USE OF FOREST INVENTORY AND ANALYSIS INFORMATION IN WILDLIFE HABITAT MODELLING: A PROCESS FOR LINKING MULTIPLE SCALES

Thomas C. Edwards, Jr. .1, Gretchen G. Moisen2, Tracey S. Frescino2, and Joshua J. Lawler3

ABSTRACT:--We describe our collective efforts to develop and apply methods for using FIA data to model forest resources and wildlife habitat. Our work demonstrates how flexible regression techniques, such as generalized additive models, can be linked with spatially explicit environmental information for the mapping of forest type and structure. We illustrated how these maps of forest structure can be used to model wildlife habitat, focusing on the prediction of suitable habitat for cavity-nesting birds in forest systems in the Intermountain West.

Landscape data often have scale-specific resolutions and extents as well as thematic content due to methods of observation, making it difficult to scale measured responses of ecological systems either upwards or downwards. For example, use of satellite-derived data such as the National Oceanic and Atmospheric Administration's 1.1 km resolution Advanced Very High Resolution Radiometer (AVHRR) for mapping animal habitat automatically limits the scale of animal study to a 1.1 km resolution. Any gains in the ability to systematically map habitat over large spatial extents are offset by a loss of resolution relating back to the animal(s) of interest. Similarly, the kinds of ecological characteristics that plants often are associated with (e.g., micro-climates, forest structure attributes) are frequently of such fine resolution that they cannot be systematically mapped or modelled over large spatial extents. As before, gains in understanding the ecological processes that may determine plant species distributions are offset by an inability to

1 Research Ecologist. USGS BRD Utah Cooperative Research Unit, Utah State University, Logan, UT 84322-5210. Phone: (435) 797 2529; Fax: (435) 797 4025; e-mail: tce@nr.usu.edu. 2 Research Forester and Forester, respectively. USDA FS Rocky Mountain Research Station, 507 25th St., Ogden, UT 84401. 3 National Research Council Post-doctoral Fellow. US EPA, 200 SW 35th St., Corvallis, OR 97333.

map these distributions over large spatial extents.

This limitation places serious constraints on the application of landscape theory to a wide range of forestry issues ranging from the development of spatially explicit, predictive maps of forest resources for use in management to wildlife habitat modeling. The full exploration of landscape relationships requires spatially explicit depictions of habitat and other variables at fine resolutions over large spatial extents. Such depictions would allow for simultaneous exploration of relationships of variables at small spatial extents (e.g., canopy closure within forest stands) and over large landscapes (e.g., pattern of canopy closure within an ecoregion). Although it is possible to model structural attributes of habitats and vegetation on small regions using satellite imagery, the regional-scale focus of many cover-mapping efforts makes it difficult to build vegetation structure into cover maps. Current efforts provide good maps of broad cover classes at landscape levels (Homer et al. 1997), but typically provide no information on the structure of the cover type, or the spatial distribution of structure within the cover type. Recently, emphasis has been placed on linking forest data with satellite-based information not only to improve the efficiency of estimates of forest population totals, but also to produce regional maps of forest class and structure, and to explore ecological relationships (Moisen and Edwards 1999, Moisen 2000, Frescino et al. 2001, Moisen and Frescino in press). Accuracy of these types of map products is reasonably high (Edwards et al. 1998, Frescino et al. 2001).

Here we describe our collective efforts to develop and apply methods for using FIA data to model forest systems, focusing on the application of these FIA-based models for wildlife management. Our process requires two steps. The first focuses on methods for modelling habitat that provide fine-grained estimations of forest habitat type and structure over large spatial extents. The second step is to use these representations of landscapes for modelling habitat use by terrestrial vertebrates at multiple scales. We illustrate how flexible regression techniques, like generalized additive models (GAM), can be linked with spatially explicit environmental information to map forest habitat structure. We next illustrate how the spatially explicit maps of forest structure can be used to model wildlife habitat, focusing on the prediction of suitable habitat for cavity-nesting birds in forest systems at landscape scales

