Southwest

Great Plains

Research Note RM-487

August 1988

USDA Forest Service

Rocky Mountain Forest and Range Experiment Station

Pattern Recognition—a Simple Model for Evaluating Wildlife Habitat

Teryl G. Grubb[1]

Pattern recognition (PATREC) uses information on the frequency of specific habitat characteristics in areas of a particular habitat suitability class. The result is a simple probability that the sample being evaluated belongs to the designated suitability class. This modeling technique, which was first developed in the medical field for computer diagnosis of disease symptoms, shows excellent potential for application in wildlife management. With habitat data obtained from aerial photography and topographic maps, a PATREC model has been developed on environmental features necessary for bald eagle (*Haliaeetus leucocephalus*) breeding occupancy in Arizona.

Keywords: Habitat, modeling, pattern recognition, wildlife management, bald eagle

Introduction

Pattern recognition (PATREC) is a probabilistic procedure originally developed in the medical profession for the diagnosis of disease (Lusted 1968). PATREC is based on Bayesian statistics whose basic premises are: (1) probabilities are orderly opinions; (2) statistics are concerned with the revision of opinions in light of new information; and (3) Bayes' theorem of probability theory is an optimal rule about how such revisions should be made (Edwards et al. 1963). As adapted to wildlife habitat evaluation by Williams et al. (1977), PATREC captures in simple mathematical form the process by which most biologists intuitively assess relative habitat conditions. PATREC's simplicity, its similarity to normal thought processes, and its statistical derivation increase its potential value as a wildlife management technique.

The purpose of this paper is twofold: first, to explain PATREC as a simple and useful technique for habitat modeling; and second, to present a first approximation model developed as an example for bald eagle breeding habitat in Arizona. For a thorough description of PATREC, see Kling (1980) and Williams et al. (1977); the former includes a step-by-step User's Guide. (See also Evans 1983 and Wilson 1983.) Limitations of PATREC related to possible inconsistencies with ecological theory and inherent subjectivity have been analyzed by Flather and Hoekstra (1985) and Flather (1982), respectively. Seitz et al. (1982) also discuss some limitations of PATREC; but they found it performed better than a nonstructured, "personal opinion approach." Bald eagle habitat data also indicate PATREC may be a valuable tool for field biologists.

[1]*Research Wildlife Biologist, Rocky Mountain Forest and Range Experiment Station, Research Work Unit in Tempe, in cooperation with Arizona State University. Station headquarters is in Fort Collins, in cooperation with Colorado State University.*

Methods

The following paragraphs quoted from Kling (1980) briefly explain PATREC and its component probabilities:

"PATREC involves use of information on the frequency with which specific habitat attributes occur among areas of a particular habitat suitability class, as well as comparable information on the frequency with which the same components occur among areas of other habitat suitability classes (e.g., the habitat attribute: 15-20% canopy cover of sagebrush (*Artemisia* sp.) may occur on 70% of the areas considered as highly suitable habitat for sage grouse and on only 20% of the areas considered as less suitable habitat). Frequencies of occurrence for the various habitat suitability classes can be called conditional probabilities and habitat attributes can be called diagnostic criteria. Diagnostic criteria and their associated conditional probabilities are used to evaluate an area of unknown quality by ascertaining the status (presence or absence) of the habitat attributes and then calculating the probability of the area being highly suitable with the use of Bayes' theorem and conditional probability values. The probability of having a highly suitable habitat can be used as an index of habitat quality.

"The PATREC approach involves formally identifying (1) the categories of habitat suitability into which an area can be classed, (2) the habitat attributes which should be examined during an evaluation, and (3) a set of probabilities which reflect the extent or frequency of association between individual habitat attributes and each suitability class. Habitat suitability of an area can then be estimated after the presence or absence of individual habitat attributes has been ascertained and a few relatively simple calculations have been completed to synthesize the information.

"A model consists of: (1) 2 or more classes indicating suitability of the habitat to support a given wildlife species (e.g., suitability and unsuitability), (2) a set of habitat attributes, which could also be termed habitat components or habitat requirements, that one looks for when evaluating habitat (e.g., 10% shrub canopy cover), and (3) a set of probabilities that reflect the degree of association between individual habitat attributes and each habitat suitability class. The probabilities included in each model are conditional probabilities, that is, they represent the probability that some condition will occur (e.g., a particular set of habitat features) given that some other condition (e.g., a high density population) also occurs.

