Identification of Prime Avian Habitat in the Y2Y Ecoregion

Judy E. Muir

April 14, 2004

Introduction

The Yellowstone to Yukon Conservation Initiative (Y2Y) is developing a reserve design to restore and maintain avian biodiversity throughout the Rocky and Mackenzie Mountains. The project uses a landscape level approach to reserve core areas of high conservation value, linked by corridors and surrounded by buffer zones that allow human use to increase with distance from core areas. My study identified areas of high quality bird breeding habitat for the reserve design; defined as being able to maintain source populations with positive population growth rates (Pulliam 1988) and having higher amounts, more stable resources and lower levels of predation, competition, parasitism and anthropogenic disturbance (Cody 1985, McLoughlin, et al. 2000, Suryan and Irons 2001). The large size and remoteness of the Y2Y ecoregion makes it logistically impossible to directly determine avian habitat quality by measuring bird habitat use and breeding success. Instead, I adopted a broad-scale modelling approach to prioritize habitat based on home range habitat selection for a group of 'umbrella' bird species, defined as those whose requirements include and protect those of sympatric species (Landres et al. 1988, Caro and O'Doherty 1999, Zacharias and Roff 2001). I also used a coarse-filter approach by identifying high quality habitat in several broad habitat cover types (Table 1) important to mountain bird communities. A previous report (November 28, 2003) described the methods and results for the process used to identify a group of umbrella bird species that represented the broad habitat cover types in Y2Y. This report describes the remaining steps to identify high quality avian habitat in Y2Y. These include development of models to predict probability of occurrence for 11 of the umbrella birds (Table 2), extrapolation of these models to predict habitat suitability for each species throughout its range in Y2Y, and the integration of the habitat suitability rankings for each species to identify areas of conservation priority for birds within Y2Y.

Methods

Bird data

A common method to census breeding birds involves point count surveys that use a standardized methodology to record all birds seen or heard within a fixed distance from widely separated locations during a specific time interval (Farnsworth et al. 2002). I collated approximately 26000 point count locations of bird survey data for the breeding season throughout the Y2Y region (see Acknowledgments). Data were predominantly from 1990 to 2000, and supplemented with data sets from 1970 to 1990 representing missing northern and high elevation habitats. I constructed a Geographic Information System (GIS) layer of the point count locations and their associated average yearly counts for the Y2Y conservation priority species (Figure 1).

Detectability

Bird survey data were collected using a variety of protocols that may have different detectability, or probability of observing a bird present, among the survey points (Hutto et al. 1995, Farnsworth et al. 2002). Species detectability differences can confound the analysis of its habitat relationships (Detttmers et al. 1999, Drapeau et al. 1999). The survey data did not provide adequate information to allow construction of distance detectability functions or comparison of detectability among different protocols (Farnsworth et al. 2002). I attempted to mitigate detectability differences by using presence/absence rather than abundance as a response variable in the models (Young and Hutto 2002), and by favouring the most detectable bird species when selecting the umbrella species (refer to report "Focal Bird Species Selection" Nov 28, 2003).

Habitat cover type data

I created a Geographic Information System (GIS) layer at 1 sq. km. resolution for each Y2Y habitat cover type using ARC/INFO software (version 8.3; ESRI 2002). Marsh and lake cover types were represented as the percent area per square kilometre calculated from digital maps of wetlands and lakes at 1:250000 scale (NTDB Canada, USGS 1994) and digital lake data at 1:100000 scale (state GIS clearinghouses). I extracted the Whitebark pine GIS layer from a habitat classification provided by Y2Y.

I derived the remaining Y2Y habitat cover types from the North America seasonal land cover region (SLCR) data (version 2.0, LP DAAC). I cross-walked the SLCR land cover classes to the Y2Y habitat cover types, allowing a single SLCR land cover class to map to multiple Y2Y habitat cover types when reasonable (Table 3). I also cross-walked to a willow habitat type for later construction of the willow riparian habitat type. I created the riparian habitat types by first merging the Y2Y cover type GIS layers into a coniferous layer (subboreal spruce, cedar hemlock, Douglas fir mix, lodgepole pine, ponderosa pine mix, mixed wood, subalpine spruce/fir) and a deciduous layer (aspen, mixed wood, northern shrubfields). I then calculated the total length of streams per square kilometre from digital stream data at 1:250000 scale (NTDB Canada), and 1:100000 scale (state GIS clearinghouses) and used this grid to identify the square kilometre cells in the coniferous, deciduous and willow habitat cover type layers that also contained stream. I extracted these cells to produce the coniferous, deciduous and willow riparian habitat cover type layers.

