Re: Contract Number #02-04-UA
Work Product #8
Judy E. Muir
April 29, 2003

Predicted Avian Habitat Quality in Y2Y

Abstract:
I created multiple regression models that used spatial location, habitat type, elevation, slope and aspect to predict species richness and numbers of species at risk for each square kilometre throughout the Yellowstone to Yukon (Y2Y) region. I found the highest predicted species richness occurred along a north-south gradient on the western portion of Y2Y, while high concentrations of species at risk were predicted in the southern part of Y2Y. I created an index of avian habitat quality for each square kilometre in Y2Y by equally weighting, and then combining, the values for species richness and numbers of rare species at that location. I ranked this index separately within each broad-scale habitat type, and then combined the rankings for all habitat types to determine relative habitat quality for birds within Y2Y. Predicted avian habitat quality is highest in the northwest of the Y2Y region. This pattern may be due to the higher moisture gradient on the west side of the Rocky Mountains and lower human population density in the north. My findings need to be treated with caution as the models I used to predict species richness and number of rare species are preliminary versions that contain a subset of the broad-scale predictors that may be important to avian habitat quality. In addition, the index of avian habitat quality created from values of species richness and number of species at risk is sensitive to the relative weightings given to these variables. This weighting is subjective, and will need to be adjusted to reflect the conservation priority placed on these factors by Y2Y.

Introduction:

Y2Y's avian conservation area design aims to identify high quality avian habitat throughout the Y2Y region. High quality habitats maintain source populations with positive population growth rates[1], have higher amounts and more stable resources, and lower levels of predation, competition, parasitism and anthropogenic disturbance[2,3,4]. Due to the large size of the Y2Y region (1.93 million square km), direct determination of avian habitat quality by measurement of bird habitat use and breeding success is logistically impossible. Hence, a predictive model for habitat quality is the only feasible option.

My study hypothesizes that avian habitat quality within Y2Y can be predicted at the regional scale using broad-scale biophysical factors (climatic and topographic conditions) and anthropogenic stressors (level and type of human activity). Biophysical factors influence habitat quality by affecting the levels and stability of resources important to survival and reproductive success. For example, higher primary productivity levels supply more food, supporting more individuals and reducing the risk of population extinction[5]. Human activities stress and degrade avian habitat quality in many ways. For example, suburban development and industrial activities contribute to higher road densities that lower wetland and stream quality, increase predation and brood parasitism rates, and increase mortality due to roadkill[6]. Human activities also result in habitat loss and fragmentation that reduce bird survival and reproductive success[7,8].

This report summarizes an analysis I performed to predict avian habitat quality throughout the Y2Y region. Since information on reproductive success is not available at this broad scale, my study uses response variables of species richness and number of species at risk as possible surrogates of habitat quality. I used Breeding Bird Survey (BBS) data to calculate these response variables throughout Y2Y. I then created a statistical model to predict each response variable for each square kilometre of the Y2Y based on the values of habitat type, UTM easting, UTM northing, elevation, slope and aspect found in that location. I imported these models into a Geographic Information System (GIS) and used them to predict species richness and number of species at risk in the Y2Y region. I combined these predictions to produce an index of avian habitat quality throughout Y2Y and then ranked this index separately for each Y2Y broad-scale avian habitat type. I combined the rankings for all habitat types to produce the final map of predicted avian habitat quality throughout Y2Y. This document references the following maps contained in the accompanying PowerPoint presentation "Predicted Y2Y Avian Habitat Quality Maps":

1. Predicted species richness throughout Y2Y
2. Predicted number of species at risk throughout Y2Y
3. Predicted avian habitat quality throughout Y2Y using a 50:50 weighting of species richness to species at risk.
4. Predicted avian habitat quality throughout Y2Y using a 75:25 weighting of species richness to species at risk.
5. Predicted relative avian habitat quality ranked within each broad-scale habitat type in Y2Y.
6. Predicted relative avian habitat quality ranked within each broad-scale habitat type in northern Y2Y
7. Predicted relative avian habitat quality ranked within each broad-scale habitat type in southern Y2Y