EXAMPLE APPLICATIONS

Study Area

The two example studies outlined here were conducted in the Uinta mountains of northern Utah, USA. The Uintas are characterized by an east-west orientation, and have an approximate length of 241 km, and a width of 48 to 64 km. Elevation ranges from ~1,700 m to ~4,000 m. The area contains conspicuously deep, V-shaped canyons on the south side of the range and less pronounced canyons on the north side of the range. The distribution of vegetation in the Uinta Mountains is highly influenced by topographic position and geographic location. Lodgepole pine (*Pinus contorta)* is the dominant vegetation type, ranging from 1,700 to 3,000 m elevation. At elevations between 2,400 m and 3,000 m, lodgepole is mixed with aspen (*Populus tremuloides*), with a few homogenous aspen stands at lower elevations. As elevation increases, lodgepole forests are gradually replaced by spruce-fir (*Picea engelmannii-Abies lasiocarpa*) forest types and are frequently interspersed with large patches of wet and dry meadows. Other forest types include pinyon-juniper (*Pinus edulis-Juniperus osteosperma*) at lower elevations on the northeastern slope, Douglas-fir (*Pseudotsuga menziesii*) on steep, protected slopes, and ponderosa pine (*Pinus ponderosa*) forests on exposed slopes on the south side of the range (Cronquist et al. 1972).

Modelling Forest Habitat Pattern and Structure

If a major objective of landscape modelling is to enhance understanding of relationships at multiple scales as a precursor for forest management, then methods for modelling scale-related ecological parameters are paramount. From a vegetation perspective, the principle question is how to accurately and efficiently model vegetation structure and patterns at multiple scales. Recent advances in statistical modelling techniques (McCullagh and Nelder 1989, Hastie and Tibshirani 1990, Hastie et al. 2001) and geographical tools, such as remote sensing and geographical information systems (GIS), have increased the opportunities for the delineation and analysis of vegetation structure

and pattern.

Readers are referred to Frescino (1998), Moisen and Edwards (1999), Moisen (2000), Frescino et al. (2001), and Moisen and Frescino (in press) for details regarding the complexities of generating spatially explicit models of forest structure using FIA data. The process is necessarily complex, and only a short overview of work in the Uinta Mountains is presented below. In this study area, five response variables collected on FIA plots were modeled as functions of a wide variety of digitally available explanatory variables through GAMs (Frescino 1998, Frescino et al. 2001). Response variables included binary forest and lodgepole presense, as well as continuous basal area, percent shrub cover, and snag density (Table 1). A variety of explanatory variables were considered, and included those related to topography, precipitaion, geology, spatial position, as well as several from TM and AVHRR platforms (Table 2) As noted above, the GAMs used for modelling purposes are nonparametric extensions of the more commonly used generalized linear models (GLM). The GAM, like the GLM, uses a link function to establish a relationship between the mean of the response variable(s) and a smoothed function of the explanatory variable(s) . The main attraction of GAMs for vegetation modelling is their ability to handle non-normal features in the data such as bimodality or asymmetry. GAMs are best described as data-driven rather than model-driven, such that the data determine the shape of the response curves rather than fitting a known function to the data. The major weakness of GAMs is the danger of over-fitting the data (Austin and Meyers 1996).

For forest and lodgepole presence, a logit link was used to transform the mean of the response to a binomial scale. For the continuous variables (basal area, percent shrubs, snag density), a Poisson link was used to transform the data to the scale of the response. A loess smoothing function (see Venables and Ripley 1997 for description) was chosen to summarize the relationship between the predictors and the response. One limitation of smoothed functions obtained from GAMs is their inability to extrapolate outside the range of the data used to build the model. To handle this problem, values of the prediction and validation data sets that were outside the range of the model-building data

set were assigned the maximum/minimum value of the respective variable in the data set.

The functional relationships between each explanatory variable and the respective response variables were analyzed for potential parametric fits following guidelines in Hastie and Tibshirani (1990) and Yee and Mitchell (1991). If a potential parametric fit existed, piecewise and second- and third-order polynomial functions were fit to the data and assessed based on the relative degree of change to the residual deviance (Cressie 1991). All explanatory variables, including all potential parametric fits, were run through a stepwise procedure to determine the best-fit model for prediction (see Chambers and Hastie 1992) using Akaike's Information Criterion (AIC). A percent deviance reduction (D^2) was also calculated for each model, representing the percent of deviance explained by the respective model (Yee and Mitchell 1991). Once the model fits were derived (see Frescino et al. 2001, Tables 3 and 4), the model was applied to all the explanatory digital layers (Table 2) and predictive map surfaces generated. The result was a series of predictive maps of forest attributes having fine resolution (~0.8 ha) and covering large spatial extents (>1 million ha) (Fig. 1).

Accuracy of the models predicting forest and lodgepole presence was high, ranging from 86% to 80%, respectively. Sixty-seven percent of the basal area validation points fell within ±15% (11.5 m^2/ha) of the true value, 75% percent of the shrub density validation points fell within ±15% of the true cover, but only 54% of the points fell within ±15% of the true snag count.