"Bayesian statistical inference is a mathematical technique commonly used when decisions must be made under conditions of uncertainty. Reduced to its most fundamental steps, the technique requires an investigator to (1) estimate the probability that some condition(s) exists or will exist in the future, (2) collect sample data related to the condition, and (3) revise the initial probability estimates to take into account the sample results. The initial probability estimates (step 1) are often referred to as prior probabilities and may be based upon historical information or intuition. They constitute the decision-maker's best estimates that a condition exists prior to learning the details surrounding the current situation. For similar reasons, the revised probabilities (products of step 3) are generally referred to as posterior probabilities."

Figure 1 depicts PATREC as described in the foregoing paragraphs. The left insert shows the actual model, and the right insert explains the Bayesian conversion from Conditional Probabilities to Posterior Probabilities, or the final habitat assessment. Prior Probabilities are a best estimate of the likelihood of occurrence for each of the Habitat Suitability Classes. These can be set at 0.5 if one simply does not know what to expect; or, Prior Probabilities can be established empirically if data, contacts, or experience permits. Using 0.5 (i.e., a 50-50 chance of occurrence for each suitability class) is satisfactory. The selection of Habitat Suitability Classes such as occupied or unoccupied, successful or unsuccessful, and the selection of Habitat Attributes, or diagnostic criteria, are straightforward. The latter process requires some insight

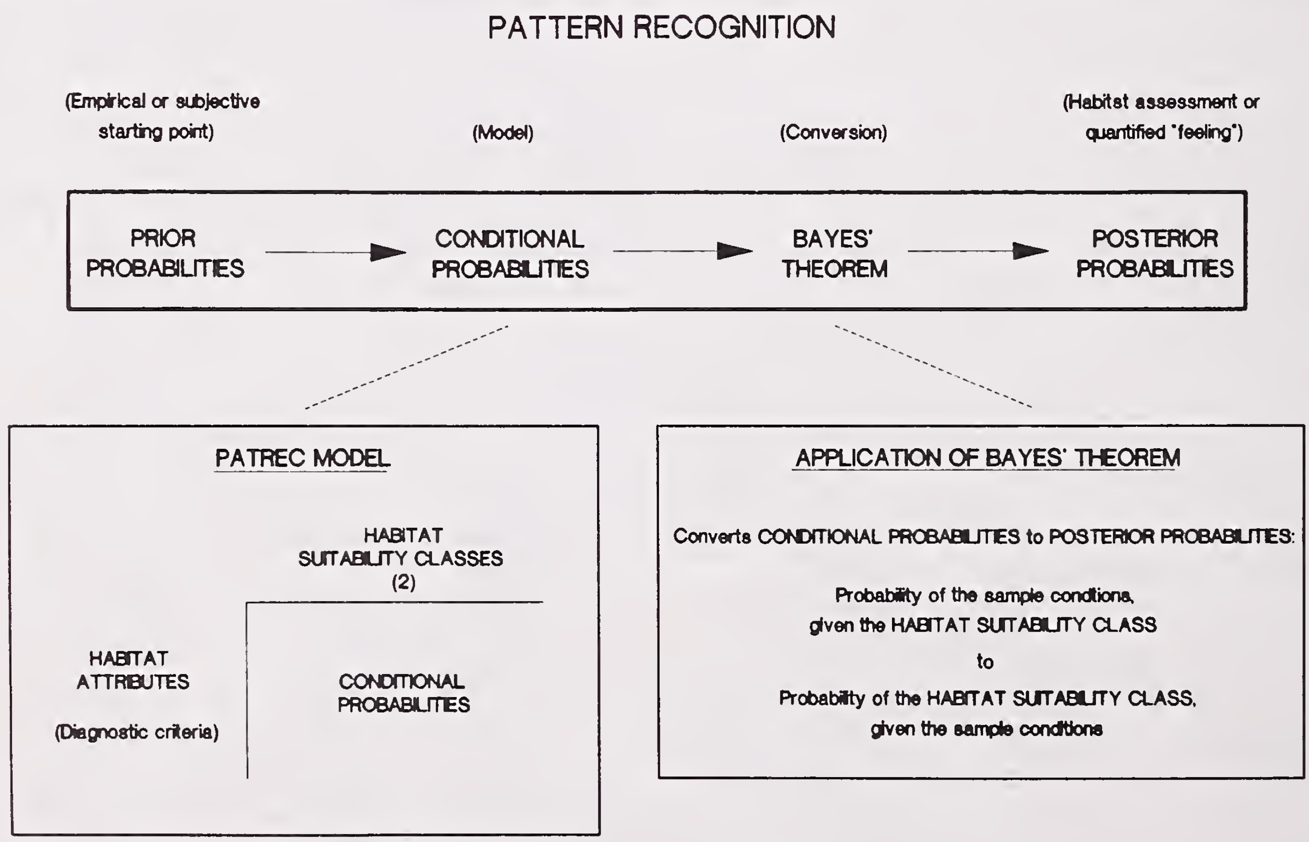

Figure 1.—A schematic of the components of the Pattern Recognition (PATREC) habitat assessment process.

