

## PUTTING A CART BEFORE THE SEARCH: SUCCESSFUL HABITAT PREDICTION FOR A RARE FOREST HERB

NORMAN A. BOURG,<sup>1,2,4</sup> WILLIAM J. MCSHEA,<sup>3</sup> AND DOUGLAS E. GILL<sup>1</sup>

<sup>1</sup>Department of Biology, University of Maryland, College Park, Maryland 20742 USA

<sup>2</sup>Smithsonian Institution National Zoological Park, Conservation and Research Center, 1500 Remount Road, Front Royal, Virginia 22630 USA

<sup>3</sup>National Zoological Park/Wildlife Conservation Society, Joint Program in Forest Ecology, Conservation and Research Center, 1500 Remount Road, Front Royal, Virginia 22630 USA

**Abstract.** The realms of rare species conservation and metapopulation biology theory are often interrelated, and hence share several basic challenges. Two of the most important are the critical and frequently difficult tasks of distinguishing a priori between habitat and nonhabitat, and then delimiting suitable habitat patches in a study area. We combined classification tree analysis, a subset of classification and regression tree (CART) modeling, with digital data layers of environmental variables in a geographic information system (GIS) to predict suitable habitat and potential new population occurrences for turkeybeard (*Xerophyllum asphodeloides*), a rare liliaceous understory herb associated with southern Appalachian pine–oak (*Pinus–Quercus*) forests, in northwestern Virginia. Sample values from eight environmental data layers and population survey data were used in the modeling process to produce a cross-validated classification tree that predicted suitable habitat in the study area. Elevation, slope, forest type, and fire frequency were the four main explanatory variables in the model. Approximately 4% of the study area was classified into five suitable habitat classes, with a misclassification error rate of 4.74%. The final 13-leaf tree correctly classified 74% of the known presence areas and 90% of the known absence areas, and ground-truthing surveys resulted in the discovery of eight new occupied habitat patches. Results of this study are important for conservation and management of *X. asphodeloides*, as well as for the applicability of the habitat modeling techniques to enhancing the study of metapopulations and disturbance regimes in Appalachian forests. In addition, they confirm the potential and value of CART and GIS-based modeling approaches to species distribution problems. Our model was successful at defining suitable habitat and discovering new populations of a rare species at the landscape scale. Similar application to other rare species could prove very useful for addressing these and other ecological and conservation issues, such as planning transplantation or reintroduction experiments, identifying metapopulation fragmentation thresholds, and formulating conservation strategies.

**Key words:** Appalachians; CART; classification tree; conservation; fire; GIS; habitat models; Melanthiaceae; metapopulations; rare plants; species distributions; *Xerophyllum asphodeloides*.

### INTRODUCTION

The realms of rare species conservation and metapopulation biology theory are often overlapping and mutually reinforcing, yet they share several basic challenges. Two of the most important are the critical and frequently difficult tasks of distinguishing a priori between habitat and nonhabitat, and then delimiting suitable habitat patches in a study area (Hanski and Simberloff 1997). As large environmental data sets in digital format have become increasingly available in ecology in recent years, the ability to analyze landscape-level variables and to include the influence of deterministic agents in modeling of population occurrences at regional scales has become feasible. In addition, the need for statistical methods less restricted

by parametric assumptions and with greater capacity for handling nonlinear interactions has grown. Such analytical techniques would be particularly useful for addressing a variety of pattern and process questions in ecology.

One particularly promising analytical method is classification and regression tree (CART) modeling (Breiman et al. 1984). First used in ecology by Verbyla (1987), this technique was utilized sparingly in the ensuing decade (Borchert et al. 1989, Lees and Ritman 1991, Moore et al. 1991, Baker 1993, Michaelsen et al. 1994, Lynn et al. 1995). More recently, CART models have proven to be powerful alternatives to traditional multiple regression-based models in a number of studies. Iverson and Prasad (1998) used regression trees to replicate successfully the current distributions and predict potential future distributions of 80 eastern U.S. tree species following climate change. In a study of the distribution of three species of California oaks

Manuscript received 2 November 2004; revised 7 March 2005; accepted 24 March 2005. Corresponding Editor: J. A. Jones.

<sup>4</sup> E-mail: nbourg@umd.edu

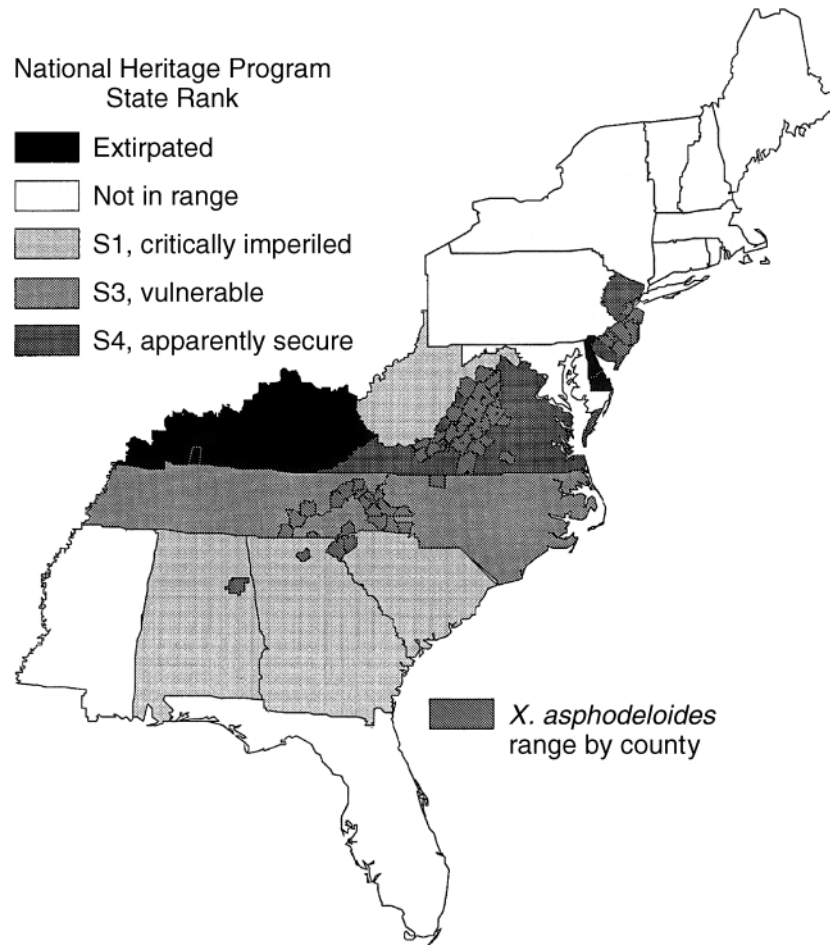


FIG. 1. Range map for turkeybeard (*Xerophyllum asphodeloides*) at the county occurrence level in the eastern United States, including current Natureserve subnational/state natural heritage program conservation status rankings (Natureserve 2005).

(*Quercus* spp.), Vayssières et al. (2000) took advantage of an extensive historical data set to compare the predictive ability of CART models and polynomial logistic regression models, and found that CART models performed significantly better in four of the six cases considered, and equally well in the remaining two cases. De'ath and Fabricius (2000) employed regression tree models to explain from 34% to 67% of the variances in the abundances of several soft coral taxa, and compared their tree results with mixed effects ANOVA and linear regression analyses of their data. In both comparisons, they found that the tree models explained nearly identical amounts of the total sums of squares, were much better at revealing patterns in the data, and were far easier to interpret due to multiple significant higher order interactions in the parametric analyses. On the other hand, Kintsch and Urban (2002) found that CART models based on environmental (physical) variables were not as effective at capturing rare species occurrences as a focal (indicator) species approach based on data from intense field surveys. Nevertheless,

the CART method did provide an important means for reducing the number of sites requiring intensive work and finer scale analysis. Finally, McKenzie et al. (2000) determined that while their regression tree models explained 20–33% more of the variation in their original data than their multiple regression models, the tree-based models were more prone to extrapolation errors when applied to broader spatial scales.