Methods and Results:

a. Predictor Variables:

I used six variables as potential predictors of species richness and number of species at risk in this preliminary version of avian habitat quality throughout the Y2Y region. These were habitat type, UTM easting, UTM northing, elevation, slope and aspect. I plan to include additional predictors including measures of primary productivity, growing degree days, road density and locations of mines in future models for avian habitat quality. The predictor variables used in this version were represented as Geographic Information System (GIS) grid layers at 1 square km resolution. Y2Y provided a GIS layer for habitat type that contained most of Y2Y's broad-scale avian habitat types (Table 1). This classification did not include riparian or wetland habitat, and only a small amount of montane habitat (0.6% of the total area). Y2Y also provided a Digital Elevation Model (DEM). I used ArcMap v8.1 software to derive GIS layers for aspect and slope from the DEM layer, and to create GIS layers representing UTM Easting and UTM Northing locations in Y2Y.

b. Response Variables:

The North American Breeding Bird Survey (BBS) was established in 1966 to monitor North American bird populations. The survey takes place along permanent BBS routes that are 39.4 km long and randomly located along secondary roads throughout North America. Each route has 50 stops located at 800 m intervals along the route. Routes are surveyed once each year during the breeding season with a 3-minute point count performed at each stop on a route. All bird species seen and heard within a 400m distance from the stop are recorded[9].

I constructed a GIS layer of all BBS route stops contained within Y2Y using ArcView 3.2 and ArcMap 8.1 software. I then used raw data from the Patuxent Wildlife Centre[10] to calculate two response variables, species richness and number of species at risk, at each BBS stop within Y2Y having at least 1 year of data from 1997 to 2002. I averaged data at a stop if more than 1 year of data were present. A detailed description of these methods is documented in my previous report submitted to Y2Y (February 2003). I calculated the response variables as follows:

i) Species richness: This is the total number of conservation priority species (Table 2) at a stop.
ii) Species at risk: These are conservation priority species that are Canadian species at risk, i.e. endangered (burrowing owl, white-headed woodpecker, sage grouse), threatened (loggerhead

shrike) and species of special concern (short-eared owl, ferruginous hawk, flammulated owl, Lewis' woodpecker, long billed curlew, tundra peregrine falcon), and American endangered/threatened species (bald eagle).

I imported the calculated response variables into the GIS and merged them with their associated BBS stops. I then extracted values of the predictor variables at each BBS stop location and combined these with the response variables at the stop to produce the data set I used in the statistical analysis. Each record in the data set contained values for species richness, number of species at risk, habitat type, UTM easting, UTM northing, elevation, slope and aspect for a unique BBS stop within Y2Y.

c. Statistical Analysis

I constructed a separate multiple regression model for each response variable using SPLUS v 6.1[11]. I randomly selected 75% of the data set to develop the models and then tested the models on the remaining data. I first examined the relationship between each predictor variable and response variable using scatterplots and spline fitting methods. I found several predictors had non-linear relationships with the response variables (e.g. Figure 1). Standard multiple regression models make some assumptions about the mathematical characteristics of the response variable (constant variance and normal distribution of errors) that are often violated with count data such as the species richness and number of species at risk response variables I used in my analysis. Generalized linear models (GLMs) are an extension of multiple regression analysis that allow these assumptions to be violated[12]. In addition, since my initial examination of the data indicated some non-linear relationships between the predictor variables and response variables I used generalized additive models (GAMs) to determine the shapes of these non-linear relationships and to test for significant non-linearities in these relationships[13]. GAMs fit cubic splines to determine the shape of the relationship between a predictor variable and response variable. Thus, the shape of the relationship is not specified beforehand, but is determined directly by the GAM modelling function. I developed a model for each response variable using the following steps[14]:

