	100
Lowland conifer	100	-	-	-	-
Northern hardwoods	100	-	-	-	-
Aspen/lowland decid.	100	-	-	-	-
Grassland	-	-	-	100	-
Agriculture	100	-	-	-	-
Wetland	-	100	-	-	-
Open water	-	-	100	-	-
Urban/barren	-	-	100	-	-

Table 1: *Lc/lu and percent each appeared present in the composite map of the pine warbler run only with a vegetation layer.*

Pine Warbler: vegetation and biomass	*Breakdown of presence in composite map (%)*						
Land-cover/land-use	0%	1-25%	26-50%	51-75%	76-99%	100%	Total (%)
Upland conifer	8.58	2.87	0.151	13.4	73.8	1.18	100
Lowland conifer	11.0	76.3	11.9	.795	-	-	100
Northern hardwoods	99.5	0.452	-	-	-	-	100
Aspen/lowland decid.	30.3	49.0	20.1	.575	-	-	100
Grassland	39.8	13.1	21.8	25.3	-	-	100
Agriculture	80.9	18.9	0.197	-	-	-	100
Wetland	27.1	19.1	38.1	15.7	-	-	100
Open water	46.1	53.9	-	-	-	-	100
Urban/barren	93.8	6.17	-	-	-	-	100

Table 2: *Lc/lu and percent each appeared present in the composite map of the pine warbler run only with the vegetation and biomass layers.*

Pine Warbler: vegetation, biomass and neighborhood	Breakdown of presence in composite map (%)						
Land-cover/land-use	0%	1-25%	26-50%	51-75%	76-99%	100%	Total (%)
Upland conifer	26.3	1.22	2.49	2.07	22.6	45.3	100
Lowland conifer	88.8	0.073	0.658	0.226	1.39	8.89	100
Northern hardwoods	92.6	0.058	0.548	0.408	1.65	4.70	100
Aspen/lowland decid.	88.3	0.140	0.549	0.513	5.23	5.18	100
Grassland	74.2	0.286	0.804	0.410	17.0	7.27	100
Agriculture	100.0	-	-	-	-	-	100
Wetland	87.5	0.187	1.17	0.840	10.4	-	100
Open water	91.4	2.57	6.07	-	-	-	100
Urban/barren	74.8	25.2	-	-	-	-	100

Table 3: *Breakdown of land-cover/land-use and percent each appeared present in the composite map of the pine warbler run with vegetation, biomass, and neighborhood layers.*

Discussion
Chi-square anlaysis of raw data

The chi-square tests for individuals and sites produced *p-values* approximately equal to 0.0. This indicates a correlation between the pine warbler and forest type where they were found. Based on these results and survey observations, the pine warbler is correlated with mature pine and not with deciduous. This is in agreement with Northwoods and with the results of GARP. There also appears to be a slight correlation between the pine warbler and young coniferous, although not as strong as the correlation with mature coniferous. However, having counts less than five may produce spurious results since this violates an assumption of the chi-square test.

Training, testing, and p-values

Both training and testing accuracies increased when more environmental layers were added. This was the case for maps modeled with all location data and ones with half the data withheld. The *p-values* for the models which utilized all location points were less than 0.05, which indicates GARP performed better than a random model. However, when GARP was run with half the location data, it did not necessarily perform better than random for the models of vegetation and vegetation with biomass. This could be a result of fewer data points because accuracies tend to steadily drop when less than 50 are used (Stockwell and Peterson 2002).

Interpretation of maps

The presence/absence map interpreted from the Northwoods map suggests the pine warbler is present only in upland conifer. The GARP model of the pine warbler run with vegetation agrees with Northwoods in that upland conifer was chosen as presence 100% of the time (Figure 3). The only difference is that grassland was chosen as potential habitat 80% of the time by GARP (Table 1). During surveys, pine warblers were sometimes found in young conifer

stands and it was noted in the Landsat image these stands were sometimes classified as *grassland*. This might explain the appearance of grassland as potential habitat when in reality these areas are young conifer.

The model run with vegetation and biomass matched Northwoods well in that areas of absolute *absence* were the same (Figure 4). However, most of the upland conifer was selected as presence between 76-99% of the time, with a very small amount being selected as presence 100% of the time (Table 2). This may be due to the introduction of the biomass layer, which represents vertical structure, as well as a surrogate for age. Since some of the young conifer was classified as *grassland*, GARP deduced based on location points found in young conifer that *grassland* was possible presence habitat. This reduced the amount of upland conifer being chosen 100% of the time.