In this study, we combined CART modeling with a geographic information system (GIS) to build a predictive model of suitable habitat for turkeybeard (*Xerophyllum asphodeloides*: Liliales, Melanthiaceae; Zomlefer et al. 2001) in the Appalachians. *X. asphodeloides* is a perennial forest understory herb that occurs in discrete mountain populations from Virginia to Alabama, as well as disjunctly in the Pine Barrens of southern New Jersey. Turkeybeard is endangered or rare in portions of its range and is in the U.S. Center for Plant Conservation's National Collection of Endangered Plants (CPC 2004; Fig. 1). Despite this spe-

cial status, its spatial distribution on the landscape is inadequately known, as is the extent of suitable habitat.

No published studies exist addressing detailed aspects of turkeybeard's habitat preferences or distribution; however, our preliminary field observations suggested that a number of environmental landscape variables might correlate well with turkeybeard population occurrences. These included: (1) elevation—although present to some extent over the elevational range of the study area (168–1360 m above sea level [asl]), occurrences predominated either at high elevation summits or mid-elevation slopes; (2) slope—populations were located mainly on either nearly flat terrain or on slightly to moderately steep slopes and spur ridges; (3) aspect—populations tended to occupy predominantly west- and north-facing sites, becoming sparse to absent as one moved onto southerly and easterly aspects; (4) forest type—most populations occurred in pine, mixed pine–oak, and xeric oak forest types; (5) fire history—visible evidence of past fires were observed in most populations; and (6) perimeters—populations were discrete and appeared to be delimited by an inability to span stream drainages and rock fields.

In addition to our observations, turkeybeard is often referred to as being a fire-adapted species dependent on disturbance by fire for its long-term population persistence (WVNH 1994, 1995a, b, Farnsworth 2003). Anecdotal support for this assertion comes from its frequent association with fire-adapted pines in New Jersey (pitch pine, *Pinus rigida*), the Appalachians (*P. rigida* and table mountain pine, *P. pungens*), and even at its southern range limit of north-central Alabama, where its sole occurrence is in rare mountain longleaf pine (*P. palustris*) forest (Thurmond and Oberholster 1996). Its only congener, beargrass (*X. tenax*), occurs abundantly in Montana, portions of the Pacific Northwest and northern California, and southwestern Canada (Hitchcock and Cronquist 1973). There are few published studies of *X. tenax*, but Maule (1959), Franklin and Dyrness (1973), and Hunter (1988) mention that it is often dominant in burned forest areas. Maule (1959) also concluded that the major factors influencing the distribution of *X. tenax* on Mount Rainier, Washington, were elevation, slope, and aspect, the last of which particularly affected soil temperature, whereas soil water content had no apparent influence. Last, our results from a previous fire and canopy alteration field experiment on a long-term marked *X. asphodeloides* population and assessment of fire histories at a number of population sites have shown that turkeybeard flowers massively in response to fire and that fire has occurred repeatedly in populations in recent history (N. A. Bourg, D. E. Gill, and W. J. McShea, *unpublished manuscript*).

Based on these observations and findings, we hypothesized that such environmental variables could be used in classification tree statistical analysis to identify important distributional explanatory variables, predict

suitable habitat, and discover new population occurrences of *X. asphodeloides*. A lack of readily available fine-scale temperature or moisture GIS layers for the study area precluded their inclusion in our modeling effort; nonetheless, previous researchers have shown that our other variables could serve as reasonable proxies for underlying variation due to temperature or moisture gradients (Whittaker 1956, Burnett et al. 1998, Nichols et al. 1998, Kintsch and Urban 2002).

#### STUDY AREA AND METHODS

The study area consisted of the three northernmost ranger districts (Deerfield, Dry River, and Lee) of the George Washington National Forest (GWNF) in western Virginia and eastern West Virginia (37°58'18" to 39°07'02" N; 78°18'04" to 79°32'32" W; USDA 1993). The total land area was 227 216 ha, all within the Ridge and Valley Physiographic Province (Harvill et al. 1977).

#### *Population surveying and mapping*

Known turkeybeard populations in the study area were surveyed by placing replicate 50 × 2 m strip transects spaced at least 100 m apart on the ground. Three transects were done in each population except for two sites, where only two transects were surveyed due to time and size/shape constraints, respectively. This resulted in areal samples of 200–300 m<sup>2</sup> in each population. Plants were counted in 5-m sections along each transect, and the number of flowering individuals and inflorescences was also tallied. In addition, total direct counts of plants were conducted in two populations of small areal extent.

Mapping of the perimeter of each population was achieved by recording the surveyor's exploratory walking path with the track and waypoint functions of a Trimble Pathfinder Basic global positioning system (GPS) unit. Digital perimeter files were then downloaded and converted into polygon shapefiles for use in ArcView GIS 3.3 (ESRI 2002).

#### *Production of fire frequency layer*

Historical fire records were collected from GWNF ranger district offices for 158 wildfires that occurred on the study area from 1983 to 2000. Exact perimeters were mapped whenever available. For those records with a known size or size range and location but lacking specific perimeter maps, circles with an area corresponding to the fire acreage or mean range size were used for mapping. Only fires at least 2 ha in size were used, and both human- and lightning-caused wildfires were included in the data set. All records were digitized onto U.S. Geological Survey digital elevation models (DEMs) of the study area.

A grid consisting of 2000 × 2000 m cells was overlaid onto the study area using the Coordinate Grid Maker extension of ArcView and a center point was placed in each cell. The number of fires/cell/year was then

calculated for each point by summing the number of fires with any part of their perimeters entering the cell, and dividing the total by 18 years. Kriging, an advanced interpolation procedure that generates an estimated surface from a scattered set of points, was then performed on this data set (Isaaks and Srivastava 1989, Stein 1999, Maclean and Cleland 2003). Ordinary kriging with a 2000-m lag distance was used to estimate the semi-variogram, with an exponential model yielding the best fit. A 30-m cell size output grid was then interpolated with the Kriging Interpolator 3.2 extension to ArcView Spatial Analyst (Boeringa 2003) to yield a continuous raster layer for the study area that was suitable for use with our other data layers (Appendix A). The resulting fire frequency index values ranged from a low of zero fires/10 years to a high of nearly one fire/10 years.

#### *Other GIS layers*

U.S. Geological Survey 30-m resolution DEMs were gathered for the study area and merged to produce the elevation layer (USGS 1993). This layer was then employed to derive the slope, aspect, planar curvature index, and profile curvature index layers using their respective command functions in ArcInfo (ESRI 2002). The aspect layer was transformed to recalculated values aligned on a northeast–southwest axis to accord with the environmental moisture gradient, similar to the standard transformation of Beers et al. (1966). The planar curvature index describes the concavity/convexity of the land surface measured perpendicular to the aspect of the slope, while the profile curvature index describes the same phenomenon but measured parallel to the slope's aspect. Their values are unitless, with positive values indicating convexity and negative indicating concavity.

The forest type layer was a modified and condensed version of the USFS continuous inventory of stand condition (CISC) digital data set (USFS 1996). A total of 44 CISC forest types were condensed into nine types by grouping into broader categories, based on combining similar CISC types and grouping under the coarser scale Southern Appalachian Assessment old growth forest types classification scheme (Appendix B; SAMAB 1996).