1. I first created a GAM model for the response variable using a spline fit. This produced the “best” fitting model which provided a baseline for comparison with potential GLM models. The GAM also indicated which predictor variables had a significant non-linear relationships with the response variable.
2. I examined plots of the GAM fit for each non-linear continuous predictor variable (UTM Easting, UTM Northing, DEM and slope) to determine which mathematical transformation (e.g. polynomial, logarithmic, exponential etc.) best fit the relationship suggested by the GAM of predictor to response variable.
3. I modelled the GLM with the transformations suggested by the GAM for all continuous variables that had significant non-linear effects, and used simple linear terms for the other predictor variables. I then compared the explanatory power of the GLM against the GAM. No significant difference between the models indicated the mathematical transformations used in the GLM were successfully approximating the relationships determined by the GAM. I repeated this step as needed, testing different transformations, until there was no significant difference between the GLM and GAM models.
4. I then reduced the GLM model’s terms by using the Cp statistic to determine which factor had the least explanatory power. I eliminated this factor from the model, and then compared the new model with the previous model. If there was no significant difference in explanatory power, I retained the simpler model. I repeated this step until removal of a term resulted in a significant difference to the previous model. I then designated the previous model as the final model for the response variable.

The final species richness model incorporated all predictor values (Table 3), while the rare species model used all factors but UTM Northing locations and aspect (Table 4). I used the final models to predict the values for each response variable in the test data set. I then determined the correlation of predicted values to the actual values calculated from the BBS data. Although the rare species model explained more deviance than the species richness model (31.1% versus 21.4%), both models had similar correlation between predicted and actual response variable values (approximately 0.4).

d. Predicted Avian Habitat Quality Throughout Y2Y

I incorporated the statistical models for the two response variables of species richness and number of species at risk into the GIS, and used values of the predictor variables in each square kilometre to calculate predicted species richness (map 1), and predicted number of species at risk (map 2) for each square kilometre within Y2Y. I normalized the species richness and number of species at risk response variables to have values from 0 to 1, and then combined them using equal weighting to produce an index of avian habitat quality at a 1 square km resolution throughout Y2Y (map 3, Table 5). The highest quality habitat was found in grassland, followed by sagebrush steppe, probably due to the high numbers of species at risk in these habitat types. Moist habitat types (cedar/hemlock and Douglas fir mix) and small lakes also showed high levels of predicted habitat quality. I tested the sensitivity of the avian habitat quality index to different weightings of species richness and number of species at risk by producing a second map using a weighting of 75% species richness and 25% number of species at risk (map 4). This map shows lower habitat quality in the southern parts of Y2Y, but is otherwise similar to the first map.

I separately ranked the habitat quality index produced by equally weighting species richness and number of species at risk (i.e. map 3) within each Y2Y broad scale habitat type into one of the following 5 classes:

- 0 – 25% (poorest habitat quality)
- 25 – 50%
- 50 – 75%
- 75 - 90%
- 90 – 100% (best habitat quality)

I then merged the rankings for all habitat types to produce the final map of predicted relative avian habitat quality (map 5). This approach ensured that each broad scale habitat type was represented in the top ranked habitat. The best avian habitat is located in the northwest part of Y2Y. I separated this map into northern and southern parts of Y2Y to display more detail (maps 6 and 7).

Conclusions and recommendations

Predicted avian habitat quality is highest in the northwest of the Y2Y region. This pattern reflects the higher moisture gradient on the west side of the Rocky Mountains and lower human population density in the north. The approach I used to rank predicted avian habitat quality within Y2Y ensured that all habitat types were represented in the highest habitat quality ranking. Simply taking the top values for habitat quality throughout Y2Y would omit less productive habitat types (e.g. northern, alpine, subalpine) that have low predicted species richness values. The broad-scale habitat types for Y2Y were selected specifically because they were believed to represent all bird communities within Y2Y. Not including all habitat types in the top ranked habitat will fail to identify important habitat for all conservation priority bird species identified in the Y2Y region.