Landscape data

I created GIS layers for several biophysical and anthropogenic variables to use as predictors in model development and subsequent extrapolation (Table 4). Biophysical factors influence avian habitat quality at a macrohabitat scale by affecting the levels and stability of resources important to survival and reproductive success. I used two measures of net primary productivity as predictors: actual evapotranspiration (AET) and normalized difference vegetation index (NDVI). I obtained a GIS layer of total AET in 1996 for Canada (Liu et al. 2003) and calculated monthly and breeding season NDVI averages and standard deviation from MODIS data (GLFC 2001-2003). Climate variables for measures of the growing season length (DAY1GROW, ENDGROW, TOTDAYSGRO, GDD), total precipitation during the growing season (PRECIP), and variation in temperature and precipitation (TEMPSEAS, PRECIPSEAS) from climate surface averages for 1971 to 2000, and elevation data from digital elevation models were provided by the Great Lakes Forestry Centre (Environment Canada 2003). I derived geographic locations (EASTING, NORTHING) for the point count locations, and the center points of each sq km. grid cell directly from the GIS layers to model potential trends in species occurrence with geographic location (Franklin 1998).

I used total length of railroad track in a square kilometre (TOTRRLEN) and distance to the nearest oil/gas well or mine (INDUSTDIS) as indicators of anthropogenic disturbance. I collected GIS layers for railroads, mine, oil and gas wells from federal, provincial, territorial and state agencies. I planned to use road density as an indicator of population density and recreational disturbance, but the survey points were positively associated with roads due to accessibility issues. Since this factor was subject to sampling bias, I removed it from the model.

I also included the percent wetland (PERWET) and total streamlength (STREAMLEN) per square kilometre GIS layers as predictors (Refer to section "*Habitat cover type data*") The streams layers were at different scales for Canada and the U.S, and this was reflected in a lower density of streams in Oregon and Montana. I only used STREAMLEN as a predictor variable for an umbrella bird when the bird's range did not include these states.

Model Development

Predictive models for species occurrence are commonly based on presence/absence data (e.g. Franklin 1998, Beard et al. 1999, Venier et al. 1999), but models may be confounded by "false negatives" due to detectability issues or unsaturated habitat, or by absences due to factors not included in the model such as inter/intra species competition (Fielding and Bell 1997, Zaniewski et al. 2002). Presence/absence models also can't make use of data sets collected on an ad-hoc or non-stratified basis with unreliable information regarding species absence. Zaniewski et al. (2002) suggested creating "pseudo" absences to allow statistical models to be used with presence only data, and this approach has been well developed with resource selection functions (RSF) by way of a "used/available" sampling protocol (Manly et al. 2002). RSF models provide a relative estimate of the probability an organism will use a site based on a statistical analysis of the association between its presence in a landscape and selected habitat attributes (Boyce and McDonald 1999, Manly et al. 2002). For example, RSF models have been used to determine habitat characteristics of marten den sites (Ruggiero et al. 1998), Eurasian lynx habitat suitability (Schadt et al. 2002) and grizzly bear distribution (Apps et al. 2004).

I used an RSF approach with a used/available sampling protocol to develop a unique model for each umbrella bird to predict the probability of its occurrence throughout the Y2Y ecoregion. Detectability issues with the bird survey data did not allow me to be confident that point count locations at which a bird was not recorded were truly unused. In addition, the survey points were biased to southern and low elevation habitats and did not sample the full gradient of predictor variables in Y2Y. Using "available" points that spanned the Y2Y region provided more confidence that the predictors were sampled over their full gradient.

I considered a point count location "used" by a species if it was detected at least once. I only allowed one point count in a square kilometre sample unit. I removed extraneous points by selecting the point count with the highest average count for a species, or in the case of a tie, by randomly selecting one of tied count locations. I allocated double the number of used points for a species as "available" points, and randomly distributed these through the bird's habitats using the same weighting system as that used to determine the umbrella species (refer to report "Focal Bird Species Selection" Nov 28, 2003). For example, American wigeon has 1 primary habitat type (marsh) and one secondary habitat (lake), so I randomly selected 67% of its available points from marsh and the remaining 33% from lake habitat within its range.