Table 1. First approximation PATREC habitat model for bald eagle breeding habitat in Arizona

Habitat attributes	Habitat suitability classes	
	Occupied	Unoccupied
Terrain		
1) Cliffs, ≥1.5 miles of ≥100 feet height	0.75	0.20
Aquatic habitat		
2) Ratio of river miles:straight miles ≥1.5	0.89	0.30
3) Miles of permanent side drainages ≥0.5	0.39	0.01
Vegetation		
4) Percent (%) riparian cover ≥5.0	0.38	0.10
5) Ratio of riparian miles:river miles ≥0.3	0.81	0.50

by the model designer and is restricted by the statistical requirement of independence between attributes. Conditional Probabilities, the key to a PATREC model, may be best explained by an example.

If 8 of 10 known occupied sites have Attribute A, the Conditional Probability of Attribute A given occupied habitat is 0.80. In the same analysis if 3 of 10 unoccupied sites have attribute A, the Conditional Probability of Attribute A given unoccupied habitat is 0.30. The PATREC model then is constructed as a list of Conditional Probabilities for each Habitat Suitability Class for all diagnostic criteria, or Habitat Attributes. The difference between the Conditional Probabilities of a particular Habitat Attribute associated with each suitability class provides the user diagnostic capability. When this difference is larger, the Habitat Attribute is more discriminatory. For example, the Conditional Probabilities of 0.80 and 0.30 for Attribute A show it to be a better diagnostic criteria than a hypothetical Attribute B with Conditional Probabilities of 0.80 and 0.70. When the Conditional Probabilities are nearly equal, there is little value in including that Habitat Attribute in the model; however, if the attribute is widely abundant (like Attribute B) it still may be important for the occurrence of the species.

Once the model is constructed (table 1), potential or sample habitat is analyzed for the presence (or absence) of each Habitat Attribute. When an attribute is present, the Conditional Probability in the model is recorded; when an attribute is not present, a value of 1 minus the Conditional Probability is entered. This procedure is followed for both Habitat Suitability Classes. When the Conditional Probabilities for all Habitat Attributes within each Habitat Suitability Class are multiplied together, the product equals the probability of the sample conditions, given that Habitat Suitability Class.

Bayes' theorem (fig. 1, right insert) allows conversion of this last probability to the more meaningful probability (called the Posterior Probability) of the Habitat Suitability Class, given the sample conditions. This is the likelihood the tract of land under evaluation is occupied, or unoccupied, by breeding bald eagles. Bayes' theorem, named after a 17th century clergyman who first proposed this form of conditional probability, actually refers to the formula with which Prior Probabilities are revised by sample information. This formula, included in the example calculations of the Results, is easily computed on any pocket calculator. Results can be used to rank habitat, or to make inferences about density potential (Williams et al. 1977). "What if" games can be played by changing deficient Habitat Attributes mathematically to test the potential benefits or impacts of habitat alterations, resulting from purposeful management or perturbations caused by resource exploitation.

Data for the PATREC model on Arizona bald eagle breeding habitat were derived from an analysis of color, 1:24,000, aerial photography and U.S.G.S. topographic quadrangle maps (Grubb 1986). Information on terrain, aquatic habitat, and vegetation was recorded within a 2-mile radius of up to 18 nest sites (aerial photos were not available for all sites) and 10 randomly selected locations scattered between the known nest sites, along the Salt and Verde Rivers. Cover was typed and area determined for cottonwood-willow, mixed broadleaf, mesquite, palo verde-mixed cacti, pinyon-juniper (Brown 1982), agricultural field, and open lake. The first 3 categories were grouped into riparian. A ratio of linear mileage of riparian vegetation to river mileage within the 2-mile radius was calculated by making the implied division. The ratio of actual river mileage within the sample plot, to the straight-line distance between the points of intersection of the drainage with the plot boundary, was calculated similarly. Prior probabilities were set at 0.5. The model was tested on a known active breeding area (the Ladders site) and on an area of unoccupied, potential habitat (Canoe Mesa).