The model run with all environmental layers (Figure 5) is also similar to Northwoods. However, there are fewer individual pixels scattered throughout, suggesting the neighborhood values did have influence on prediction. The introduction of neighborhoods smoothed the image more and reduced the amount of pixels surrounded by completely different vegetation. Also, only 45.3% of upland conifer was chosen as habitat 100% of the time (Table 3). Surprisingly, over a quarter of upland conifer was never chosen as habitat. This could be a result of biomass aiding in the selection of upland conifer areas that were too young for the pine warbler or neighborhood layers selecting out upland conifers which were not surrounded by other upland conifers.

Conservation implications

This analysis has shown that GARP can produce finer niche maps and incorporate more detailed information about landscape structure into its prediction. When modeling for birds, the landscape structure must be considered in addition to other environmental factors as a main component of habitat preference. Although vegetation type is important, the vertical aspect of vegetation along with the horizontal aspect of the landscape (the areas surrounding preferred habitat) are major factors contributing to habitat choice. The models which involved vertical and horizontal structure provided greater variability when compared with the model that used vegetation only. In this way, GARP maps can produce finer geographic distributions of species. The models incorporating structural aspects of the landscape provided variability among vegetation types chosen and at different frequencies. This information can contribute to suitability ratings of certain habitat for species that vegetation type alone cannot offer. For example, it is the author's personal observation that pine warblers sometimes occur in young conifers. With the introduction of vertical structure into the model, GARP suggests this possibility as well. Young conifers, then, could be described as secondary habitat for the pine warbler based on GARP and validated by field work.

This analysis demonstrates research being conducted that involves modeling and GIS at a landscape level, thus aiding in the prediction of more precise ecological niches. The results suggest that niche modeling through the use of GARP is an exciting new tool which can assist in locating areas of unknown habitation, as well as focusing conservation efforts for endangered and threatened species. GARP models have already been used to design reserves for rare species in Mexico (Peterson et al. 2000) and to predict species invasions (Peterson and Vieglais 2001). This analysis demonstrates the capabilities of GARP at a regional level and the production of finer niche maps for the pine warbler through the incorporation of landscape spatial structure.

Future analysis

The results of this paper are part of a NSF funded project titled "BDEI: Radar Remote Sensing of Habitat Structure for Biodiversity Informatics" at the University of Michigan. Currently, four other birds are under study and analysis is underway with them as well. Further detail of GARP modeling will be examined by observing the rule-sets chosen to produce individual output maps. By examining these rules, our aim is to better understand why GARP is choosing certain vegetation as presence and others as absence.

References

Anderson, R.P., Gomez-Laverde, M., Peterson, A.T., 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. Global Ecology and Biogeography 11, 131-141.

Austin, G.E., Thomas, C.J., Houston, D.C., Thompson, D.B.A., 1996. Predicting the spatial distribtuion of buzzard Buteo buteo nesting areas using a Geographical Information System and remote sensing. Journal of Applied Ecology 33, 1541- 1550.

Azevedo, J.C.M, Jack, S.B., Coulson, R.N., Wunneburger, D.F., 2000. Functional heterogeneity of forest landscapes and the distribution and abundance of the red-cockaded woodpecker. Forest Ecology and Management 127, 271-283.

Benyus, J.M., Buech, R.R., Nelson, M.D., 1992. Wildlife in the Upper Great Lakes Region: A Community Profile, NC-301. U.S. Department of Agriculture, North Central Forest Experiment Station, St. Paul, MN.

Bergen, K.M., Dobson, M.C., Sharik, T.L., Brodie, I., 1995. Final report: structure, composition, and above-ground biomass of SIR-C/X-SAR and ERS-1 Forest Test Stands 1991-1994, Raco Michigan Site, 026511-7-T. Radiation Laboratory, Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI.

Bergen, K.M., Dobson, M.C., Pierce, L.E., Ulaby, F.T., 1998. Characterizing carbon in a northern forest by using SIR-C/X-SAR imagery. Remote Sensing of Environment 63, 24-39.

Blake, J.G., Karr, J.R., 1987. Breeding birds of isolated woodlots: area and habitat relationships. Ecology 686, 1724-1734.

Bolger, D.T., Scott, T.A., Rotenberry, J.T., 1997. Breeding bird abundance in an urbanizing landscape in coastal southern California. Conservation Biology 112, 406-421.