Modelling Cavity Bird Nesting Habitat in Forested Systems

Once the maps of forest attributes are generated, the next step is to generate models of bird presence based partly on the spatially explicit forest maps. As an example in the Uinta Mountains, habitat associations based on landscape patterns were modelled for four species of cavity-nesting birds nesting in aspen . (Lawler 1999, and Lawler and Edwards in press). These species include red-naped sapsuckers (*Sphyrapicus nuchalis*), northern flickers (*Colaptes*

auratus), tree swallows (*Tachycineta bicolor*), and mountain chickadees (*Parus gambeli*).

Habitat models were built for each of the four species using classification trees (Breiman et al. 1984, Venables and Ripley 1997). Classification trees are a flexible and simple tool for modelling complex ecological relationships (De'ath and Fabricius 2000). Classification trees work by recursive partitioning of the data into smaller and more homogenous groups with respect to the response variable. Each split is made by the explanatory variable and the point along the distribution of that variable that best divides the data. See De'ath and Fabricus (2000) and Lawler and Edwards (in press) for a more thorough discussion of the use of classification trees in ecological modelling.

The four species models included a number of variables pertaining to the amount and configuration of forest and open area (Fig. 2) Spatially explicit predictions for each of the four species were produced from these models (Fig. 3). The spatial configuration of forest that was predicted as suitable nesting habitat differed among the four species, and the models varied in their ability to correctly predict nests at the new sites (Lawler and Edwards in press). The northern flicker model was the most accurate (84% of nests correctly classified). The red-naped sapsucker and tree swallow models were also relatively accurate (80%, and 75% of the nests correctly classified, respectively). The mountain chickadee model was far less accurate, correctly predicting only 50% of the nests at the test sites. These estimates are within ranges of accuracy reported elsewhere (Edwards et al. 1996).

DISCUSSION

The ability to create spatially explicit depictions of vegetation type and structure is dependent, in part, on the flexibility and capability of the models used to predict vegetation characteristics. GAMs, in contrast to some analytical procedures (e.g., ordination and linear regression models), do not make *a priori* assumptions about underlying relationships, thus allowing the data to drive the fit of the model instead of the model driving the data. The graphical nature of GAMs also allows a visualization of the additive contribution of each variable to the respective response

using smoothed functions. One limitation of GAMs is the uncertainty associated with extrapolation of the smoothed functions, particularly at the tails of the distribution. As suggested by Hastie and Tibshirani (1990) and Yee and Mitchell (1991), parametric functions were to the model whenever statistically allowable, thus constraining the behavior of the functions in the extreme ranges of the data. See Moisen and Frescino (in press) for a comparison of alternative modelling techniques for predictive mapping applications.

Once the vegetation type and structure is modelled, the resultant maps can be linked with wildlife models and used to create predictive maps. Although predictive models based on landscape patterns may prove to be a promising technique in light of their ease of use and relative accuracy, like all models they have distinct shortcomings. The ability to build such models depends on having access to remotely sensed data. Fortunately, remotely sensed data is not only becoming more diverse but it is also more widely available. Because the field of landscape ecology is relatively young, associations between given species and landscape patterns are not as prevalent in the literature (Karl et al. 1999) as are associations with the composition and structure of vegetation at relatively fine spatial scales (e.g., Cody 1985). Thus many of the basic habitat associations related to landscape patterns will need to be determined in the field for the first time.

Selecting the scales at which to measure landscape patterns is difficult when modelling several different species. Different species are likely to respond to their environment at different spatial scales (Wiens 1989). Models built solely at coarse spatial scales and using only vegetation type are likely to be less accurate when fine-scale associations with structural attributes are strong. Our approach, which employs techniques capable of modelling fine-scale attributes (e.g., canopy closure, stem density) at fine resolutions, overcomes this issue and generally increases model predictive capabilities. The use of new, more flexible modelling techniques such as classification trees (De'ath and Fabricus 2000) may further improve the predictive capability of models of forest resources, and the wildlife dependent on these resources, as well as the ease of model building and interpretation. Although our results indicate that our approach may not work equally well for all species, we found that when tested and refined, models that rely on

landscape patterns derived from FIA data may provide a reliable alternative to traditional wildlife models which require the collection of habitat data in the field and have no spatial resolution.