Results

Table 1 is the PATREC model derived from this analysis. To illustrate the application of this PATREC model, 2 test applications are detailed:

Example 1, Ladders –	1) Cliffs, 5.9 mi $\geq$ 100 ft	Yes
	2) River mi:straight mi, 2.5	Yes
	3) Mi permanent side drainages, 0	No
	4) % riparian cover, 1%	No
	5) Riparian mi:river mi, 0.31	Yes

Calculation 1 – Product of the probabilities under each of the habitat classes

Probability of these inventory data (ID) given area is occupied (O)

$$P(ID/O) = (0.75)(0.89)(1-0.39)(1-0.38)(0.81) = 0.20448$$

Probability of these inventory data (ID) given area is unoccupied (U)

$$P(ID/U) = (0.20)(0.30)(1-0.01)(1-0.10)(0.50) = 0.02673$$

Calculation 2 – Bayes' theorum for probability of occupied habitat (O) given these inventory data (ID)

$$P(O/ID) = \frac{P(O)\ P(ID/O)}{P(O)\ P(ID/O) + P(U)\ P(ID/U)}$$

$$= \frac{(0.5)(0.20448)}{(0.5)(0.20448) + (0.5)(0.02673)}$$

$$= 0.88$$

Conclusion – The probability of this location being occupied given the set of inventory data is 0.88; and the probability of this location being unoccupied given the set of inventory data is 0.12.

Example 2, Canoe Mesa –	1) Cliffs, 1.4 mi $\geq$ 100 ft	No
	2) River mi:straight mi, 2.1	Yes
	3) Mi permanent side drainages, 0	No
	4) % riparian cover, 4%	No
	5) Riparian mi:river mi, 0.96	Yes

Calculation 1 – Product of the probabilities under each of the habitat classes

Probability of these inventory data (ID) given area is occupied (O)

$$P(ID/O) = (1-0.75)(0.89)(1-0.39)(1-0.38)(0.81) = 0.06816$$

Probability of these inventory data (ID) given area is unoccupied (U)

$$P(ID/U) = (1-0.20)(0.30)(1-0.01)(1-0.10)(0.5) = 0.10692$$

Calculation 2 – Bayes' theorum for probability of occupied habitat (0) given these inventory data (ID)

$$P(O/ID) = \frac{P(O)\ P(ID/O)}{P(O)\ P(ID/O) + P(U)\ P(ID/U)}$$

$$= \frac{(0.5)(0.06896)}{(0.5)(0.06896) + (0.5)(0.10692)}$$

$$= 0.39$$

Conclusion – The probability of this location being occupied given the set of inventory data is 0.39; and the probability of this location being unoccupied given the set of inventory data is 0.61.

Discussion

The Ladders test site was evaluated as very likely occupied (0.88). Canoe Mesa, in contrast, appears to be unlikely breeding habitat based on these limited attributes. However, if the percent of riparian cover is increased to some undetermined level above the 5% point, the probability of Canoe Mesa being occupied becomes 0.78, indicating the possible importance and critical level of this attribute. Similarly, the model also indicates the Ladders site is deficient in the amount of riparian cover, as well as in miles of permanent side drainages. Managing habitat to create new drainages is not feasible; but, based on these results, management to improve riparian vegetation could improve the quality of eagle breeding habitat at the Ladders site (revised Posterior Probability = 0.96). However, the 0.88 value may suggest to managers that management dollars could be more effec-

tively spent improving a poorer site with greater potential for improvement (i.e., Canoe Mesa, 0.39 to 0.78, rather than Ladders, 0.88 to 0.96).

For now, this model can only be considered a first generation model. It was stated earlier that Habitat Attributes should be independent of each other. The diagnostic criteria of the eagle breeding habitat model presented here are obviously not entirely independent. However, the seriousness of not conforming to this standard is still being debated by investigators in the medical field (Lusted 1968).

At present, the model is consistent with field experience, although it is not recommended for applied habitat evaluation until actual field data can be analyzed, incorporated, and tested. As later generation PATREC models are developed through the addition of attributes such as flow rates, prey densities, human disturbance, etc., and refined to maximize sensitivity, it will be possible to rank habitat quality at all known or potential eagle breeding sites in Arizona. It also will be possible to determine the probable effects of proposed habitat improvements by adjusting levels of deficient or marginal habitat attributes that could be improved by management by recalculating a revised probability of habitat suitability. The importance of individual attributes also can be evaluated in this way. Impacts of proposed environmental development in riparian areas can be assessed, and sites that have little or no management potential can be diagnosed in order to efficiently direct limited management funds.