I developed a logistic regression RSF model for each umbrella bird with available and used samples assigned values of 0 and 1 respectively. All statistical analysis was performed using S-PLUS v6.1 (Insightful Corporation 2002). I first tested for correlation among the predictor variables for each umbrella bird as two highly correlated significant predictors can both appear non-significant if modelled together (Guisan et al. 2002). I then defined a set of models *a priori* for the umbrella that allowed only one of a group of highly correlated predictor variables (Pearson's $r >= 0.7$) to be used in a model, but included all predictors in the full set of *a priori* models. Generalized linear models (GLMs) are extensions of standard multiple regression analysis that address violations in assumptions when the response is proportion data representing relative numbers of available and used sites (Guisan and Zimmerman 2000, Crawley 2002). GLMs also keep the predictions within the range of feasible values for the response variable (e.g. probability values from 0 to 1 for a 1/0 presence/available response) (Guisan and Zimmerman 2000) and easily model non-linear species responses to environmental gradients (Austin 2002, Oksanen and Minchin 2002). I developed each *a priori* model for an umbrella species using the following steps:

1) I created a generalized additive model (GAM) with spline smoothers for all continuous variables to test for significant non-linear responses in these variables. I also examined plots of the GAM fit for each significant non-linear predictor to determine which mathematical transformation (i.e. second or third order polynomial, logarithmic, exponential or piecewise linear) best fit the response curve (Franklin 1998, Guisan et al. 2002, Miller and Franklin 2002). GAMs are semi-parametric extensions of GLMs that use the data to determine the shape of the relationship between the response and the set of explanatory variables rather than assuming some form of parametric relationship (Guisan et al. 2002).
2) I modelled a GLM with the transformations suggested by the GAM for all variables with significant non-linear effects, and used simple linear terms for the other predictor variables. If the plot for a significant non-linear predictor suggested more than one possible transformation, I modelled each transformation and selected the one that minimized the residual deviance of the model.
3) I built the final, reduced model with a backwards stepwise technique using Mallow's 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 a significant difference in explanatory power, I retained the previous model, otherwise I repeated this step with the simpler model (Roland et al. 2000).

Finally, I compared the residual deviance and number of parameters used in the final models retained for the set of *a priori* models, and selected the model that had the best trade-off between minimizing the residual deviance and number of parameters.

Spatial Autocorrelation

Species occurrence or abundance is often positively autocorrelated, with samples close to each other being more similar than expected due to chance (Lichstein et al. 2002). This may result from spatial structure in environmental variables, or biotic factors such as conspecific attraction and dispersal (Lichstein et al. 2002). Spatial autocorrelation violates statistical assumptions of sample point independence and independently distributed errors, and in regression analyses may result in predictors being kept mistakenly in the model (Legendre and Fortin 1989). I attempted to reduce potential spatial autocorrelation by only allowing one sample per square kilometre. I also examined model residuals in semivariograms to test for spatial correlation (Legendre and Fortin 1989).

Model Validation

I validated the models by evaluating model fit i.e. how well the predictors explained the response variable (Guisan et al. 2002), and by testing how well model predictions discriminated, or classified, used and available points (Pearce and Ferrier 2000). GLMs are fit by maximum likelihood rather than the least squares method of standard regression models, with "deviance" rather than R^2 being the measure of the goodness of model fit (Crawley 2002). I used the percent deviance explained by a model as a measure of its fit.

I tested the discrimination ability of the final model for each umbrella species using a K-fold partitioning technique that allows a model to be tested without independent data. The data for a model is divided into K partitions, then a training model is developed on all combinations of K-1 partitions of the data and used to predict the probability of occurrence for test data consisting of the unused partition (Fielding and Bell 1997, Pearce and Ferrier 2000). I used Huberty's rule of thumb to determine the number of partitions (Fielding and Bell 1997). I developed each training model for an umbrella species by fitting the umbrella's final model to the training data.

