Wildlife Biometrics Unit
Carlos Avery Wildlife management Area
5463 – C West Broadway
Forest Lake, Minnesota  55025
(651) 296-5200
COMPARATIVE INTERPRETATION OF COUNT, PRESENCE-ABSENCE AND POINT METHODS FOR SPECIES DISTRIBUTION MODELS

Geert Aarts, John Fieberg, and Jason Matthiopoulos

ABSTRACT

1. The need to understand the processes shaping population distributions has resulted in a vast increase in the diversity of spatial wildlife data, leading to the development of many novel analytical techniques that are fit-for-purpose. One may aggregate location data into spatial units (e.g. grid cells), and model the resulting counts or presence-absences as a function of environmental covariates. Alternatively, the point data may be modeled directly, by combining the individual observations with a set of random or regular points reflecting habitat availability, a method known as a use-availability design (or, alternatively a presence–pseudo-absence or case-control design).

2. Although these spatial point, count and presence-absence methods are widely used, the ecological literature is not explicit about their connections and how their parameter estimates and predictions should be interpreted. The objective of this study is to recapitulate some recent statistical results and illustrate that under certain assumptions, each method can be motivated by the same underlying spatial Inhomogeneous Poisson point-process (IPP) model in which the intensity function is modeled as a log-linear function of covariates.

3. The Poisson likelihood used for count data is a discrete approximation of the IPP likelihood. Similarly, the presence-absence design will approximate the IPP likelihood, but only when spatial units (i.e., pixels) are extremely small (Baddeley et al., 2010). For larger pixel sizes, presence-absence designs do not differentiate between 1 or multiple observations within each pixel, hence leading to information loss.

4. Logistic regression is often used to estimate the parameters of the IPP model using point data. Although the response variable is defined as 0 for the availability points, these 0s do not serve as true absences as is often assumed; rather, their role is to approximate the integral of the denominator in the IPP likelihood (Warton and Shepherd 2010). Due to this common misconception, the estimated exponential function of the linear predictor (i.e., the resource selection function) is often assumed to be proportional to occupancy. Like IPP and count models, this function is proportional to the expected density of observations.

5. Understanding these (dis-)similarities between different species distribution modeling techniques should improve biological interpretation of spatial models, and therefore advance ecological and methodological cross-fertilization.
A BAYESIAN HIERARCHICAL OCCUPANCY MODEL FOR TRACK SURVEYS CONDUCTED IN A SERIES OF LINEAR, SPATIALLY CORRELATED SITES

Chrisna Aing², Sarah Halls², Kiva Oken², Robert Dobrow², and John Fieberg.

ABSTRACT

1. Natural resource agencies often rely on surveys of animal sign (e.g., scat, scent marks, tracks) for population assessment, with repeat surveys required to model and account for uncertain detection. Using river otter Lontra canadensis snow track survey data as a motivating example, we develop a 3-level occupancy model with parameters that describe (1) site-level occupancy probabilities, (2) otter movement (and thus, track availability), and (3) recorded presence-absence of tracks (conditional on the availability of tracks for detection).

2. We incorporated several recent developments in occupancy modeling, including the presence of both false negatives and false positives, spatial and temporal correlation, and repeated sampling across distinct observers.

3. We investigated optimal allocation of sampling effort (e.g., within and among snowfall events) using simulations. We also compared models that allowed site-level occupancy and track laying processes to be spatially correlated to models that assumed independence among sites.

4. Both types of models (independence and spatial) performed well across a range of simulated parameter values, but the spatial model resulted in more accurate point estimates for detection parameters and credibility intervals with better coverage rates when data were spatially correlated. When applied to real data, the spatial model resulted in a higher estimate of the occupancy rate ( ) than the baseline model (0.82 versus 0.59). A minimum of 15-20 helicopter flights, distributed among at least three unique snow events, were needed to meet precision goals (standard error < 0.05).

5. Synthesis and applications. We describe a flexible and robust occupancy modeling framework that accounts for heterogeneous detection rates in surveys of animal sign. The method allows for spatially correlated sites, and should have broad relevance to surveys conducted by many natural resource agencies.