The maps of predicted avian habitat quality within Y2Y are preliminary and the results need to be treated with caution. While there are many Breeding Bird Survey routes within the Y2Y region, they do not provide a representative sample of habitat types or geographic region within Y2Y. In particular, montane habitat was not sampled, while higher elevation (alpine/tundra and subalpine spruce/fir) and northern habitats (boreal spruce and northern shrubfields) are poorly sampled. The roadside survey nature of the BBS contributes to this negative sampling bias as sparsely populated regions in the north and in higher elevations have few roads. In addition, 36 BBS routes in the northern half of Y2Y have not been surveyed in the last 6 years. This lack of data means there is less certainty for model predictions in alpine tundra, subalpine, northern shrubfields and boreal spruce habitats, and in northern regions of Y2Y.

It is important to realize that the ranking of avian habitat quality throughout Y2Y is sensitive to the relative weightings given to the response variables of species richness and number of species at risk. This weighting is subjective, and will need to be adjusted to reflect the conservation priority placed on these factors by Y2Y.

I plan to enhance these models over the remainder of 2003 to include additional predictor variables (measures of primary productivity, growing degree days, road density and locations of mines), and use additional data sets to better represent alpine, subalpine and northern habitats. I also plan to perform a spatial autocorrelation analysis of the data as this affects regression results. Spatial autocorrelation occurs when samples are not independent, but are influenced by samples at other locations in the analysis[15]. Independence is a fundamental assumption of statistical tests whose violation leads to increased type I errors (finding significant results when they don't exist). In regression analyses, this may result in predictors being kept mistakenly in the regression model[16] (Legendre and Fortin 1989).

Literature Cited

[1] Pulliam, H.R. 1988. Sources, sinks, and population regulation. American Naturalist 132:652-661.
[2] Cody, M. 1985. editor. Habitat Selection in Birds. Academic Press, Inc. London.
[3] McLoughlin, P.D., Ferguson, S.H. and Messier, F. 2000. Intraspecific variation in home range overlap with habitat quality: a comparison among brown bear populations. Evolutionary Ecology 14:39-60.
[4] Suryan, R.M. and Irons, D.B. 2001. Colony and population dynamics of black-legged kittiwakes in a heterogeneous environment. The Auk 118(3):636-649.
[5] Gaston, K.J. 2000. Global patterns in biodiversity. Nature 405:220-227.
[6] Trombulak, S.C. and Frissell, C.A. 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conservation Biology 14(1):18-30.
[7] Robinson, S.K., Thompson III, F.R., Donovan, T.M. Whitehead, D.R. and Faabord, J. 1995. Regional forest fragmentation and the nesting success of migratory birds. Science 267:1987-1989.
[8] Weinberg, H.J. and Roth, R.R. 1998. Forest area and habitat quality for nesting wood thrushes. The Auk 115(4):879-889.
[9] North American Breeding Bird Survey http://www.mp2-pwrc.usgs.gov/bbs/about
[10] Patuxent Wildlife Research Center ftp://pwrcftpr.er.usgs.gov/mp/bbs/datafiles
[11] S-Plus 6.1 for Windows. Insightful Corporation, Seattle, USA.
[12] Crawley, M.J. 2002. Statistical Computing An Introduction to Data Analysis using S-PLUS. John Wiley & Sons Ltd. Chicester, U.K.
[13] Guisan, A., Edwards Jr., T.C. and Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157(2-3): 89-100.
[14] Roland, J., Keyghobadi, N. and Fownes, S. 2000. Alpine *Parnassius* butterfly dispersal: effects of landscape and population size. Ecology 81(6): 1642-1653.
[15] Legendre, P. and Legendre, L. 1998. Numerical Ecology 2nd English Edition. Elsevier Scientific B.V. Amsterdam.
[16] Legendre, P. and Fortin, M. J. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107-138.