I visually compared frequency distribution density plots and box plots of predicted values for used versus available samples. Models with good discrimination ability show little overlap between the predicted probability distributions, and the mean predicted probability will be higher for the used sites (Pearce and Ferrier 2000). Discrimination can also be evaluated by determining how well a model's predictions correctly classify the site as used or unused/available (Guisan and Zimmerman 2000). Receiver operating characteristic (ROC) curves provide a measure of classification accuracy by testing a range of probability thresholds to determine the classification (Miller and Franklin 2002). The measure of classification accuracy is the area under the ROC curve (AUC) with AUC values over 0.9 indicating very good discrimination, 0.7 to 0.9 reasonable and values less than 0.7 indicating poor discrimination (Pearce and Ferrier 2000). I used SPSS v11.5.0 (2002) to produce the ROC curves for each umbrella's predicted probabilities, and to calculate the associated AUC and its standard error.

Model Extrapolation

I used the final model for each umbrella bird to predict the probability of its occurrence for each square kilometre within its breeding range in Y2Y that was also of a habitat type used by the bird. I then created GIS maps showing relative habitat quality for each umbrella bird by ranking its predicted probabilities of occurrence into 5 categories. I also created maps showing the most suitable habitat within each Y2Y habitat cover type by allowing a species to contribute 10% of its habitat area that was allocated among its habitat types using the same weighting system used to determine the umbrella species (refer to report "Focal Bird Species Selection" Nov 28, 2003). For example, White-crowned Sparrow has 1 primary habitat type (subalpine) and two secondary habitats (alpine/tundra and northern shrubfields), so I weighted subalpine by 0.5 and the secondary habitats by 0.25 each. I then took the top 5% of White-crowned Sparrow's predicted probabilities for subalpine, and the top 2.5% probabilities for each of its secondary habitat types.

I developed maps separately for the north and south parts of Y2Y as probabilities of species' occurrence tended to increase from north to south due to warmer temperatures, longer growing seasons and higher primary productivity in southern regions that bias high quality habitat to the southern parts of Y2Y. As well, northern planning is of interest to Y2Y so it was important to identify the highest quality habitat within this region. I used the southern boundary of a Y2Y ecological priority area "Peace River Break" to demark the boundary separating the two halves.

Validation of Most Suitable Habitat

I calculated the total area of the "best" quality habitat that fell within a map of existing protected areas to assess how much of the "best" quality habitat within each cover type was protected. I also tested if locations of RAMSAR sites and important bird areas corresponded to the "best" quality habitat within each cover type. The methods used to construct the maps of protected areas, RAMSAR sites and IBAs were described in a previous report "Map of Existing Protected Areas Important to Birds in Y2Y", December 19, 2002. I updated the protected areas map with a more recent version from Miistakis Institute.

Results

Probability of Occurrence Models

The models explained from 19.6% (CLNU) to 78.5% (GCRF) of the deviance, and included from 3 to 8 significant predictors ($p < 0.05$) (Table 5). Most predictor relationships were non-linear, with second order polynomials predominating. Piecewise linear relationships (thresholds) were also common, particularly for PRECIP and elevation. PRECIP, Easting, variation in primary productivity, a measure of growing season length (DAY1GROW, ENDDAYGROW or PER3GDD) and elevation were important predictors in several models. No anthropogenic factor was found to be a significant predictor in any model. Several variables were highly correlated and placed in separate *a priori* models for an umbrella bird. Semivariograms constructed with model residuals showed no spatial autocorrelation.

Probability density and box plots for predictions produced by K-fold validation showed that most models discriminated well (CLNU - Figure 2, YEWA – Figure 3). The area under the ROC curve ranged from 0.720 (BLPW) to 0.949 (YEWA) (Table 5), and was highly correlated with the model's percent deviance explained (Pearson's correlation coefficient, $r = 0.83$).

.

Species and Habitat Suitability Maps

The species habitat ranking maps (PowerPoint presentation "Species Habitat Suitability Maps") showed the most suitable habitat tended to be patchy and spread throughout a species' range, although concentrations of high quality habitat could be seen for several species (AMDI, ATSP, BLPW, RUGR, WCSP, WIWA and YEWA). Species generally showed little overlap among the most suitable habitat, with the exception of 3 areas in the northern part of the Y2Y region. These were the extreme northwest corner of Y2Y in the Yukon Territory (AMDI, ATSP, GCRF and WCSP), the eastern edge of Y2Y in the Northwest Territories (AMDI, RUGR, WCSP and YEWA) and north-central British Columbia (BLPW, COLO, RUGR, WIWA, YEWA).