ACKNOWLEDGMENTS

Funding for this research was provided by the USDA Forest Service, Rocky Mountain Research Station, Ogden, Utah, in cooperation with the USGS Biological Resources Division, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, and the USGS Biological Resources Division Gap Analysis Program. We thank J. Bissonette, D. Major, and R. Schultz for their helpful reviews and, once again, acknowledge our ever-constant research companion, T. B. Murphy.

LITERATURE CITED

Austin, M.P.; Meyers, J.A. 1996. Current approaches to modelling the environmental niche of eucalypts: implication for management of forest biodiversity. Forest Ecology and Management 85:95-106.

Breiman, L.; Friedman J.H.; Olshen, R.A.; Stone, C.J. 1984. Classification and regression trees. Pacific Grove, CA: Wadsworth and Brooks/Cole.

Chambers, J.M,; Hastie, T. (eds.). 1992. Statistical models in S. Pacific Grove, CA: Wadsworth and Brooks/Cole.

Cody, M.L. 1985. Habitat selection in birds. San Diego, CA: Academic Press.

Cressie, N.A. C. 1991. Statistics for spatial data. New York, NY: John Wiley and Sons.

Cronquist, A.; Holmgren, A.H.; Holmgren, N.H.; Reveal, J.L. 1972. Intermountain flora volume 1: vascular plants of the Intermountain West, U.S.A. New York, NY: Hafner Publishing Company, Incorporated.

De'ath, G.; Fabricius, K.E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81: 3178-3192.

Edwards, T.C. Jr.; Deshler, E.T.; Foster, D.; Moisen, G.G. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. Conservation Biology 10: 263-270.

Edwards, T.C., Jr.; Moisen, G.G.; Cutler, R. 1998. Assessing map uncertainty in remotely-sensed, ecoregion-scale cover-maps. Remote Sensing of Environment 63:73-83.

Frescino, T.S. 1998. Development and validation of forest habitat models in the Uinta Mountains, Utah. Unpublished M.S. Thesis, Utah State University, Logan, UT.

Frescino, T.S.; Edwards, T.C., Jr.; Moisen, G.G. 2001. Modelling spatially explicit forest structural variables using generalized additive models. Journal of Vegetation Science 12:15-26.

Hastie, T.; Tibshirani, R.J. 1990. Generalized additive models. London, UK: Chapman and Hall.

Hastie, T.; Tibshirani, R.J.; Freidman, J. 2001. The elements of statistical learning. Data mining, inference, and prediction. New York, NY: Springer-Verlag.

Hintze, L.F. 1980. Geologic map index of Utah. Utah Geological and Mineralogical Survey, Salt Lake City, Utah.

Homer, C.G.; Ramsey, R.D.; Edwards, T.C., Jr.; Falconer, A. 1997. Landscape cover-type modelling using a multi-scene thematic mapper mosaic. Photogrammetric Engineering and Remote Sensing 63:59-67.

Lawler, J.J. 1999. Modelling habitat attributes of cavity-nesting birds in the Uinta Mountains, Utah: a hierarchical approach. Unpublished Ph.D. Dissertation, Utah State University, Logan, UT.

Lawler, J.J.; Edwards, T.C., Jr. In press. Landscape patterns as predictors of nesting habitat: a test using four species of cavity-nesting birds. Landscape Ecology.

McCullagh, P.; Nelder, J.A. 1989. Generalized linear models. London, UK: Chapman and Hall.

Moisen, G.G. 2000. Comparing nonlinear and nonparametric modelling techniques for mapping and stratification in forest inventories of the Interior Western USA. Unpublished Ph.D. Dissertation, Utah State University, Logan, UT.

Moisen, G.G.; Edwards, T.C., Jr. 1999. Use of generalized linear models and digital data in a forest inventory of northern Utah. Journal of Agricultural, Biological and Environmental Statistics 4:164-182.

Moisen, G.G.; Frescino, T.S. Comparing five modelling techniques for predicting forest characteristics. In press, Ecological Modelling.

Roberts, D.W.; Cooper, S.V. 1989. Concepts and techniques of vegetation mapping. In: Ferguson, D.; Morgan, P.; Johnson, F.D., (eds.). Land classifications based on vegetation: applications for resource management. USDA For. Serv. Gen. Tech. Rep. INT-257, Ogden, UT: 90-96.

Swift, L.W., Jr. 1976. Algorithm for solar radiation on mountain slopes. Water Resources Research 12:108-112.

Venables, W.N.; Ripley, B.D. 1997. Modern applied statistics with S-plus. New York, NY: Springer-Verlag.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Yee, T.W.; Mitchell, N.D. 1991. Generalized additive models in plant ecology. Journal of Vegetation Science 2:587-602.