Modification and condensation of soil types from Soil Survey Geographic Database (SSURGO) digital county soil maps (NRCS 2003) were also undertaken to arrive at the final soil type GIS layer. Soil types were grouped into seven final classes based on their similarity in six relative criteria gathered from county soil survey publications: described composition, depth, drainage, pH, fertility, and available water capacity (Appendix C).

#### *CART model production*

A data set for use in modeling was obtained by placing 90 m diameter sampling circles randomly across the study area, both within and outside of mapped tur-

keybeard population occurrences. After elimination of those circles that either overlapped or had the majority of their area located outside the study area boundary, the final data set consisted of 633 sample locations, 132 of which were known occupied suitable circles and 501 of which were probable absence (unsuitable) circles. Given the discrete nature and abrupt boundaries of turkeybeard populations observed in the field, as well as the fact that we had mapped all of the known locations identified either during our fieldwork or from consultations with natural resource managers and botanists, we felt justified in categorizing the absence samples as such. We subsequently tested this categorization by evaluating the model's classification accuracy on smaller known absence areas and the known presence areas. Mean data values of each circle for each of the eight environmental variables were extracted from their respective GIS layers with the GIS utilities function of the ERDAS Imagine image processing software program (ERDAS 1997). Distribution of the data values with respect to the dependent variable for each of the GIS layers showed that there were significant differences between classes for all continuous variables except the landform indices, and that they represented the actual proportions of the categories of the dependent variable well (Appendix D).

Classification tree analysis was then performed on this data set in the S-PLUS 6.0 statistical package (Insightful Corporation 2001) using the RPART version 3 (Recursive PARTitioning; Therneau and Atkinson 1997, Mayo Foundation 2002) library addition. RPART allows one to conduct  $v$ -fold cross-validation runs on the data to enable the determination of the optimally sized tree, a process called pruning. We performed 10-fold cross-validation, where each run consisted of 10 random divisions of the data into 90% learning and 10% test sets, using the default "Gini" index impurity measure as the splitting index and the following control parameter settings for the fitting function: minimum number of observations in a node before attempting a split = 5, minimum number of observations in a leaf (terminal node) = 2, and the default threshold complexity parameter value = 0.001. Output from the summary function of RPART was examined to evaluate competitor and surrogate split variables. In all cases, the variable that yielded the greatest improvement to deviance was chosen as the splitting variable at a given node, and surrogate variable splits were not employed because there were no missing values for any of the data set observations.

RPART employed the learning set to construct 10 classification trees, and each test set was then percolated through its respective tree to calculate a table of cross-validation error values for various tree lengths. We conducted 100 of these 10-fold cross-validation runs for a total of 1000 simulations and then tallied the optimal tree size values from their cross-validation error tables based on the two evaluation criteria of Brei-

man et al. (1984): the 1 SE rule and the minimum cross-validation error rule. The 1 SE rule states that the best tree is that which is smallest where its estimated error rate is within one standard error of the minimum. The minimum cross-validation error rule states simply that the optimal tree is that which minimizes the cross-validation error in a given run. Under either rule, the modal tree size over all the simulations is then selected as the optimal tree (Breiman et al. 1984, De'ath and Fabricius 2000). This final classification tree model was then entered into the Knowledge Engineer function of ERDAS Imagine, and pixel assignment to classes for the entire study area was performed by using the resulting classification tree file along with the eight digital raster environmental layers in ERDAS Imagine's Knowledge Classifier utility.

#### Model evaluation

We evaluated the classification of the study area in two ways: (1) by calculating the percentage of known turkeybeard presence and absence areas classified correctly (i.e., included as suitable and unsuitable habitat respectively), and (2) by ground-truthing a subset of the predicted habitat patches. We digitized 23 small to large-sized (2700–524 700 m<sup>2</sup>, mean area = 108 117 m<sup>2</sup>) known absence areas that we had traversed during terrain explorations in the population-mapping phase of the fieldwork to use along with the mapped known population occurrence areas for the first evaluation. For the second evaluation, we selected only those patches at least 5 ha in size as potential ground-truthing sites, and included patches that contained known mapped turkeybeard populations. A “naïve but competent” field technician was then employed to conduct walking transect surveys through as many of these patches as possible from November 2002 to January 2003, with time, road, and weather conditions permitting. By “naïve but competent” we mean that we selected a person who was experienced in plant surveys, plant identification, and GPS use but had no prior experience either with turkeybeard or in the study area, and did not know which patches contained mapped turkeybeard populations beforehand.

#### RESULTS

Twenty-three known turkeybeard populations were surveyed and mapped during the summer of 2000, with one additional population assessed in the summer of 2001, for a total of 24 populations. We found considerable range in the mean plant density, areal extent, and estimated population sizes among the populations, although the density of plants in flower was uniformly low across all sites (Appendix E).

Compilation of the results of the cross-validation runs yielded a five-leaf tree as the modal size under the 1 SE rule while the minimum cross-validation error criterion produced a 13-leaf tree as the optimum (Fig. 2). The 13-leaf tree, however, was the most frequently

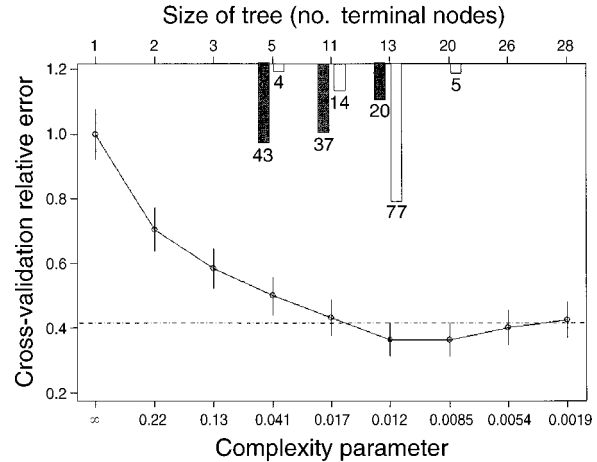


FIG. 2. Cross-validation relative error for the classification tree modeling of *X. asphodeloides* habitat. The plot is for a single representative 10-fold cross-validation and includes SE estimates for each tree size. The bar chart at the top of the plot shows the numbers of trees of each size selected under the 1 SE rule (gray) and minimum error rule (white) from a series of 100 cross-validations. The dashed line indicates the 1 SE cutoff above the minimum error value. The bottom x-axis indicates the complexity parameter values associated with each tree size.

represented size overall, achieving moderate frequency under the 1 SE rule and scoring as the best tree in 77 of the 100 runs under the minimum cross-validation error rule. Additionally, the 13-leaf tree yielded an overall misclassification error rate of only 4.74% when applied to the full model data set, whereas the five-leaf tree misclassified 7.90% of this data set. Sensitivity (i.e., true positive rate = the proportion of observations correctly identified as suitable) of this tree size, at  $0.811 \pm 0.03$ , was also considerably greater than that of the five-leaf tree ( $0.652 \pm 0.04$ ), while specificity (i.e., true negative rate) of both tree models was high and virtually identical ( $0.990 \pm 0.004$  for 13-leaf tree,  $0.992 \pm 0.004$  for five-leaf tree). Finally, use of the larger tree to classify the study area resulted in correctly classifying 74.4% of the total known presence area and 89.7% of the total known absence area, while the smaller tree was less accurate at classifying the known presence area (66.3% correct) and no better at classifying the known absence area (90.2% correct). Given these performance differences and the richer information content of the 13-leaf tree, we chose it as the optimal tree size and used it to perform the final classification of the study area (Fig. 3). Using the misclassification error terminology of De'ath and Fabricius (2000), the 4.74% misclassification error rate of this tree compares very favorably with a 50% error rate for classification of the data based on “blind guessing” and a 20.9% error rate using the “go with the majority rule” of the null model, which in this case would constitute classifying none of the sampling circles as having *X. asphodeloides* present in them.