The maps for the best quality habitat in each habitat cover type indicated similar patterns to the species habitat maps (PowerPoint presentation "Habitat Quality Maps", legend field "value" indicates the number of species using a square kilometre grid cell). The high quality habitat in the northwest part of Y2Y (Yukon Territory) was due to high quality alpine and northern shrubfields habitat. Aspen, all riparian habitats, mixed wood and northern shrubfields all had good quality habitat along the eastern edge of Y2Y in the Northwest Territories. Alpine/tundra, bog, deciduous and willow riparian, boreal, and northern shrubfields also showed a concentration of high quality habitat in north-central British Columbia. In addition, Alpine/tundra had a patch of high quality habitat in northern B.C. and there was a concentration of high quality habitat for mixed wood, deciduous riparian and marsh on the eastern border of Y2Y between Fort Nelson and Fort St. John, B.C.

The regions of high quality habitat in the Yukon and NWT were characterized by high seasonality in precipitation and temperature, moderate precipitation, low elevation and high average primary productivity (NDVI). The Yukon region had a moderate number of growing degree days, while the NWT region had a high number of growing degree days and a longer growing season. The area in British Columbia had low to moderate elevation, high precipitation, a high number of growing degree days, a longer growing season and moderate to high levels of primary productivity (NDVI).

The amount of highest quality habitat that was covered by protected areas ranged from 1.2% for Mixed Wood to 35.6% for Bog (Table 6). Alpine/Tundra, Boreal, Ponderosa Pine and subalpine were represented quite well by protected areas having > 25% of their prime habitat protected. Aspen, Coniferous and Deciduous Riparian, habitats were poorly protected with less than 10% of their prime habitat overlapped by protected areas.

Creston Valley RAMSAR site in southern BC was located within a couple of km of several patches of high quality lake, marsh, deciduous and coniferous riparian habitat. Only two IBAs coincided with prime habitat. Lock Katrine Wetland in Wyoming overlapped with high quality marsh habitat, and Skookumchuk Prairie in southern BC was associated with high quality mixed wood and aspen habitat.

Recommendations

High Quality Avian Habitat

The models for 11 umbrella bird species identified three large regions of high quality avian habitat in the northern part of Y2Y: the extreme northwest corner, along the eastern edge in the NWT and in north-central British Columbia. High quality habitat was patchily distributed for many birds, possibly reflecting the patchy nature of many habitat types that were modelled (e.g. marsh, bog, riparian, lakes), and the variation in topography and associated climatic conditions that occur over short distances in mountainous terrain. It is important to realize that this habitat has been identified by predictive models based on several assumptions and subject to sources of error. Furthermore, models predictions have not been tested on independent data. While all models showed good discriminatory power, and hence predictive ability, it is imperative to do some level of ground-truthing to verify that high quality habitat identified by the models does indeed correspond to good bird breeding season habitat.

Several habitat types (alpine, bog, boreal forest, ponderosa pine and subalpine) were well covered by existing protected areas in Y2Y. However, my analysis did not assess how this protection was distributed across Y2Y or partitioned among Y2Y ecological priority areas. This would be worth analyzing to verify that habitat types are protected throughout their range in Y2Y. High quality avian habitat was not identified for cedar/hemlock, grassland, lodgepole pine, moist Douglas fir or sagebrush steppe as birds representing these habitats were not modelled. Some of these umbrella birds also represent subalpine, aspen, alpine and riparian habitat and their models may identify additional high quality habitat within these habitat types. I recommend that models be developed for the remaining umbrella species to complete the identification of high quality avian habitat.

Coniferous, deciduous and willow riparian, aspen, wetlands and mixed wood habitat types were poorly covered by existing protected areas in Y2Y and some planning effort should be directed at conserving these areas.

Model performance

Predictive models aim to estimate the probability of species occurrence at unsampled locations by using the statistical relationship between the actual species presence and a limited number of predictor variables (Guisan et al. 2002). This type of model is particularly applicable to continental scale planning in vast and poorly sampled regions such as Y2Y. A broad scale modelling approach does have limitations though; particularly in predicting distributions for microhabitat specialists whose habitat needs cannot be modelled at a broad scale.