¹Abstract from paper provisionally accepted for publication in the Journal of Applied Ecology
²Department of Mathematics, Carleton College, Northfield, Minnesota 55057
SIGHTABILITY MODELS ARE BINARY LOGISTIC-REGRESSION MODELS USED TO ESTIMATE AND ADJUST FOR VISIBILITY BIAS IN WILDLIFE-PopULATION SURVEYS. LIKE MANY MODELS IN WILDLIFE AND ECe, SIGHTABILITY MODELS ARE TYPICALLY DEVELOPED FROM SMALL OBSERVATIONAL DATA SETS WITH MANY CANDIDATE PREDICTORS. AGGRESSIVE MODEL-SELECTION METHODS ARE OFTEN EMPLOYED TO CHOOSE A ‘BEST’ MODEL FOR PREDICTION AND EFFECT ESTIMATION, DESPITE EVIDENCE THAT SUCH METHODS CAN LEAD TO OVERFITTING (i.e., SELECTED MODELS MAY DESCRIBE RANDOM ERROR OR NOISE RATHER THAN TRUE PREDICTOR-RESPONSE CURVES) AND POOR PREDICTIVE ABILITY. WE USED MOOSE-SIGHTABILITY DATA FROM NORTHEASTERN MINNESOTA (2005–2007) AS A CASE STUDY TO ILLUSTRATE AN ALTERNATIVE APPROACH, WHICH WE REFER TO AS DEGREES-OF-FREEDOM (df) SPENDING: SAMPLE-SIZE GUIDELINES ARE USED TO DETERMINE AN ACCEPTABLE LEVEL OF MODEL COMPLEXITY AND THEN A PRE-SPECIFIED MODEL IS FIT TO THE DATA AND USED FOR INFERENCE. FOR COMPARISON, WE ALSO CONSTRUCTED SIGHTABILITY MODELS USING AIC STEP-DOWN PROCEDURES AND MODEL AVERAGING (BASED ON A SMALL SET OF MODELS DEVELOPED USING DF-SPENDING GUIDELINES). WE USED BOOTSTRAP PROCEDURES TO MIMIC THE PROCESS OF MODEL-FITTING AND PREDICTION, AND TO COMPUTE AN INDEX OF OVERFITTING, EXPECTED PREDICTIVE ACCURACY, AND MODEL-SELECTION UNCERTAINTY. THE INDEX OF OVERFITTING INCREASED 13% WHEN THE NUMBER OF CANDIDATE PREDICTORS WAS INCREASED FROM 3 TO 8 AND A ‘BEST’ MODEL WAS SELECTED USING STEP-DOWN PROCEDURES. LIKewise, model-selection uncertainty increased when the number of candidate predictors increased. Model averaging (based on $R = 30$ models with 1–3 predictors) effectively “shrunk” regression coefficients toward zero and produced similar estimates of precision to our 3-df pre-specified model. As such, model averaging may help to guard against overfitting when too many predictors are considered (relative to available sample size). The set of candidate models will influence the extent to which coefficients are shrunk toward 0, which has implications for how 1 might apply model averaging to problems traditionally approached using variable-selection methods. We often recommend the df-spending approach in our consulting work, because it is easy to implement and it naturally forces investigators to think carefully about their models and predictors. Nonetheless, similar concepts should apply whether 1 is fitting 1 model or using multi-model inference. For example, model-building decisions should consider the effective sample size, and potential predictors should be screened (without looking at their relationship to the response) for missing data, narrow distributions, collinearity, potentially overly influential observations, and measurement errors (e.g., via logical error checks).
Researchers employing resource selection functions (RSFs) and other related methods aim to detect correlates of space-use and mitigate against detrimental environmental change. However, an empirical model fit to data from 1 place or time is unlikely to capture species responses under different conditions, because organisms respond nonlinearly to changes in habitat availability. This phenomenon, known as a functional response in resource selection, has been debated extensively in the RSF literature, but continues to be ignored by practitioners for lack of a practical treatment. We therefore extend the RSF approach to enable it to estimate generalized functional responses (GFRs) from spatial data. GFRs employ data from several sampling instances characterized by diverse profiles of habitat availability. By modeling the regression coefficients of the underlying RSF as functions of availability, GFRs can account for environmental change and thus predict population distributions in new environments. We formulate the approach as a mixed-effects model so that it is estimable by readily available statistical software. We illustrate its application using (1) simulation and (2) wolf home-range telemetry. Our results indicate that GFRs can offer considerable improvements in estimation speed and predictive ability over existing mixed-effects approaches.
INTEGRATED POPULATION MODELING OF BLACK BEARS IN MINNESOTA: IMPLICATIONS FOR MONITORING AND MANAGEMENT

John Fieberg, Kyle W. Shertzer, Paul B. Conn, Karen V. Noyce, and Dave L. Garshelis.

ABSTRACT

Background: Wildlife populations are difficult to monitor directly because of costs and logistical challenges associated with collecting informative abundance data from live animals. By contrast, data on harvested individuals (e.g., age and sex) are often readily available. Increasingly, integrated population models are used for natural resource management, because they synthesize various relevant data into a single analysis.

