UNIVERSITY OF ALBERTA

HABITAT SELECTION MODELS FOR GRASSLAND BIRDS AT CANADIAN FORCES BASE SUFFIELD

by

TREVOR S. WIENS ©

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF EARTH AND ATMOSPHERIC SCIENCES

EDMONTON, ALBERTA
SPRING 2006

Your file Votre référence
ISBN: 0-494-13909-9
Our file Notre référence
ISBN: 0-494-13909-9

Dedication

This thesis is dedicated to my wife, Hiroko, who patiently supported me for the duration of studies, and my daughter, Yuriko, whose happiness to see me at the end of the day always made me smile. Without their support this thesis would not have been possible.

Abstract

K-fold cross-validation was used to determine the predictive ability of logistic regression estimated resource selection function (RSF) models. Models were evaluated and selected based on their general, spatial, and temporal predictive ability (3-way RPI or 3-way RSF Plot Index). This method was used to evaluate which remotely sensed and GIS-based predictor variables, acting as proxies for structural habitat characteristics, were effective for modelling habitat selection of eleven grassland bird species.

Five years of bird point count data from an area of native prairie were used. The 3-way RPI method is dependent on the assignment of output to arbitrary suitability classes. Methods to partially ameliorate the threshold dependency created by this class assignment were developed. The use of random, temporal, and spatial partitions of the data to evaluate general, temporal, and spatial model robustness was demonstrated to be superior to standard methods of general testing.

Acknowledgements

I am grateful to Brenda Dale of the Canadian Wildlife Service for allowing me to work with her data and giving constructive criticism to the core work of this thesis. I want to thank my supervisor, Peter Kershaw, who was accessible, supportive, and source of good advice. Without the help and excellent suggestions of Mark Boyce the quality of this work would be much lower and I would have learned much less than I did. I also want to thank Michael Guenter who took the time to copy edit and give suggestions on the writing of this thesis. I also need to thank Barry Adams of Alberta Sustainable Resource Development, Brent Smith of CFB Suffield, and Brian Wiens of the Meteorological Service of Canada for supplying data and assistance. I also want to thank Arturo Sanchez for his assistance.

Although many assisted the final product is my own and all errors and omissions are my responsibility alone.

Table of Contents

List of Tables

List of Figures

Abbreviations and Glossary

ρ_{blue} Blue reflectance. For Landsat this is band 1. For MODIS this is channel 3.

ρ_{NIR} Near-infrared reflectance. For Landsat this is band 4. For MODIS this is channel 2.

ρ_{red} Red reflectance. For Landsat this is band 3. For MODIS this is channel 1.

AIC Akaike's Information Criteria. Founded on advanced theories of entropy and information, this method proposes to select a model from within a group of models that is the best. This is not a measure of its absolute fit, only relative to the other models. BIC or Bayesian Information Criteria is similar but tends to select more parsimonious models. These methods are founded on the concept of the existence of a true model.

AGRASID Agricultural Region of Alberta Soil Inventory Database. This information consists of a detailed use guide, a GIS layer of polygons and a series supporting databases.

ANPP Above-ground Net Primary Production or Productivity.

AUC Area Under the Curve. This refers to the area under a ROC graph. Values of 0.5 indicate a random relationship. Values higher than 0.5 indicate a positive relationship and models scoring 0.7 or higher are considered useful.

AWC Available Water Capacity. Water retained within soil of which a proportion is available for plant use.

Brightness A index derived from the taselled-cap transform. Brightness is a measure of soil reflectance.

CFB Canadian Forces Base.

Cook's Distance This is a measure of the impact of an individual observation on a regression equation.

CSM Conserved Soil Moisture. An weighted average of precipitation over two years, designed to approximate the amount of plant available water in soil.

CWS Canadian Wildife Service.

CTI Compound Topographic Index. A measure of position on a slope with low values indicating near the top and large values indicating near the bottom.

DEM Digital Elevation Model.

ERS Ecological Range Site. A range management focused plant community classification system.

EVI Enhanced Vegetation Index. A MODIS specific vegetation index.

$EVI = 2\frac{\rho_{NIR}-\rho_{red}}{L+\rho_{NIR}+C_1\rho_{red}+C_2\rho_{blue}}$

L is a canopy background adjustment term. C_1 and C_2 affect the aerosol correction by ρ_{blue} of ρ_{red}.

FSS Fold Sort Score. A index ranging from 0 to 1 indicating how many k-fold Spearman-rank correlation test scores were equal or higher than the critical value to be considered stronger than could be attributed to random chance.

GIS Geographic Information System.

GVI Green Vegetation Index or greenness. A vegetation index derived from the tasselled-cap transform. GVI is a measure of reflectance from green vegetation.

MODIS Moderate Resolution Imaging Spectroradiometer. A NASA satellite with 36 spectal bands and spatial resolutions of 250m, 500m, and 1000m.

MSAVI$_2$ Modified Soil Adjusted Vegetation Index 2. It's performance is similar to SAVI, but is self calibrating and thus easier to use.

$MSAVI_2 = \frac{2\rho_{NIR}+1\sqrt{(2\rho_{NIR}+1)^2-8(\rho_{NIR}-\rho_{red})}}{2}$

NDVI Normalized Difference Vegetation Index. $NDVI = \frac{\rho_{NIR}-\rho_{red}}{\rho_{NIR}+\rho_{red}}$

NPP Net Primary Production or Productivity.

NWA National Wildlife Area.

Remote Sensing Gathering data from a distance. For the purposes of this thesis, this would mean use of satellite or airplane mounted sensors.

ROC Receiver Operating Characteristics. A method of testing model performance based on a matrix of true positives, false negatives, false positives, and true negatives.

RPI RSF Plot Index. Either an individual Spearman-rank value or mean value over series of k-folds. This index is an indicator of the strength of the predicted / observed relationship and is bound by the limitations of the Spearman-rank correlation test.

RSF Resource Selection Function. A function proportional to the probability of use by an organism.

SAVI Soil-adjusted Vegetation Index. $SAVI = \frac{\rho_{NIR}-\rho_{red}}{\rho_{NIR}+\rho_{red}+L}(L+1)$
L is a correction factor that ranges from 0 – 1. Usually it is set to 0.5.

Tasselled-cap transform A method of processing reflectance data from multiple bands via means similar to principle components analysis. It was originally created for Landsat MSS data and later adapted to other data sources.

VCT Vegetation Cover Type. A physiognomic plant community classification scheme for the Suffield NWA.

Wetness A index derived from the tasselled-cap transform. Wetness is a measure of reflectance from water, either in the soil or in the vegetation.

1 Models for management

1.1 History and goals

As someone starting a graduate program after more than 10 years of practical experience, my interest was in doing research of practical value. I had worked for a number of years with different organisations on bird atlas and checklist programs, so the question of species distribution was one of particular interest. When speaking with Brenda Dale of the Canadian Wildlife Service (CWS), I became aware of a mature data set for grassland birds from the Suffield National Wildlife Area (NWA) within the Canadian Forces Base (CFB) Suffield in southern Alberta, and their interest in looking at the utility of using remotely-sensed and GIS data to model habitat selection for these species.

The interest of CWS in having someone do research with their data provided the means to do a Masters of Science research project that was not the classic one field season master's project. Instead, my project would be firmly grounded in real management problems and would provide the experience and knowledge I hoped to gain from a graduate program.

The bird data had been gathered as part of a long-term monitoring effort by CWS and CFB Suffield. This monitoring was conducted as part of the management strategy for cattle grazing in the NWA. As there are many bird species that show marked preference for native grasslands, monitoring the abundance and distribution of these species became part of that larger management plan. Species chosen for this project were those neither too common or too rare, and which depended on native grasslands (Table 2-1).

The research project objectives were clear. Test and confirm the utility of remotely-sensed data for constructing habitat selection models and create usable models for as many of the target species as possible. As the research unfolded and course work was undertaken, the shape of the project evolved and became focused on two basic questions. First, what methods were available to evaluate model performance generally, spatially, and temporally? Second, what types of predictor variables were needed to model habitat selection over space and time with a minimum of cost and effort?

1.2 Study area

The NWA is a smaller area embedded in CFB Suffield that covers approximately 2600km^2. The northern portion of the NWA is not subject to cattle grazing, so it was not part of the ongoing monitoring program even though field surveys were conducted in 1994 and 1995 as part of the federal effort to assess this unique area.

The project study area was the south block of the NWA. The south block is approximately 190km^2 or 7% of CFB Suffield. Topographically within CFB Suffield the elevation range is 583–846m, but in the south block the range is 619–756m (Natural Resources Canada 2004). Thus, it was clear that although the south block was representative of a portion of CFB Suffield, models developed within this area would have limited application to the entire base if variables were dependent on characteristics unique to this region. With this in mind, I attempted to use variables that would be transferable to areas outside the south block and the NWA. In the areas surrounding CFB Suffield, much of that land is under cultivation whereas most of the NWA has never been cultivated, or hasn't been cultivated for many years. This difference also limited the direct application of models developed within the NWA to adjacent areas.

The grasslands of southern Alberta are dominated by herbaceous vegetation. Perennial grasses dominate net primary production (NPP), but forbs are the primary contributors to species diversity (Bragg and Steuter 1996). Trees are generally limited to areas with greater moisture, such as riparian areas. As a semi-arid landscape, topography plays an important role in moisture availability (Bragg and Steuter 1996) and thus the distribution of plants. Related to the issue of topography are wetlands of various classes from permanent to ephemeral which are preferred habitats of some species (Lowther et al. 2001).

With relevance to avifauna grasslands consist of three structural layers. First, there is litter, or dead vegetation fallen on the ground. Second, there is grass cover, meaning the density and height of living grass and standing dead. Third there are shrubs. These three elements strongly affect reproductive and feeding behaviour and thus are important classifiers of grassland bird habitat (Gill and Poole 1993–2004).

CFB Suffield is within the dry mixedgrass natural subregion (Adams et al. 2004). The south block is mostly native mixed-grass prairie. Although at first glance the area appears to be in a natural state, there is actually a tremendous amount of oil and gas activity in the base, though wells are all subsurface. The one obvious impact of this activity that is linear disturbance such as roads, trails, and pipelines. Along many roadways there is Crested wheatgrass (*Agropyron desertorum*) which was either seeded or brought in on tires. Another alteration to the landscape has been the placement of dugouts and wells for watering cattle. In areas close to the watering areas, grazing is more intense, causing changes in the amount of litter, standing dead, and green plant cover.

1.3 Statistical context

Researchers often find themselves with data that are not simple use / non-use data sets. Instead, they have data that are listings of used habitats and listings of available habitats that may or may not be in use by the species in question. Such data conform to a use / available, presence only, or presence / available study design. This is further complicated when the same locations might be recorded as used sites as well as available sites (Johnson et al. 2006). This type of study design is likely to occur in situations where field data are compared against data extracted randomly in a GIS environment. Of critical importance when using these study designs is that traditional methods of evaluation for logistic regression models are not appropriate (Boyce et al. 2002).

The problems with use / available study designs is that all traditional model evaluation techniques for these types of data are dependent upon the use of a confusion matrix. This necessitates knowledge of true positives, true negatives, false positives, and false negatives. With use / available data and designs, true positives are known, but true negatives are not.

K-fold cross-validation to measure model predictive ability is an appropriate means to evaluate use / available models (Boyce et al. 2002). Although this technique was explicitly developed for use / available designs, as it is a measure of model predictive ability, its expansion to any RSF model is appropriate as the utility of an RSF model is dependent on its ability to predict. For this study I employed a use / non-use design with bird point-count survey data collected over five years and over a heterogeneous landscape. This situation permitted the expansion of the k-fold technique to use k-folds based on year or spatial location to evaluate model performance in a robust fashion.

1.4 Unused and unavailable data

During the course of the project, a draft of the new Ecological Range Site (ERS) system by Adams et al. (2004) was made available. These data were an amalgamation of soils, plant communities, and topography. Since this layer was available for the NWA as well as areas outside of CFB Suffield it was considered because it could enable the application of constructed models to areas outside the NWA and CFB Suffield. Plant communities within the ecological range sites were classified by dominant species which provided some indication of some of the critical structural factors thought to be important for birds.

The Vegetation Cover Type (VCT) system based on the report of Adams et al. (1999) was also available for the NWA. The VCT data was physiognomic in nature, describing the vegetation in terms of structural elements rather than species associations, as used in the ERS system. For birds, the VCT classification, although rather general, was more intuitive and potentially more useful as it defined communities in terms similar to those used to describe bird habitat.

To examine these two systems thoroughly, I first examined the two systems, both in terms of classifications and the distribution of the polygons to try to understand their commonalities and differences. I subsequently attempted to use these systems to construct models for the study species with varying degrees of success. Finally, I examined some of the source field data on shrub cover used to generate the VCT classifications and compared this against the final classifications for both the VCT and ERS system. In the end these data sources were dropped from consideration because the polygon sizes of these areas were very large, so they didn't have sufficient spatial resolution to model habitat selection at the scale of the study area. Associated with this scale issue was the fact that there were very few classes to consider in the study area and a few of those constituted the majority of the study area.

As discussed in the description of the study area, cattle grazing was ongoing during the study and its associated activities affected the abundance of relevant habitat factors. Measures of cattle use were collected by CWS but not for all avian sample points. Cattle will travel only limited distances from water, so measures of distance to water have potential as indirect indices of grazing. Unfortunately, neither CFB Suffield nor CWS possessed an accurate and properly classified layer of wetlands and dugouts. As a result, a measure of cattle impact was not available during the course of the project. This was unfortunate, as grazing can be an important factor for most of the species under consideration (Gill and Poole 1993–2004). In recent discussions with Brenda Dale (October 2005), CWS and CFB Suffield are now in the process of creating such a GIS layer.

1.5 Temporal trends

During the course of the field program, there were large variations in seasonal precipitation (Figure 3-1) and in bird numbers from year-to-year (Table 3-7). This variability created a challenge in creating multi-year models that would accurately predict some of the dramatic changes in species relative abundance.

1.6 Summary

There were two questions explored in this thesis. First, what methods exist or could be enhanced to robustly evaluate habitat selection models for grassland birds? Second, what remotely-sensed or GIS-based predictor variables were effective for modelling this selection? These questions are addressed in Chapters 2 and 3 respectively. Chapter 4 provides a summary of this work and outlines some possible directions for future research in this area.

References

Adams, B. W., Poulin-Klein, L., Moisey, D., McNeil, R., 2004. Rangeland plant communities and range health assessment guidelines for the dry mixedgrass natural subregion of Alberta, Pub. No. T/040 Edition. Rangeland Management Branch, Public Lands Division, Alberta Sustainable Resource Development, Lethbridge, Alberta.

Adams, G. D., Trottier, Gary, C., Strong, Wayne, L., MacDonald, I. D., Barry, S. J., Gregoire, P. G., Babish, G. W., Weiss, G., 1999. Vegetation component report. In: Canadian Forces Base Suffield National Wildlife Area Wildlife Inventory. Canadian Wildlife Service, Edmonton, AB, unpublished report.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., Schmiegelow, F. K. A., 2002. Evaluating resource selection functions. Ecological Modelling 157, 281–300.

Bragg, T. B., Steuter, A. A., 1996. Prairie ecology — the mixed prairie. In: Nelson, E. B., Samson, F. B., Knopf, F. L. (Eds.), Prairie Conservation. Island Press.