Table 1. Summary of response variables for modeling forest attributes in the Uinta Mountains, Utah, USA. See Frescino et al. (2001) for additional details. P = proportion of model-building points defined as forest and lodgepole pine, respectively.

Forest attribute	Type	Description	Distribution
Forest presence	Binomial	>10% tree cover	P = 0.77
Lodgepole pine presence	Binomial	Majority of forest cover	P = 0.31
Basal Area (m^2/ha)	Continuous	Area of trees at 1.37 m basal ht. (Trees > 2.5 cm DBH)	Range: 0 to 70 Median: 16
Shrubs (%)	Continuous	Sum of total cover from upper, mid, and lower layers	Range: 0 to 92 Median: 15
Snag Density	Continuous	Total salvable and non-salvable (Snags > 10.2 cm DBH)	Range: 0 to 248 Median: 5

Table 2. Summary of explanatory variables used to model forest attributes in the Uinta Mountains, Utah, USA. See Frescino et al. (2001) for additional details.

Variable	Type	Resolution	Source
Elevation (m)	Continuous	90 m	DMA
Aspect (°)			Derived from DMA
	Continuous	90 m	Relative annual solar radiation (Swift 1976)
	Discrete	90 m	9 categories
	Continuous	90 m	Radiation/wetness index (Roberts and Cooper 1989)
Slope (%)	Continuous	90 m	Derived from DMA
Precipitation	Continuous	90 m	Downscaled from PRISM; yearly precipitation climate maps
Geology			Hintze (1980)
	Discrete	1:500,000	Timeframe (1-Precambrian, 2-Mississippian to Euocene, 3-Alluvium)
	Discrete	1:500,000	Nutrients (1-sandstone and limestone, 2-sedimentary, 3-alluvial)
	Discrete	1:500,000	Rock Type (1-sedimentary, 2-alluvial)
Easting	Continuous	-	UTM Easting coordinates
Northing	Continuous	-	UTM Northing coordinates
District	Discrete	-	7 National Forest Ranger Districts
TM-classified	Discrete	90 m	GAP Analysis (Homer et al. 1997)
AVHRR	Continuous	1000 m	NOAA (June 1990)
TM			Landsat TM (June 1990/August 1991)
	Continuous	30 m	TM Band 3 (Red)

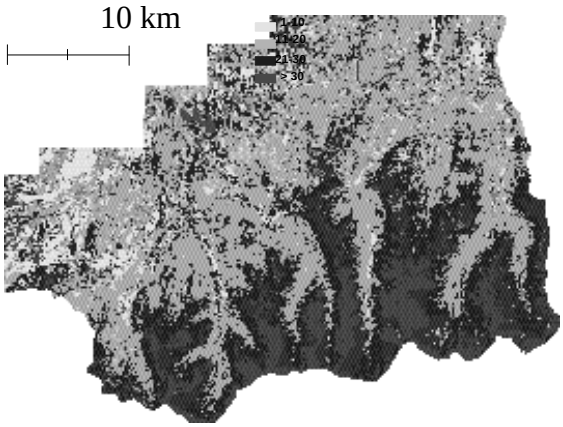
Continuous	30 m	TM Band 4 (Near-infrared)
Continuous	30 m	TM Band 5 (Mid-infrared)

Figure 1. *Example maps of nominal (lodgepole presence) and continuous (basal area) responses generated for an ~100,000 ha region of the Uinta Mountains, Utah (from Frescino 1998).*

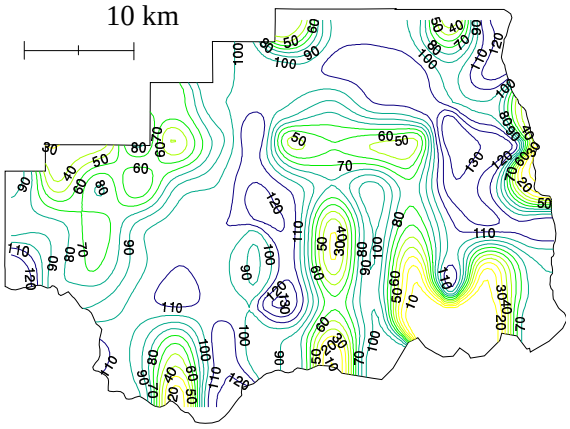
Figure 2. *Classification and regression tree model predicting nesting habitat for red-naped sapsuckers. Models for the other species were similar in structure, varying only in the predictor variables and tree complexity (see Lawler and Edwards, in press).*