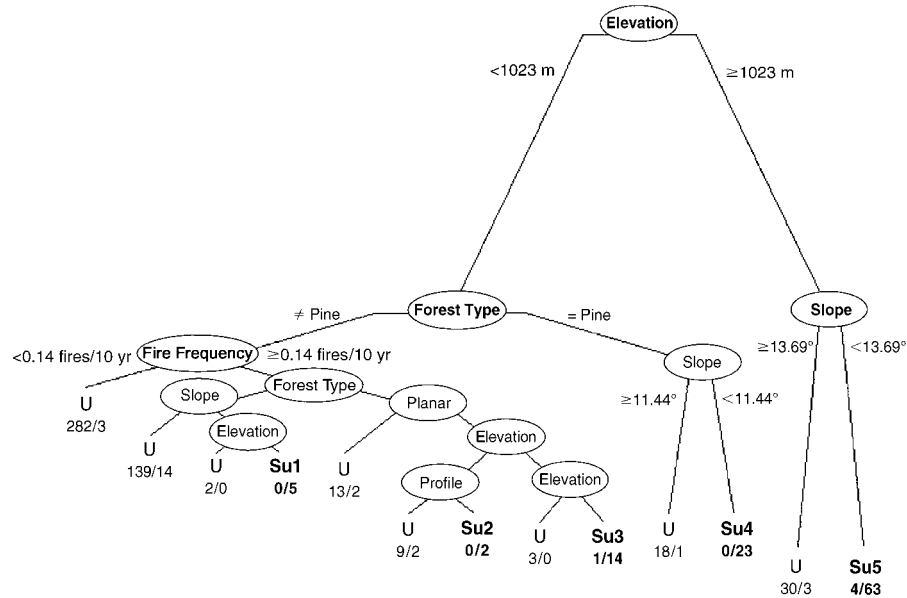


FIG. 3. Final cross-validated classification tree that served as the basis for the model of predicted habitat for turkeybeard (*X. asphodeloides*) on the George Washington National Forest study area in northwestern Virginia, USA. Splitting variables at each decision node of the tree are enclosed in ovals, and the values for each variable in the upper portion of the tree are labeled on the left and right branches emanating from each split. Values for the splits in the lower portion of the tree are as follows. Forest type: left branch = COSO, DMO, HWPNH, MM, and RW; right branch = XPO, OC, and BLB (see Appendix B). Slope: left branch,  $\geq 4.55^\circ$ ; right branch,  $< 4.55^\circ$ . Elevation: left branch,  $< 664\text{ m}$ ; right branch,  $\geq 664\text{ m}$ . Planar Index: left branch,  $< -0.025$ ; right branch,  $\geq -0.025$ . Elevation: left branch,  $< 742\text{ m}$ ; right branch,  $\geq 742\text{ m}$ . Profile Index: left branch,  $\geq -0.30$ ; right branch,  $< -0.30$ . Elevation: left branch,  $\geq 876\text{ m}$ ; right branch,  $< 876\text{ m}$ . Branch lengths below each split are proportional to the amount of variance explained by the classification variable at the split. The end nodes or "leaves" of the tree are labeled with the two classes of the dependent variable: U, unsuitable habitat; Su, suitable habitat. Numbers below the end node labels refer to the number of sample points classified into that node; the first number indicates the number of U samples placed into that leaf, and the second indicates the number of Su samples in the leaf. The five Su habitat classes are in bold type and are numbered (Su1–Su5) to correspond to the predicted suitable habitat classes in Fig. 4 and Appendix F.

Application of this 13-leaf tree model to the study area resulted in placing 3.9% (8753 ha) of the terrain into five suitable habitat classes (Figs. 3 and 4). This tree size was most successful at achieving the multiple goals of: (1) producing a moderately sized tree containing ecologically meaningful explanatory variables while obtaining a low misclassification error rate for the model data set; (2) placing the highest proportion of known population areas into the suitable habitat classes while at the same time minimizing the total amount of the study area categorized as suitable; and (3) maintaining high fidelity in known absence area classification.

The final classification tree model identified the following four variables as major determinants for explaining the distribution of turkeybeard populations and identifying suitable habitat patches: elevation, slope, forest type, and fire frequency index (Fig. 3). One hundred of 633 sampling circles in the model data set (15.8%) occurred at high elevation ( $>1023\text{ m asl}$ ), and turkeybeard was found almost exclusively on fairly gentle slopes ( $<13.7^\circ$ ) here. This high elevation category contained 50% (66 of 132) of the known presence sampling circles, and only three of these were misclassified. The remaining 533 sampling circles were

split into two major groups by the forest type and fire frequency index variables. While only 7.9% (42 of 533) of these samples fell into pine-dominated forest types, 18.2% (24 of 132) of the known presence samples occurred here, with only one misclassified and located on more steeply sloped terrain ( $\geq 11.4^\circ$ ). The fire frequency index variable was able to split out 285 of the final 491 samples (58%) into the very low to nil fire frequency category. Three of these were misclassified known presences, representing only 2.3% (3 of 132) of these data. The rest of the known presence data (39 samples or 29.5%) occurred in areas of higher fire frequency ( $\geq 0.14\text{ fires/10 years}$ ). In addition to elevation reentering into the lower branches of the tree as a classifying variable, the two measures of topographic curvature at the local scale also appeared as important factors. Suitable habitat classes in this portion of the tree model were characterized by mainly convex planar topography ( $\geq -0.025$ ; 18 of 132 presence samples, or 13.6%), with a small subset of these (4 of 132, or 3.0%) occurring at lower elevation sites ( $<742\text{ m asl}$ ) with some additional concavity in their profile topography. Last, the greatest number of misclassified samples (14 of 132 presence samples, or 10.6%) appeared in the classification node that was mainly defined as more