The "used/available" RSF sampling protocol handled the data limitations well as despite limited surveys in many parts of Y2Y, the use of "available" points meant that models could be developed using samples over the entire study area. GAMs and GLMs worked well to model species occurrence as many predictors had non-linear relationships with species occurrence.

The species occurrence models varied considerably in explanatory power, ranging from 19.6% (CLNU) to 78.5% (GCRF) of deviance explained. This is a reasonable range as habitat selection occurs over multiple scales and most bird-habitat models explain only a portion of the variance (Wiens et al. 1987, Orians and Wittenberger 1991, Young and Hutto 2002). Important predictors in the models were amount of precipitation and length of the growing season, elevation, geographic location and variation in primary productivity. It is important to recognize that most of these factors had non-linear relationships with bird occurrence, and high levels of these factors do not necessarily correspond to a high probability of species occurrence. Simply looking for high levels of these factors within Y2Y will not identify high quality bird habitat.

Anthropogenic factors were not significant predictors in the models. This is partly due to sampling bias, as bird surveys are not likely to be conducted close to mines, oil/gas wells or railroads unless these effects are of particular interest in the study. The distance to industrial sites tended to be very large, and few point count locations were associated with railroads. It is also possible that the effects of these factors occur at a local scale and were not detectable in the models. Studies could be designed to test these anthropogenic factors explicitly.

Model assumptions

Multiple regression models provide quantitative predictions for given values of environmental measures and a model usually can't be generalized to other species (Wiens and Rotenberry 1981). Using regression models of umbrella species occurrence to predict high quality avian habitat assumes that the response curves for an umbrella species are the same as those for the species it represents. It is imperative to test this assumption by modelling some species assumed to be protected by an umbrella and then comparing the response curves. Predictions of high quality habitat for a "protected" species can also be used to test how effectively the high quality habitat for an umbrella overlaps that for species it aims to protect.

The species-habitat models made other assumptions and were subject to sources of error that weakened the models. In particular, models assumed that the species distribution is in equilibrium with the environment, species-habitat relationships were consistent throughout the study area and the full gradient of the species-response relationship was sampled (Guisan et al. 2002, Miller and Franklin 2002). In addition, RSF models assumed that the locations available to the animals were correctly identified (Manly et al. 2002). These assumptions were only partially supported. The bird survey data spanned different time periods and showed temporal variation as surveys had inconsistent species detections at a given location over multiple survey years. Some predictors such as elevation and geographic location did not change over the study period. However, climate and primary productivity varied and their values at a sample point with a species' detection may not have been measured in the same year as the detection. Using 30 year averages for the climate data may have lessened this error.

A species may have responded differently to environmental predictors in different parts of the large Y2Y region (Osborne and Suarez-Seoane 2002). As well, regression models based on empirical data model a species' realized niche and implicitly incorporate biotic interactions and negative stochastic effects that can change from one region to another (Guisan et al. 2002).

The "available points" were identified subject to being within a bird's breeding season range and constrained to the habitat types it used, so points were likely correctly identified and represented the full gradient of the predictor variables. However, the "used" points likely did not sample the full ranges of the response curves as the bird survey data was sparse in northern and high elevation habitats. I performed an exhaustive search to locate data in these habitats, but there is little available. I recommend that Y2Y try to establish collaborations with conservation groups and researchers working in high elevation and northern regions to obtain more representative survey data and redo the models with these additional data.

Other sources of error

I didn't find evidence of spatial autocorrelation that may have confounded the models. My sample distance (at least 1 km between points) is larger than autocorrelated distances for birds found in literature. For example, Lichstein et al. (2002) found autocorrelation occurred over distances less than 500 m for 19 bird species, with the exception of Veery that was autocorrelated over distances greater than 2 km. Koenig (1998) tested spatial autocorrelation at distances up to 1.2 km for 88 California bird species and found only 1 species (mourning dove) showed spatial autocorrelation during the breeding season. The species he tested included 3 that I modelled: American dipper, Wilson's warber and yellow warbler.