Gill, F., Poole, A. (Eds.), 1993–2004. The Birds of North America. The Birds of North America, Inc., Philadelphia, PA.

Johnson, C. J., Nielsen, S. E., Merrill, E. E., McDonald, T. E., Boyce, M. S., 2006. Resource selection functions based on use-availability data: theoretial motivation and evaluation methods. Journal of Wildlife Management In press.

Lowther, P. E., Douglas, III, H. D., Gratto-Trevor, C. L., 2001. Willet (*Catoptrophorus semipalmatus*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 579. The Birds of North America, Inc., Philadelphia, P.A.

Natural Resources Canada, 2004. GeoBase Canadian Digital Elevation Data. URL `http://www.geobase.ca`

2 Three way k-fold cross-validation of logistic regression models applied to grassland birds [1]

2.1 Introduction

Resource selection functions (RSF) are defined to be proportional to the probability of use of a resource unit by an organism (Manly et al. 2002). The utility of these models is dependent on their ability to predict. If a model has a spatial and temporal component, prediction will vary in space, in time, and as a function of its stability. Boyce et al. (2002) point out that many apparently adequate models fail in new areas or time periods. To use RSF models effectively, an evaluation of their predictive ability in all three dimensions is needed. We put forward a method to evaluate model performance in each of these dimensions and to select models based on their predictive ability. Equally important with this method is not only a means to select the model that can predict more accurately, but to understand the limitations of selected models so that they can be used appropriately.

For use / available (also known as presence only or presence / available) study designs, locations where species are observed, can also appear in the model as available locations (Johnson et al. 2006). This situation makes traditional methods for model evaluation inappropriate (Boyce et al. 2002). Boyce et al. (2002) put forward the concept of using k-fold cross-validation as a means to evaluate presence / available models using plots of expected vs. observed and the Spearman-rank correlation to generate an index value (hereafter referred to as RSF Plot Index or RPI). The RPI statistic provides a measure of a models predictive ability. Considering that the utility of RSF models lies in their ability to predict, the application of RPI to any RSF model, regardless of study design, is appropriate. In this paper we expanded the RPI technique to examine general, spatial, and temporal model stability. We also developed a method to help ameliorate the difficulty of the RPI's sensitivity to number and placement of suitability classes (bins) when conducting k-fold cross-validation.

We constructed a series of logistic regression models for grassland birds using a use /

[1] A version of this chapter is being prepared for submission and review as: Wiens, T.S., Boyce, M.S., Dale, B.C., Kershaw, G.P., 2006. Three way k-fold cross-validation of logistic regression models. Ecological Modelling

non-use study design (Manly et al. 2002). We selected models using RPI values from the set of potential models that passed the goodness-of-fit (GOF) test of le Cessie and van Houwelingen (1991) for logistic regression.

2.2 Methods

We modelled habitat selection for eleven species of grassland birds (Table 2-1) in the Canadian Forces Base (CFB) Suffield National Wildlife Area (Suffield NWA), in south-eastern Alberta, Canada (50°14′N, 110°37′W), using five years of point-count survey data. Over the course of the study there were large variations in precipitation. Drought conditions existed in the area in 2001 and in the early spring of 2002. Late spring of 2002, as we were conducting our field surveys, higher than normal precipitation occurred but was insufficient to offset the previous year's deficit. Other years were relatively normal (Figure 3-1). To take advantage of these variable conditions, we selected bird species from the pool of those detected that were known to be sensitive to year-to-year variation in plant growth (Gill and Poole 1993–2004) or which we observed to vary in abundance and distribution over the course of our study. We also avoided species that were very common or those which were extremely rare in our study area (Table 3-1). For a summary of top model performance by species see table 2-6.

Table 2-1: Common and scientific names (American Ornithologist's Union) for modelled species

Common Name	Scientifi c Name
Willet	*Catoptrophorus semipalmatus*
Upland Sandpiper	*Bartramia longicauda*
Marbled Godwit	*Limosa fedoa*
Sprague's Pipit	*Anthus spragueii*
Clay-colored Sparrow	*Spizella pallida*
Vesper Sparrow	*Pooecetes gramineus*
Lark Bunting	*Calamospiza melanocorys*
Savannah Sparrow	*Passerculus sandwichensis*
Baird's Sparrow	*Ammodramus bairdii*
Grasshopper Sparrow	*Ammodramus savannarum*
McCown's Longspur	*Calcarius mccownii*

2.2.1 Model Components

Bird data were collected using the point-count procedure outlined in Dale et al. (1999) in late May and early June of 2000–2004. Survey sites (263) were placed on 1km spaced transects at 500m intervals to give a representative sample of the entire study area. At each survey location, birds were recorded by song or sight inside or outside of a 100-m-radius for 5 minutes.

We categorised predictor components into temporal and vegetation groups. The vegetation variables were used to provide the spatial component of our models. Temporal variables consisted of precipitation data which would provide information on inter-annual variation in plant growth. The effect of previous year precipitation and above-ground net primary productivity (ANPP) on current year ANPP was studied by Oesterheld et al. (2001). Regression models with r^2 values of 0.58 and 0.60 resulted from current year precipitation and previous year ANPP or current and previous year precipitation respectively. The relationship between inter-annual plant growth and precipitation is however not linear (Flanagan et al. 2002); wet years produce a much greater effect on increased production than dry years affect decreased production. The relationship between ANPP and grassland canopy structure was studied by Lane et al. (2000) who found ANPP increases were positively correlated with increases in canopy height and Leaf Area Index.

Based on this research we posited that the use of precipitation data, together with measures of spatial variation in plant growth, would be sufficient to form good predictive models for at least some of the birds under study. Elsewhere we explored the relative performance of different combinations of temporal and spatial predictor variables (Chapter 3).

2.2.2 Data Processing

Data were managed using PostgreSQL 8.0. This was chosen because of its ease of use with other selected tools. Geographic Resources Analysis Support System (GRASS) version 6.0 was used for raster and vector analyses. Statistical analysis was conducted using R (R Development Core Team 2004). The GRASS `v.sample.buffer` module was used to extract raster attributes within a 250-m-radius buffer around survey sites and save them in a PostgreSQL database.

2.2.3 Predictor Variables

Environment Canada provided precipitation data for the five climate stations surrounding the study area within a 60-km-radius. This information was merged to form regional average precipitation by month for the period of 1998 through 2004. These regional averages were then wrapped up into current year, previous year, and two years previous totals for use in the models.

Price et al. (2002) found that discrimination between different grassland plant communities and management practices was best done using a single image from July rather than May or September. We thus used a cloud-free Landsat 7 image from 29 July 2000 which was a slightly dry, but close to normal year. The time of year and year chosen provided some confidence that this image would provide a good measure spatial variation of plant growth at a 30m spatial resolution for the study area. From this image the Normalized Difference Vegetation Index (NDVI), Modified Soil Adjusted Vegetation Index 2 ($MSAVI_2$) (Qi et al. 1994), and the tasselled cap Brightness, Greenness or Green Vegetation Index (GVI), and Wetness indices (Crist et al. 1986, Huang et al. 2001, Kauth and Thomas 1976) were created.

Table 2-2: Model predictor variables

Variable Group	Variable	Source	Rationale or Comment
Vegetation	NDVI	Landsat 7	Reflectance of green vegetation
	$MSAVI_2$	Landsat 7	Reflectance of green vegetation
	GVI	Landsat 7	Reflectance of green vegetation
	bright	Landsat 7	Brightness is a measure of reflectance mostly attributable to soil
	wet	Landsat 7	Wetness is a measure of moisture in vegetation and soil.
	vegetation index variance	Landsat 7	Variance within the 250m radius survey location is a indicator of heterogeneity. Spatial heterogeneity is of varying importance to different birds (Gill and Poole 1993–2004).
Temporal	$precip_t$	Environment Canada	Current year precipitation from January 1 to May 31
	$precip_{t-1}$	Environment Canada	Previous year precipitation
	$precip_{t-2}$	Environment Canada	Two years previous precipitation
	CSM	Environment Canada	Weighted average of past 2 years of precipitation

Determining how best to combine predictor variables based on *a priori* assumptions was difficult due to their general and somewhat ambiguous nature. For instance, similar combinations of precipitation and vegetation variables could conceivably model a number of different species depending on model beta weights.

Having many possible models presented us with a number of undesirable options. One option would have been to consider the use of backward step Akaike's Information Criteria (AIC) model selection with a series of nested model groups. This technique however, is generally unacceptable to statisticians, so we rejected this approach. A second option would have been to run all possible combinations, but this could only be perceived as fishing. Therefore we took a number of measures to reduce the number of models.

With the large number of vegetation indices, the number of conceivable combinations was very large. Since different vegetation indices have different sensitivities to moisture and soil effects, we wanted to include multiple indices to capture this variability. However, since the indices were all derived from a single source and thus were correlated we wanted to avoid collinearity problems. We calculated a correlation matrix for NDVI, $MSAVI_2$, GVI, Brightness, and Wetness for the entire Landsat scene. We conservatively choose a cut point of 0.7 or higher to remove combinations from consideration (Mason and Perreault 1991). Combinations passing this criteria were $MSAVI_2$ with GVI and Wetness and NDVI or $MSAVI_2$ with Brightness and Wetness. The reasons why these combinations are reasonable are discussed in Chapter 3.

Research into ANPP indicated that antecedent moisture conditions were very important for plant growth, thus current year, previous year, and two years previous precipitation was calculated. To limit the number of possible combinations we decided to use precipitation data in only three ways. First, we used Conserved Soil Moisture (CSM) (Williams and Robertson 1965); an estimate of soil moisture on May 1^{st} using a weighted combination of precipitation data from the previous two years. Second, we used the current and previous year's precipitation. Third, we used precipitation from the current year, previous year, and two years previous.

These combinations of variables were arranged into four groups of six models each, for each species (24 models / species).

2.2.4 Model Considerations

The number of individuals observed at a single location was small. With such small numbers, the modelling of abundance would be difficult. It was decided to reduce the data to recorded or not recorded and use logistic regression to model habitat selection.

There has been some recent debate concerning the severity of spatial autocorrelation on model accuracy (Diniz-Filho et al. 2003, Lennon 1999, 2000, Nielsen et al. 2002). The bird survey technique employed here, of closely spaced survey locations on transects, is a possible cause for concern. The issue of spatial autocorrelation in RSF models is related to bias in hypothesis testing (Lennon 2000) caused by reduced variance in the model. The RPI statistic is not concerned with, or appropriate for, hypothesis testing; it is a measure of model predictive ability. Issues related to spatial autocorrelation do not affect the RPI statistic and thus, were not relevant.

2.2.5 Model Processing

As recommended by Burnham and Anderson (2002), a series of ecologically reasonable models were drafted for each species. These models were then defined in a configuration file to be read by a small program which would then run the k-folds and other tests.

During processing, models were trained using the complete data set. From these full models the GOF values (using the *Design* library for R or S-Plus) were computed. All models failing the GOF test ($P \leq 0.10$) were excluded from consideration.

At this point, three separate, five-fold cross-validations were conducted. Source data were partitioned into 5 equal-sized random blocks using the `sample` function in R. Subsequently 4 blocks were selected to train a model and then a fifth partition was predicted. This was repeated five times such that all five parts had been predicted by the other four. Subsequently, k-fold cross-validations were conducted by splitting the data apart by field year and by five spatially distinct blocks within the study area. These k-fold cross-validations were used to measure the temporal and spatial stability of the models.

Since the observed data were binary, data had to be grouped into a series of suitability classes or bins in order to generate a plot of observed vs. predicted likelihoods. Ten equal-area bins using normalised scores and a moving-window average were employed for this purpose (see next section for details). The Spearman-rank correlation test or Spearman r_s value was calculated for each plot, giving a RPI value for each fold. The RPI values for the random, temporal, and spatial components were recorded in tabular format by their respective folds as well as a mean value. For the spatial and temporal folds, fold scores were compared against the critical value for the Spearman-rank test with a n=10 and an alpha of 0.05 ($\geq$ 0.564 Wagner (1992)). Folds passing this test were counted and divided by the number of folds. Therefore, each model had a fold sort score (FSS) based on the combined performance of the spatial and temporal k-fold RPI values. Recent work by Bengio and Grandvalet (2004) point out that there are no unbiased estimators of k-fold cross-validation variance. Although some adaptation of boot-strapping might be feasible to overcome this, this was not explored. Observationally, different runs of the same models gave quite different individual fold and mean RPI values. The objective of the FSS was to generate a consistent measure of a models performance to be used for model selection. Thus, random k-folds were not included.

After all models were estimated, a sorted list by FSS was created to identify strong models quickly and easily. After manual examination of the lists and models, a single best or group of top models were chosen for each species. Residuals plots and Cook's distance plots were examined to identify data points exerting an inordinate influence on the model or models. Prior to modelling, the source data had been examined for errors and outliers, yet all models had some data points identified as overly influential. These potential offending data points ($< 2\%$) were trimmed and the top candidate models were re-run to evaluate their performance. If the resulting models were similar in performance after the removal of these data, the effects of the removed data points were classified as non-problematic. If however the models changed substantially in their predictive ability, they were excluded from consideration. In this fashion we could examine the top candidate models thoroughly and select a single best model for each species.

2.2.6 Binning Methods

Boyce et al. (2002) and Hirzel et al. (2006) point out that one of the difficulties with the RPI method, is the arbitrary nature of the bins. The results generated by this method can be very sensitive to the number and placement of the bins.

This sensitivity is seen in considering two hypothetical situations. For the sake of argument, we will assume that the model considered is a true RSF function. We will also assume that the bins have been evenly distributed from 0 to 1. If the distribution of our model's predicted probabilities were skewed right, then the resulting plot would have a rise, a peak, and a long tail. Now this would not necessarily mean that the function was not a true RSF, but rather it could indicate that a very small percentage of the landscape consisted of optimal habitat as defined by the model. Unfortunately, the placement of the bins, prevents proper detection of this, and is in fact, insensitive the ecological reality of the study area and might generate an artificially low RPI value.

To deal with variation in the distribution of predicted values, equal area bins should be used. We constructed equal area bins based on the results of the full model. Predicted values were generated and sorted and then bins with equal numbers of records were created. Corrections in the case of an uneven number of bins were done in favour of the last bin to try to prevent artifactual dips at the upper end of the graph. When k-folds were conducted, their distributions were compared against the base distribution and an adjustment factor was calculated to force all predicted data into bins as if there had been an even distribution of points across all bins. This method ensured that the bins were a reflection of the actual probability distribution, and a true measure of the observed vs. predicted relationship.

A second problem situation is the number of bins used. As the number of bins increases, eventually zero value bins will appear in the plot. These zero value bins are a function of data sparseness, not of an invalid model. However, as above, artificially low RPI values would be reported. We started with 10 bins and assessed our data to see if we were likely to encounter data sparseness induced zero value bins. In our case, the Clay-colored Sparrow was the species with the least data, with only 40 records in total. Divided by 5, this indicated that if records were distributed evenly there would be only 8 records per fold. We predicted that this would not be a problem as it was likely that the lower 2 or 3 bins might often be empty. This prediction was borne out in observation, with using 10 bins providing no difficulties related to sparsity of data.

In our situation however, a third problem existed. We had a multi-year model with variable numbers of birds from year-to-year (Table 2-5), and we were testing the model over diverse spatial blocks. Under these conditions it was reasonable to expect that the amount of suitable habitat available from year-to-year, and from spatial block to spatial block, would vary greatly. We had determined the break points for the equal area bins based on the full model. Thus, we expected variation in the data distribution over the individual k-folds, as a reflection of the real variation in the spatial and temporal landscapes. To address this issue, we applied a normalisation technique from remote sensing. In order to compare images from one year to the next, images are often standardised to account for differences in lighting conditions, etc. This standardisation allows for meaningful comparison among images. We applied this concept to our k-fold process by transforming all values to a range between 0 and 1 using minimum and maximum values. This was done before determining our break points based on the full model as well as before bin assignment for the individual k-folds. In this manner, we could be assured of these non-random plots, combining in a meaningful way.

Hirzel et al. (2006) suggest the use of a moving window to overcome sensitivity to the number of bins. We applied this method to our normalised equal area bins by including one bin above, and one bin below, into an average value plotted for each bin. Thus for bin zero, it was an average between bin 0 and bin 1. For bin 1, it was an average between bins 0, 1, and 2, and so on.

2.3 Results and Discussion

2.3.1 Binning Results

The example in table 2-3 illustrates the value of the methods we employed. The use of equal area bins over fixed bins resulted in stronger relationships from the data as predicted. In addition the use of normalised bins improved the performance for the spatial and temporal k-folds and the use of the moving window average resulted in further improvement. When all three were combined, the relationships were the strongest and remained significant, although weaker, as the bin number increased causing data scarcity problems.

The use of normalisation with random k-folds could be inappropriate. The random sampling from the data set was intended to provide a representative sample of the population. In this context, it could be more theoretically sound not to use the normalisation. The relative stability in the RPI values generated from the random k-fold for all different binning techniques indicated no loss in apparent performance from this conservative action.

For practical usage of the RPI, it is important to determine what minimum value should be considered valid. In our case we calculated the Spearman r_s values for each bin separately and then provided those along with a mean of all folds. This value could be calculated in a single step, but in that case, the resulting P values would be artificially low because the individual k-folds were not independent. Thus, we chose conservatively and used 0.564 (based on n=10, alpha=0.05) as the cutoff value for which relationships could be considered real and not attributable to random chance. For reference, this value would be equivalent to an receiver operating characteristics (ROC) graph area under the curve (AUC) value of 0.5. At this point however, we are unsure about what value to recommend as an analogue to the AUC value of 0.7 or higher for considering a model useful (Boyce et al. 2002).

Table 2-3: Comparison of different binning methods

Baird's Sparrow k-fold Spearman r_s values (Critical values for alpha 0.05: n=10, 0.564; n=15, 0.441; n=20, 0.377) See table 2-4 for model details.			
Bins	Random Fold Mean	Year Fold Mean	Block Fold Mean
Normalised equal area bins with moving window average			
10	0.9103	0.8066	0.9228
15	0.9064	0.6971	0.8833
20	0.8620	0.6387	0.7976
Equal area bins with moving window average			
10	0.8289	0.7042	0.9402
15	0.9109	0.5668	0.9119
20	0.8727	0.5332	0.8790
Normalised equal area bins			
10	0.8086	0.5841	0.7319
15	0.7371	0.4492	0.6560
20	0.7439	0.4060	0.5605
Equal area bins			
10	0.8390	0.5398	0.8102
15	0.7038	0.3046	0.6987
20	0.7213	0.2966	0.6282
Fixed bins with moving window average			
10	0.0517	0.4605	0.1615
15	0.0543	0.4113	0.0962
20	0.0351	0.3572	0.0786

2.3.2 3-way validation

The behaviour of an animal using an ecosystem is incredibly complex, which, quite remarkably, often can be modelled with only a few variables. Unlike artificial systems such as cellular automata, in the real world, we don't have access to the underlying mechanisms from which these complex behaviours arise. Considering this fact, is it reasonable to assume that a single measure will provide all the information needed to assess the stability and appropriateness of a model? Our research suggests that measures of a model's performance in all dimensions of variance are necessary to have confidence in its validity. We were able to generate at least one model for each species that would be considered adequate using the GOF and general RPI criteria. Yet only 4 of these models survived scrutiny of their temporal and spatial performance by extending the k-fold validation to those dimensions (Table 2-6). If the individual temporal and spatial k-folds are examined, the status of these models becomes more complicated. We use two species to illustrate this.

Table 2-4: Example Models

Baird's Sparrow (GOF 0.6089)								
$X\hat{\beta} = -23.2364 - 0.02867\ \mathrm{precip}_t + 0.0304\ \mathrm{precip}_{t-1} - 58.4714\ \mathrm{ndvi} + 0.0071\ \mathrm{bright} - 0.0034\ \mathrm{wet} + 411.1721\ \mathrm{ndvi}_{var}$								
k-fold Spearman r_s values (Critical value for alpha 0.05: n=10, 0.564)								
Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9758	< 0.0001	2000	0.7774	0.0007	1	0.9521	< 0.0001
2	1.0000	< 0.0001	2001	0.8788	< 0.0001	2	0.9225	< 0.0001
3	0.9758	< 0.0001	2002	0.7399	0.0014	3	0.7638	0.0009
4	0.9698	< 0.0001	2003	0.6431	0.0053	4	1.0000	< 0.0001
5	0.6303	0.0061	2004	0.9939	< 0.0001	5	0.9758	< 0.0001
Mean	0.9103	0.0012	Mean	0.8066	0.0015	Mean	0.9228	0.0002
Marbled Godwit (GOF 0.1606)								
$X\hat{\beta} = -2.1833 - 0.0146\ \mathrm{CSM} + 3.2831\ \mathrm{ndvi} - 0.1221\ \mathrm{bright} - 0.2182\ \mathrm{wet} + 389.6775\ \mathrm{ndvi}_{var}$								
k-fold Spearman r_s values (Critical value for alpha 0.05: n=10, 0.564)								
Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9054	< 0.0001	2000	0.6497	0.0049	1	0.6497	0.0049
2	0.7038	0.0024	2001	0.7723	0.0008	2	0.9759	< 0.0001
3	0.3042	0.0984	2002	0.9521	< 0.0001	3	0.2626	0.1299
4	0.9521	< 0.0001	2003	0.9286	< 0.0001	4	0.6303	0.0061
5	0.9286	< 0.0001	2004	0.9521	< 0.0001	5	0.6694	0.0038
Mean	0.7588	0.0202	Mean	0.8510	0.0011	Mean	0.6376	0.0289

Table 2-5: Number of bird records by year and spatial block

Baird's Sparrow						
Year	Block 1	Block 2	Block 3	Block 4	Block 5	Total
2000	19	12	1	16	11	53
2001	2	3	0	0	1	5
2002	2	0	0	1	0	3
2003	9	7	0	32	23	71
2004	18	12	4	28	40	102
Marbled Godwit						
Year	Block 1	Block 2	Block 3	Block 4	Block 5	Total
2000	10	7	2	4	5	28
2001	17	11	1	20	12	61
2002	23	12	2	15	9	61
2003	12	5	1	11	10	39
2004	11	2	5	6	9	33

Baird's Sparrow is a species of mixed-grass or fescue grasslands (Green et al. 2002). Although locally abundant, this species is generally rare throughout its range. Its tendency to rapidly shift its distribution and abundance from year to year, make determining the status of this species difficult (Green et al. 2002, McGillivray and Semenchuk 1998). This species prefers areas of mixed grass with intermediate height and density, residual vegetation, and little or no scattered low shrubs. Although preferring native grasslands, this species will use non-native grasslands with similar structural factors to its preferred habitat (Dale et al. 1997, Owens and Myres 1973). In prairie Canada structural factors are shrub cover less than 20%, litter of 0.1–4cm deep, and average grass height between 10–30cm (Dale 1983). A series of models were drafted for this species consisting of different combinations of precipitation data, different vegetation indices, and vegetation index variance within the site survey area. This species' best model is summarised in table 2-4.

For Baird's Sparrow, the spatial blocks scored quite high, but temporally, much of the data was marginal. In 2003, the RPI value was low. The species almost disappeared from the study area in 2002 (Table 2-5). This information in conjunction with the weather pattern (section 2.2), shows how the model has difficulty predicting the sudden rise in numbers after the drought conditions ended. This information provides insight into the limitations of the model in regard to moisture cycles. Examination of random RPI values or the mean values for the random, temporal, and spatial folds can not provide this information. Practically, this model likely could be extended outside the study area to the entire NWA if conjoined with an appropriate field program to confirm and improve the model.

The Marbled Godwit nests in grasslands with low to moderate vegetative cover and is often observed in prairie wetlands, sloughs, and shallow lakes (McGillivray and Semenchuk 1998). It feeds on insects, aquatic plants, leeches, and small fish using its long bill to probe into the benthos (Gratto-Trevor 2000). Nests are usually within 300–350m from a wetland. Use of non-native habitats as well as wetlands or ephemeral wetlands is reported. Tall grass cover is avoided and this species prefers large blocks of grassland (Gratto-Trevor 2000). A series of models were drafted for this species consisting of different combinations of precipitation data, different vegetation indices, and vegetation index variance within the site survey area. This species' best model is summarised in table 2-4.

For the Marbled Godwit, temporally the model has a lower RPI value in 2000, but has higher values in other years. Spatially, the model is poor, utterly failing in block 3 and having low values in all but block 2. The failure in block 3 could relate to data sparsity and a general temporal pattern that is not synchronised with the other blocks or the general trend (Table 2-5). The pattern in block 4 is also in contrast to the other blocks and the general trends. These complexities however do not suggest bad field data, but rather, there are important predictor variables missing from the model. Given the species requires at least some wetland habitat, a measure of wetland location and permanency would have likely made a better predictor than our Landsat based wet variable. Overall, the model predicts well temporally, but is marginal spatially. From a practical point of view, this means the model could be used for temporal trends within the study area, but extending its use to the entire NWA could not be recommended.

In this comparative example, it is clear that general measures of model stability are insufficient. Further, examination of the individual spatial and temporal RPI values revealed difficulties in these dimensions that were not obvious from examining the mean RPI values.

To reiterate the message of Burnham and Anderson (2002), building models based on data dredging is a dangerous business, so we did not examine the temporal and spatial distribution of the bird records prior to model construction. Instead, these models were built upon the simple, but defensible, ecological premise about the spatial and temporal variation of vegetation detected and predicted with vegetation indices and climate data. Examining data in detail after the models were constructed, prompted by the three-way testing, provided understanding of model limitations. Our approach to model testing provides the means to better understand model limitations so that they can be applied within the parameters of those limitations and not beyond.

Table 2-6: Model Summary

Species	.5 Mean Spearman r_s values and RPI Fold Sort Scores (FSS)					GOF
(strong models)	Random	Temporal	FSS	Spatial	FSS	
Willet	0.6701	0.6275	0.8	0.5796	0.6	0.2202
Upland Sandpiper	0.7289	0.6661	0.6	0.5762	0.6	0.1062
Marbled Godwit	0.7422	0.8510	1.0	0.6376	0.8	0.1606
Sprague's Pipit	0.9579	0.9260	1.0	0.9137	1.0	0.6807
Clay-colored Sparrow	0.6874	0.6383	0.6	0.7283	0.8	0.5030
Vesper Sparrow	0.8512	0.7650	0.8	0.5852	0.8	0.2901
Lark Bunting	0.9717	0.3972	0.4	0.7208	0.8	0.1294
Savannah Sparrow	0.9288	0.7541	0.8	0.4374	0.4	0.8792
Baird's Sparrow	0.9103	0.8066	1.0	0.9228	1.0	0.6089
Grasshopper Sparrow	0.9763	0.9344	1.0	0.9212	1.0	0.8260
McCown's Longspur	0.8935	0.9469	1.0	0.8382	1.0	0.1430

2.4 Conclusion

The utility of RSF models depends crucially on their predictive ability. Prediction can vary in space, in time, and as a function of the stability of the model. Therefore, a measure of the predictive ability of an RSF in all of these dimensions is required. If properly tested, RSF models then can be applied within the bounds of their limitations and no further. Use of improperly tested RSF models could lead to serious errors in management, which in turn could have grave consequences for the species of concern. Therefore, we strongly recommend the application of this technique in all possible cases.

The comparative example of Baird's Sparrow and the Marbled Godwit clearly demonstrates the value of the 3-way approach in selecting the most robust model from a set of alternatives and in identifying the limitations and appropriate uses of those models. Our summary results (Table 2-6) show that the Baird's Sparrow / Marbled Godwit example was not unique, but representative of our entire study.

The use of normalised equal area bins was shown to be superior both empirically and theoretically to fixed or simple equal area bins. The addition of the moving-window average put forward by Hirzel et al. (2006), improved performance further and helped to ameliorate threshold dependency issues.

Managers and researchers employing RSF models need to evaluate the predictive ability of their models. Because models can vary temporally, spatially, as well as generally, the predictive ability of models must be evaluated in all these dimensions. Our method of using a 3-way k-fold cross-validation with the RPI method of Boyce et al. (2002) was demonstrated to be equal to the task. To our knowledge, no other method exists that can provide as much useful information or test the robustness of RSF models as thoroughly as has been demonstrated here.

Many researchers may have to write their own routines for doing custom k-fold cross-validations as this functionality is not available in all software packages. For researchers using S-plus or R, copies of our k-fold R scripts are available upon request to the corresponding author.

References

Bengio, Y., Grandvalet, Y., 2004. No unbiased estimator of the variance of k-fold cross-validation. Journal of Machine Learning Research 5, 1089–1105.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., Schmiegelow, F. K. A., 2002. Evaluating resource selection functions. Ecological Modelling 157, 281–300.

Burnham, K. P., Anderson, D. R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd Edition. Springer-Verlag New York, Inc.

Crist, E. P., Laurin, R., Cicone, R. C., 1986. Vegetation and soils information contained in transformed thematic mapper data. In: Proceedings of IGARSS' 86 Symposium. European Space Agency, Paris, pp. 1465–70, ref. ESA SP-254.

Dale, B. C., 1983. Habitat relationships of seven species of passerine birds in Last Mountain Lake, Saskatchewan. Master's thesis, Univeristy of Regina, Regina, SK.

Dale, B. C., Goosen, J. P., Taylor, P., 1999. Avifauna component report. In: Canadian Forces Base Suffield National Wildlife Area Wildlife Inventory. Canadian Wildlife Service, Edmonton, AB, unpublished report.

Dale, B. C., Martin, P., Taylor, P., 1997. Effects of hay management on grassland songbirds in Saskatchewan. Wildlife Society Bulletin 25, 616–626.

Diniz-Filho, J. A. F., Bini, L. M., Hawkins, B. A., 2003. Spatial autocorrelation and red herrings in geographical ecology. Global Ecology and Biogeography 12, 53–64.

Flanagan, L. B., Wever, L. A., Carlson, P. J., 2002. Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland. Global Change Biology 8, 599–615.

Gill, F., Poole, A. (Eds.), 1993–2004. The Birds of North America. The Birds of North America, Inc., Philadelphia, PA.

Gratto-Trevor, C. L., 2000. Marbled godwit (*Limosa fedoa*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 492. The Birds of North America, Inc., Philadelphia, P.A.

Green, M. T., Lowther, P. W., Jones, S. L., Davis, S. K., Dale, B. C., 2002. Baird's sparrow (*Ammodramus bairdii*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 638. The Birds of North America, Inc., Philadelphia, P.A.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling In press.

Huang, C., Wylie, B., Yang, L., Homer, C., Zylstra, G., 2001. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. Tech. rep., United States Geological Survey.

Johnson, C. J., Nielsen, S. E., Merrill, E. E., McDonald, T. E., Boyce, M. S., 2006. Resource selection functions based on use-availability data: theoretial motivation and evaluation methods. Journal of Wildlife Management In press.

Kauth, R. J., Thomas, G. S., 1976. The tasselled cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: Proceedings on the symposium on machine processing in remotely sensed data. West Lafayette, Indiana: LARS, Purdue University, pp. 4B–41–4B–51.

Lane, D. R., Coffin, D. P., Lauenroth, W. K., 2000. Changes in grassland canopy structure across a precipitation gradient. Journal of Vegetation Science 11, 359–368.

le Cessie, S., van Houwelingen, J. C., 1991. A goodness-of-fit test for binary regresssion models, based on smoothing methods. Biometrics 47 (4), 1267–1282.

Lennon, J., 1999. Resource selection functions: taking space seriously. Trends in Ecology and Evolution 14, 399–400.

Lennon, J., 2000. Red-shifts and red herrings in geographical ecology. Ecography 23, 101–113.

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., Erickson, W. P., 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies, 2nd Edition. Kluwer Academic Publishers, Inc., Boston.

Mason, C. H., Perreault, Jr, W. D., 1991. Collinearity, power, and interpretation of multiple regression. Journal of Marketing Research 28 (3), 268–280.

McGillivray, W. B., Semenchuk, G. P., 1998. The Federation of Alberta Naturalists Field Guide to Alberta Birds. Federation of Alberta Naturalists.

Nielsen, S. E., Boyce, M. S., Stenhouse, Gordon, B., Munro, R. H., 2002. Modeling Grizzly Bear habitats in the Yellowhead ecosystem of Alberta: Taking autocorrelation seriously. Ursus 13, 45–56.

Oesterheld, M., Loreti, J., Semmartin, M., Sala, O. E., 2001. Inter-annual variation in primary production of semi-arid grassland related to previous-year production. Journal of Vegetation Science 12, 137–142.

Owens, R. A., Myres, M. T., 1973. Effects of agriculture upon populations of native passerine birds in an alberta fescue grassland. Canadian Journal of Zoology 51, 697–713.

Price, K. P., Guo, X., Stiles, J. M., 2002. Optimal landsat tm band combinations and vegetation indices for discrimination of six grassland types in eastern kansas. International Journal of Remote Sensing 23 (23), 5031–5042.

Qi, J., Chebouni, A., Huete, A. R., Kerr, Y. H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48, 119–126.

R Development Core Team, 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL `http://www.R-project.org`

Wagner, S. F., 1992. Introduction to Statistics. Harpercollins College Outline Series. HapperCollins.

Williams, G. D. V., Robertson, G. W., 1965. Estimating most probable prairie wheat production from precipitation data. Canadian Journal of Plant Science 45, 34–47.

3 Predictors of grassland bird habitats in Southeastern Alberta, Canada [2]

3.1 Introduction

Throughout ornithological literature, description of habitats as plant species associations has been, and remains common (Gill and Poole 1993–2004). The importance of structural elements, however, has been recognised for some time (MacArthur and MacArthur 1961) with the development of the concept for grassland birds first gaining full expression in the work of Wiens (1969). Wiens described the habitats of grassland birds using the height and density of plant cover in various life form categories. Since then many studies, including a number in prairie Canada (Dale 1983, Davis et al. 1999, McMaster and Davis 2001) have utilised the Wiens method to describe bird habitats. Associations were found between individual grassland bird species and litter depth or distribution, vegetation height and thickness, amount of residual cover, and the presence or absence of shrub.

Using these established relationships to create models useful for predicting bird occurrence suffers from several limitations. The first is that, although a given grassland bird species might be consistent in using areas with low cover or little litter, the absolute values of the structural measures in areas occupied often varies widely within a season or between years (Dale 1983). The birds appear to make decisions on a relative rather than absolute basis within certain extremes of unacceptably high or low values. This inconsistency in structural values makes creation of models that are temporally robust a challenge. In addition to the variability in the environment and in the behaviour of the birds in question, sampling many of these structural measures requires detailed field work. For management purposes, this is not practical. Our purpose was to assemble a set of robust models for use in the management of priority species in the Canadian Forces Base (CFB) Suffield National Wildlife Area (Suffield NWA) using remotely sensed and GIS based data as proxies for these structural measures.

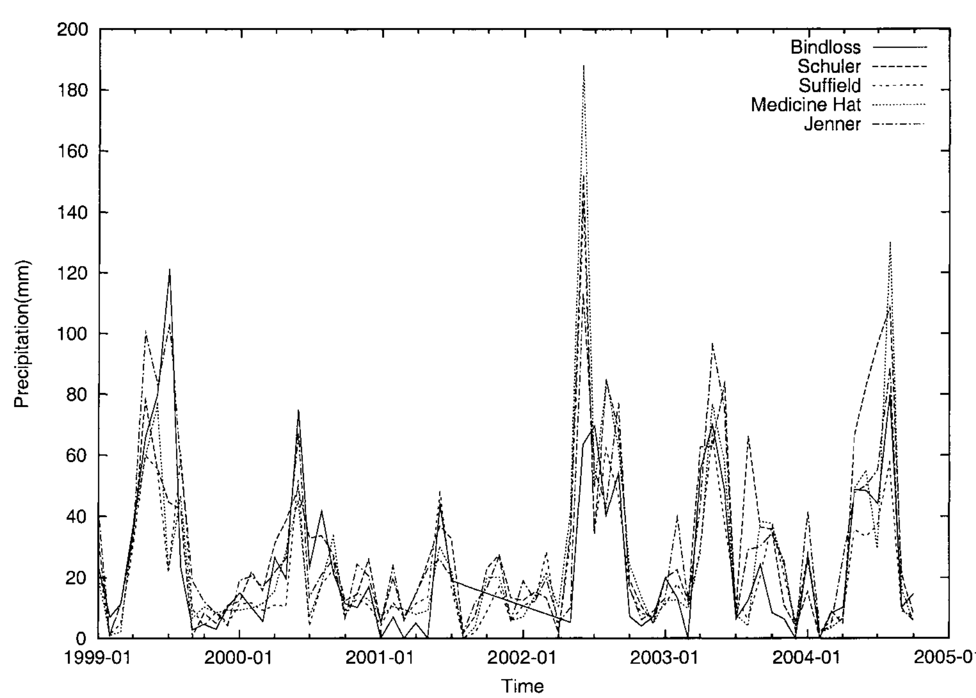
Manly et al. (2002) defines a resource selection function (RSF) as a function proportional to the probability of use of resource unit by an organism. Models having spatial

[2] A version of this chapter is being prepared for submission and review as: Wiens, T.S., Dale, B.C., Kershaw, G.P., Boyce, M.S., 2006. Predictors of grassland bird habitats. Ecological Applications

and temporal components will vary in space and time, and as a function of model stability. Since the utility of these models is found in their ability to predict, we applied the method outlined in (Chapter 2) to evaluate and select models based on their predictive ability. In this method k-fold cross-validation is employed to generate index values, or RSF Plot Index (RPI) scores, for random, temporal, and spatial folds which are indicators of the models predictive ability generally, temporally, and spatially. RPI values for fixed temporal and spatial k-folds are then used to assign a value between 0 and 1 describing how many of the folds passed the critical value for the RPI test statistic (in this case the Spearman-rank correlation test). In this way, models with the best predictive ability can be selected, and the predictive limitations of the best models can be clearly understood.

Using five years of point-count survey data we modelled habitat selection for eleven species of grassland birds in the Suffield NWA, in south-eastern Alberta, Canada (50°14′N, 110°37′W). Using the 3-way RPI method, we evaluated several variables known to affect the spatial distribution and growth of plants from year-to-year, and their ability to predict habitat selection for grassland birds.

Figure 3-1: Monthly precipitation from weather stations within 60km radius of study area.



3.2 Methods

During our study there was large variation in precipitation (Figure 3-1). In 2001, precipitation was low creating drought conditions in the area. In 2002, precipitation was low during the first part of the year and higher than normal in late spring as field surveys were being completed. This precipitation was not enough to offset the previous year's deficit so drought-like conditions persisted. Other years were much closer to normal precipitation years. To take advantage of these variable conditions, we selected bird species known to be sensitive to year-to-year variation in plant growth (Gill and Poole 1993–2004) or that we observed to vary in abundance and distribution over the course of our study. We also avoided species that were ubiquitous or nearly so, as well as those that were extremely rare in our study area (Table 3-1).

Table 3-1: Bird species names (American Ornithologist's Union) and occurrence

Common Name	Scientifi c Name	Percent Occurrence over all Survey Sites
Willet	*Catoptrophorus semipalmatus*	35.36
Upland Sandpiper	*Bartramia longicauda*	36.12
Marbled Godwit	*Limosa fedoa*	55.13
Sprague's Pipit	*Anthus spragueii*	80.23
Clay-colored Sparrow	*Spizella pallida*	10.65
Vesper Sparrow	*Pooecetes gramineus*	55.13
Lark Bunting	*Calamospiza melanocorys*	47.53
Savannah Sparrow	*Passerculus sandwichensis*	69.20
Baird's Sparrow	*Ammodramus bairdii*	55.51
Grasshopper Sparrow	*Ammodramus savannarum*	77.95
McCown's Longspur	*Calcarius mccownii*	53.61

3.2.0.1 Bird Surveys

During late May and early June of 2000–2004, bird data were collected using the point-count procedure outlined in Dale et al. (1999). In total there were 263 survey sites systematically placed to give a representative sample of the entire study area. Every 500m along east/west transects spaced 1 km apart, birds were recorded by song or sight inside or outside of a 100m radius during a sampling period of 5 minutes. Time of day, observer, and time of year bias were minimised through planned variation of the sample order. Bird numbers and inside or outside the 100m radius, or in-transit observations were recorded. Models used all observations associated with each survey site.

Table 3-2: Predictor variables

Variable Group	Variable	Source	Rationale or Comment
Temporal	$precip_t$	Environment Canada	Current year precipitation from January 1 to May 31
	$precip_{t-1}$	Environment Canada	Previous year precipitation
	$precip_{t-2}$	Environment Canada	Two years previous precipitation
	CSM	Environment Canada	Weighted average of past 2 years of precipitation
	$MSAVI_2$	MODIS	Reflectance of green vegetation at survey time each year
	NDVI	MODIS	Reflectance of green vegetation at survey time each year
	EVI	MODIS	Reflectance of green vegetation at survey time each year
Vegetation	LsNDVI	Landsat 7 TM	Reflectance of green vegetation
	$LsMSAVI_2$	Landsat 7 TM	Reflectance of green vegetation
	GVI	Landsat 7 TM	Reflectance of green vegetation
	bright	Landsat 7 TM	Brightness is a measure of reflectance mostly attributable to soil
	wet	Landsat 7 TM	Wetness is a measure of moisture in vegetation and soil.
	vegetation index variance	Landsat 7 TM	Variance within the 250 m radius survey location might be an indicator of habitats heterogeneity.
Landscape	CTI	GeoBase DEM	Compound Topographic Index provides a indicator of a locations position on a slope
	soil	AGRASID	A relative index derived from a weighted average soil texture of each soil horizon in profile.

3.2.1 Model Components

Topographic and climatic factors affect the spatial distribution and year-to-year variation in vegetation characteristics known to be important to grassland birds. We categorised our model components into temporal, vegetation, and landscape groups (Table 3-2). Temporal components consisted of climate and annual remote sensing images from the time of the bird surveys. The vegetation and landscape variables constituted the spatial component of our models.

We first constructed simple models using precipitation data and vegetation indices from

Landsat 7 TM data (30m pixels) as our temporal and spatial measures respectively. We subsequently compared precipitation against Moderate Resolution Imaging Spectroradiometer (MODIS) data (250m pixels) to determine which data source was a better predictor of the temporal variation of our study species. Last, we added the spatial variables related to landscape into our top model groups for each species to assess if this additional information would yield stronger models. Survey locations were spaced 500m apart so buffers with a radius of 250m were used to avoid overlap. Mean or variance values of vegetation indices and landscape variables within these buffers were used.

3.2.1.1 Temporal

In large-scale studies of vegetation response to moisture, grasslands have been shown to be highly variable (Huxman et al. 2004). Factors contributing to this variation include precipitation, temperature, wind, and solar radiation. These data have been used with soils, topography, and species associations to estimate the rate of potential evapotranspiration, actual evapotranspiration, and ultimately changes in soil moisture and plant available water. Oesterheld et al. (2001) studied the effect of previous-year precipitation and above-ground net primary productivity (ANPP) on current year ANPP. Current-year precipitation generally accounted for about 40% of the variance in ANPP. When they added previous-year ANPP or previous-year precipitation into their regression models, r^2 values to 0.58 and 0.60 respectively were obtained. Smoliak (1986) analysed a 50-year-record from southeastern Alberta and calculated an r^2 values of 0.54 for annual precipitation vs ANPP. The authors point out that nutrient availability is limited by moisture, which also can limit decomposition. These combined effects along with structural changes in response to moisture availability in a previous year could have some effect on current year ANPP not directly related to conserved soil moisture.

Flanagan et al. (2002) examined inter-annual variation in plant growth. They found that although precipitation was a major contributing factor to inter-annual variation, the relationship was not linear; wet years produced a much greater effect on increased production than dry years affected decreased production. Lane et al. (2000) examined the relationship between ANPP and grassland canopy structure. Increased ANPP was positively correlated with increases in canopy height as well as Leaf Area Index. Because vegetation height and thickness are commonly useful predictors for grassland birds, we used readily available precipitation data as a surrogate for ANPP.

Research into precipitation interpolation suggested that surface interpolation of monthly data was not problematic (Shen et al. 2001). Based on the literature about ANPP, it seemed that precipitation over the course of the year would be all that was needed. Since however we were interested in the vegetation state at the beginning of June we used precipitation in the current year up to the first of June which coincided with our field surveys. Upon consideration of these facts and the relatively small size of our study area, precipitation was based on a regional average from the weather stations within a 60km radius of the Suffield NWA. Precipitation data was obtained from Environment Canada for the climate stations in the region.

Precipitation data were included in models in three ways. First we used Conserved Soil Moisture (CSM) (Williams and Robertson 1965); an estimate of soil moisture on May 1^{st} using a weighted combination of precipitation data from the previous two years (Equation 3.1). Second we used the current and previous year's precipitation. Third, we used precipitation from the current year, previous year, and two years previous. These calculations resulted in common values for all point-count locations for each survey year.

$$\begin{aligned} CSM &= 0.36A+(0.37B-0.2(0.36A))+0.13C+ \\ &\quad (0.30D-0.2(0.36A+(0.37B-0.2(0.36A))+0.13C)) \end{aligned} \tag{3.1}$$

Where:
A = total precipitation during August, Sept. Oct in year t-2
B = total precipitation during November of t-2 to April of t-1
C = total precipitation during May through October of t-1
D = total precipitation during November of t-1 through April of year t

Another method of measuring year-to-year variation due to climate was through the use of yearly remotely sensed data that coincided with the bird survey from MODIS. The MODIS 16-day vegetation index composite (data product MOD13A1 IV) tiles for the Suffield region were downloaded for 2000–2004 for the end of May and the beginning of June. Although problems with the MODIS leaf area index algorithm have been identified (Shabanov et al. 2005) this did not affect our use of the reflectance and vegetation index products. Rahman et al. (2005) evaluated the use of MODIS Enhanced Vegetation Index (EVI) as a measure of gross primary production (GPP) and found it to be useful. Problems with MODIS related to our use, appear to be limited to registration variation over multiple scenes (Kawamura et al. 2005) which are usually less than 1 pixel, but up to 2 pixels. Since values used in the model were mean and variance measures of an area surrounding individual survey locations and would include multiple pixels, we did not smooth or resample the MODIS imagery to account for this potential error.

The MODIS images were tiled and reprojected using the MODIS Reprojection Tool (MRT) (U.S. Geological Survey 2004) and saved in GeoTIFF format. This imagery was then imported into GRASS 6.0. Image quality was assessed using the `r.bitpattern` module to extract information in the provided bit pattern layers on the Normalized Difference Vegetation Index (NDVI) and EVI. All pixels were designated as useable, so no data were removed at this point. In addition to the NDVI and EVI vegetation indices, the Modified

Soil Adjusted Vegetation Index 2 ($MSAVI_2$) (Qi et al. 1994) was calculated. $MSAVI_2$ was selected due its demonstrated good performance in areas with low plant cover.

3.2.1.2 Vegetation

Satellite reflectance is commonly used to generate classification maps based on field surveys or within-image similarities. This technique allows researchers to construct *a priori* models based on previously observed habitats correlations. Plant communities are abstractions of reality and thus represent artificial groupings across a gradient of biological reality. Useful as the classification approach can be, it is our opinion that birds select areas for breeding from the range of existing heterogeneity that exists in the real world at any given time. In consideration of this, we felt it was more appropriate to model bird habitats selection using the full range of information available to us. Therefore, we used vegetation index values as proxies for plant cover, bare ground, and moisture. If successful, this approach could provide a starting point for detailed field surveys to determine the physical factors measured by the vegetation indices. From a practical perspective however, this approach provides a working model for identifying areas of potential use by study species, without reliance on possibly irrelevant habitats classifications.

Price et al. (2002) evaluated which Landsat TM band combinations and times of year were optimal for discriminating between different grassland plant communities in tallgrass prairie in Kansas. They found that for single image employing standard vegetation indices, an image from July performed much better than one from May or September. Because we wanted to characterise our landscape with a single image, we used a Landsat 7 ETM+ source image from 29 July 2000. During our study period, 2000 was close to a normal precipitation year, so the image was likely to provide a reasonable indicator of the spatial variation of growth at a 30m spatial resolution for the study area. NDVI and $MSAVI_2$ vegetation indices (VIs) as well as the tasselled-cap Brightness, GVI or Green Vegetation Index (GVI), and Wetness indices (Crist et al. 1986, Huang et al. 2001, Kauth and Thomas 1976) were calculated. NDVI, $MSAVI_2$, and GVI were thus used as proxies for grass cover and density. Brightness was used as a proxy for percent bare soil cover, which as we had no measure of litter, acted as its inverse. Wetness was used to provide information on spatial variation in soil moisture conditions.

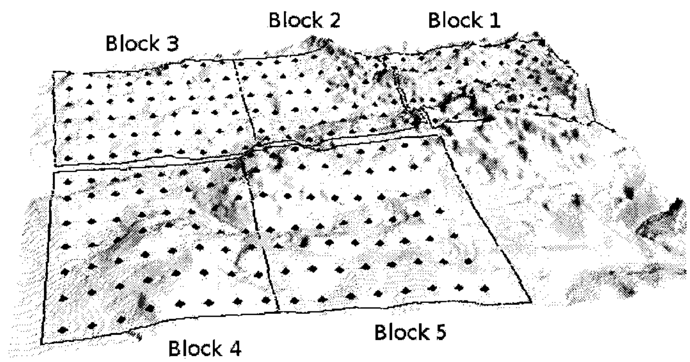
When predictor variables are strongly correlated, the standard errors of the regression coefficients can be inflated (Harrell 2001). This collinearity can then make it difficult to interpret regression coefficients but does not affect the joint influence of highly correlated variables (Harrell 2001). Because the different vegetation indices have different sensitivity to soil and moisture effects, where possible we wanted to use multiple indices to better describe the spatial variation of habitats. A correlation matrix for NDVI, $MSAVI_2$, GVI, Brightness, and Wetness for the entire Landsat scene was produced. A conservative value of lower than 0.7 acceptable correlation between pairs (Mason and Perreault 1991) of variables was used because this data was all derived from a single source and was thus correlated. Acceptable combinations were $MSAVI_2$ with GVI and Wetness and NDVI or $MSAVI_2$ with Brightness and Wetness. The theoretical foundation for these combinations relates to the difference among these indices. $MSAVI_2$, like the Soil Adjusted Vegetation Index (SAVI), was designed to reduce soil variation effects in the vegetation index signal. $MSAVI_2$ is very similar in performance to SAVI (Qi et al. 1994) and is in fact a self calibrating version of SAVI. Many other vegetation indices like NDVI and GVI have been demonstrated to vary more with the soil substrate than SAVI (Huete 1989) and thus $MSAVI_2$ also. Therefore, combining $MSAVI_2$ with GVI provides a means to identify areas of variation in soil and vegetation reflectance. Brightness and Wetness are representative of soil reflectance and moisture. These combinations were included in model comparisons with index mean and variance for each of the vegetation indices. Spatial heterogeneity is a factor of known importance to many grassland birds (Gill and Poole 1993–2004, Wiens 1974). Therefore, variance of $MSAVI_2$, GVI, or NDVI within the 250m radius buffer around each survey site was included in models for some species as a measure of spatial heterogeneity.

3.2.1.3 Landscape

Topography plays an important role in moisture availability in a semi-arid landscape and can influence the occurrence and productivity of plants (Bragg and Steuter 1996). Soil texture influences drainage and also is likely to influence plant growth. Both general discussions (Scott 1995) and classification system of grassland ecosystems (Adams et al. 2004) included topographic and soil texture elements in their descriptions. Based on this knowledge we included these elements in our models to enhance spatial discrimination. Landscape components consisted of a Digital Elevation Model (DEM) derived attribute and soil texture.

The DEM data from GeoBase (Natural Resources Canada 2004) with an approximate resolution of 25m was downloaded, merged, and reprojected to make a layer for CFB Suffield at 30m resolution. The challenge with using topographic features in the prairies is that they are subtle. However, within the study area, there was considerable variation in topography (Figure 3-2). Simple attributes such as slope and elevation do not provide information about where a location is in relationship to other elements. In simple terms, these variables can't inform the model if the point is at the top of a hill or in a drainage channel. We considered the use of morphometric features, such as planar, pit, channel, etc, but because such classifications are dependent on an arbitrary window size for calculation, it was not pursued. Instead, we calculated the Compound Topographic Index (CTI) (Equation 3.2) described by Gessler et al. (2000) using `r.topidx`. The original application of the

Figure 3-2: Digital elevation model view of study area with survey points and spatial blocks used for k-fold cross-validation shown on surface



CTI was in hydrological modelling and was called the wetness index which was renamed and applied to soil science by Gessler et al. (1995). The CTI is a measure of position on a slope, with lower values indicating position near the top of a slope, middle values indicating a position in the middle of a slope and high values indicating places near the bottom of a slope.

$$CTI = \ln(\frac{A_s}{\tan\beta}) \tag{3.2}$$

where:
A_s = specific catchment area (area per unit width orthogonal to flow direction)
β = slope angle

CFB Suffield provided a 1:50,000 soils map following the Agricultural Region of Alberta Soil Inventory Database (AGRASID) (Alberta Soil Information Centre 2001) conventions. The AGRASID data has lab-estimated Available Water Capacity (AWC) values, which are crucial for soil moisture tracking, vegetation growth modelling, or the calculation of drought indices. Research by De Jong and Loebel (1982), Oosterveld and Chang (1980) has shown that if calibrated locally, soil texture can be used to calculate AWC for a soil. Thus, we used soil texture as a component that was indicative of soil moisture available for plant growth. To simplify, a relative soil texture or particle size index was created using a weighted mean by profile contribution to the entire horizon. This relative soil texture index was calculated for each soil type with small values indicating fine texture or small particles and large values indicating course texture or large particles.

3.2.2 Data Processing

All data were managed using PostgreSQL 8.0. Raster and vector analyses were conducted using the Geographic Resources Analysis Support System (GRASS) version 6.0. Statistical analysis was conducted using R (R Development Core Team 2004) and the *Rpy* and *pyPgSQL* libraries to enable Python 2.4 scripting of R processes and the production of $\LaTeX\,2_\varepsilon$ documentation. Raster attributes were extracted within a 250m radius buffer around survey sites using the `v.sample.buffer` module which can calculate maximum, minimum, mean, variance, diversity, and mode and save the attributes into a PostgreSQL database.

3.2.3 Statistical Methods

The number of individuals observed at a single location was small, so the modelling of abundance would be difficult. Therefore, data were reduced to recorded or not recorded and logistic regression was used to model habitat selection employing a use / non-use study design Manly et al. (2002).

Although the application of k-fold cross-validation was developed to deal specifically with use / available data and study designs (Boyce et al. 2002), its use with use / non-use study designs is not precluded. In fact, since the utility of RSF functions depends on their predictive ability, the application of the 3-way RPI method is recommended (Chapter 2).

The RPI method is based on the application of k-fold cross-validation. Individual k-folds are used to measure the relationship between predicted likelihoods and observed data. As a general measure, we used five equally sized random k-folds. The five years of field data was used for temporal k-folds. The study area was divided into five geographically-distinct blocks (Figure 3-2). Once divided, four parts were used to predict the fifth, which was then repeated five times removing a different part each time. These data were grouped into 10 bins, or suitability classes for each k-fold. From these data, Spearman-rank correlation was used to compare the relationship between predicted and observed. Both individual and average values were recorded for random, temporal, and spatial folds. Individual folds from the temporal and spatial folds were compared against the critical value for Spearman r_s for which the relationship is considered to be higher than can be attributed to random chance; for example, with n=10, and an alpha of 0.05, the critical value of r_s is 0.564 (Wagner 1992). The number of folds passing this test was divided by the total number of folds to give an index between 1.0 (for all) to 0.0 (for none). Therefore, each model had a fold sort score (FSS) based on the combined performance of the spatial and temporal k-fold RPI values. Concurrent with this procedure, the goodness-of-fit (GOF) test of le Cessie and van Houwelingen (1991), as implemented in the *Design* library for R or S-Plus, was conducted and models failing this test ($P \leq 0.10$) were excluded. This test was chosen over the commonly employed Hosmer-Lemeshow GOF test because in their review of a series of GOF tests for logistic regression, Hosmer et al. (1997) concluded that test of le Cessie and van Houwelingen (1991) was as good as any they examined and did not suffer from some of the theoretical problems of the Hosmer-Lemeshow test. This, combined with the fact that the le Cessie and van Houwelingen (1991) test was readily available in R and the Hosmer-Lemeshow test was not, made the choice obvious.

After all models were estimated, a single best or group of top models were chosen for each species. This choice was made both on RPI scores and through examination of the model beta values in relation to their predictor variables. Models that were ecologically unreasonable were discarded. Residuals plots and Cook's distance plots were examined to identify data points exerting an inordinate influence on the model or models. Prior to modelling, the source data had been examined for errors and outliers, yet all models had some data points identified as overly influential. These potential offending data points ($<$ 2%) were removed and the top candidate models were re-run to evaluate their performance. If the resulting models were similar in performance after the removal of these questionable data, the effects of the removed data points were classified as non-problematic. If however, the models changed substantially in their predictive ability, they were excluded from consideration. In this fashion, we could examine the top candidate models thoroughly and select a single best model for each species.

3.2.4 Limitations

In modern computing environments, when researchers are presented with a large number of possible models, the temptation exists to try them all. Burnham and Anderson (2002) strongly argue against this practice and we agree that this is an approach that is likely to cause more problems than it will solve. Determining how best to combine predictor variables based on *a priori* assumptions was difficult in this study due to their general and somewhat ambiguous nature compared to the specific structural values informing our choices and the different influence of these variables on individual bird species. For instance, similar combinations of precipitation and vegetation variables could conceivably model a number of different species depending on model beta weights.

Vegetation indices are built upon the differences in the strength of reflectance of materials at different spectral frequencies. They do not however, provide a means to deal with the fact that individual pixels are composed of heterogeneous landscapes. As a result, all vegetation indices are sensitive to variations in substrate soils and moisture to varying degrees (Huete 1989). This limitation has stimulated many research efforts to explore the use of spectral unmixing as a way of providing better detection of vegetation and other landscape features. Despite the limitations of vegetation indices, they can still be of great utility in habitats selection modelling if their lack of precision is understood.

Making ecological inference from our models is unwise because of the limitations of our approach and the predictor variables used. Our goal however, was to create predictive models for management and conservation purposes within the NWA and surrounding areas. For other researchers and land managers, we expect that similar models will also be found to be effective, yet we recommend caution in drawing conclusions about a species biology based on models using these types of criteria. A summary of the birds' ecology is provided (Table 3-5) to assist in the understanding of the best models presented in table 3-6 and in understanding the limitations of the predictor variables used.

Table 3-3: Top MODIS and precipitation models. Scores listed are Mean RSF k-fold index values (Mean RPI) which is a Spearman-rank correlation value indicating predictive ability, Fold Sort Scores (FSS) which are the proportion of folds with scores higher than random, and the le Cessie and van Houwelingen (1991) goodness-of-fit (GOF) test for logistic regression.

Species	MODIS and Landsat			Precipitation and Landsat		
	Mean RPI	FSS	GOF	Mean RPI	FSS	GOF
Willet	0.8488	0.6	0.2472	0.6701	0.7	0.2202
Upland Sandpiper	0.8879	1.0	0.1016	0.7289	0.7	0.1062
Marbled Godwit	0.7541	0.9	0.6542	0.7422	0.9	0.1606
Sprague's Pipit	0.9143	0.8	0.1085	0.9579	1.0	0.6807
Clay-colored Sparrow	0.6890	0.7	0.2048	0.6874	0.7	0.5030
Vesper Sparrow	0.8353	0.7	0.2239	0.8512	0.9	0.2900
Lark Bunting	n/a[1]	n/a	n/a	0.9717	0.8	0.1294
Savannah Sparrow	0.7832	0.7	0.3289	0.9288	0.6	0.8792
Baird's Sparrow	0.9703	0.9	0.7444	0.9510	1.0	0.6089
Grasshopper Sparrow	0.9196	0.9	0.1106	0.9763	1.0	0.8260
McCown's Longspur	0.8444	0.8	0.3127	0.8935	1.0	0.1430

[1]All models failed GOF test indicating the predicted model output differed significantly from the observed values

3.3 Results and Discussion

3.3.1 Annual remote sensing

For the MODIS set of models, these were divided into 4 groups of 6 models each giving a total of 24 possible models. This MODIS model set was compared against our climate-based models which consisted of different combinations of vegetation indices and precipitation data which were also divided into 4 groups of 6 giving a total of 24 models. The relative performance by species is listed with Mean RPI, FSS, and GOF values (Table 3-3). Examination of this table reveals that for most species, climate-based models had better predictive capability than their MODIS counterparts. Detailed ground measurements of structure have found considerable year-to-year variation in the values associated with a given species (Dale 1983). Although MODIS is spatially crude (250m pixels), it did provide a spatially explicit temporal measure, unlike the precipitation data, so this result was somewhat surprising. For most species these additional spatial data appeared to have acted as confounding variables. For Upland Sandpiper and Savannah Sparrow however, the MODIS-based models were superior.

3.3.2 Landscape variables

Soil texture and CTI were added to the Landsat and precipitation models to determine if model performance could be enhanced. In the cases of Upland Sandpiper and Savannah Sparrow, MODIS and Landscape models were tested also. For many of the species, the changes were subtle and required detailed examination of the RPI values for each fold to assess which models were superior. However, in general, the inclusion of CTI and soil texture improved model performance (Table 3-4). For Clay-colored Sparrow there was a dramatic improvement with the inclusion of landscape variables. That is probably related to their dependence on shrubs for nesting. Woody vegetation occurrence could not be predicted by the vegetation indices utilised in this study, but slope position or drainage would provide some information in this regard.

Table 3-4: Top Landsat models with and without landscape variables. Scores listed are Mean RSF k-fold index values (Mean RPI) which is a Spearman-rank correlation value indicating predictive ability, Fold Sort Scores (FSS) which are the proportion of folds with scores higher than random, and the le Cessie and van Houwelingen (1991) goodness-of-fit (GOF) test for logistic regression.

Species	With Landscape			Without Landscape		
	Mean RPI	FSS	GOF	Mean RPI	FSS	GOF
Willet	0.8370	0.7	0.1143	0.6701	0.7	0.2202
Upland Sandpiper	0.8103	1.0	0.1318	0.8879	1.0	0.1016
Marbled Godwit	0.7113	0.9	0.7831	0.7422	0.9	0.1606
Sprague's Pipit	0.9720	1.0	0.5157	0.9579	1.0	0.6807
Clay-colored Sparrow	0.8581	0.8	0.7241	0.6874	0.7	0.5030
Vesper Sparrow	0.9431	0.9	0.6059	0.8512	0.9	0.2900
Lark Bunting	0.9107	0.7	0.1293	0.9717	0.8	0.1294
Savannah Sparrow	0.7471	0.8	0.9267	0.9288	0.6	0.8792
Baird's Sparrow	0.9855	0.9	0.4897	0.9510	1.0	0.6089
Grasshopper Sparrow	0.9855	1.0	0.8093	0.9763	1.0	0.8260
McCown's Longspur	0.9763	1.0	0.8763	0.8935	1.0	0.1430

Table 3-5: Bird Ecology Summary

Species	Cover	Moisture	Heterogeneity	Source
Willet	short, low cover	breed near wetlands	low	Lowther et al. (2001)
Upland Sandpiper	low to moderate grass, moderate to high litter	dry	3 types of habitats needed	Houston and Bowen (2001)
Marbled Godwit	low to moderate	ephemeral wetlands	low	Gratto-Trevor (2000)
Sprague's Pipit	intermediate height grasses, with moderate litter	well-drained areas	moderate	Robbins and Dale (1999)
Clay-colored Sparrow	shrubby grasslands	moderate to high	moderate	Knapton (1994)
Vesper Sparrow	moderate, some bare ground, and some shrub or edge	moderate	high	Jones and Cornely (2002)
Lark Bunting	moderate with shrub or sagebrush	moderate to dry	high	Shane (2000)
Savannah Sparrow	dense, some shrub	moderate	low	Wheelwright and Rising (1993)
Baird's Sparrow	moderate with residual vegetation	moderate	high	Green et al. (2002)
Grasshopper Sparrow	moderate with patchy bare ground	moderate to moist	moderate to high	Vickery (1996)
McCown's Longspur	very sparse	dry	homogeneous	With (1994)

3.3.3 Model Strengths and Weaknesses

Due to the somewhat inexact nature of the predictor variables, understanding the likely ecological causes for model failures and successes is not immediately obvious. We will use three species to examine the relative strength and weakness of these models and areas that are in need of improvement.

Sprague's Pipit prefers well-drained areas in open grasslands with grasses of intermediate height and thickness with moderate litter, but areas with shrubs are avoided (Robbins and Dale 1999). These habitat requirements are congruent with the observation that this bird is intolerant of heaving grazing. This bird breeds in native grasslands and tends to avoid tame pasture (McGillivray and Semenchuk 1998).

Table 3-6: Best models by species with mean RPI scores for random, temporal, and spatial k-folds and fold scort scores (FSS) indicating the proportion of folds with Spearman r_s values equal to or above critical value.

Species (strong models)	Mean RPI values and Fold Sort Scores (FSS)					GOF	Model (without beta weights values)
	Random	Temporal	FSS	Spatial	FSS		
Willet	0.5118	0.6275	0.8	0.5796	0.6	0.2202	$-a - b_1precip_t + b_2precip_{t-1} - b_3precip_{t-2} + b_4LsNDVI + b_5bright - b_6wet + b_7LsNDVI_{var}$
Upland Sandpiper	0.6327	0.7619	1.0	0.7447	1.0	0.1016	$a - b_1MSAVI_2 + b_2LsNDVI - b_3LsNDVI_{var}$
Marbled Godwit	0.7985	0.8510	1.0	0.6376	0.8	0.1606	$-a - b_1CSM + b_2LsNDVI - b_3bright - b_4wet + b_5LsNDVI_{var}$
Sprague's Pipit	0.9579	0.9260	1.0	0.9137	1.0	0.6807	$a - b_1precip_t + b_2precip_{t-1} - b_3GVI + b_4wet + b_5GVI_{var}$
Clay-colored Sparrow	0.7075	0.7184	0.8	0.7216	0.8	0.7241	$a + b_1CTI + b_2soil + b_3precip_t + b_4precip_{t-1} + b_5precip_{t-2} + b_6GVI + b_7wet + b_8GVI_{var}$
Vesper Sparrow	0.9475	0.9259	1.0	0.7609	0.8	0.6059	$a + b_1CTI + b_2soil - b_3CSM + b_4LsMSAVI_2 - b_5GVI + b_6wet - b_7LsMSAVI_2var$
Lark Bunting	0.8351	0.3350	0.4	0.8104	1.0	0.1293	$a - b_1CTI + b_2soil - b_3precip_t + b_4precip_{t-1} + b_5LsMSAVI_2 - b_6GVI + b_7wet - b_8LsMSAVI_2var$
Savannah Sparrow	0.7645	0.8686	1.0	0.4843	0.6	0.9267	$-a + b_1CTI - b_2soil + b_3MSAVI_2 - b_4LsMSAVI_2 + b_5GVI - b_6wet + b_7LsMSAVI_2var$
Baird's Sparrow	0.9103	0.8066	1.0	0.9228	1.0	0.6089	$-a - b_1precip_t + b_2precip_{t-1} - b_3LsNDVI + b_4bright - b_5wet + b_6LsNDVI_{var}$
Grasshopper Sparrow	1.0000	0.9334	1.0	0.9952	1.0	0.8093	$a + b_1CTI + b_2precip_t + b_3precip_{t-1} + b_4precip_{t-2} + b_5GVI - b_6GVI_{var}$
McCown's Longspur	0.8935	0.9469	1.0	0.8382	1.0	0.1430	$-a + b_1precip_t - b_2precip_{t-1} - b_3LsMSAVI_2 + b_4GVI - b_5wet$

The model for the Sprague's Pipit with the best predictive power consisted of five predictor variables. Current year precipitation was negative and previous year precipitation was positive. This would indicate that antecedent moisture conditions are very important. Also, this is logical considering this species need for moderate litter which is the product of the previous growing season. This is also congruent with the observed variation in bird number for this species. In 2001, which was the first drought year, the number of records for Sprague's Pipit dropped to 56% of the number observed in 2000 (Table 3-7). In 2002, as the drought continued, the numbers dropped to 22% of 2000 values. In 2003, with litter being affected by the summer precipitation in 2002, numbers rebounded to 68% of 2000 values and in 2004, the number of observations for Sprague's Pipit were 108% of 2000 values. This pattern clearly shows the litter and previous year production effect on the relative abundance of this species.

Table 3-7: Number of species records by year

Species	2000	2001	2002	2003	2004
Willet	22	35	35	9	10
Upland Sandpiper	13	42	26	32	21
Marbled Godwit	28	61	61	39	33
Sprague's Pipit	112	63	25	76	121
Clay-colored Sparrow	5	6	5	14	10
Vesper Sparrow	44	73	88	60	42
Lark Bunting	43	32	21	65	83
Savannah Sparrow	80	94	41	86	80
Baird's Sparrow	59	6	3	71	102
Grasshopper Sparrow	97	89	43	115	130
McCown's Longspur	7	77	122	38	30

GVI was weighted negatively and wetness positively, as well as GVI variance. As described earlier, GVI is known to be sensitive to soil effects (Huete 1989). Given this sensitivity, it would appear that the reflectance indices selected areas that are wet, but without the most vegetation, and areas with some spatial heterogeneity. If an area is wet, but doesn't have dense plant cover, this seems to indicate good drainage. These observations about the model are in agreement with the general ecology of this bird as it relates to grass cover, litter, and shrub. Although this model performed well, it clearly could benefit from a better measure of litter.

McCown's Longspur uses areas such as moderately grazed short-grass prairie, or structurally similar habitats such as overgrazed pastures (With 1994). This bird breeds in open areas with sparse plant cover. In Alberta its distribution is localised (McGillivray and Semenchuk 1998). Structural studies report a preference of about 25% bare ground and an average plant height of 5 cm, with nests usually placed beside clumps of grass or beneath shrubs (With

1994).

The best predictive model for McCown's Longspur has five predictor variables. Current year precipitation is weighted positively and previous year precipitation is weighted negatively. This combination worked well in this model to accurately tracking the rise and fall of this bird's relative abundance during and after the drought (Table 3-7). This pattern is expected with the documented affinity of this species for dry and sparsely vegetated habitats. The negative weighting of the wetness index is also in agreement with this. The combination of negative $MSAVI_2$ and positive GVI demonstrates avoidance of the greenest areas but selection of areas where the two indices disagree. GVI is more sensitive to soil effects than $MSAVI_2$ which suggests selection of areas without litter. Models for this bird which included brightness performed poorly suggesting that it was a poor proxy for litter. As with Sprague's Pipit, a better litter proxy would likely have improved this models performance.

The Grasshopper Sparrow prefers moderately open grasslands with areas of patchy bare ground. This species is easily overlooked and has a broad range with an uneven local distribution where it is often absent from seemingly suitable habitat (McGillivray and Semenchuk 1998). This bird tends to select areas that are more lush and have shrub cover, but avoids areas with extensive shrub (Vickery 1996). It forages exclusively on the ground for insects and seeds and thus requires sites with bare ground as it is a visual predator. Besides native prairie, it is also found in abandoned pastures and other tall-grass dominated areas in southeastern Alberta (McGillivray and Semenchuk 1998). Nests are domed with overhanging grasses and have a side entrance making them difficult to locate (Vickery 1996).

The best model for Grasshopper Sparrow has six predictor variables. CTI was weighted positively indicating a positive relationship to areas near the bottom of slopes, and thus wetter areas. Current, previous year, and two years previous precipitation were all weighted positively as would be expected of a species requiring moister conditions and which dropped to 50% of its 2000 observations in 2002, the second year of the drought. GVI was also positively weighted as expected, but GVI variance was negatively related. Because Grasshopper Sparrow is known to need areas with a spatial heterogeneity, the negative weighting of vegetation index variance raises doubts about its validity as a measure of spatial heterogeneity that is meaningful to this species or other grassland birds. Thus this model would have likely performed better with a better measure of spatial heterogeneity. To address this question, the use of imagery with finer spatial resolution is the next logical step.

The examination of these models clearly demonstrates the need for better measures of litter and spatial heterogeneity. Not discussed above were the models of Willet and Upland Sandpiper that both performed poorly spatially, probably due to the absence of a properly classified wetland layer. For species like these, that are dependent on wetlands at different times of the year or Marbled Godwit which is known to chose ephemeral wetlands (Gratto-Trevor 2000), this additional layer would be very useful.

On the positive side, the use of precipitation data was very useful and provided predictive capability that was easily understood in ecological terms. The ease of use over annual remote sensing imagery and superior performance provides some confidence in recommending its use for other grassland models with a temporal component. In addition to precipitation, soil texture and CTI had a positive effect on the performance of many models in a manner that was easy to understand, so they too can be recommended.

3.4 Conclusion

The models constructed for the eleven study species were based on the simple concept that changes in vegetation in space and time directly effect changes in bird distributions over space and time. We found that simple models based on vegetation indices and precipitation data were often quite usable, but the addition of soil texture and CTI contributed to better models in many cases. Examination of a series of models hightlighted the need for better measures of spatial heterogeneity, litter, and wetland types and distribution. The application of this approach outside grassland ecosystems is uncertain because many other ecosystems are slower to respond to year-to-year changes in climate.

Annual images from MODIS did not appear to be as useful for temporal predictors as precipitation. However since our study area was relatively small, it is likely that for larger areas, MODIS data could be a very powerful tool when combined with finer scale data such as Landsat 7 TM imagery, and landscape characteristics derived from DEM data.

References

Adams, B. W., Poulin-Klein, L., Moisey, D., McNeil, R., 2004. Rangeland plant communities and range health assessment guidelines for the dry mixedgrass natural subregion of Alberta, Pub. No. T/040 Edition. Rangeland Management Branch, Public Lands Division, Alberta Sustainable Resource Development, Lethbridge, Alberta.

Alberta Soil Information Centre, 2001. Agrasid30: Agricultural region of alberta soil inventory database. Edited by Brierley, J.A., Martin, T.C., and Spiess, D.J., agriculture and Agri-Food Canada, Research Branch; Alberta Agricultrue, Food and Rural Development, Conservation and Development Branch. Version 3.0.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., Schmiegelow, F. K. A., 2002. Evaluating resource selection functions. Ecological Modelling 157, 281–300.

Bragg, T. B., Steuter, A. A., 1996. Prairie ecology — the mixed prairie. In: Nelson, E. B., Samson, F. B., Knopf, F. L. (Eds.), Prairie Conservation. Island Press.

Burnham, K. P., Anderson, D. R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd Edition. Springer-Verlag New York, Inc.

Crist, E. P., Laurin, R., Cicone, R. C., 1986. Vegetation and soils information contained in transformed thematic mapper data. In: Proceedings of IGARSS' 86 Symposium. European Space Agency, Paris, pp. 1465–70, ref. ESA SP-254.

Dale, B. C., 1983. Habitat relationships of seven species of passerine birds in Last Mountain Lake, Saskatchewan. Master's thesis, Univeristy of Regina, Regina, SK.

Dale, B. C., Goosen, J. P., Taylor, P., 1999. Avifauna component report. In: Canadian Forces Base Suffield National Wildlife Area Wildlife Inventory. Canadian Wildlife Service, Edmonton, AB, unpublished report.

Davis, S., Duncan, D., Skeel, M., 1999. Distribution and habitat associations of three endemic grassland songbirds in southern Saskatchewan. Wilson Bulletin 111, 389–396.

De Jong, R., Loebel, K., 1982. Empirical relations between soil components and water retention at 1/3 and 15 atmospheres. Canadian Journal of Soil Science 62, 343–350.

Flanagan, L. B., Wever, L. A., Carlson, P. J., 2002. Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland. Global Change Biology 8, 599–615.

Gessler, P. E., Chadwick, O. A., Chamran, F., Althouse, L., Holmes, K., 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal 64 (6), 2046–2056.

Gessler, P. E., Moore, I. D., McKenzie, N. J., Ryan, P. J., 1995. Soil-landscape modelling and spatial precition of soil attributes. International Journal of Geographical Information Systems 9 (4), 421–432.

Gill, F., Poole, A. (Eds.), 1993–2004. The Birds of North America. The Birds of North America, Inc., Philadelphia, PA.

Gratto-Trevor, C. L., 2000. Marbled godwit (*Limosa fedoa*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 492. The Birds of North America, Inc., Philadelphia, P.A.

Green, M. T., Lowther, P. W., Jones, S. L., Davis, S. K., Dale, B. C., 2002. Baird's sparrow (*Ammodramus bairdii*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 638. The Birds of North America, Inc., Philadelphia, P.A.

Harrell, F. E., 2001. Regression modeling strategies. Springer, New York.

Hosmer, D. W., Hosmer, T., Le Cessie, S., Lemeshow, S., 1997. A comparison of goodness-of-fit tests for the logistic regression model. Statistics in Medicine 16, 965–980.

Houston, S. C., Bowen, Jr, D. E., 2001. Sprague's pipit (*Anthus spragueii*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 580. The Birds of North America, Inc., Philadelphia, P.A.

Huang, C., Wylie, B., Yang, L., Homer, C., Zylstra, G., 2001. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. Tech. rep., United States Geological Survey.

Huete, A., 1989. Soil influences in remotely sensed vegetation-canopy spectra. In: Asrar, G. (Ed.), Theory and Applications of Optical Remote Sensing. A Wiley-Interscience publication, pp. 107–141.

Huxman, T. E., Smith, M. D., Fay, P. A., Knapp, A. K., Shaw, M. R., Loik, M. E., Smith, S. D., Tissue, D. T., Zak, J. C., Weltzin, J. F., Pockman, W. T., Sala, O. E., Haddad, B. M., Harte, J., Koch, G. W., Schwinning, S., Small, E. E., Williams, D. G., 2004. Convergence across biomes to a common rain-use efficiency. Nature 429, 651–654.

Jones, S. L., Cornely, J. E., 2002. Vesper sparrow (*Pooecetes gramineus*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 624. The Birds of North America, Inc., Philadelphia, P.A.

Kauth, R. J., Thomas, G. S., 1976. The tasselled cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: Proceedings on the symposium on machine processing in remotely sensed data. West Lafayette, Indiana: LARS, Purdue University, pp. 4B–41–4B–51.

Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, G., Wang, S., 2005. Monitoring of forage conditions with MODIS imagery in the xilingol steppe, inner mongolia. International Journal of Remote Sensing 26 (7), 1423–1436.

Knapton, R. W., 1994. Clay-colored sparrow (*Spizella pallida*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 120. The Birds of North America, Inc., Philadelphia, P.A.

Lane, D. R., Coffin, D. P., Lauenroth, W. K., 2000. Changes in grassland canopy structure across a precipitation gradient. Journal of Vegetation Science 11, 359–368.

le Cessie, S., van Houwelingen, J. C., 1991. A goodness-of-fit test for binary regresssion models, based on smoothing methods. Biometrics 47 (4), 1267–1282.

Lowther, P. E., Douglas, III, H. D., Gratto-Trevor, C. L., 2001. Willet (*Catoptrophorus semipalmatus*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 579. The Birds of North America, Inc., Philadelphia, P.A.

MacArthur, R. H., MacArthur, J. W., 1961. On bird species diversity. Ecology 42 (3), 594–598.

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., Erickson, W. P., 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies, 2nd Edition. Kluwer Academic Publishers, Inc., Boston.

Mason, C. H., Perreault, Jr, W. D., 1991. Collinearity, power, and interpretation of multiple regression. Journal of Marketing Research 28 (3), 268–280.

McGillivray, W. B., Semenchuk, G. P., 1998. The Federation of Alberta Naturalists Field Guide to Alberta Birds. Federation of Alberta Naturalists.

McMaster, D., Davis, S., 2001. An evaluation of Canada's Permanent Cover Program: habitat for grassland birds? Journal of Field Ornithology 72, 195–210.

Natural Resources Canada, 2004. GeoBase Canadian Digital Elevation Data. URL `http://www.geobase.ca`

Oesterheld, M., Loreti, J., Semmartin, M., Sala, O. E., 2001. Inter-annual variation in primary production of semi-arid grassland related to previous-year production. Journal of Vegetation Science 12, 137–142.

Oosterveld, M., Chang, C., 1980. Empirical relations between laboratory determinations of soil texture and moisture retention. Canadian Agricultural Engineering 22 (2), 149–151.

Price, K. P., Guo, X., Stiles, J. M., 2002. Optimal landsat tm band combinations and vegetation indices for discrimination of six grassland types in eastern kansas. International Journal of Remote Sensing 23 (23), 5031–5042.

Qi, J., Chebouni, A., Huete, A. R., Kerr, Y. H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48, 119–126.

R Development Core Team, 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL `http://www.R-project.org`

Rahman, A. F., Sims, D. A., Cordova, V. D., El-Masri, B. Z., 2005. Potential of MODIS EVI and surface temperature for directly esitmating per-pixel C fluxes. Geophysical Research Letters 32, L19404.

Robbins, M. B., Dale, B. C., 1999. Sprague's pipit (*Anthus spragueii*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 439. The Birds of North America, Inc., Philadelphia, P.A.

Scott, G. A., 1995. Canada's vegetation: A world perspective. McGill-Queen's University Press.

Shabanov, N. V., Huang, D., Yang, W., Tan, B., Knyazikhin, Y., Myneni, R. B., Ahl, D. E., Gower, S. T., Huete, A. R., Eduardo, L., Aragã, O., Shimabukuro, Y. E., 2005. Analysis and optimization of the MODIS leaf area index algorithm retrievals over broadleaf forests. IEEE Transactions on Geoscience and Remote Sensing 43 (8), 1855–1865.

Shane, T. G., 2000. Lark bunting (*Calamospiza melanocorys*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 542. The Birds of North America, Inc., Philadelphia, P.A.

Shen, S. S., Dzikowski, P., Guilong, L., Griffith, D., 2001. Interpolation of 1961–97 daily temperature and precipitation data onto Alberta polygons of Ecodistrict and Soil Landscapes of Canada. Journal of Applied Meteorology 40, 2162–2177.

Smoliak, S., 1986. Influence of climatic condition on production of *Stipa-Bouteloua* prairie over a 50-year period. Journal of Range Management 39, 100–103.

U.S. Geological Survey, 2004. Modis reprojection tool. Electronic software available for download.

Vickery, P. D., 1996. Grasshopper sparrow (*Ammodramus savannarum*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 239. The Birds of North America, Inc., Philadelphia, P.A.

Wagner, S. F., 1992. Introduction to Statistics. Harpercollins College Outline Series. HapperCollins.

Wheelwright, N. T., Rising, J. D., 1993. Savannah sparrow (*Passerculus sandwichensis*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 45. The Birds of North America, Inc., Philadelphia, P.A.

Wiens, J., 1969. An approach to the study of ecological relationships among grassland birds. Ornithological Monographs (8), 93pp.

Wiens, J. A., 1974. Habitat heterogeneity and avian community structure in north american grasslands. American Midland Naturalist 91 (1), 195–213.

Williams, G. D. V., Robertson, G. W., 1965. Estimating most probable prairie wheat production from precipitation data. Canadian Journal of Plant Science 45, 34–47.

With, K. A., 1994. Mccown's longspur (*Calcarius mccownii*). In: Poole, A., Gill, F. (Eds.), The Birds of North America. No. 96. The Birds of North America, Inc., Philadelphia, P.A.

4 Application of research

4.1 Summary

In this thesis I developed a method for selection and robust evaluation of Resource Selection Function (RSF) models and applied it to the evaluation of remotely sensed and GIS-based data sources for modelling the habitat selection of grassland birds. Models were built using five years of point-count data. Low numbers of birds were observed at each survey location so the data was reduces to presense / absence and logistic regression was used to estimate an RSF. RSF's are defined as proportional to the probability of use of a resource unit by an organism (Manly et al. 2002). In this context, Boyce et al. (2002) put forward a method of using k-fold cross-validation to create predicted vs. observed line graphs using a set of suitability classes. Spearman-rank correlation tests were calculated on these results were summarised by an RSF Plot Index or RPI. This method allowed for the evaluation of the predictive ability of an RSF.

One challenging facet of this research was finding ways to ameliorate the RPI's threshold dependency. Threshold dependency is created through the selection of cut or break points for binning the results because this selection is arbitrary and can significantly affect the resulting RPI score. Hirzel et al. (2006) suggested using a moving window to create a pseudo-continuous function. This was shown to be a good concept, but failed to address variations in the predicted probability distributions for each fold. I demonstrated that a method using normalised scores within equal area bins was more stable and resulted in stronger RPI scores. To this method, I added the moving-window average to provide more consistent score values and help liberate the RPI method from its threshold-dependency problems.

The core of my work with the RPI method was applying it to show how model predictions varied in time and space. This was done by using temporal folds based on the five seasons of field study and spatial folds using five distinct spatial blocks within the study area. Examination of RPI values for individual spatial and temporal divisions proved to be an easy method of identifying potential weakness in models and understanding how they might be used appropriately within and beyond the spatial and temporal borders of a study. To my knowledge this is the first study of its kind and represents an advance in robust model testing methods.

Chapter 3 of this thesis examined the utility of remotely sensed and GIS data in creating habitat selection models for grassland birds. Remotely sensed data sources were Landsat 7 TM and MODIS. Additional data sources were DEM, a soils map, and precipitation data. A series of base models were constructed using vegetation indices derived from the Landsat 7 TM data and precipitation data to model the spatial and temporal variation in habitat respectively. Some of these simple models performed quite well. A second step was to compare the performance of annual remote sensing images from MODIS against precipitation data. Precipitation data performed better in most cases. Lastly, soil texture and CTI were added to the top model groups and yielded improved models for most species. The general results from this were that simple precipitation and single year vegetation index models can be effective for some species, but both soil texture and CTI are useful attributes to improve on these models. The study area was not large, so MODIS did not perform well in most cases. MODIS performance was good enough however to raise interest in its application in models for larger areas.

4.2 Future work

During the research in RPI model selection and evaluation, I observed a great deal of variance in the random k-folds. Recent work by Bengio and Grandvalet (2004) has demonstrated that there are no unbiased estimators of variance for k-fold cross-validation. Although the RPI method proved to be a useful tool for fixed temporal and spatial k-folds, the issue of variance among folds in cross-validation needs to be addressed.

The work on remotely sensed predictors for grassland bird habitat selection models clearly indicated that issues of scale need to be explored. There are two main issues. First, habitat heterogeneity is a fine-scale issue from the perspective of remote sensing, so although vegetation index variance was used as a proxy for this with some success, finer spatial resolution imagery is needed to capture this information. SPOT imagery (2.5m pixels) is now available for the Suffield region and this would be an excellent data source for evaluating this question. The other side of the scale question could be addressed with Breeding Bird Survey data and MODIS. Both data sources are freely available, which would make the perfect compliment to the SPOT work and would allow for exploration of scale dependency in our models, such as that conducted by Davidson and Csillag (2003).

Finally, further research is needed to evaluate the effectiveness of grazing information as a predictor variable. The distance-to-water index discussed in chapter 1, based on a properly classified GIS layer of wetlands and dugouts, would be a good starting point. Such a layer would improve models for species affected by grazing and would also be invaluable for some species, such as the Willet, Upland Sandpiper, or Marbled Godwit, which are known to have strong associations with wetlands at different times during their breeding cycle.

References

Bengio, Y., Grandvalet, Y., 2004. No unbiased estimator of the variance of k-fold cross-validation. Journal of Machine Learning Research 5, 1089–1105.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., Schmiegelow, F. K. A., 2002. Evaluating resource selection functions. Ecological Modelling 157, 281–300.

Davidson, A., Csillag, F., 2003. A comparison of three approaches for predicting C4 species cover of northern mixed grass prairie. Remote Sensing of Environment 86, 70–82.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling In press.

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., Erickson, W. P., 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies, 2nd Edition. Kluwer Academic Publishers, Inc., Boston.

A Top models

In this appendix the details on the single best model for each species are presented. The format of this presentation is one species per page with the common name, the model formula, Goodness-of-fit statistics, Wald statistics, and RPI statistics.

A.1 Willet

$$X\hat{\beta} = -10.1597 - 0.0223\,\mathrm{precip_t} + 0.0019\,\mathrm{precip_{t-1}} - 0.0042\,\mathrm{precip_{t-2}} + 6.0487\,\mathrm{Lsndvi} + 0.0378\,\mathrm{bright} - 0.0454\,\mathrm{wet} + 132.0096\,\mathrm{Lsndvi_{var}}$$

Table A-1: Willet goodness of fit

Sum of squared errors	95.7601
Expected value—H0	95.2682
SD	0.4012
Z	1.2260
P	0.2201

Table A-2: Willet Wald statistics

	χ^2	$d.f.$	P
precip_t	7.60	1	0.0058
precip_{t-1}	0.53	1	0.4647
precip_{t-2}	2.26	1	0.1331
Lsndvi	0.31	1	0.5749
bright	0.78	1	0.3760
wet	1.07	1	0.3001
Lsndvi_{var}	1.31	1	0.2525
TOTAL	44.96	7	< 0.0001

Table A-3: Willet RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.5028	0.0217	2000	0.8788	< 0.0001	1	0.3747	0.0600
2	0.8116	0.0004	2001	0.9225	< 0.0001	2	0.7937	0.0005
3	0.9698	< 0.0001	2002	0.6636	0.0041	3	0.2273	0.1635
4	0.2066	0.1869	2003	0.0121	0.7625	4	0.7723	0.0008
5	0.0679	0.4671	2004	0.6606	0.0043	5	0.7302	0.0016
Mean	0.5118	0.1352	Mean	0.6275	0.1542	Mean	0.5796	0.0453

A.2 Upland Sandpiper

$$X\hat{\beta} = 10.5216 - 3.2369\,\mathrm{MSAVI}_2 + 53.3722\,\mathrm{Lsndvi} - 1221.084\,\mathrm{Lsndvi}_{\mathrm{var}}$$

Table A-4: Upland Sandpiper goodness of fit

Sum of squared errors	113.0529
Expected value—H0	112.2815
SD	0.4711
Z	1.6373
P	0.1016

Table A-5: Upland Sandpiper Wald statistics

	χ^2	$d.f.$	P
MSAVI_2	3.44	1	0.0636
Lsndvi	41.54	1	< 0.0001
Lsndvi_{var}	7.41	1	0.0065
TOTAL	45.49	3	< 0.0001

Table A-6: Upland Sandpiper RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.7126	0.0021	2000	0.6324	0.0060	1	0.8547	0.0001
2	0.6303	0.0061	2001	0.6694	0.0038	2	0.7879	0.0006
3	0.7937	0.0005	2002	0.8825	< 0.0001	3	0.6894	0.0029
4	0.2608	0.1315	2003	0.7937	0.0005	4	0.6431	0.0053
5	0.7663	0.0009	2004	0.8315	0.0002	5	0.7484	0.0012
Mean	0.6327	0.0282	Mean	0.7619	0.0021	Mean	0.7447	0.0020

A.3 Marbled Godwit

$$X\hat{\beta} = -2.1833 - 0.0146\,\mathrm{CSM} + 3.2832\,\mathrm{Lsndvi} - 0.1222\,\mathrm{bright} \\ - 0.2182\,\mathrm{wet} + 389.6775\,\mathrm{Lsndvi}_{\mathrm{var}}$$

Table A-7: Marbled Godwit goodness of fit

Sum of squared errors	163.4240
Expected value—H0	162.6065
SD	0.5826
Z	1.4032
P	0.1606

Table A-8: Marbled Godwit Wald statistics

	χ^2	$d.f.$	P
CSM	7.41	1	0.0065
Lsndvi	0.15	1	0.7026
bright	13.10	1	0.0003
wet	38.45	1	< 0.0001
Lsndvi_{var}	11.94	1	0.0005
TOTAL	82.33	5	< 0.0001

Table A-9: Marbled Godwit RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9759	< 0.0001	2000	0.6497	0.0049	1	0.6497	0.0049
2	0.6303	0.0061	2001	0.7723	0.0008	2	0.9759	< 0.0001
3	0.7937	0.0005	2002	0.9521	< 0.0001	3	0.2626	0.1299
4	1.0000	< 0.0001	2003	0.9286	< 0.0001	4	0.6303	0.0061
5	0.5924	0.0092	2004	0.9521	< 0.0001	5	0.6694	0.0038
Mean	0.7985	0.0032	Mean	0.8510	0.0011	Mean	0.6376	0.0289

A.4 Sprague's Pipit

$$X\hat{\beta} = 4.0256 - 0.0275\,\mathrm{precip_t} + 0.0196\,\mathrm{precip_{t-1}} - 0.1890\,\mathrm{GVI} + 0.2187\,\mathrm{wet} + 0.0031\,\mathrm{GVI_{var}}$$

Table A-10: Sprague's Pipit goodness of fit

Sum of squared errors	195.8105
Expected value—H0	196.2231
SD	1.0026
Z	−0.4116
P	0.6807

Table A-11: Sprague's Pipit Wald statistics

	χ^2	$d.f.$	P
precip_t	80.03	1	< 0.0001
precip_{t-1}	118.98	1	< 0.0001
GVI	13.97	1	0.0002
wet	64.82	1	< 0.0001
GVI_{var}	0.58	1	0.4462
TOTAL	185.63	5	< 0.0001

Table A-12: Sprague's Pipit RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9759	< 0.0001	2000	1.0000	< 0.0001	1	1.0000	< 0.0001
2	1.0000	< 0.0001	2001	0.9879	< 0.0001	2	0.8155	0.0003
3	0.9521	< 0.0001	2002	0.9698	< 0.0001	3	0.8704	< 0.0001
4	1.0000	< 0.0001	2003	0.8788	< 0.0001	4	1.0000	< 0.0001
5	1.0000	< 0.0001	2004	0.7937	0.0005	5	0.8825	< 0.0001
Mean	0.9856	< 0.0001	Mean	0.9260	0.0001	Mean	0.9137	< 0.0001

A.5 Clay-colored Sparrow

$$
\begin{aligned}
X\hat{\beta} =& 25.6012 + 0.8254\,\mathrm{CTI} + 0.0611\,\mathrm{soil} + 0.0106\,\mathrm{precip_t} \\
& + 0.0019\,\mathrm{precip_{t-1}} + 0.0024\,\mathrm{precip_{t-2}} + 0.3420\,\mathrm{GVI} \\
& + 0.0872\,\mathrm{wet} + 0.0048\,\mathrm{GVI_{var}}
\end{aligned}
$$

Table A-13: Clay-colored Sparrow goodness of fit

Sum of squared errors	36.0923
Expected value—H0	35.9490
SD	0.4062
Z	0.3530
P	0.7241

Table A-14: Clay-colored Sparrow Wald statistics

	χ^2	$d.f.$	P
CTI	1.03	1	0.3099
soil	0.45	1	0.5007
precip$_t$	0.58	1	0.4478
precip$_{t-1}$	0.16	1	0.6932
precip$_{t-2}$	0.41	1	0.5227
GVI	6.01	1	0.0142
wet	1.27	1	0.2591
GVI$_{var}$	0.46	1	0.4993
TOTAL	42.10	8	< 0.0001

Table A-15: Clay-colored Sparrow RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.8727	< 0.0001	2000	0.6606	0.0043	1	0.6896	0.0029
2	0.9152	< 0.0001	2001	0.6606	0.0043	2	0.8297	0.0002
3	0.4472	0.0345	2002	0.8545	0.0001	3	0.7723	0.0008
4	0.3632	0.0652	2003	0.4526	0.0330	4	0.8255	0.0003
5	0.9394	< 0.0001	2004	0.9638	< 0.0001	5	0.4909	0.0240
Mean	0.7075	0.0200	Mean	0.7184	0.0083	Mean	0.7216	0.0057

A.6 Vesper Sparrow

$$
\begin{aligned}
X\hat{\beta} =& 2.0677 + 0.4615\,\mathrm{CTI} + 0.2116\,\mathrm{soil} - 0.0179\,\mathrm{CSM} \\
& + 24.7991\,\mathrm{Lsmsavi_2} - 0.4913\,\mathrm{GVI} + 0.3200\,\mathrm{wet} \\
& - 44.7246\,\mathrm{Lsmsavi_2var}
\end{aligned}
$$

Table A-16: Vesper Sparrow goodness of fit

Sum of squared errors	188.1633
Expected value—H0	188.5817
SD	0.8109
Z	−0.5160
P	0.6059

Table A-17: Vesper Sparrow Wald statistics

	χ^2	$d.f.$	P
CTI	1.72	1	0.1895
soil	26.25	1	< 0.0001
CSM	12.95	1	0.0003
$\mathrm{Lsmsavi_2}$	70.13	1	< 0.0001
GVI	34.46	1	< 0.0001
wet	60.61	1	< 0.0001
$\mathrm{Lsmsavi_2var}$	4.16	1	0.0415
TOTAL	143.13	7	< 0.0001

Table A-18: Vesper Sparrow RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9521	< 0.0001	2000	0.9759	< 0.0001	1	0.6112	0.0075
2	0.9054	< 0.0001	2001	1.0000	< 0.0001	2	0.9521	< 0.0001
3	0.9521	< 0.0001	2002	0.7723	0.0008	3	0.8598	0.0001
4	0.9759	< 0.0001	2003	0.9054	< 0.0001	4	0.4526	0.0330
5	0.9521	< 0.0001	2004	0.9759	< 0.0001	5	0.9286	< 0.0001
Mean	0.9475	< 0.0001	Mean	0.9259	0.0002	Mean	0.7609	0.0081

A.7 Lark Bunting

$$X\hat{\beta} = 17.0859 - 0.2635\,\mathrm{CTI} + 0.1824\,\mathrm{soil} - 0.0093\,\mathrm{precip}_t + 0.0120\,\mathrm{precip}_{t-1} + 27.6232\,\mathrm{Lsmsavi}_2 - 0.3254\,\mathrm{GVI} + 0.3037\,\mathrm{wet} - 31.2577\,\mathrm{Lsmsavi}_2\mathrm{var}$$

Table A-19: Lark Bunting goodness of fit

Sum of squared errors	141.6894
Expected value—H0	143.4961
SD	1.1911
Z	−1.5168
P	0.1293

Table A-20: Lark Bunting Wald statistics

	χ^2	$d.f.$	P
CTI	0.41	1	0.5212
soil	15.03	1	0.0001
precip_t	7.94	1	0.0048
precip_{t-1}	44.11	1	< 0.0001
$\mathrm{Lsmsavi}_2$	67.52	1	< 0.0001
GVI	11.60	1	0.0007
wet	40.49	1	< 0.0001
$\mathrm{Lsmsavi}_2\mathrm{var}$	2.52	1	0.1122
TOTAL	182.71	8	< 0.0001

Table A-21: Lark Bunting RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	1.0000	< 0.0001	2000	0.0133	0.7514	1	0.9521	< 0.0001
2	0.9759	< 0.0001	2001	0.1032	0.3655	2	0.7511	0.0012
3	0.5201	0.0186	2002	0.5924	0.0092	3	0.8375	0.0002
4	0.9759	< 0.0001	2003	0.0606	0.4929	4	0.8015	0.0005
5	0.7038	0.0024	2004	0.9054	< 0.0001	5	0.7097	0.0022
Mean	0.8351	0.0042	Mean	0.3350	0.3238	Mean	0.8104	0.0008

A.8 Savannah Sparrow

$$X\hat{\beta} = -2.8212 + 0.2117\,\mathrm{CTI} - 0.0117\,\mathrm{soil} + 1.8398\,\mathrm{MSAVI}_2 \\ - 14.7542\,\mathrm{Lsmsavi}_2 + 0.2283\,\mathrm{GVI} - 0.1019\,\mathrm{wet} + 91.9517\,\mathrm{Lsmsavi}_2\mathrm{var}$$

Table A-22: Savannah Sparrow goodness of fit

Sum of squared errors	240.1365
Expected value—H0	240.1611
SD	0.2673
Z	−0.0919
P	0.9267

Table A-23: Savannah Sparrow Wald statistics

	χ^2	$d.f.$	P
CTI	0.41	1	0.5205
soil	0.12	1	0.7254
MSAVI_2	2.27	1	0.1317
$\mathrm{Lsmsavi}_2$	34.43	1	< 0.0001
GVI	10.45	1	0.0012
wet	9.16	1	0.0025
$\mathrm{Lsmsavi}_2\mathrm{var}$	21.00	1	< 0.0001
TOTAL	46.72	7	< 0.0001

Table A-24: Savannah Sparrow RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.1957	0.2004	2000	0.6497	0.0049	1	0.0290	0.6383
2	0.9521	< 0.0001	2001	0.9054	< 0.0001	2	0.9054	< 0.0001
3	0.9521	< 0.0001	2002	0.9759	< 0.0001	3	0.7065	0.0023
4	0.7937	0.0005	2003	0.8598	0.0001	4	0.2066	0.1869
5	0.9286	< 0.0001	2004	0.9521	< 0.0001	5	0.5739	0.0111
Mean	0.7645	0.0402	Mean	0.8686	0.0010	Mean	0.4843	0.1677

A.9 Baird's Sparrow

$$X\hat{\beta} = -23.2364 - 0.0287\,\mathrm{precip_t} + 0.0304\,\mathrm{precip_{t-1}} - 58.4714\,\mathrm{Lsndvi} + 0.0071\,\mathrm{bright} - 0.0034\,\mathrm{wet} + 411.1721\,\mathrm{Lsndvi_{var}}$$

Table A-25: Baird's Sparrow goodness of fit

Sum of squared errors	149.3114
Expected value—H0	148.7249
SD	1.1463
Z	0.5117
P	0.6089

Table A-26: Baird's Sparrow Wald statistics

	χ^2	*d.f.*	*P*
precip_t	35.50	1	< 0.0001
precip_{t-1}	69.86	1	< 0.0001
Lsndvi	38.76	1	< 0.0001
bright	0.04	1	0.8385
wet	0.01	1	0.9256
Lsndvi_{var}	11.78	1	0.0006
TOTAL	124.11	6	< 0.0001

Table A-27: Baird's Sparrow RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.9758	< 0.0001	2000	0.7774	0.0007	1	0.9521	< 0.0001
2	0.9759	< 0.0001	2001	0.8788	< 0.0001	2	0.9225	< 0.0001
3	0.9517	< 0.0001	2002	0.7399	0.0014	3	0.7638	0.0009
4	0.9939	< 0.0001	2003	0.6431	0.0053	4	1.0000	< 0.0001
5	0.9759	< 0.0001	2004	0.9939	< 0.0001	5	0.9758	< 0.0001
Mean	0.9746	< 0.0001	Mean	0.8066	0.0015	Mean	0.9228	0.0002

A.10 Grasshopper Sparrow

$$X\hat{\beta} = 13.9587 + 0.7973\,\mathrm{CTI} + 0.0006\,\mathrm{precip_t} + 0.0103\,\mathrm{precip_{t-1}} + 0.0063\,\mathrm{precip_{t-2}} + 0.2977\,\mathrm{GVI} - 0.0121\,\mathrm{GVI_{var}}$$

Table A-28: Grasshopper Sparrow goodness of fit

Sum of squared errors	206.4078
Expected value—H0	206.1962
SD	0.8765
Z	0.2413
P	0.8093

Table A-29: Grasshopper Sparrow Wald statistics

	χ^2	$d.f.$	P
CTI	5.43	1	0.0198
precip_t	0.01	1	0.9188
precip_{t-1}	28.72	1	< 0.0001
precip_{t-2}	16.18	1	0.0001
GVI	101.91	1	< 0.0001
GVI_{var}	8.38	1	0.0038
TOTAL	208.49	6	< 0.0001

Table A-30: Grasshopper Sparrow RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	1.0000	< 0.0001	2000	0.9054	< 0.0001	1	0.9759	< 0.0001
2	1.0000	< 0.0001	2001	0.9759	< 0.0001	2	1.0000	< 0.0001
3	1.0000	< 0.0001	2002	0.9759	< 0.0001	3	1.0000	< 0.0001
4	1.0000	< 0.0001	2003	0.9044	< 0.0001	4	1.0000	< 0.0001
5	1.0000	< 0.0001	2004	0.9054	< 0.0001	5	1.0000	< 0.0001
Mean	1.0000	< 0.0001	Mean	0.9334	< 0.0001	Mean	0.9952	< 0.0001

A.11 McCown's Longspur

$$X\hat{\beta} = -11.8256 + 0.0077\,\mathrm{precip_t} - 0.0111\,\mathrm{precip_{t-1}} - 11.2871\,\mathrm{Lsmsavi_2} + 0.2438\,\mathrm{GVI} - 0.2471\,\mathrm{wet}$$

Table A-31: McCown's Longspur goodness of fit

Sum of squared errors	165.4682
Expected value—H0	164.0818
SD	0.9466
Z	1.4646
P	0.1430

Table A-32: McCown's Longspur Wald statistics

	χ^2	$d.f.$	P
precip_t	6.72	1	0.0095
precip_{t-1}	85.65	1	< 0.0001
$\mathrm{Lsmsavi}_2$	14.09	1	0.0002
GVI	9.29	1	0.0023
wet	42.31	1	< 0.0001
TOTAL	175.10	5	< 0.0001

Table A-33: McCown's Longspur RPI values

Fold	Random	P	Year	Temporal	P	Block	Spatial	P
1	0.7723	0.0008	2000	0.9515	< 0.0001	1	0.9759	< 0.0001
2	0.9939	< 0.0001	2001	1.0000	< 0.0001	2	0.9759	< 0.0001
3	1.0000	< 0.0001	2002	0.9697	< 0.0001	3	0.5997	0.0085
4	0.9759	< 0.0001	2003	0.8375	0.0002	4	0.6694	0.0038
5	1.0000	< 0.0001	2004	0.9759	< 0.0001	5	0.9698	< 0.0001
Mean	0.9484	0.0002	Mean	0.9469	< 0.0001	Mean	0.8382	0.0025