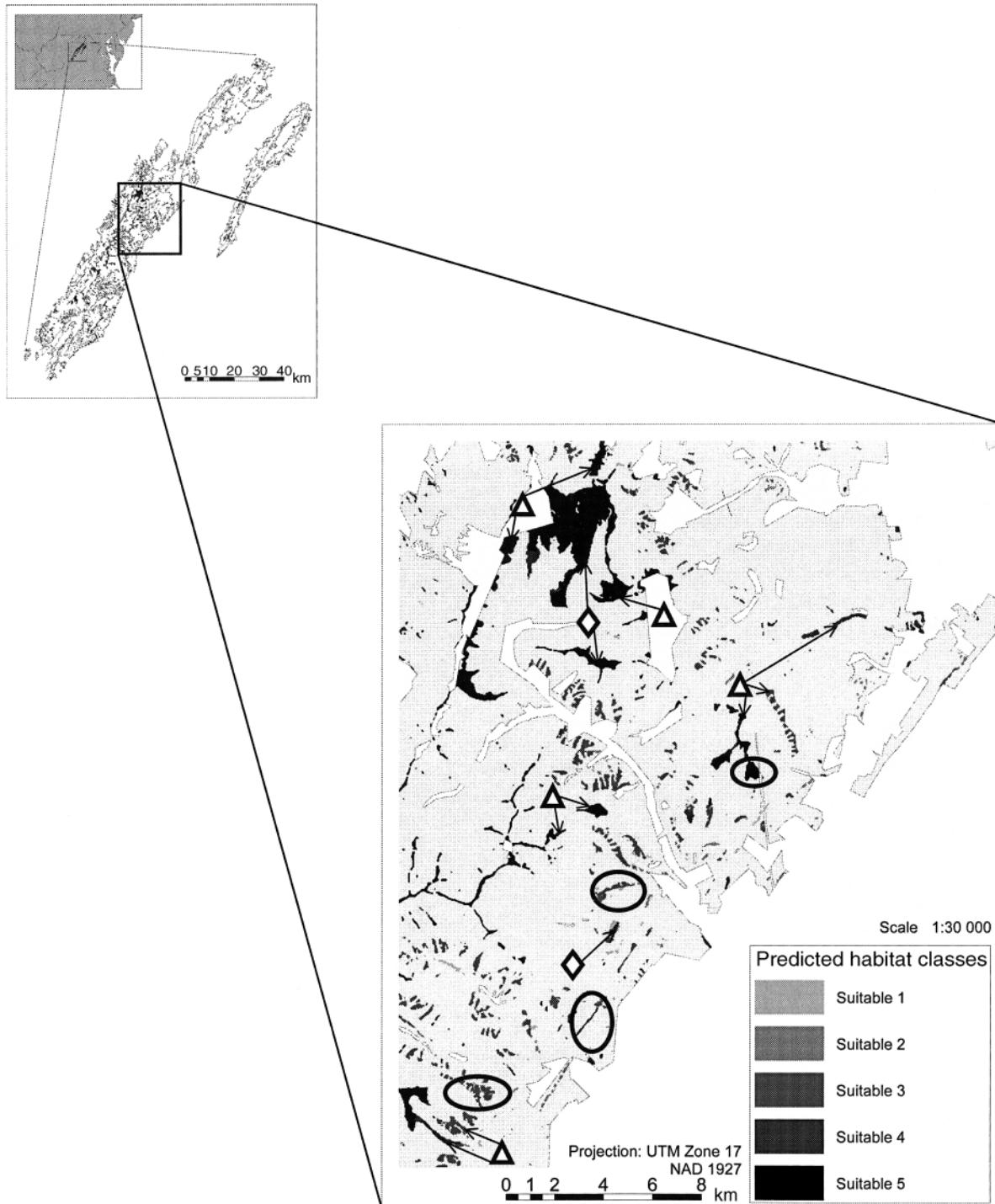


FIG. 4. Predicted habitat model map for turkeybeard (*X. asphodeloides*) on the George Washington National Forest study area in northwestern Virginia, USA (227 216 ha, inset upper left) produced from the classification tree model. Suitable habitat grayscale-coded categories in the magnified view correspond to the Su1–Su5 leaves in the classification tree in Fig. 3. Ovals encircle new population patches discovered during ground-truthing; triangles with their corresponding arrows indicate unoccupied ground-truthed patches, and diamonds with their associated arrows indicate prior known population ground-truthed patches. See Appendix F for a color version of this figure.

than gently sloped areas ( $\geq 4.55^\circ$ ) in predominantly deciduous forest types, where species such as *Quercus prinus*, *Q. coccinea*, *Q. rubra*, *Q. alba*, and to a lesser extent *Acer rubrum* were most common.

Forty-six predicted suitable habitat patches were surveyed during the ground-truthing exercises. Total distances covered in the walking transects were 54 and 69 km in unsuitable and suitable habitat, respectively. Our ground-truthing technician was 100% successful (6 of 6 patches) in finding turkeybeard in those patches harboring known, previously mapped populations that were covertly included in his search list. These surveys also resulted in the discovery of eight new, occupied suitable habitat patches, equal to a 20% occupancy rate for the remaining 40 ground-truthed patches (Fig. 4, Appendix F). All of the five suitable habitat classes were represented at least twice in these new populations, although suitable habitat class 2 covered only a small portion (<5%) of their total area. Additionally, six false negatives were found, but these all occurred near to predicted suitable habitat harboring known or new populations (mean distance from nearest suitable habitat = 171 m).

#### DISCUSSION

Our modeling effort was successful at defining suitable habitat for and discovering new populations of a rare species in a topographically complex environment at the landscape scale. The results of our approach compare favorably with the performance achieved by CART models in several recent studies (Iverson and Prasad 1998, De'ath and Fabricius 2000, McKenzie et al. 2000, Vayssieres et al. 2000, Kintsch and Urban 2002). In addition, the model's identification of elevation, slope, and fire as three of the four main explanatory variables is consistent with the findings of Maule (1959) and Franklin and Dyrness (1973) for turkeybeard's western congener, *X. tenax*. Regardless of whether elevation and slope may have served as proxies for underlying causal factors in our study, they, as well as fire frequency and forest type, played an important role in yielding a predictive model that performed well. Further investigation, GIS layer production, and model redevelopment would be needed to assess whether inclusion of other variables more directly related to factors such as temperature or moisture would improve the outcome we obtained (Vayssieres et al. 2000, Kintsch and Urban 2002).

Although we do not present evidence herein to characterize *X. asphodeloides* as having a metapopulation structure, the results are pertinent to questions of habitat definition at the metapopulation level (Husband and Barrett 1996, Freckleton and Watkinson 2002, Murphy and Lovett-Doust 2004). Over the past decade interest in the influence of spatial structure on the ecological and evolutionary dynamics of populations has increased markedly. Indeed, advocacy for a metapopulation approach to addressing this issue has become

extremely popular in both ecology and conservation biology (Husband and Barrett 1996, Hanski and Simberloff 1997, Hanski 1999:1–21). This is a powerful and justifiable approach because it encourages process- and scale-oriented research and conservation strategies that focus attention on among-population processes (Thrall et al. 2000). Such processes, the main examples of which are environmental stochasticity and deterministic threats like natural disturbance and succession, are the most likely ultimate causes of local population extinction (Thomas 1994, Harrison and Taylor 1997). Plants are particularly vulnerable to deterministic extinction threats due to their immobility and restricted capacity for dispersal.

In their review of 44 recent papers on the regional population dynamics of more than 33 plant species, Freckleton and Watkinson (2002) listed only two studies where potential suitable habitat was defined quantitatively (Lesica 1992, Giles and Goudet 1997). The remainders were either easily delimited (e.g., aquatic plants in transient pools), used qualitative or subjective definitions, or did not define potential habitat patches. This is not a trivial issue, as the occurrence of suitable habitat in discrete patches that may be occupied by local reproducing populations is one of the four necessary conditions for characterizing the regional dynamics of a species as a metapopulation (Hanski 1997). Our effort demonstrates a powerful and readily interpretable GIS-compatible modeling approach for satisfying this requirement in regional-level population investigations. In cases where a species does not conform as readily to suitable/unsuitable habitat delineations, predictive habitat gradient models, as advocated by Murphy and Lovett-Doust (2004), can still be undertaken using CART modeling in combination with grid-based GIS data layers.

Our findings also have significant conservation and management implications for the role of fire in the Appalachians as well as in the life history of *X. asphodeloides* in particular. Research on fire ecology in Appalachian forest communities has traditionally focused on effects on overstory trees and their regeneration (Zobel 1969, Barden and Woods 1973, 1976, Komarek 1974, Harmon 1982, Van Lear and Waldrop 1989, Williams and Johnson 1990, Abrams 1992, Sutherland et al. 1993, Delcourt and Delcourt 1997, Harrod et al. 1998, Shumway et al. 2001); little research has been done on the understory plant component. Fire, however, may have a major influence on the population ecology of understory species in fire-influenced communities (Bond and van Wilgen 1996, Curtis 1998, Quintana-Ascencio et al. 2003, Menges and Quintana-Ascencio 2004). Additionally, studies of understory species such as *X. asphodeloides* are valuable, given that the Appalachian forest cover of today is almost entirely mature secondary regrowth from the extensive deforestation of the 18th and 19th centuries (Shands



1992), which obliterated much of any historical fire evidence contained in the canopy trees.

No detailed ecological studies of *X. asphodeloides* had been undertaken until this research effort. Additionally, the full extent of turkeybeard population occurrences in the Appalachians is currently unknown, constituting a critical information gap that our modeling approach can now begin to fill. *X. asphodeloides* is broadly distributed within the larger forest matrix in isolated patches across a range of drier forest types (hardwood to pine-dominated) that may be subject to natural succession and human-induced habitat loss. This raises the question of whether population occurrences in turkeybeard are the result of a dynamic process of individuals tracking rare patches of suitable habitat in a metapopulation context (Harrison and Taylor 1997) or whether their distribution is a relict of ancient fragmentation processes independent of edaphic and ecologically induced patchiness (Whittaker 1956). Insights into these issues and their relation to fire occurrence are particularly relevant for conservation and management of mountain populations of turkeybeard because most occur on national forest and park lands, where fire suppression policies have been in place for much of the past century (USDA 1993).

Equally important, the predictive habitat modeling effort described herein has the potential for linkage to regional genetic studies through its ability to efficiently delineate and identify areas harboring new populations. Cruzan (2001) used stepwise regression models and genetic diversity data from multiple populations to estimate the critical number of neighboring populations (fragmentation threshold) and metapopulation diameter needed for the regional maintenance of genetic diversity in the large-flowered skullcap (*Scutellaria montana*). However, this was possible only because of the availability of population occurrence information from extensive and time-intensive prior field survey data collected by state resource managers. Applying our modeling approach could substantially enhance the feasibility of estimating such fragmentation thresholds for other species whose spatial distributions are poorly known. The results of this study are thus important not only for conservation and management of *X. asphodeloides* in the Appalachians, but also as confirmation of the potential and value of CART and GIS-based modeling approaches for addressing species distribution problems and related questions in ecology.

There are some limitations to our current habitat model. First, as one would expect, its classification accuracy decreased as population occurrence area decreased. Fifty-three percent of the known mapped small (<5 ha) turkeybeard populations were missed by the 13-leaf tree model. Adding additional sample points from other small populations to the model data set would likely improve classification accuracy for such areas. Second, the inclusion of more and better fire history data would improve the fire frequency layer

and potentially lead to fire appearing as an even stronger explanatory variable in the tree model. This assertion is supported by the fact that fire frequency was the predominant splitting variable used to classify the major remaining group of misclassified presence samples (14) in the next largest tree model (15-leaf tree). We were limited to using only the most recent 18 years of fire records because of a lack of older records covering a longer coincident time period on two of the districts of the study area. However, it appears evident that many *X. asphodeloides* sites have experienced repeated fire over at least the past 200 years, based on analyses of fire-scarred tree cross-sections collected from a subset of our turkeybeard populations on the GWNF (N. A. Bourg, D. E. Gill, and W. J. McShea, *unpublished manuscript*). It is interesting to note that even with the short time period of the available fire history data, the highest value of the fire frequency index (nearly 1 fire/10 years) corresponds well with longer term fire frequencies (8–13 years) that have been reported for Appalachian pine and oak forest types (Harmon 1982, Abrams 1992, 2000, 2003, Sutherland et al. 1993, Shumway et al. 2001). Additionally, the entry of the fire variable into the tree model at elevations below 1023 m agrees with the known patterns of lightning-caused fires on the GWNF, which occurred predominantly in the mid-elevation range (610–914 m; USFS 1997). Finally, application of the model to other potential study areas must be restricted to those having similar topographic, forest, and soil type ranges and categories, such as those found in other parts of the southern Appalachians. The basic modeling approach would be valid in other regions but would by necessity require testing and reevaluation in the new multivariate space.

Even though *X. asphodeloides* is a long-lived perennial and one of only two species in its genus, its distribution throughout the elevational range of the central and southern Appalachians and its relationship with fire argues against it being considered a relict species similar to a number of other Appalachian plants (Whittaker 1956, Godt et al. 1995, 1996, Kintsch and Urban 2002). Recent work on *X. tenax* in Oregon by Vance et al. (2004) has shown that this species has an early-acting self-incompatibility system, and our own pollination studies of *X. asphodeloides* have documented a similar self-incompatible breeding system (N. A. Bourg, *unpublished data*). This suggests that in the longer term, populations may be vulnerable to deterministic extinction brought about by natural succession and persistent low flowering levels in the absence of fire, given the present altered fire regimes in the Appalachians. If periodic fire enables population maintenance and enhances the potential for occasional gene flow among populations, then there may be validity in characterizing turkeybeard's population dynamics as a "habitat-tracking metapopulation" (Harrison and Taylor 1997).

Additional research is needed to fully evaluate this hypothesis.

Our modeling effort was successful at defining suitable habitat and discovering new populations of a rare species at the landscape scale. The model is relevant to metapopulation-level questions, and has potential for linkage to population genetic studies. Application of similar modeling efforts to other rare species could be very useful for defining suitable habitat, discovering new populations, planning transplantation or reintroduction experiments, identifying metapopulation fragmentation thresholds, and addressing a variety of other ecological and conservation questions.

#### ACKNOWLEDGMENTS

A National Fish and Wildlife Foundation General Conservation Challenge Grant to W.J. McShea, N.A. Bourg, and D.E. Gill, a Smithsonian Institution Pre-doctoral Fellowship to N.A. Bourg, the NSF Research Training Grant in the Biology of Small Populations at the University of Maryland-College Park (UMCP), and the UMCP Graduate School, provided funding for this research. We thank Terry Slater, Ecosystem Management Officer for the Dry River Ranger District of the GWNF, and Steve Croy, GWNF Ecologist, for their important assistance with this study. Dr. Peter Leimgruber and Melissa Songer provided access and guidance at the Smithsonian Conservation and Research Center (CRC) GIS laboratory. Dr. Mary C. Christman offered beneficial statistical and CART modeling input. We thank Kevin Koy for his advice on model preparation and Lyndon Estes for his assistance with ground-truthing. Kari Veblen, Carrie O'Brien, Wrenn Martin, Deb Rowan, Jenny Gagnon, Dan Kelly, Dan Guertin, Julian Bock, and other CRC interns provided valuable assistance and camaraderie in the laboratory and/or field.