The SLCR land cover data classification is estimated to be about 75% accurate (Scepan 1999). As well, the whitebark pine layer I used was of uncertain origin and accuracy. I used these land cover classes to delineate the spatial extent for the "available" species points, and model extrapolation throughout Y2Y. I don't know how severely the inaccuracies in habitat cover type impacted the models or identification of high quality avian habitat.

Error propagation in the GIS is another source of error that is difficult to quantify. My GIS data was from multiple sources that used different geographic projections and mapping scales. Usually a substantial amount of reformatting and recalculation was needed to create the final GIS layers for the predictor variables that undoubtedly introduced inaccuracies in these layers.

Acknowledgements

Bird survey data were kindly provided by: Alberta Fish and Wildlife; BC Ministry of Water, Land and Air Protection; Canadian Wildlife Service (Andre Breault, Mike Gill, Craig Machtans, Wendy Nixon, Mike Norton and Pamela Sinclair); Dan Farr; Sandra Kinsey and Laird Law; MacKenzie Monitoring Station; Mike and Diane McIvor; Northern Region Landbird Monitoring Program; Parks Canada; Patuxent Wildlife Center; Yukon Department of Environment (Cam Eckert), Scott Wilson/Kathy Martin, UBC; Bird Studies Canada; U.S. Fish and Wildlife.

Literature Cited

Apps, Clayton D., McLellan, Bruce N., Woods, John G. and Michael F. Proctor. 2004. Estimating grizzly bear distribution and abundance relative to habitat and human influence. Journal of Wildlife Management **68**(1):138-152.

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

Andrews, P. and O'Brien, E.M. 2000. Climate, vegetation, and predictable gradients in mammal species richness in southern Africa. Journal of Zoology **251**:205-231.

Beard, Karen H., Hengartner, Nicolas and David K. Skelly. 1999. Effectiveness of predicting breeding bird distributions using probabilistic models. Conservation Biology **13**(5):1108-1116.

Boyce, Mark S. and Lyman L. McDonald. 1999. Relating popultions to habitats using resource selection functions. Trends in Ecology and Evolution **14**(7):268-272.

Boyce, Mark S., Vernier, Pierre R., Nielsen, Scott E. and Fiona K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling **157**:271-399.

Caro, T. M. and G. O'Doherty. 1999. On the use of surrogate species in conservation biology. Conservation Biology **13**(4):805-814.

Cody, M. 1985. editor. Habitat Selection in Birds. Academic Press, Inc. London. 558 p.

Crawley, M.J. 2002. Statistical Computing An Introduction to Data Analysis using S-PLUS. John Wiley & Sons Ltd. Chicester, U.K.

Dettmers, Randy, Buehler, David A., Bartlett, John G. and Nathan A. Klaus. 1999. Influence of point count length and repeated visits on habitat model performance. Journal of Wildlife Management **63**(3):815-823

Drapeau, Pierre, Leduc Alain and Raymond McNeill. 1999. Refining the use of point counts at the scale of individual points in studies of bird-habitat relationships. Journal of Avian Biology **30**:367-382.

Farnsworth, George L., Pollock, Kenneth H., Nichols, James D., Simons, Theodore R., Hines, James E. and John R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. The Auk **119**(2):414-425.

Fielding, Alan H. and John F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation 24(1):38-49.

Flather, C. H., K. R. Wilson, D. J. Dean and W.C. McComb. 1997. Identifying gaps in conservation networks: of indicators and uncertainty in geographic-based analyses. Ecological Applications **7**(2):531-542.

Franklin, J.1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. Journal of Vegetation Science **9**:733-748.

GLCF – Global Land Cover Facility. University of Maryland. http://umiacs.umd.edu/index.html

Great Lakes Forestry Centre. 2003. Environment Canada.

Guisan, Antoine and Niklaus E. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**(2-3):147-186.

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.

Hutto, Richard L., Hell, Sallie J., Kelly, Jeffrey F. and Sandra M. Pletschet. 1995. A comparison of bird detection rates derived from on-road versus off-road point counts in northern Montana. USDA Forest Service Gen. Tech. Report. PSW-GTR-149.

Landres, P. B., J. Verner and J. W. Thomas. 1988. Ecological uses of vertebrate indicator species: a critique. Conservation Biology **2**(4):316-328.