Methodology/Principal Findings: We investigated the performance of integrated population models applied to black bears (Ursus americanus) in Minnesota, USA. Models were constructed using sex-specific age-at-harvest matrices (1980–2008), data on hunting effort and natural food supplies (which affects hunting success), and statewide mark–recapture estimates of abundance (1991, 1997, 2002). We compared this approach to Downing reconstruction, a commonly used population monitoring method that utilizes only age-at-harvest data. We first conducted a large-scale simulation study, in which our integrated models provided more accurate estimates of population trends than did Downing reconstruction. Estimates of trends were robust to various forms of model mis-specification, including incorrectly specified cub and yearling survival parameters, age-related reporting biases in harvest data, and unmodeled temporal variability in survival and harvest rates. When applied to actual data on Minnesota black bears, the model predicted that harvest rates were negatively correlated with food availability and positively correlated with hunting effort, consistent with independent telemetry data. With no direct data on fertility, the model also correctly predicted 2-point cycles in cub production. Model-derived estimates of abundance for the most recent years provided a reasonable match to an empirical population estimate obtained after modeling efforts were completed.

Conclusions/Significance: Integrated population modeling provided a reasonable framework for synthesizing age-at-harvest data, periodic large-scale abundance estimates, and measured covariates thought to affect harvest rates of black bears in Minnesota. Collection and analysis of these data appear to form the basis of a robust and viable population monitoring program.


2 Southeast Fisheries Science Center, National Marine Fisheries Service, Beaufort, North Carolina
CORRELATION AND STUDIES OF HABITAT SELECTION: PROBLEM, RED HERRING OR OPPORTUNITY?¹

John Fieberg, Jason Matthiopoulos², Mark Hebblewhite³, Mark S. Boyce⁴ and Jacqueline L. Frair⁵

ABSTRACT

With the advent of new technologies, animal locations are being collected at ever finer spatiotemporal scales. We review analytical methods for dealing with correlated data in the context of resource selection, including post hoc variance inflation techniques, ‘two-stage’ approaches based on models fit to each individual, generalized estimating equations and hierarchical mixed-effects models. These methods are applicable to a wide range of correlated data problems, but can be difficult to apply and remain especially challenging for use–availability sampling designs, because the correlation structure for combinations of used and available points are not likely to follow common parametric forms. We also review emerging approaches to studying habitat selection that use fine-scale temporal data to arrive at biologically based definitions of available habitat, while naturally accounting for autocorrelation by modeling animal movement between telemetry locations. Sophisticated analyses that explicitly model correlation rather than consider it a nuisance, like mixed effects and state-space models, offer potentially novel insights into the process of resource selection, but additional work is needed to make them more generally applicable to large data sets based on the use–availability designs. Until then, variance inflation techniques and 2-stage approaches should offer pragmatic and flexible approaches to modeling correlated data.

²NERC Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY16 8LB, United Kingdom
³Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, Montana 59812
⁴Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9
⁵State University of New York, College of Environmental Science and Forestry, Syracuse, New York 13210
RESOLVING ISSUES OF IMPRECISE AND HABITAT-BIASED LOCATIONS IN ECOLOGICAL ANALYSES USING GPS TELEMETRY DATA

Jacqueline L. Frair, John Fieberg, Mark Hebblewhite, Francesca Cagnacci, Nicholas J. DeCesare, and Luca Pedrotti

ABSTRACT

Global positioning system (GPS) technologies collect unprecedented volumes of animal location data, providing ever greater insight into animal behaviour. Despite a certain degree of inherent imprecision and bias in GPS locations, little synthesis regarding the predominant causes of these errors, their implications for ecological analysis or solutions exists. Terrestrial deployments report 37 per cent or less non-random data loss and location precision 30 m or less on average, with canopy closure having the predominant effect, and animal behaviour interacting with local habitat conditions to affect errors in unpredictable ways. Home range estimates appear generally robust to contemporary levels of location imprecision and bias, whereas movement paths and inferences of habitat selection may readily become misleading. There is a critical need for greater understanding of the additive or compounding effects of location imprecision, fix-rate bias, and, in the case of resource selection, map error on ecological insights. Technological advances will help, but at present, analysts have a suite of ad hoc statistical corrections and modeling approaches available—tools that vary greatly in analytical complexity and utility. The success of these solutions depends critically on understanding the error-inducing mechanisms, and the biggest gap in our current understanding involves species-specific behavioural effects on GPS performance.

2 State University of New York, College of Environmental Science and Forestry, