available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/ecolinf

Multi-dimensional vegetation structure in modeling avian habitat

Kathleen M. Bergen, Amy M. Gilboy, Daniel G. Brown*

University of Michigan, United States

ARTICLE INFO

Article history:
Received 31 December 2006
Accepted 2 January 2007

Keywords:
Bird habitat
Forest structure
Landscape structure
Radar
Landsat
GARP

ABSTRACT

The goal of this study was to evaluate the contributions of forest and landscape structure derived from remote sensing instruments to habitat mapping. Our empirical data focused at the landscape scale on a test site in northern Michigan, using radar and Landsat imagery and bird-presence data by species. We tested the contributions of multi-dimensional forest and landscape structure variables using GARP (Genetic Algorithm for Rule-Set Production), a representative modeling methodology used in biodiversity informatics. For our multi-dimensional variables, radar data were processed to derive forest biomass maps and these data were used with a Landsat-derived vegetation type classification and spatial neighborhood analyses. We collected field data on bird species presence and habitat for northern forest birds known to have a range of vegetation habitat requirements. We modeled and tested the relationships between bird presence and 1) vegetation type, 2) vegetation type and spatial neighborhood descriptions, 3) vegetation type and biomass, and 4) all variables together, using GARP, for three bird species. Modeled results showed that inclusion of biomass or neighborhoods improved the accuracy of bird habitat prediction over vegetation type alone, and that the inclusion of neighborhoods and biomass together generally produced the greatest improvement. The maps and model rules resulting from the multiple factor models were interpreted to be more precise depictions of a particular species habitat when compared with the models that used vegetation type only. We suggest that for bird species whose niche requirements include forest and landscape structure, inclusion of multi-dimensional information may be advantageous in habitat modeling at the landscape level. Further research should focus on testing additional variables and species, on further integration of newer radar and lidar remote sensing capabilities with multi-spectral sensors for quantifying forest and landscape multi-dimensional structure, and incorporating these in biodiversity informatics modeling.

1. Introduction

Biodiversity and ecosystem informatics is an emerging field that has been defined as “the creation, integration, analysis, and understanding of information regarding biological diversity” (Biodiversity Informatics, 2004). A core need of biodiversity informatics is the capability to produce maps not only of known locations of biological species occurrences, but also potential locations of these the same species based on similar habitat properties (Pennisi, 2000). To accomplish this, biodiversity informatics increasingly relies on inductive methods and models to map habitat and range of species. Several modeling approaches, including GARP (Genetic Algorithm for Rule-Set Production) use museum or field species presence data in conjunction with environmental layers. To construct realistic models and output habitat maps, input environmen-

* *Corresponding author.*
E-mail address: kbergen@umich.edu (K.M. Bergen).

1574-9541/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2007.01.001

tal layers describing essential habitat requirements or range characteristics are required.

At landscape to regional scales, landscape structure is increasingly believed to be a primary factor determining the habitat preferences of species (Dunning et al., 1992; Wiens, 1995; Boulinier et al., 2001). In addition to vegetation type, landscape patch metrics such as shape, size, and edge, are among the variables contributing to a quantitative definition of landscape structure (McGarigal and Marks, 1995), but spatially continuous geostatistical methods and neighborhood-based descriptions for categorical or continuous data may also be employed (Gustafson, 1998). This component of landscape structure has a largely horizontal definition. Quantifying vertical or volumetric structure creates a multi-dimensional description. For example, tree canopy height, biomass, density, understory presence, and/or canopy layering are also important structural variables for many forest bird species (Morgan and Freedman, 1986).

Biodiversity informatics uses spatial data of land cover and vegetation derived from remotely sensed datasets (Gottschalk et al., 2005). The capabilities of multi-spectral passive optical sensors such as Landsat, ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer), SPOT (Systeme Pour l'Observation de la Terre), or MODIS (MODerate Resolution Imaging Spectroradiometer), are useful for discriminating vegetation type and horizontal structure. Newer radar and lidar sensors have the capability to directly quantify vertical and volumetric dimensions of vegetation structure. The fusion of radar or lidar capabilities with widely available optical data, such as that from Landsat, and interpreting them for multi-dimensional structural characteristics is generating significant interest for describing forest and landscape structure (Bergen et al., 2006).

1.1. *Study objectives*

Our goal was to assess the value of integrating radar and optical remote sensing data to model multi-dimensional habitat space. We used empirical data at the landscape scale for a test site in Northern Michigan, using SIR-C radar and Landsat imagery and bird species data observed in the field to map bird habitat characteristics. Our primary modeling methodology was based on GARP, a representative modeling method used in biodiversity informatics. Our specific objectives were to: 1) create multi-dimensional vegetation structure and bird species datasets; 2) develop models and test for the influence of forest and landscape structure on habitat predictions, 3) assess output model accuracies, and 4) interpret the usefulness of forest and landscape structure environmental layers for habitat mapping.

1.2. *Background*

Birds can be particularly responsive to characteristics of forest multi-dimensional structure. While some bird species are generalists, many species have narrower ecological niches. Along with the composition of vegetation, niche discriminating characteristics can include the amount and configuration of horizontal patches (James and Wamer, 1982; McGarigal and McComb, 1995; Flather and Sauer, 1996), patch area (Howe, 1984; Boecklen, 1986; Freemark and Merriam, 1986), edge effects (Flaspohler et al., 2001; Chalfoun et al., 2002), and forest cover and fragmentation (Trzcinski et al., 1999).

Observations have also shown that suitable habitat for a bird species may be based on volumetric and vertical characteristics of the vegetation (Morgan and Freedman, 1986). These include stand age, height or biomass (Probst and Weinrich, 1993; Green and Griffiths, 1994; Nelson and Buech, 1996; Buchanan et al., 1999), shrub versus forest structure (Goransson, 1994), structure of a shrub layer within forests (Reid et al., 2004), and effects of forest thinning (Siegel and DeSante, 2003). One prior study focused on radar-based mapping of vegetation structure and bird diversity, where bird species and abundance were observed to change across both vegetation type and structural gradients. The authors concluded that some measure of vegetation structure is needed to understand how birds perceive habitat (Imhoff et al., 1997).

Both optical and radar image data are potentially suitable for classifying land cover and vegetation and providing maps of horizontal landscape structure. However, classifications derived from optical remote sensing instruments such as Landsat at fine (30 or 60 m) spatial resolutions (Vogelmann et al., 2001), and MODIS at coarser (1 km) spatial resolutions (Friedl et al., 2002; Hansen et al., 2005) are more widely available. Of these, the sensor used is determined by the spatial scale of the question, e.g. landscape, regional or global. Vegetation classifications derived from 30 m Landsat data are appropriate for landscape to regional-scale analysis and, in forested areas, are often produced at a level of detail equivalent to forest communities (Anderson Level II), and sometimes to species or species groups (Anderson Level III; (Anderson et al., 1976). Land-cover and vegetation classifications may also serve as the basis for calculating metrics of landscape horizontal structure.

Of all sensor types, the active sensors radar and lidar have the greatest capabilities for directly describing vegetation vertical and volumetric structure. Active sensors transmit and receive their own energy rather than relying on reflected sunlight to form an image, and radar sensors do so in the longer microwave portion ($\sim$1 mm to $\sim$1 m wavelengths) of the electromagnetic spectrum. Radars are described by their wavelengths (e.g. C-band at approximately 6 cm, L-band at approximately 23 cm); transmit-receive polarizations (horizontal or vertical propagation of radar waves); energy incidence angles with respect to Earth's surface; and spatial resolutions (Lillesand et al., 2004). Radar reflection (called backscatter coefficient, or $\sigma°$, in decibels) at a given wavelength, polarization, and incidence angle is determined by earth terrain structural properties and electrical properties. In the case of forest vegetation, structural properties are the dominant factor; where the specific contributing structural properties of vegetation canopies are 1) size distribution of components (for trees main stem, branches, and foliage) relative to wavelength, 2) orientation of components, and 3) number of reflecting components (Ulaby et al., 1986).

This dependence of radar backscatter on the structural properties of vegetation, in addition to the capability of long off-nadir wavelengths to penetrate through the vegetation canopy, is the basis for radar's ability to provide direct estimates of vegetation structure (Pierce et al., 1998).

Numerous studies have now demonstrated the relationship between vegetation structural parameters and radar backscatter. These studies show that typically there is a positive relationship between radar backscatter coefficient (in dBs) and field-measured dry biomass (in kg/m^2), meaning that forests of higher biomass will have greater radar backscatter. Different vegetation structural types, for example, conifer or deciduous, exhibit somewhat different forms of this relationship. Extinction (leveling off of backscatter) may occur at very high biomass quantities of a particular forest structural type and varies by type. Some studies have used empirical methods to derive parameters such as biomass, and to map these as continuous data over regional landscapes (Dobson et al., 1995; Bergen, 1997). In addition to radar backscatter methods, radar interferometry can be used to directly estimate forest height and, like lidar, may have potential for describing within-canopy vertical structure (Treuhaft and Siqueira, 2004).

Modeling methods for habitat distribution (Guisan and Zimmermann, 2000) based on species presence data have included range and envelope methods (e.g. BIOCLIM) (Busby, 1991); regression-based generalized linear models or logistic regression (GLM) and generalized additive models (GAM) (Austin, 2002); genetic algorithms (GAs) such as GARP (Payne and Stockwell, 1996; Anderson et al., 2003); and several novel machine-learning methods such as neural networks, regression trees or entropy models (Elith et al., 2006; Garzón et al., 2006). Algorithms have been combined, for example the GARP Modeling System (GMS) has sometimes been called a "super-algorithm" in that it uses several of the methods above in its formation of species distribution models by combining sets of rules using a GA approach (Stockwell and Peters, 1999; Peterson et al., 2002). The GA approach refers to the idea that solutions to modeling problems in a machine environment evolve the same way organisms evolve through natural selection. A set of possible solutions are formed and, through a series of iterations that include, for example, mutations, deletions, and crossing over, the solutions are modified and tested until the best solution is found (Stockwell and Peters, 1999).

2. Methods and materials

2.1. Study area

The study was carried out in the Michigan Forests Test Site (MFTS). The MFTS is located in the eastern part of Michigan's Upper Peninsula largely within the Hiawatha National Forest (Fig. 1). The MFTS was established as a NASA test site for the SIR-C (Shuttle Imaging Radar-C) instrument flown on the space shuttle in 1994. The site is approximately 20 km wide, and centered on 46.39° N, 84.88° W. GIS datasets of vegetation type and radar-derived forest structure are available for the site, as are field data of composition and structure for 70 four-hectare forest test stands (Bergen et al., 1995).

The forests of the MFTS are characterized by pole to mature northern hardwoods and red and jack pine plantations, interspersed with younger even-aged conifer and aspen stands, plus lowland communities. The mostly forested landscape is stable with relatively little disturbance, i.e. limited timber harvest primarily in conifer plantations, gap dynamics in older northern hardwoods stands, and overall slow northern latitude growth.

2.2. Remote sensing data preparation

The land-cover and vegetation type environmental layer (hereafter called vegetation type layer) of the study site was acquired from the widely-used U.S. National Land Cover Data (NLCD) dataset (http://www.epa.gov/mrlc/). These data were originally derived from Landsat Thematic Mapper (Landsat TM) scenes taken 1992–1994. The positional accuracy of this 30 m raster data was designed to be ±0.5 pixels or ±15 m (Lillesand et al., 1998). NLCD thematic (vegetation class) overall accuracy for the U.S. Great Lakes region was 83% at Level I and 64% at Level II (with low accuracies in agricultural classes lowering the overall accuracy and with higher accuracies in forest classes) (Wickham et al., 2004). The vegetation type layer was recoded from Anderson Level III classes to nine Anderson Level II classes to be congruent with our field and biomass layer data (Fig. 2a; Table 1). The nine classes include four forest type classes, three non-forest vegetation classes, and two (negligible at <1%) non-vegetated classes.

To represent vegetation spatial structure, two different 30 m raster layers – majority and variety – were created in ArcGIS (ESRI, 2002) using neighborhood statistics on the vegetation type layer. For each cell in the input vegetation type layer, neighborhood functions compute a statistic based on the values of the surrounding cells within a specified distance and this value is then sent to the output grid. Our neighborhoods approximated a circle with a 6-cell radius, slightly greater than the bird survey plots. The majority statistic is determined by the vegetation type that occurs most often in the neighborhood. In this case, the cells of the output grid were converted to the type of vegetation most often occurring within the specified neighborhood, which may be the same value as the input grid (Fig. 2b). The variety statistic determines the number of unique vegetation types within a neighborhood; therefore, the cells in the output map have a number value that represents how many (1–9 in this case) different kinds of vegetation were within the specified distance (Fig. 2c).

Biomass is a particularly useful integrative variable for vegetation age, height and density when combined with vegetation type. Biomass was also our best dataset in terms of previous validation, and we used biomass to test for the influence of forest volumetric structure. The original data for our radar-derived biomass layer were compiled in a project using C- and L-band SIR-C radar data taken over the MFTS which had been orthorectified to a positional accuracy of better than ± < one 25 m pixel). In the SIR-C project to quantify and map biomass, we first measured field forest composition and structure in 70 four-hectare geo-referenced (positional accuracy better than 10 m) forest test stands (Bergen et al., 1995). These field estimates of stand height, basal area, crown biomass, trunk biomass and total biomass had error rates of 15% or better (Bergen et al., 1995). Total aboveground biomass for the 70 forest test stands ranged from <2.5 kg/m^2 for clear-cut seedling stands to 27.3 kg/m^2 for a mature northern hardwoods stand (Bergen et al., 1995). These figures are also representative of the range of total biomass in the overall study site.

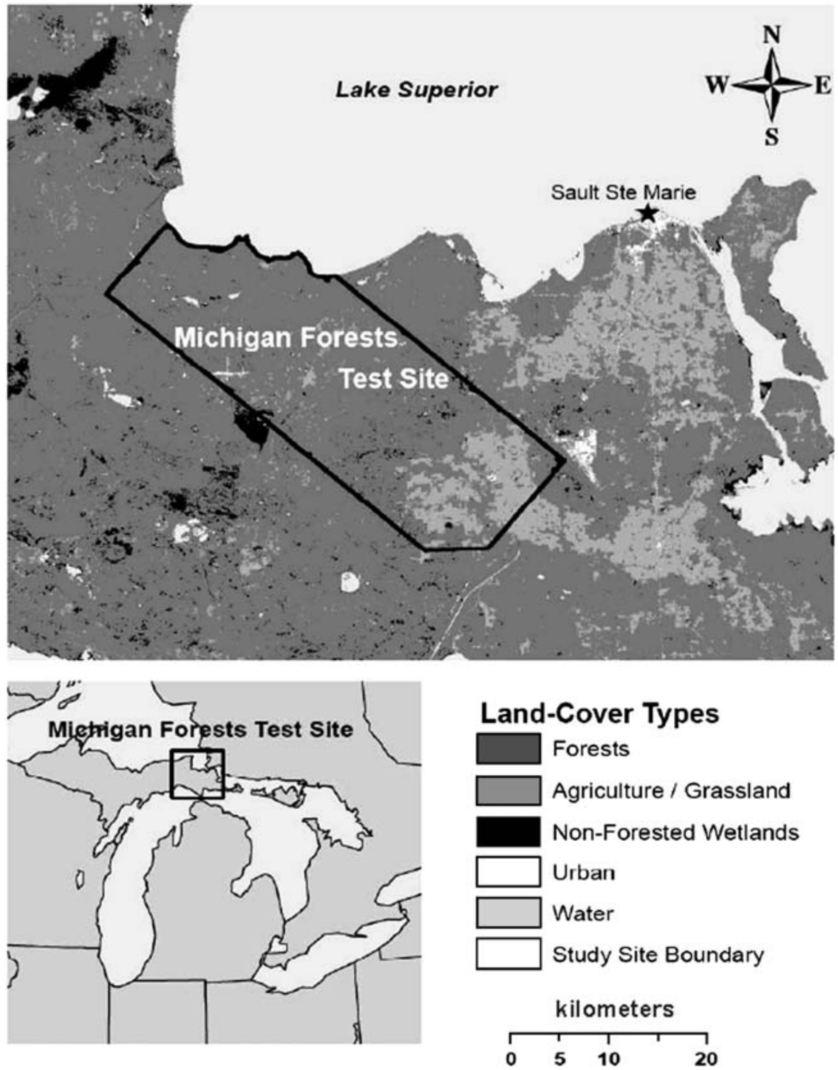

Fig. 1 – Michigan Forests Test Site (MFTS). The inset map shows the location of the study site in the upper Great Lakes region, USA. The main map background is derived from Landsat data used in this study and depicts general land-cover types of the region. The outlined rectangle corresponds to the path and swath of the radar sensor and to the test site boundaries.

Next, vegetation-class-specific inversion algorithms were developed by least squares regression of the radar image backscatter for the test stands on field-measured test stand structural parameters. Inversion algorithms were developed for estimated height, basal area, crown biomass, and trunk biomass and then these were combined to produce an estimate of total biomass (Fig. 3). Analysis showed that aboveground biomass over the study site had been estimated with an overall RMSE of better than 1.4 kg/m^2 dry wt (biomass range 0 to 30 kg/m^2) (Dobson et al., 1995). Application of inversion algorithms was extended from the test stands, upon which they were developed, to the entire radar image resulting in a spatially continuous map of biomass quantities at a 25 m spatial resolution (Bergen and Dobson, 1999).

Some additional work was needed on the radar-derived biomass layer for use in this project. Available vegetation type and biomass data were both from approximately 1994 providing congruence in the spatial datasets. However, our bird data were from 2002. For this reason, while our study and output maps can be interpreted as to dependence of bird species habitat selection on forest and landscape structure in terms of vegetation type and biomass, they should not be interpreted with respect to absolute biomass values. The MFTS is a relatively stable remote area, but we did not use any areas of significant intervening disturbance as training or testing sites. Other processing included resampling the pixel size of the biomass image from 25 m to 30 m for co-registration with the 30 m Landsat data. Nearest neighbor resampling was used to preserve original data values. The site biomass range of approximately 0–30 kg/m^2 is typical of the broader region, and biomass layer values were linearly scaled to 0–255 (Fig. 2d).

2.3. Bird data collection and analysis

Bird presence surveys were conducted July 2002 in the same forest test stands used to develop the biomass layer (Gilboy,

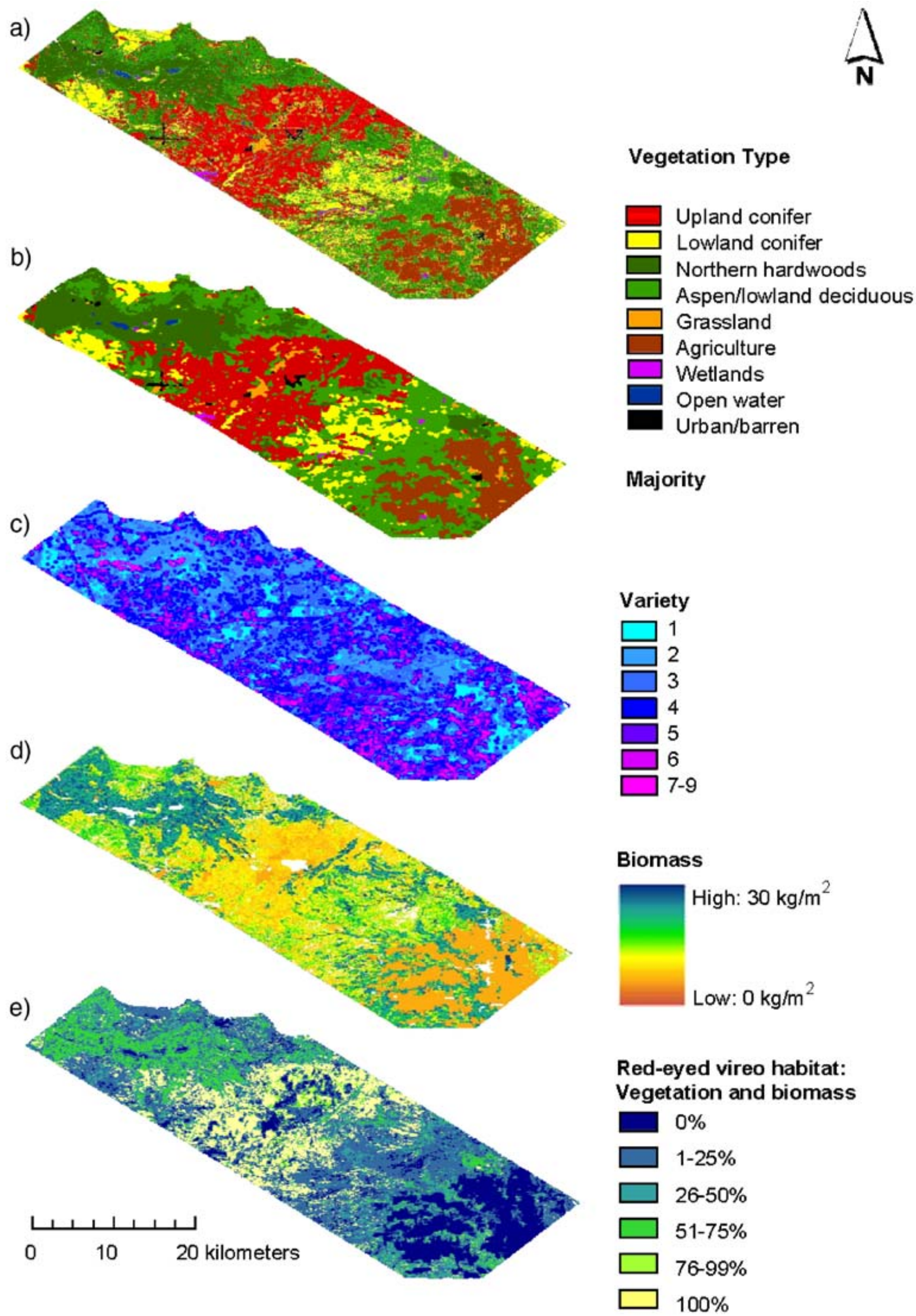

Fig. 2 – Images used or created through this study: a) vegetation type layer, b) majority layer, c) variety layer, d) biomass layer, e) output composite habitat map for red-eyed vireo modeled with vegetation and biomass.

2003). A total of 61 stands were selected to represent the range of forest composition and structure in the study site. Early morning surveys were conducted within the stands using the circular-plot method (Reynolds et al., 1980) to detect bird presence and absence. Bird surveys were also conducted by the U.S. Forest Service, Hiawatha National Forest during June/July 2002 (Langstaff, 2002) using the same measurement protocol on 71 additional sites.

Of the 57 bird species recorded during the field surveys, ten had enough individuals to allow reliable analysis with GARP (Stockwell and Peterson, 2002). Three of these were nominated for analysis based on their differences in habitat requirements. A chi-square test was run on our field data for the three potential bird species and vegetation types to confirm dependence on vegetation type. A significance value of 0.05 was selected and p-values less than this indicated a relationship

Table 1 – The distribution in percent of forest and other land-cover types in the MFTS (column one) and the distribution of forest types within just the forested component of the MFTS (column two)

Vegetation type	Percent of total area (%)	Percent of forested area (%)
Upland conifer	21.7	26.0
Lowland conifer	11.0	13.2
Northern hardwoods	20.2	24.2
Aspen/lowland deciduous	30.6	36.6
Grassland	4.4	–
Agriculture	9.6	–
Wetland	1.2	–
Urban/barren	0.9	–
Water	0.4	–
	100	100

Data is derived from the vegetation type layer used in this study.

between bird species and vegetation type. Sixty pine warblers (*Dendroica pinus*), 75 chipping sparrows (*Spizella passerina*), and 55 red-eyed vireos (*Vireo olivaceus*) were detected and these geo-located observations formed the bird point location data for input into GARP models. Each geolocated bird survey point was plotted in ArcView (ESRI, 2002) software in a geographic (latitude/ longitude) projection for compatibility with Desktop GARP.

To support selection of input variables and assessment of results we collected information about known habitat characteristics of the three bird species. The USFS NORTHWOODS database contains summary habitat requirements for 389 species of birds and other animals in the upper Great Lakes region USA in 20 aquatic and terrestrial habitat types (Benyus et al., 1992). The database was derived from verified field-sighting data combined from regional National Forests and contains species names, their habitats, season of occurrence, abundance, and versatility rating (1–12), the latter being the number of different types of region-wide habitat the species can use. If 50% or more of the reporting Forests listed a species as using a habitat type then the species was listed as using that habitat in NORTHWOODS.

The pine warbler is often considered the most characteristic breeding bird in the pine forests of Eastern North America (Evers, 1991). NORTHWOODS reports a versatility rating of 1 for the PIWA and lists mature upland coniferous as its only habitat. The chipping sparrow is one of the most widespread and abundant sparrows breeding in the United States and Canada. The chipping sparrow prefers borders of natural forest openings, edges of water bodies, open woodlands and weedy fields and NORTHWOODS reports a versatility rating of 6 for this bird (Middleton, 1987). The red-eyed vireo is one of the most common songbirds within the woodlands of Eastern North America and the most common vireo species found breeding in Michigan (Payne, 1983) with a NORTHWOODS versatility rating of 7. The red-eyed vireo prefers mature deciduous and mixed deciduous–coniferous forests, and is sometimes found in younger forests of the same types.

Habitat suitability maps (hereafter referred to as NWDVEG maps) were constructed for each of the three bird species using the habitat vegetation type preferences listed in NORTHWOODS matched with our Landsat-derived vegetation type layer. This represents a deductive approach to habitat mapping using only vegetation type spatial data as input. These deductive maps were constructed for eventual statistical and visual comparison with the vegetation-only predictions and output maps generated by GARP.

We derived several metrics to describe the match between bird species point locations and suitable areas in the NWDVEG maps, including number of bird point locations that fell into suitable habitat areas on the NWDVEG maps, the percent of the test site mapped as suitable habitat, percent improvement of the NWDVEG map over a chance model, and the vegetation types NORTHWOODS indicated as suitable habitat.

2.4. *Modeling with GARP*

The GARP program uses the georeferenced species point locations in combination with environmental data layers to create

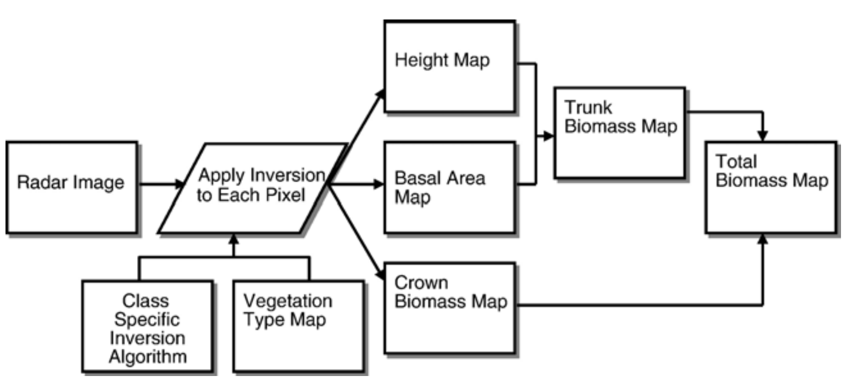

Fig. 3 – Methods for deriving biophysical parameters height and biomass from radar. In the radar backscatter method used in this study L- and C-band image backscatter were regressed on field-measured samples to create inversion algorithms which were then applied to the entire image. Separate inversion algorithms and mapped estimates were made of height, basal area, crown biomass, and trunk biomass, and then these were combined to produce an estimate of total biomass mapped as a continuous quantity.

models of species habitat and range (Stockwell and Noble, 1992; Payne and Stockwell, 1996; Stockwell and Peters, 1999). The algorithm software has been made broadly available through Desktop GARP, available at http://www.lifemapper.org/desktopgarp (Scachetti-Pereira, 2002). GARP modeling can be grouped into four main steps: 1) data preparation, 2) model development, 3) model application and validation, and 4) communication and output (Payne and Stockwell, 1996; Gilboy, 2003). We ran these sequences of steps separately for each of the bird species in combination with different environmental grid layers: 1) vegetation only, 2) vegetation and biomass, 3) vegetation and neighborhoods, and 4) all environmental variables.

In the data preparation step, the modules Rasterize and Presample prepare the input species data for use in GARP. Rasterize converts species point data into raster layers. Duplicate species points within one cell do not provide more information, so are removed by absorption into a single observation (Stockwell and Peters, 1999). If one or more points fall within a cell, the cell takes a presence value; otherwise it remains zero. Presample takes the newly created raster layers and creates split training and testing data sets (we chose a 50/50 split) by randomly sampling the raster data set of all data points for a bird species, 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 input species location input data. The other 1250 are re-sampled from the total geographic space to replicate absence data, termed "background". Presample thus creates large sets of presence and background data regardless of how many location points were input into GARP (Stockwell and Peters, 1999).

The heart of the GA process is in the model development steps. After the training set is generated by Presample, it is input into the next program Initial. This creates an initial model that is the starting point for the GARP algorithm. The initial model is a set of rules that influence the development of the subsequent models. Within GARP, 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 where only a single variable value within the precondition of the rule is used. An example atomic rule would be: *if the vegetation is upland conifer, then the species is present*. The second format of rules is the BIOCLIM rules. A BIOCLIM rule 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. Only absence can be inferred because falling within the range of tolerance does not guarantee the presence of a species. The third type of rule, a range rule, is a generalization of the BIOCLIM rule that allows for negation. The final class of rules, logit rules, is based on logistic multiple regression models where there is a positive dependence between species presence and model variable/s.

Continuing the model development step, the fourth module Explain applies the GA to improve the initial models, iterates, and then outputs the best of these models. In our procedures, each bird species-variable layer/s combination was run 20 times with 1000 maximum iterations per run. During iterations, GARP continuously tests the utility of the current set of rules, modifies rules, and terminates when the rule archive no longer changes, or reaches 1000 iterations, whichever comes first. Until this termination point, the program continues to create new populations by modifying archived rules with genetic recombination. The three heuristic operators are join, crossover, and mutate. Join is simply the joining of two rules to produce a longer rule. The crossover operation mimics the genetic exchange of real genes when two structures in the population 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 of the rule-sets.

Model application and validation begins with the Verify module. This program tests the predictive accuracy of the model developed from the training set on the original training set and then also on the reserve testing set that was created in Presample. The results of these are confusion matrices, which record the proportions of errors and accuracies made by the model (Stockwell and Peters, 1999). When applied to the testing data, accuracy is independent of the data used to formulate the rules and thus considered a more reliable estimate of how well the rules worked. Training and testing statistics were output as contingency matrices for each of our models, percent accuracies were calculated, and chi-square p-values provided. The second part of model application and validation is carried out through the module Predict which takes the newly created model and forms a prediction for each value and cell as applied to the environmental layers.

Model communication and output includes both spatial data and text of model rules. In this final step the module Image converts the calculations produced in Predict into image formats for visualization as predicted suitability layers. Finally, the Translate function screens the rule-sets and eliminates rules that were not used to make predictions. By doing this, only the most influential rules are presented to the user and the user is provided with files containing information on the rule-sets (or model) used to generate the spatial layers.

We ran each model 20 times (with 1000 iterations per run) for a given combination of bird species and environmental layers, resulting in 20 separate output maps. Each set of 20 maps was combined to form a composite map that indicates the number of runs out of 20 that a cell was predicted to contain suitable habitat for the species in question. These composite maps were made for each of the overall model runs — four environmental layer combinations for each of three bird species resulting in 12 composite maps of predicted bird species habitat.

2.5. *GARP model accuracies and evaluation*

Models were evaluated in several ways. First, training and testing accuracies and p-values output by the Verify module were observed and compared. The p-value calculated by GARP is the result of a chi-square test, which uses the test points set aside by GARP and the area predicted presence (e.g. predicted as habitat) by the output model. Values less than 0.05 indicate that GARP performed significantly better than a random model in terms of the accuracies. The training and testing accuracies represent how well the model did in predicting habitat for the bird location points used to construct the model (training) and for the reserved points not used in model construction (testing).

In a different form of model evaluation, bird point locations were overlaid on both NWDVEG and GARP output maps to compare the proportion of points that fell within the presence area in the maps to the expected proportion if the model were random. This resulted in statistics of the percent improvement of the NWDVEG and GARP models over a random model. In the case of the vegetation only model runs, the GARP maps were further compared to the NWDVEG maps as a test of the influence of the modeling method on results.

If environmental layers are relevant to habitat requirements, use of more than one layer generally increases accuracy of predictions and a sill is often reached at four or five layers (Peterson and Cohoon, 1999). As a final step, we examined both the output maps and the model rule-sets output by GARP to further examine why GARP predicted certain areas as possible habitat using the different multi-dimensional environmental layer combinations, and to evaluate if they conformed to our expectations based on published ecological niche descriptions and our field observations of bird species habitat use.

3. Results

3.1. *Bird data relationships to vegetation*

For all three species, chi-square analysis of field-observed bird locations and vegetation produced p-values less than 0.001. Interpretation of the chi-square results showed that the pine warbler was highly correlated with mature upland conifer ($p<0.001$) and not with deciduous. This is in agreement with NORTHWOODS and the literature. Our field data show that there was also a slight correlation between the pine warbler and young conifer. For the chipping sparrow, NORTHWOODS lists mature upland coniferous, mature upland deciduous and shrub-sapling opening as three possible habitats. In our field data, the chipping sparrow was correlated with mature upland conifer, young conifer (which could also be described as shrub-sapling openings; $p<0.001$), and not with deciduous. Chi-square results showed that the red-eyed vireo was highly correlated with northern hardwoods ($p<0.001$), consistent with NORTHWOODS. Chi-square analysis also showed a correlation with mature upland conifer which was not one of the several habitats listed in NORTHWOODS, although mature mixed (mixed conifer–deciduous overstory) was listed as possible habitat. The mature upland conifer areas we surveyed did not have a mixed overstory, but some did have a significant deciduous understory.

The NWDVEG maps also served as an indicator of the relationship of the three bird species to vegetation type. Two of the three NWDVEG maps with actual data points overlaid performed somewhat better than random, and the NWDVEG map for the red-eyed vireo performed worse than random. The latter can be attributed to the red-eyed vireos in our site being found in mature upland conifers (not mapped as habitat based on NORTHWOODS descriptions) with substantial deciduous undergrowth (Table 2).

3.2. *GARP models training and testing accuracy*

Training accuracies for the GARP models of the three bird species for the different input vegetation and structural data combinations ranged from 62–85%; testing accuracies ranged from 59–84% (Table 3). For 10 of the 12 models, p-values were less than 0.05, indicating that GARP performed significantly better than a random model for most of the runs. The two cases in which the p-values were slightly greater than 0.05, were the two red-eyed vireos with only 1–2 data layers (~0.08). All of the GARP model training and testing values increased between 1 to 22 percentage points when more environmental layers of multi-dimensional habitat were added (Table 3). The highest training and testing values of the 12 models occurred when all four of the environmental layers were used.

3.3. *GARP map output*

Fig. 2e and Table 4 provide an example of one output map and accompanying summary table. Shown are results for the red-eyed vireo modeled with vegetation and biomass. As an example of how this output is interpreted, this GARP model using vegetation and biomass selected 75.2% of the cells of upland conifer of as habitat 100% of the time and selected 51.3% of the cells of northern hardwoods as habitat between 51 and 75% of the time. Maps and summary tables were constructed for all species-variable combinations (Gilboy, 2003), and compared with model rule-sets and published habitat descriptions.

Table 2 – NWDVEG map analysis

Species	Points in presence areas (%)	Image area designated as habitat (%)	Improvement over chance model (%)	Vegetation layer types
Pine warbler	50.0	21.7	28.3	Upland conifer
Chipping sparrow	81.3	67.0	14.3	Upland conifer Lowland conifer Northern hardwoods Grassland/shrubland Agriculture
Red-eyed vireo	49.1	50.8	1.4	Northern hardwoods Aspen/lowland deciduous

Shown in columns 1–5 are 1) bird species, 2) the percent of bird sample locations that fell into suitable habitat areas on the NWDVEG maps, 3) the percent of the test site mapped as suitable habitat, 4) the percent improvement of the NWDVEG map over a chance model, and 5) the Landsat vegetation layer types that NORTHWOODS indicated as suitable habitat.

Table 3 – GARP accuracy average statistical results of the 20 individual presence/absence maps modeled with bird species sample points and different combinations of environmental layers

	Training accuracy (%)	Testing accuracy (%)	P-value
Pine warbler			
Vegetation only	62	59	0.002
Vegetation and biomass	73	64	0.035
Vegetation and neighborhoods	82	80	0.000
All environmental layers	84	81	0.000
Chipping sparrow			
Vegetation only	77	77	0.000
Vegetation and biomass	83	79	0.000
Vegetation and neighborhoods	85	80	0.000
All environmental layers	85	84	0.000
Red-eyed vireo			
Vegetation only	63	60	0.083
Vegetation and biomass	69	61	0.081
Vegetation and neighborhoods	71	63	0.015
All environmental layers	75	66	0.007

When the GARP composite maps of different variable combinations and for each species were compared to each other several trends were observed (Table 5). First, the addition of the biomass layer tended to reduce the amount of the image area designated as suitable habitat as compared with e.g. vegetation only. The inclusion of neighborhoods layer generally increased the amount of area designated as suitable habitat. The inclusion of all structural layers increased the amount of area designated as suitable habitat over vegetation alone. In almost all cases, the inclusion of the additional structural variables increased the percent improvement of the model of a chance model.

4. Discussion

Further interpretation of composite maps together with output rule-sets illuminated more specifically why certain multi-dimensional environmental layer combinations improved a particular bird species habitat map. Some of the most important relationships are highlighted in the discussion below for each bird with respect to its four GARP models.

4.1. *Pine warbler*

In the vegetation only GARP model for the pine warbler, the upland conifer vegetation type was selected as the primary habitat and all upland conifer cells were considered presence in 100% of the composite maps. For the pine warbler, this is the most appropriate habitat according to the literature. Some grassland/shrubland was also chosen as potential habitat by GARP. During surveys, ten pine warblers were detected in young conifer stands, and these areas could have been mapped on the Landsat-derived vegetation type layer as either upland conifer or grassland/shrubland. Although not listed as habitat by NORTHWOODS, pine warblers have been previously observed in young pine when their preferred habitat is not available (Rodewald et al., 1999).

In the pine warbler vegetation and biomass model, the upland conifer cells were divided into different presence categories instead of just the 100% category as in the vegetation only model. Very low biomass value cells were rarely selected. Since the pine warbler's preferred habitat is mature pine, the GARP model typically selected out the youngest pines as inappropriate, thus predicting a more precise distribution map than classifications based only on vegetation type. However, some Landsat-classified grassland/shrubland cells were selected as habitat. Even though this is not the pine warbler's preferred habitat, our surveys confirmed that they did occur there. Some young conifer areas, then, could be described as secondary habitat for the pine warbler based on GARP vegetation and biomass layers and fieldwork.

In the vegetation and neighborhoods model for the pine warbler, more than half of the upland conifer was selected as habitat 100% of the time, compared to only 1.18% of upland conifer being chosen as habitat 100% of the time in the vegetation and biomass model. This reflects influence of neighboring cells. In order for a cell to be considered presence habitat, the rules stated that it usually needed to be upland conifer, but cells that were classed as lowland conifer, northern hardwoods, and aspen/lowland deciduous in the vegetation type layer AND surrounded by upland conifer cells resulted in a majority value for upland conifer and were thereby given a presence prediction. This means that even if a certain 30 m vegetation type cell may not be appropriate habitat by itself, if it was surrounded by upland conifer, the chances of it being suitable habitat increased, thus increasing the area of coniferous forests for use as habitat for the pine warbler.

For the pine warbler model with all layers, most atomic and BIOCLIM rules suggested a range of biomass values for presence prediction, yet required the vegetation type and majority value to be upland conifer. In light of this, either the GARP predictions modeled with neighborhood layers, or with

Table 4 – Summary table of the GARP model for red-eyed vireo using vegetation and biomass

Vegetation type	Percent of presence cells of each vegetation type by presence category: Presence category 0%	1–25%	26–50%	51–75%	76–99%	100%	Total (%)
Upland conifer	1.0	10.2	1.1	2.1	10.4	75.2	100
Lowland conifer	8.7	72.5	18.8	–	–	–	100
Northern hardwoods	3.5	31.2	13.9	51.3	–	–	100
Aspen/lowland deciduous	12.3	60.3	27.4	–	–	–	100
Grassland	42.8	57.0	0.2	–	–	–	100
Agriculture	97.0	3.0	–	–	–	–	100
Wetland	71.6	28.4	–	–	–	–	100
Open water	34.6	65.2	0.2	–	–	–	100
Urban/barren	–	44.2	55.8	–	–	–	100

Table 5 – GARP map and rule-set analysis

	Points in presence areas (%)	Image area designated habitat (%)	Better than chance model (%)*	Rules used	Vegetation types selected >50% of time
Pine warbler					
Vegetation only	66.7	26.1	40.6/12.3	All 4	Upland conifer Grassland/shrubland
Vegetation and biomass	65.0	20.7	44.3	All 4	Upland conifer
Vegetation and neighborhoods	88.3	36.8	51.5	logit, BIOCLIM, range	Upland conifer
All layers	88.3	34.6	53.7	logit, BIOCLIM, range	Upland conifer
Chipping sparrow					
Vegetation only	78.7	22.5	56.2/41.9	BIOCLIM, atomic	Upland conifer
Vegetation and biomass	73.3	12.4	60.9	BIOCLIM, range, atomic	Upland conifer
Vegetation and neighborhoods	94.7	23.7	71.0	BIOCLIM, logit	Upland conifer, northern hardwoods, lowland conifer
All layers	89.3	21.0	68.3	logit, BIOCLIM	Upland conifer, northern hardwoods, lowland conifer grassland/shrubland
Red-eyed vireo					
Vegetation only	45.5	21.7	23.8/25.5	All 4	Upland conifer
Vegetation and biomass	67.3	29.4	37.9	All 4	Upland conifer, northern hardwoods
Vegetation and neighborhoods	54.5	22.3	32.2	All 4	Upland conifer, northern hardwoods
All layers	74.5	24.9	46.6	All 4	Upland conifer, northern hardwoods

Shown in columns 1–6 for each species are 1) model type, 2) the number of bird sample locations that fell into suitable habitat areas on the GARP maps, 3) the percent of the test site mapped as suitable habitat, 4) the percent improvement of the GARP map over a chance model, 5) the rules (logit, BIOCLIM, range, and/or atomic) used in the model, and 6) the Landsat vegetation layer types that the GARP model chose as suitable habitat greater then 50% of the time.
*For vegetation only, the second number is the % improvement of the GARP modeled map over the NWDVEG MAP.

all the layers, offered the most precise regional maps of the pine warbler distribution in the study site area, and pine warbler appears to be most dependent on vegetation type and neighborhoods. Predictions were made more precise within these areas by adding biomass information.

4.2. *Chipping sparrow*

For the chipping sparrow model with vegetation only, upland conifer was selected as the only suitable habitat. This is not in complete agreement with NORTHWOODS or the literature, which states that although upland conifer is habitat for the chipping sparrow, it should prefer weedy fields, edges of forests, and coniferous areas that have been modified by humans. However, selection of upland conifer probably occurred in the models because the upland conifer areas in our site were in fact the most disturbed (some logging and salvage activities) and fragmented in the otherwise fairly undisturbed site.

When modeled with vegetation and biomass, chipping sparrows were more likely found within upland conifer areas with lower biomass according to our GARP models. Chipping sparrows thrive in human-modified landscapes and are actually found in more abundance there than in mature forests (Middleton, 1987), so the identification of the more mature pines as inappropriate habitat creates a more precise geographic distribution. This model, however, still did not account for the diversity of habitats the chipping sparrow uses. Not all of the possible habitats listed by NORTHWOODS for the chipping sparrow were present or surveyed, and the lack of presence in certain habitats (e.g. agriculture) was probably due to a lack of survey points in our study, which focused primarily on forested habitat.

The incorporation of the neighborhood layers in the model did not as significantly improve the accuracy of predictions for the chipping sparrow which uses more fragmented and disturbed habitats. This was seen most strikingly in the agricultural southeast portion of the site. In the models run with vegetation only (and vegetation and biomass), individual cells of presence are scattered throughout this agricultural area. However, when the models were run with neighborhood layers, this area of agriculture mostly appeared as absence. Because of the chipping sparrow's preference for edges, open woods, and weedy fields, those individual cells of vegetation that were smoothed over in the majority layer could possibly have served as habitat for the chipping sparrow. For this reason, GARP did not predict a more realistic distribution for the chipping sparrow when modeled with neighborhood layers.

The chipping sparrow model using all environmental layers differed from the previous models in that more lowland conifer, northern hardwoods, and grassland/shrubland showed up in the higher categories of presence. Although these percentages were small, this model suggested the possibility of the chipping sparrow living in these vegetation types that were also listed in NORTHWOODS. Also, as determined by the rule-sets, most biomass values above 153 (scale 0–255) caused an absence prediction, indicating that chipping sparrows, as should be expected, are selecting against dense, mature forests. However, since this model included neighborhood layers, individual cells of presence have been eliminated throughout the image, but particularly among agriculture. Upland conifer was also distributed more evenly among the presence categories. For this reason, of the four models predicted by GARP, the vegetation and biomass model was probably the most realistic in terms of distribution for the chipping sparrow in the study site.

5. Red-eyed vireo

The selection of upland conifer as habitat in the vegetation only model for the red-eyed vireo conflicts, in part, with NORTHWOODS. We believe GARP predicted upland conifer as habitat because of deciduous undergrowth in many stands of mature upland conifer. Twenty-five out of 55 red-eyed vireos were detected in upland conifer stands that had considerable deciduous undergrowth. In the several mature upland conifer stands that did not have any deciduous undergrowth, red-eyed vireos were not found. Also, NORTHWOODS listed mature mixed forest as possible red-eyed vireo habitat, and these areas could have been classed as upland conifer in the Landsat image. The vegetation only model also conflicted with NORTHWOODS by not predicting northern hardwood as red-eyed vireo habitat. Typically, red-eyed vireos use deciduous vegetation (Robinson, 1981; Ehrlich et al., 1988; Benyus et al., 1992), but GARP predicted presence in northern hardwoods in only 25% of the vegetation only output maps.

In the red-eyed vireo model using vegetation type and biomass, three-fourths of upland conifer was predicted as habitat 100% of the time. One atomic rule stated that presence should occur if the vegetation type was upland conifer. This, combined with rules stating that biomass must be between 60 and 190 (high to very high relative values for conifer), caused some of the upland conifer to be selected as habitat. This appears realistic because for upland conifer to have substantial deciduous undergrowth to support red-eyed vireos, it is likely these pine forests are older and taller with higher biomass values. Another atomic rule stated that presence should also occur if the vegetation was northern hardwoods. This rule, combined with the previous rule indicating biomass range for presence prediction, also explains why more than half of the northern hardwoods were predicted as habitat 51–75% of the time, more so than when modeled with just vegetation. Also, 27.4% of aspen/lowland deciduous was chosen as habitat 26–50% of the time — more than was predicted when modeled with vegetation only. The aspen/lowland deciduous areas chosen as habitat have relatively high biomass values for that type, ranging from approximately 100–180 and are somewhat structurally similar to northern hardwoods.

In the red-eyed vireo model using vegetation and neighborhoods, upland conifer was selected as habitat among 76% or more of the output maps. Also, northern hardwoods were selected, divided among all presence categories. Two atomic rules stated that in order for presence to be assigned to a cell, the vegetation had to be upland conifer or northern hardwoods, with the majority value also upland conifer or northern hardwoods, respectively. Only large, contiguous areas of vegetation were selected as habitat. Also, based on range and BIOCLIM rules, most cells were predicted as absence if the variety value was greater than five. This could also be thought of in terms of heterogeneity. The habitat distributions predicted with the aid of neighborhood layers were large groups of cells, or the less diverse areas. Since the red-eyed vireo is a habitat sensitive species and will not inhabit small patches of area (Robbins et al., 1989), this modeling including neighborhoods is most likely a more realistic version of distribution in this area.

In the GARP map for the red-eyed vireo modeled with all environmental layers, northern hardwoods has the highest presence predictions, with 34% of northern hardwoods appearing as habitat in more than half of the separate distribution maps, more in agreement with NORTHWOODS. Upland conifer also showed up in the higher categories of presence (again a contradiction to NORTHWOODS). Atomic rules stating that vegetation must be either upland conifer or northern hardwoods affected the distribution prediction. The range of approximate biomass values that corresponded with presence in conifer usually fell between 60 and 217 (high to very high biomass values for mostly planted pines). Presence northern hardwoods had approximate biomass values that ranged approximately from 150 to 230 (mature biomass values). Variety values for a presence prediction typically ranged between 2 and 4, which could be interpreted as a fairly non-fragmented area, and majority values had to be either upland conifer or northern hardwoods. All of these factors combined allowed for higher prediction of northern hardwoods as habitat for the red-eyed vireo than in this model including all layers than in other models. This combination of vegetation, biomass, and neighborhood values as well as the significant appearance of northern hardwoods and upland conifer as substantial habitat area makes this distribution probably the most realistic of the predicted distributions for the red-eyed vireo. This also suggests that for some bird species understory composition and structure is important for a total habitat description and that our inductive modeling combined with structural data brought to light habitat preferences and uses not previously brought out.

5.1. *Summary*

When modeled with vegetation only, the three bird models shared very similar output maps, in particular the maps of the pine warbler and chipping sparrow. The introduction of biomass resulted in the division of forest types into all the presence categories, not just the 100% category. Certain areas and cells were predicted as inappropriate habitat (e.g. high biomass conifer forest in chipping sparrow models) or as appropriate habitat (e.g. mature forest in pine warbler models). Therefore biomass resulted in greater discernment of appropriate habitat in the cases where this is important to a particular species. The most notable influence of the neighborhood layers used in this study on the bird models was selection of large contiguous areas. This especially aided models of birds whose habitat requirements are for large contiguous areas (e.g. red-eyed vireo) and did not aid models of birds whose habitat requirements included fine spatial heterogeneity (e.g. chipping sparrow). Overall, evaluation of GARP rule-sets showed ecologically logical rules for prediction with respect to vegetation, biomass and neighborhood characteristics of bird habitat.

6. Conclusions

Multi-dimensional forest and landscape structure is known to be important in habitat selection of major taxonomic groups, including birds. We investigated the influence of selected multi-dimensional landscape and forest structure variables derived from Landsat and radar and used them in a representative inductive habitat modeling method. The results of this study showed that GARP models that included forest and landscape structure achieved higher accuracies based on training and testing data, and resulting maps and rule-sets could be interpreted to more realistic or precise depictions of a particular species habitat when compared with the models that used vegetation type only.

Our studies of just three common forest birds suggest that, in addition to biomass, it would be useful to test for influence of other structure variables. For example, the red-eyed vireo was often found in habitat that prior research had not indicated to be habitat. Our work showed that this was because of the presence of understory deciduous layers in otherwise mature pure red pine overstories. We believe that in addition to volumetric descriptors such as biomass, it would be especially useful to develop structural variables from remote sensing that include within-canopy vertical structure and overstory and understory configurations. Development of this type of structural data will advance in parallel with the wider availability of radar and lidar data.

Previous use of GARP has typically been at regional to continental scales. Our study is one of the first studies using GARP inductive modeling software at a landscape level using fine spatial resolution data and demonstrated its applicability at that scale. An interesting and useful aspect of GARP was its incorporation of more than one algorithm type in its overall methodology, and its output of rule-set archives generating useful ecological information. With respect to the important consideration of scale, while structural variables are clearly useful at the landscape level, we suggest that core variables such as vegetation biomass and landscape spatial structure may also be meaningful at regional to continental scales.

Building on the important biodiversity efforts of museums and wildlife and biodiversity organizations, advances in remotely sensed data and appropriate modeling methodologies facilitate realistic habitat and range mapping for biodiversity informatics. In addition, utility of the methods demonstrated here extends beyond the development of a habitat or range model. Once habitat has been identified, models can aid in prediction of the likely effect of future land-use change on species habitat and range.

Acknowledgements

This study was supported by the U.S. National Science Foundation, Biodiversity and Ecosystem Informatics Program, award NSF EIA-0131281. The authors would like to acknowledge the following individuals and organizations: Eric Gustafson and the U.S. Forest Service North Central Research Station for cooperative use of the NORTHWOODS database; Craig Dobson and the University of Michigan Radiation Laboratory SIR-C radar program; Steve Sjogren, Susan Emery, and Luke Langstaff of the U.S. Forest Service, Hiawatha National Forest for contributions during our field data collection; Ricardo Scachetti-Pereira, Centro de Referência em Informação Ambiental, Brazil for GARP technical expertise; and Shannon Brines and Neil Carter, University of Michigan Environmental Spatial Analysis Laboratory.

REFERENCES

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., 1976. A land use and land cover classification system for use with remote sensor data, vol. 964. U.S. Geological Survey, Reston, VA.

Anderson, R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecological Modelling 162, 211–232.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. Ecological Modelling 157, 101–118.

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., 1997. Classification, Biomass Estimation, and Carbon Dynamics of a Northern Forest Using SIR-C/X-SAR Imagery. Ph.D. Thesis, University of Michigan, Ann Arbor, 168 pp.

Bergen, K.M., Dobson, M.C., 1999. Integration of remotely sensed radar imagery in modeling and mapping of forest biomass and net primary production. Ecological Modelling 122 (3), 257–274.

Bergen, K.M., Knox, R.G., Saatchi, S., 2006. Multi-Dimensional Forested Ecosystem Structure: Requirements for Remote Sensing Observations. NASA/CP-2005-212778. NASA GSFC, Washington, D.C.

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. 036511-7-F, Radiation Laboratory, EECS Dept. The University of Michigan, Ann Arbor, MI.

Biodiversity Informatics, 2004. Online journal, University of Kansas Biodiversity Research Center, http://jbi.nhm.ku.edu.

Boecklen, W.J., 1986. Effects of habitat heterogeneity on the species-area relationships of forest birds. Journal of Biogeography 13, 59–68.

Boulinier, T., et al., 2001. Forest fragmentation and bird community dynamics: inference at regional scales. Ecology 82, 1159–1169.

Buchanan, J.B., Lewis, J.C., Pierce, D.J., Forsman, E.D., Biswell, B.L., 1999. Characteristics of young forests used by spotted owls on the western Olympic Peninsula, Washington. Northwest Science 73, 255–263.

Busby, J.R., 1991. BIOCLIM — a bioclimatic analysis and prediction system. In: Margules, C.R., Austin, M.P. (Eds.), Nature Conservation: Cost Effective Biological Surveys and Data Analysis. CSIRO, Canberra, pp. 64–68.