Legendre, P. and Fortin, M. J. 1989. Spatial pattern and ecological analysis. Vegetatio **80**:107-138.

Lichstein, Jeremy W., Simons, Theodore R., Shriner, Susan A. and Kathleen E. Franzreb. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecological Monographs **72**(3):445-463.

Liu, J., Chen, J.M. and Cihlar, J. 2003. Mapping evapotranspiration based on remote sensing: an application to Canada's landmass. Water Resources Research **39**(7):Art. No. 1189.

LP DAAC. Land Processes Distributed Active Archive Center, USGS EROS Data Center. http://LPDAAC.usgs.gov.

Manly, Bryan F.J., McDonald, Lyman L., Thomas, Dana L., McDonald, Trent L. and Wallace P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies. 2nd Edition. Kluwer Academic Publishers. Dordrecht.

McLoughlin, P.D., S. H. Ferguson and F. Messier. 2000. Intraspecific variation in home range overlap with habitat quality: a comparison among brown bear populations. Evolutionary Ecology **14**:39-60.

Miller, Jennifer and Janet Franklin. 2002. Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. Ecological Modelling **157**(2-3): 227-247.

NTDB Canada. National Topographic Data Base. Natural Resources Canada.

Oksanen, Jari and Peter R. Minchin. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling **157**(2-3):119-129.

Orians, Gordon H. and James F.Wittenberger. 1991. Spatial and temporal scales in habitat selection. The American Naturalist **137**(Supplement):S29-S49.

Pearce, Jennie and Simon Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling **133**:225-245

Pulliam, H.R. 1988. Sources, sinks, and population regulation. American Naturalist **132**:652-661.

Roland, J., Keyghobadi, N. and Fownes, S. 2000. Alpine Parnassius butterfly dispersal: effects of landscape and population size. Ecology **81**(6): 1642-1653.

Ruggiero, Leonard F., Pearson, Dean E. and Stephen E. Henry. 1998. Characteristics of American marten den sites in Wyoming. Journal of Wildlife Management **62**(2):663-673.

Schadt, S., Revilla, E., Wiegand, T., Knauer, F., Kaczensky, P., Breitenmoser, U., Bufka, L., Cerveny, J., Koubek, P., Huber, T., Stanisa, C. and L. Trepl. 2002. Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. Journal of Applied Ecology **39**:189-203

Scepan, J. 1999. Thematic validation of high-resolution global land-cover data sets. Photogrammetric Engineering and Remote Sensing **65**(9):1051-1060.

Suarez, A.V., Pfennig, K.S. and Robinson, S.K. 1997. Nesting success of a disturbance-dependent songbird on different kind of edges. Conservation Biology **11**(4):928-935.

Suryan, R.M. and D. B. Irons. 2001. Colony and population dynamics of black-legged kittiwakes in a heterogeneous environment. The Auk **118**(3):636-649.

U.S.G.S. 1994. United States Environmental Protection Agency land use/land cover digital data.

Venier, L.A., McKenney, D.W., Wang, Y. and J. McKee. 1999. Models of large-scale breeding-bird distribution as a function of macro-climate in Ontario, Canada. Journal of Biogeography **26**:315-328.

Wiens, John A. and Rotenberry, John T. 1981. Censusing and the evaluation of avian habitat occupancy. Pp 522-532 in Studies in Avian Biology No 6. Ralph, C. John and Scott, Michael editors. Cooper Ornithological Society.

Wiens, John A., Rotenberry, John T. and Beatrice Van Horne. 1987. Habitat occupancy patterns of North Americn shrubsteppe birds. OIKOS **48**:132-147.

Young, Jock S. and Richard L. Hutto. 2002. Use of regional-scale exploratory studies to determine bird-habitat relationships.pp 107-119 in: Predicting species occurrences: issues of accuracy and scale. Scott, J.M. et al. editors. Island Press, Covello. California.

Zacharias, M.A. and Roff, J.C. 2001. Use of focal species in marine conservation and management: a review and critique. Aquatic Conservation: Marine and Freshwater Ecosystems **11**:59-71

Zaniewski, A.Elizabeth, Lehmann, Anthony and Jacob McC. Overton. 2002. Prediction species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecological Modelling **157**(2-3):261-280.