Figure 3. *Vegetation and spatially explicit prediction maps for northern flicker nesting habitat. Medium gray in the vegetation map represents suitable nesting habitat, and is based on classical WHR approaches (see text). Note how the amount and distribution of gray is reduced under the refined vegetation models, which then are incorporated in the wildlife models as described in the text. Nests are represented as circles with cross-hairs.*

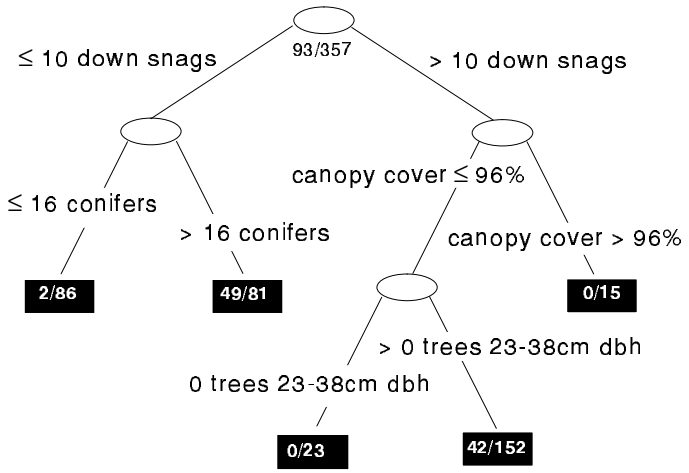
Probability map of lodgepole, Evanston Ranger District, Uinat Mountains, Utah, USA.

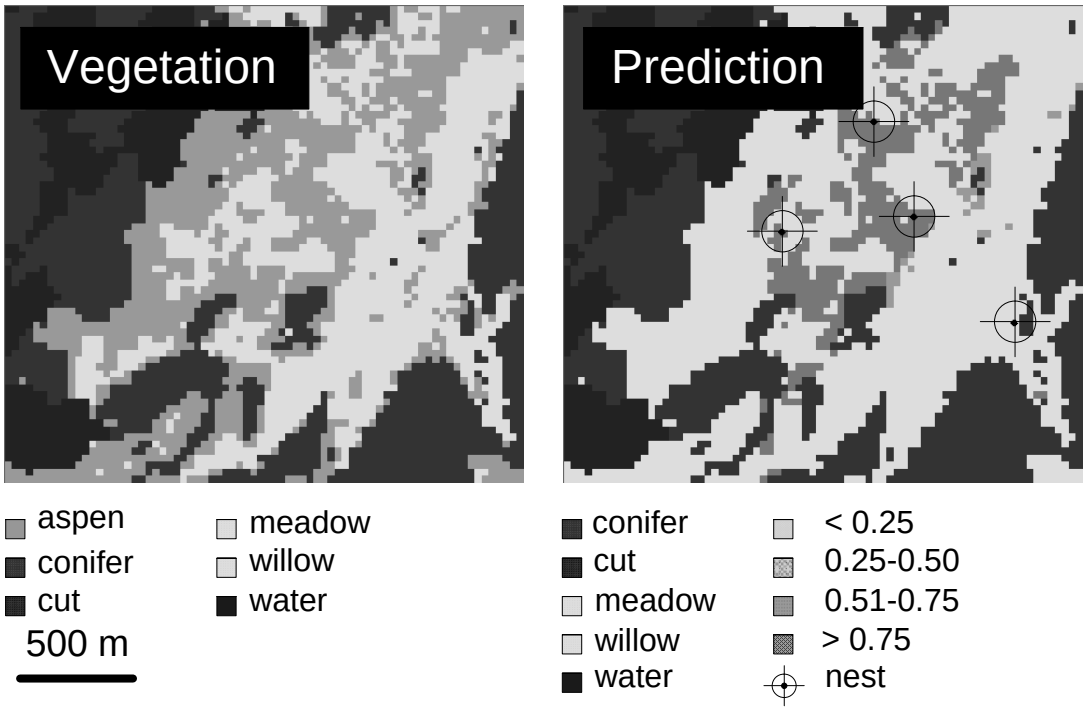


Contour map of predicted basal area, Evanston Ranger District, Uinta Mountains, Utah, USA.



CART model
Red-naped sapsucker




Vegetation
Prediction
aspen
conifer
cut
meadow
willow
water
500 m
conifer
cut
meadow
willow
water
< 0.25
0.25-0.50
0.51-0.75
> 0.75
nest