#### LITERATURE CITED

- Abrams, M. D. 1992. Fire and the development of oak forests. *BioScience* **42**:346–353.
- Abrams, M. D. 2000. Fire and the ecological history of oak forests in the eastern United States. Pages 46–55 in D. A. Yaussy, compiler. *Proceedings: Workshop on Fire, People, and the Central Hardwoods Landscape* (Richmond, Kentucky, 2000). USDA Forest Service, Northeastern Research Station, General Technical Report NE-274, Newtown Square, Pennsylvania, USA.
- Abrams, M. D. 2003. Where has all the white oak gone? *BioScience* **53**:927–939.
- Baker, F. A. 1993. Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. *Plant Disease* **77**:136–139.
- Barden, L. S., and F. W. Woods. 1973. Characteristics of lightning fires in southern Appalachian forests. Pages 345–361 in *Proceedings, 13th Tall Timbers Fire Ecology Conference*. Tall Timbers Research Station, Tallahassee, Florida, USA.
- Barden, L. S., and F. W. Woods. 1976. Effects of fire on pine and pine-hardwood forests in the southern Appalachians. *Forest Science* **22**:399–403.
- Beers, T. W., P. E. Dress, and L. C. Wensel. 1966. Aspect transformation in site productivity research. *Journal of Forestry* **64**:691–692.
- Boeringa, M. 2003. Kriging interpolator 3.2 for ArcView spatial analyst. Nieuwland Automatisering, Wageningen, The Netherlands. (<http://www.nieuwland.nl>)
- Bond, W. J., and B. W. van Wilgen. 1996. *Fire and plants. Population and community biology series 14*. Chapman and Hall, New York, New York, USA.
- Borchert, M., F. W. Davis, J. Michaelsen, and L. D. Oyler. 1989. Interactions of factors affecting seedling recruitment of blue oak (*Quercus douglasii*) in California. *Ecology* **70**:389–404.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. G. Stone. 1984. *Classification and regression trees. The Wadsworth statistics/probability series*. Chapman and Hall, New York, New York.
- Burnett, M. R., P. V. August, J. H. Brown, and K. T. Kilgbeck. 1998. The influence of geomorphological heterogeneity on biodiversity: I. a patch-scale perspective. *Conservation Biology* **12**:363–370.
- CPC. 2004. *Species list of the National Collection of Endangered Plants*. Center for Plant Conservation, Missouri Botanical Garden, St. Louis, Missouri, USA.
- Cruzan, M. B. 2001. Population size and fragmentation thresholds for the maintenance of genetic diversity in the herbaceous endemic *Scutellaria montana* (Lamiaceae). *Evolution* **55**:1569–1580.
- Curtis, N. P. 1998. A post-fire ecological study of *Xanthorrhoea australis* following prescribed burning in the Warby Range State Park, north-eastern Victoria, Australia. *Australian Journal of Botany* **46**:253–272.
- De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* **81**:3178–3192.
- Delcourt, H. R., and P. A. Delcourt. 1997. Pre-Columbian Native American use of fire on southern Appalachian landscapes. *Conservation Biology* **11**:1010–1014.
- ERDAS. 1997. *ERDAS field guide*. Fourth edition, revised and expanded. ERDAS, Atlanta, Georgia, USA.
- ESRI. 2002. *ArcView and ArcInfo GIS*. Environmental Systems Research Institute, Redlands, California, USA.
- Farnsworth, E. J. 2003. CPC national collection plant profile, national collection of endangered plants, *Xerophyllum asphodeloides*. Center for Plant Conservation, Missouri Botanical Garden, St. Louis, Missouri, USA. ([http://www.centerforplantconservation.org/NC\\_Choice.html](http://www.centerforplantconservation.org/NC_Choice.html))
- Franklin, J. F., and C. T. Dyrness. 1973. *Natural vegetation of Oregon and Washington*. General Technical Report PNW-8. USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, USA.
- Freckleton, R. P., and A. R. Watkinson. 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology* **90**:419–434.
- Giles, B. E., and J. Goudet. 1997. A case study of genetic structure in a plant metapopulation. Pages 429–454 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, New York, New York, USA.
- Godt, M. J. W., J. L. Hamrick, and S. Bratton. 1995. Genetic diversity in a threatened wetland species, *Helonias bullata* (Liliaceae). *Conservation Biology* **9**:596–604.
- Godt, M. J. W., B. R. Johnson, and J. L. Hamrick. 1996. Genetic diversity and population size in four rare southern Appalachian plant species. *Conservation Biology* **10**:796–805.
- Hanski, I. 1997. Metapopulation dynamics: from concepts and observations to predictive models. Pages 69–91 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, New York, New York, USA.
- Hanski, I. 1999. *Metapopulation ecology*. Oxford University Press, New York, New York, USA.
- Hanski, I., and D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. Pages 5–26 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, New York, New York, USA.

- Harmon, M. E. 1982. Fire history of the westernmost portion of the Great Smoky Mountains National Park. *Bulletin of the Torrey Botanical Club* **109**:74–79.
- Harrison, S., and A. D. Taylor. 1997. Empirical evidence for metapopulation dynamics. Pages 27–42 in I. A. Hanski and M. E. Gilpin, editors. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, New York, New York, USA.
- Harrod, J., P. S. White, and M. E. Harmon. 1998. Changes in xeric forests in western Great Smoky Mountains National Park, 1936–1995. *Castanea* **63**:346–360.
- Harvill, A. M., Jr., C. E. Stevens, and D. M. E. Ware. 1977. *Atlas of the Virginia Flora: Part 1, Pteridophytes through Monocotyledons*. Virginia Botanical Associates, Farmville, Virginia, USA.
- Hitchcock, C. L., and A. Cronquist. 1973. *Flora of the Pacific Northwest*. University of Washington Press, Seattle, Washington, USA.
- Hunter, J. E. 1988. Prescribed burning for cultural resources. *Fire Management Notes* **49**(2):8–9.
- Husband, B. C., and S. C. H. Barrett. 1996. A metapopulation perspective in plant population biology. *Journal of Ecology* **84**:461–469.
- Insightful Corporation. 2001. S-Plus 6 for Windows user's guide and software program. Insightful Corporation, Seattle, Washington, USA.
- Isaaks, E. H., and R. M. Srivastava. 1989. *An introduction to applied geostatistics*. Oxford University Press, New York, New York, USA.
- Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* **68**:465–485.
- Kintsch, J. A., and D. L. Urban. 2002. Focal species, community representation, and physical proxies as conservation strategies: a case study in the Amphibolite Mountains, North Carolina, U.S.A. *Conservation Biology* **16**:936–947.
- Komarek, E. V. 1974. Effects of fire on temperate forests and related ecosystems: southeastern United States. Pages 251–277 in T. T. Kozlowski and C. E. Ahlgren, editors. *Fire and ecosystems*. Academic Press, New York, New York, USA.
- Lees, B. G., and K. Ritman. 1991. Decision-tree and rule-induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments. *Journal of Environmental Management* **15**:823–831.
- Lesica, P. 1992. Autecology of the endangered plant *Howellia aquatilis*; implications for management and reserve design. *Ecological Applications* **2**:411–421.
- Lynn, H., C. L. Mohler, S. D. DeGloria, and C. E. McCulloch. 1995. Error assessment in decision-tree models applied to vegetation analysis. *Landscape Ecology* **10**:323–335.
- Maclean, A. L., and D. T. Cleland. 2003. Determining the spatial extent of historical fires with geostatistics in northern lower Michigan. Pages 289–299 in N. Omi and L. A. Joyce, technical editors. *Fire, fuel treatments, and ecological restoration*. Conference Proceedings (Fort Collins, Colorado, 2002). USDA Forest Service, Rocky Mountain Research Station, Proceedings RMRS-P-29.
- Maule, S. M. 1959. *Xerophyllum tenax*, squawgrass, its geographic distribution and its behaviour on Mount Rainier, Washington. *Madroño* **15**:39–48.
- Mayo Foundation. 2002. RPART3 for S-PLUS 6 for Windows, as released March 2002. Mayo Foundation for Medical Education and Research, Rochester, Minnesota, USA.
- McKenzie, D., D. L. Peterson, and J. K. Agee. 2000. Fire frequency in the interior Columbia River basin: building regional models from fire history data. *Ecological Applications* **10**:1497–1516.
- Menges, E. S., and P. F. Quintana-Ascencio. 2004. Population viability with fire in *Eryngium cuneifolium*: deciphering a decade of demographic data. *Ecological Monographs* **74**:79–99.
- Michaelsen, J., D. S. Schimel, M. A. Friedl, F. W. Davis, and R. C. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science* **5**:673–686.
- Moore, D. E., B. G. Lees, and S. M. Davey. 1991. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Journal of Environmental Management* **15**:59–71.
- Murphy, H. T., and J. Lovett-Doust. 2004. Context and connectivity in plant metapopulations and landscape mosaics: does the matrix matter? *Oikos* **105**:3–14.
- Natureserve. 2005. NatureServe Explorer: an online encyclopedia of life [web application]. Version 4.2. NatureServe, Arlington, Virginia, USA. (<http://www.natureserve.org/explorer>)
- Nichols, W. F., K. T. Killingbeck, and P. V. August. 1998. The influence of geomorphological heterogeneity on biodiversity: II. A landscape perspective. *Conservation Biology* **12**:371–379.
- NRCS (Natural Resources Conservation Service). 2003. Soil survey geographic (SSURGO) database fact sheet. U. S. Dept. of Agriculture, Natural Resources Conservation Service, National Cartography and Geospatial Center, Fort Worth, Texas, USA. (<http://soils.usda.gov/survey/geography/>)
- Quintana-Ascencio, P. F., E. S. Menges, and C. W. Weekley. 2003. A fire-explicit population viability analysis of *Hypericum cumulicola* in Florida rosemary scrub. *Conservation Biology* **17**:433–449.
- SAMAB. 1996. Southern Appalachian Assessment (SAA) GIS database. In K. A. Herman, editor. *Southern Appalachian Man and the Biosphere Program*. Norris, Tennessee, USA.
- Shands, W. E. 1992. The lands nobody wanted: the legacy of the eastern national forests. Pages 19–44 in H. K. Steen, editor. *Origins of the national forests: a centennial symposium*. Forest History Society, Durham, North Carolina, USA.
- Shumway, D. L., M. D. Abrams, and C. M. Ruffner. 2001. A 400-year history of fire and oak recruitment in an old-growth oak forest in western Maryland, USA. *Canadian Journal of Forest Research* **31**:1437–1443.
- Stein, M. L. 1999. *Interpolation of spatial data: some theory for kriging*. Springer Series in Statistics. Springer, New York, New York, USA.
- Sutherland, E. K., H. Grissino-Mayer, C. A. Woodhouse, W. W. Covington, S. Horn, L. Huckaby, R. Kerr, J. Kush, M. Moore, and T. Plumb. 1993. Two centuries of fire in a southwestern Virginia *Pinus pungens* community. Proceedings of the IUFRO Conference on Inventory and Management in the Context of Catastrophic Events: Altered States of the Forest (Pennsylvania State University, Center for Statistical Ecology and Environmental Statistics, University Park, Pennsylvania, 1993). USDA Forest Service Northeastern Research Station, RWU NE-4153, Delaware, Ohio, USA.
- Therneau, T. M., and E. J. Atkinson. 1997. An introduction to recursive partitioning using the RPART routine. Technical Report 61, Mayo Clinic, Section of Biostatistics, Rochester, Minnesota, USA.
- Thomas, C. D. 1994. Extinction, colonization and metapopulations: environmental tracking by rare species. *Conservation Biology* **8**:373–378.
- Thrall, P. H., J. J. Burdon, and B. R. Murray. 2000. The metapopulation paradigm: a fragmented view of conservation biology. Pages 75–95 in A. G. Young and G. M. Clarke, editors. *Genetics, demography and viability of fragmented populations*. Cambridge University Press, New York, New York, USA.

- Thurmond, D., and S. Oberholster. 1996. Longleaf restoration in Northeast Alabama: Shoal Creek Ranger District, Talladega National Forest. Pages 154–155 in J. S. Kush, compiler. Proceedings First Longleaf Alliance Conference. Longleaf Pine: a Regional Perspective of Challenges and Opportunities (Mobile, Alabama, 1996). Longleaf Alliance Report Number 1. The Longleaf Alliance, Andalusia, Alabama, USA.
- USDA. 1993. Final revised land and resource management plan, George Washington National Forest. USDA Forest Service Southern Region. Supervisor's Office, George Washington National Forest, Harrisonburg, Virginia, USA.
- USFS. 1996. Continuous inventory of stand condition (CISC) forest types database. USDA Forest Service Southern Region. George Washington and Jefferson National Forests, Bridgewater, Virginia, USA.
- USFS. 1997. Fire history of the George Washington National Forest: 1915–1995. Internal report, George Washington National Forest, U.S. Dept. of Agriculture, Roanoke, Virginia, USA.
- USGS. 1993. Digital elevation models—data users guide 5. U.S. Geological Survey, Reston, Virginia, USA. (<http://edc.usgs.gov/geodata>)
- Vance, N. C., P. Bernhardt, and R. M. Edens. 2004. Pollination and seed production in *Xerophyllum tenax* (Melanthiaceae) in the Cascade Range of central Oregon. *American Journal of Botany* **91**:2060–2068.
- Van Lear, D. H., and T. A. Waldrop. 1989. History, uses, and effects of fire in the Appalachians. USDA Forest Service General Technical Report SE-54, Southeastern Forest Experiment Station.
- Vayssières, M. P., R. E. Plant, and B. H. Allen-Diaz. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science* **11**:679–694.
- Verbyla, D. L. 1987. Classification trees: a new discrimination tool. *Canadian Journal of Forest Research* **17**:1150–1152.
- Whittaker, R. H. 1956. Vegetation of the Great Smoky Mountains. *Ecological Monographs* **26**:1–80.
- Williams, C. E., and W. C. Johnson. 1990. Age structure and the maintenance of *Pinus pungens* in pine–oak forests of southwestern Virginia. *American Midland Naturalist* **124**:130–141.
- WVNHP (West Virginia Natural Heritage Program). 1994. *Xerophyllum asphodeloides* element stewardship abstract. West Virginia Natural Heritage Program, Elkins, West Virginia, USA.
- WVNHP (West Virginia Natural Heritage Program). 1995a. *Xerophyllum asphodeloides* element state ranking form as of 25 January 1995. West Virginia Natural Heritage Program, Elkins, West Virginia, USA.
- WVNHP (West Virginia Natural Heritage Program). 1995b. *Xerophyllum asphodeloides* plant characterization abstract (Global). West Virginia Natural Heritage Program, Elkins, West Virginia, USA.
- Zobel, D. B. 1969. Factors affecting the distribution of *Pinus pungens*, an Appalachian endemic. *Ecological Monographs* **39**:303–333.
- Zomlefer, W. B., N. H. Williams, W. M. Whitten, and W. S. Judd. 2001. Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae), with emphasis on *Zigadenus*: evidence from ITS and *TRNL-F* sequence data. *American Journal of Botany* **88**:1657–1669.

#### APPENDIX A

A map showing the study area location on the George Washington National Forest (inset) and kriged fire frequency raster layer for the study area, created from 158 wildfire occurrences during the period of 1983–2000, used in the classification tree and GIS-based predictive habitat modeling is available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A1.

#### APPENDIX B

A table showing condensed forest types used in the habitat modeling effort and the corresponding continuous inventory of stand condition (CISC) type codes included in them is available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A2.

#### APPENDIX C

A table showing condensed soil classes based on SSURGO data used in the habitat modeling effort and their descriptive rankings for the six relative county soil survey grouping criteria is available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A3.

#### APPENDIX D

Box–whisker plots (top three rows) and bar graphs (bottom row) showing the distribution of the modeling data set values for the eight environmental GIS layers with respect to the dependent variable (unsuitable [U] and suitable [Su]) are available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A4.

#### APPENDIX E

A table showing plant density, areal extent, and population size estimates for the 24 mapped and surveyed populations of *X. asphodeloides* in the study area, in descending order of areal extent is available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A5.

#### APPENDIX F

A color version of Fig. 4 is available in ESA's Electronic Data Archive: *Ecological Archives* E086-151-A6.