Unfortunately, the more sophisticated models become, the less they tend to be used by, or available to, most field biologists. PATREC has potential for bridging this gap, by being a simple, comprehensible, yet powerful modeling tool that can assist managers in recognizing and quantifying patterns of wildlife habitat use. The ultimate value of such a technique is effectively summarized by Seitz et al. (1982):

> Perhaps in the final analysis the test of a model's 'goodness' is the degree to which it improves a land-use decision (Thesen 1974). The decision made regarding the way a tract of land is used is the important thing, and habitat evaluation models are only important in how they help during the decision-making process. Whether anyone should use the model discussed depends on whether it can help in making a better decision for wildlife. If biologists consciously consider what they are doing during habitat evaluations, then approaches like PATREC will help identify important habitat attributes, increase our understanding of what constitutes good habitat, and improve our ability to communicate that understanding. This increased understanding and improved communication should improve our ability to make intelligent decisions and recommendations when managing wildlife.

Acknowledgments

Wade L. Eakle and Joseph L. Ganey analyzed maps and aerial photography for the habitat data included in the bald eagle model. Eakle also assisted in developing that model. Appreciation is extended to Larry A. Forbis, David R. Patton, and William K. Seitz for very helpful manuscript reviews. This research was jointly funded by USDA Forest Service, U.S. Fish and Wildlife Service, and U.S. Bureau of Reclamation.

Literature Cited

Brown, D. E. 1982. Biotic communities of the American Southwest—United States and Mexico. Desert Plants. 4:1–342.

Edwards, W.; Lindman, H; Savage, L. J. 1963. Bayesian statistical inference for psychological research. Psychological Review. 70:193–242.

Evans, L. C. 1983. Impact assessment and mitigation planning with habitat evaluation models. Fort Collins, CO: Colorado State University. 299 p. M.S. thesis.

Flather, C. H. 1982. Use of ecological theory to evaluate pattern recognition: implications to wildlife assessments. Fort Collins, CO: Colorado State University. 109 p. M.S. thesis.

Flather, C. H.; Hoekstra, T. W. 1985. Evaluating population-habitat models using ecological theory. Wildlife Society Bulletin. 13:121–130.

Grubb, T. G. 1986. Arizona bald eagle research, 1983–1985. Unpubl. rep. Tempe, AZ: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 172 p.

Kling, C. L. 1980. Pattern recognition for habitat evaluation. Fort Collins, CO: Colorado State University. 244 p. M.S. thesis.

Lusted, L. B. 1968. Introduction to medical decision making. Springfield, IL: Charles C. Thomas. 271 p.

Seitz, W. K.; Kling, C. L.; Farmer, A. H. 1982. Habitat evaluation: a comparison of three approaches on the northern Great Plains. Transactions of the North American Wildlife and Natural Resources Conference. 47:82–95.

Thesen, A. 1974. Some notes on systems models and modeling. International Journal of Systems Science. 5:145–152.

Williams, G. L.; Russell, K. R.; Seitz, W. K. 1977. Pattern recognition as a tool in the ecological analysis of habitat. In: Classification inventory, and analysis of fish and wildlife habitat: proceedings of a national symposium; Phoenix, AZ. Washington, DC: U.S. Fish and Wildlife Service, Office of Biological Service: 521–531.

Wilson, J. J. 1983. Pattern recognition habitat evaluation models for southeastern Idaho. Fort Collins, CO: Colorado State University. 104 p. M.S. thesis.

Rocky
Mountains

Southwest

Great
Plains

U.S. Department of Agriculture
Forest Service

Rocky Mountain Forest and Range Experiment Station

The Rocky Mountain Station is one of eight regional experiment stations, plus the Forest Products Laboratory and the Washington Office Staff, that make up the Forest Service research organization.

RESEARCH FOCUS

Research programs at the Rocky Mountain Station are coordinated with area universities and with other institutions. Many studies are conducted on a cooperative basis to accelerate solutions to problems involving range, water, wildlife and fish habitat, human and community development, timber, recreation, protection, and multiresource evaluation.

RESEARCH LOCATIONS

Research Work Units of the Rocky Mountain Station are operated in cooperation with universities in the following cities:

Albuquerque, New Mexico
Flagstaff, Arizona
Fort Collins, Colorado*
Laramie, Wyoming
Lincoln, Nebraska
Rapid City, South Dakota
Tempe, Arizona

*Station Headquarters: 240 W. Prospect St., Fort Collins, CO 80526