Collins, S.L., 1983. Geographic variation in habitat structure for the wood warblers in Maine and Minnesota. Oecologia 59, 246-252.

Debinski, D.M., Kindscher, K., Jakubauskas, M.E., 1999. A remote sensing and GIS-based model of habitats and biodiversity in the Greater Yellowstone Ecosystem. International Journal of Remote Sensing 2017, 3281-3291.

Dobson, M.C., Ulaby, F.T., Pierce, L.E., Sharik, T.L., Bergen, K.M., Kellndorfer, J., Kendra, J.R., Li, E., Lin, Y.C., Nashashibi, A., Sarabandi, K., Siqueira, P., 1995. Estimation of forest biophysical characteristics in northern Michigan with SIR-C/X-SAR. IEEE Transactions on Geoscience and Remote Sensing 33, 877-895.

Feria, A.T.P., Peterson, A.T., 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. Diversity and Distributions 8, 49-56.

Godown, M.E., Peterson, A.T., 2000. Preliminary distributional analysis of US endangered bird species. Biodiversity and Conservation 9, 1313-1322.

Hepinstall, J.A., Sader, S.A., 1997. Using Bayesian statistics, thematic mapper satellite imagery, and breeding bird survey data to model bird species probability of occurrence in Maine. Photogrammetric Engineering and Remote Sensing 6310, 1231-1237.

Herr, A.M., Queen, L.P., 1993. Crane habitat evaluation using GIS and remote sensing. Photogrammetric Engineering and Remote Sensing 5910, 1531-1538.

Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22, 415-427.

Langstaff, L., 2002. Raco Plains Budworm Project. U.S. National Forest Service, Hiawatha National Forest, St. Ignace, MI.

Lyon, J.G., 1983. Landsat-derived land-cover classifications for locating potential kestrel nesting habitat. Photogrammetric Engineering and Remote Sensing 492, 245-250.

MacArthur, R.H., Wilson, E.O., 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, NJ.

Mazerolle, M.J., Villard, M.A., 1999. Patch characteristics and landscape context as predictors of species presence and abundance: a review. Ecoscience 61, 117-124.

McGarigal, K., Marks, B.J., 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. U.S. Department of Agriculture, U.S. Forest Service, Portland, OR.

Nix, H.A., 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore, R., (Ed.). Atlas of Elapid Snakes of Australia. Australian Government Publishing Service, Canberra.

Osborne, P.E., Alonso, J.C., Bryant, R.G., 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. Journal of Applied Ecology 382, 458 - 471.

Payne, K., Stockwell, D.R.B., 1996. GARP modeling system user guide and technical reference. URL: http://biodi.sdsc.edu/Doc/GARP/Manual/manual.html (Last accessed Dec. 14, 2002).

Peterson, A.T., Egbert, S.L., Sanchez-Cordero, V., Price, K.P., 2000. Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. Biological Conservation 93, 85-94.

Peterson, A.T., 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103, 599-605.

Peterson, A.T., Vieglais, D.A., 2001. Predicting species invasions using ecological niche modeling. BioScience 51, 363-371.

Peterson, A.T., Ball, L., Cohoon, K.P., 2002. Predicting distributions of Mexican birds using ecological niche modelling methods. Ibis 144 (Electronic version), 27-32.

Reynolds, R.T., Scott, J.M., Nussbaum, R.A., 1980. A variable circular-plot method for estimating bird numbers. Condor 82, 309-313.

Scachetti-Pereira, R., 2002. Desktop GARP: users manual. URL: http://www.lifemapper.org/desktopgarp (Last accessed March 17, 2003).

Stockwell, D.R.B., Noble, I.R., 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. Mathematics and Computers in Simulation 33, 385-390.

Stockwell, D., Peters, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 132, 143-158.

Stockwell, D.R.B., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148, 1-13.

Trzcinski, M.K., Fahrig, L., Merriam, G., 1999. Independent effects of forest cover and fragmentation on the distribution of forest breeding birds. Ecological Applications 92, 586-593.

Wiens, J.A., 1995. Landscape mosaics and ecological theory. In: Hansson, L., Fahrig, L., Merriam, G. (Eds.). Mosaic Landscapes and Ecological Processes. Chapman and Hall, London.

With, K.A., King, A.W., 2001. Analysis of landscape sources and sinks: the effect of spatial pattern of avian demography. Biological Conservation 100, 75-88.