Chalfoun, A.D., Ratnaswamy, M.J., Thompson, F.R.I., 2002. Songbird nest predators in forest-pasture edge and forest interiors in a fragmented landscape. Ecological Applications 12 (3), 858–867.

Dobson, M.C., et al., 1995. Estimation of forest biophysical characteristics in northern Michigan with SIR-C/X-SAR. IEEE Transactions on Geoscience and Remote Sensing 33 (4), 877–895.
Dunning, J.B., Danielson, B.J., Pulliam, H.R., 1992. Ecological processes that affect population in complex landscapes. Oikos 65 (1), 169–175.
Ehrlich, P.R., Dobkin, D.S., Wheye, D., 1988. The Birder's Handbook: A Field Guide to the Natural History of North American Birds. Simon and Schuster Inc., New York, New York.
Elith, J., et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29 (2), 129–151.
ESRI, 2002. Redlands, CA.
Evers, D.C., 1991. Pine warbler. In: Brewer, R., McPeek, G.A., R.J.A. Jr. (Eds.), Atlas of Breeding Birds of Michigan. Michigan State University Press, East Lansing, MI, pp. 412–413.
Flaspohler, D.J., Temple, S.A., Rosenfield, R.N., 2001. Species-specific edge effects on nest success and breeding bird density in a forested landscape. Ecological Applications 11 (1), 32–46.
Flather, C.H., Sauer, J.R., 1996. Using landscape ecology to test hypotheses about large-scale abundance patterns in migratory birds. Ecology 77 (1), 28–35.
Freemark, K.E., Merriam, H.G., 1986. Importance of area and habitat heterogeneity to bird assemblages in temperate forest fragments. Biological Conservation 36, 115–141.
Friedl, M.A., et al., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83 (1–2), 287–302.
Garzón, M.B., et al., 2006. Predicting habitat suitability with machine learning models: the potential area of *Pinus sylvestris* L. in the Iberian Peninsula. Ecological Modelling 197 (3–4), 383–393.
Gilboy, A., 2003. Effects of Landscape Spatial Structure and Composition on Models of Bird Habitat Selection. M.S. Thesis. University of Michigan, Ann Arbor, 143 pp.
Goransson, G., 1994. Bird fauna of cultivated energy shrub forests at different heights. Biomass and Bioenergy 6 (1–2), 49–52.
Gottschalk, T.K., Huettmann, F., Ehlers, M., 2005. Thirty years of analysing and modelling avian habitat relationships using satellite imagery data: a review. International Journal of Remote Sensing 26 (12), 2631–2656.
Green, R.E., Griffiths, G.H., 1994. Use of preferred nesting habitat by stone curlews *Burhinus oedicnemus* in relation to vegetation structure. Journal of Zoology 233, 457–471.
Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135 (2–3), 147–186.
Gustafson, E.J., 1998. Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1, 143–156.
Hansen, M.C., Townshend, J.R., DeFries, R.S., Carroll, M., 2005. Estimation of tree cover using MODIS data at global, continental and regional/local scales. International Journal of Remote Sensing 26 (19), 4359–4380.
Howe, R.W., 1984. Local dynamics of bird assemblages in small forest habitat islands in Australia and North America. Ecology 65 (5), 1585–1601.
Imhoff, M.L., Sisk, T.D., Milne, G., Morgan, G., Orr, T., 1997. Remotely sensed indicators of habitat heterogeneity: use of synthetic aperture radar in mapping vegetation structure and bird habitat. Remote Sensing Environment 60, 217–227.
James, F.C., Wamer, N.O., 1982. Relationships between temperate forest bird communities and vegetation structure. Ecology 63 (1), 159–171.
Langstaff, L., 2002. Raco Plains Budworm Project, U.S. National Forest Service. Hiawatha National Forest, St. Ignace, Michigan.
Lillesand, T., Kiefer, R., Chipman, J., 2004. Remote Sensing and Image Interpretation. Wiley, New York. 763 pp.
Lillesand, T., et al., 1998. Upper Midwest Gap Analysis Program Image Processing Protocol. EMTC 98-G001, U.S. Geological Survey, Environmental Management Technical Center, Onalaska, Wisconsin.
McGarigal, K., McComb, W.C., 1995. Relationships between landscape structure and breeding birds in the Oregon Coast Range. Ecological Monographs 65, 235–260.
McGarigal, K., Marks, B., 1995. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. PNW-GTR-351, U.S. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.
Middleton, A.L.A., 1987. Chipping sparrow. In: Cadman, M.D., Eagles, P.F.J., Helleiner, F.M. (Eds.), Atlas of the Breeding Birds of Ontario. University of Waterloo Press, Waterloo, Ontario, Canada, p. 440.
Morgan, K., Freedman, B., 1986. Breeding bird communities in a hardwood forest succession in Nova Scotia. Canadian Field-Naturalist 100, 506–519.
Nelson, M.D., Buech, R.R., 1996. A test of 3 models of Kirtland's warbler habitat suitability. Wildlife Society Bulletin 24, 89–97.
Payne, K., Stockwell, D.R.B., 1996. GARP Modelling System User's Guide and Technical Reference.
Payne, R.B., 1983. A Distributional Checklist of the Birds of Michigan. Ann Arbor, University of Michigan Museums of Zoology Miscellaneous Publication, No. 164.
Pennisi, 2000. Taxonomic revival. Science 289, 2306.
Peterson, A.T., Cohoon, K.P., 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modelling 117, 159–164.
Peterson, A.T., Ball, L., Cohoon, K.P., 2002. Predicting distributions of Mexican birds using ecological niche modelling methods. Ibis 144, 27–32.
Pierce, L.E., Bergen, K.M., Dobson, M.C., Ulaby, F.T., 1998. Multitemporal land-cover classification using SIR-C/X-SAR imagery. Remote Sensing of Environment 64 (1), 20–33.
Probst, J.R., Weinrich, J., 1993. Relating Kirtlands Warbler population to changing landscape composition and structure. Landscape Ecology 8, 257–271.
Reid, S., Diaz, I.A., Armesto, J.J., Willson, M.F., 2004. Importance of native bamboo for understory birds in Chilean temperate forests. Auk 121 (2), 515–525.
Reynolds, R.T., Scott, J.M., Nussbaum, R.A., 1980. A variable circular-plot method for estimating bird numbers. Condor 82 (3), 309–313.
Robbins, C.S., Dawson, D.K., Dowell, B.A., 1989. Habitat area and requirements of breeding forest birds of the Middle Atlantic States. Wildlife Monographs 103, 1–34.
Robinson, S.K., 1981. Ecological relations and social interactions of Philadelphia and Red-eyed Vireos. Condor 83, 16–26.
Rodewald, P.G., Withgott, J.H., Smith, K.G., 1999. Pine Warbler. The Birds of North America 438, 1–27.
Scachetti-Pereira, R., 2002. Desktop GARP: Users Manual.
Siegel, R.B., DeSante, D.F., 2003. Bird communities in thinned versus unthinned Sierran mixed conifer stands. Wilson Bulletin 115 (2), 155–165.
Stockwell, D., 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 13 (2), 143–158.
Stockwell, D., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148, 1–13.
Treuhaft, R.N., Siqueira, P.R., 2004. The calculated performance of forest structure and biomass estimates from interferometric radar. Waves in Random Media 14 (2), S345–S358.
Trzcinski, M.K., Fahrig, L., Merriam, G., 1999. Independent effects of forest cover and fragmentation on the distribution of forest breeding birds. Ecological Applications 9 (2), 586–593.
Ulaby, F.T., Moore, R.K., Fung, A.K., 1986. Microwave Remote Sensing: Active and Passive: Volume III: From Theory to Applications. Artech House Inc., Norwood, MA.

Vogelmann, J.E., et al., 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67, 650–652.
Wickham, J.D., Stehman, S.V., Smith, J.H., Yang, L., 2004. Thematic accuracy of the 1992 National Land-Cover Data for the western United States. Remote Sensing of Environment 91 (3–4), 452–468.
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, pp. 1–26.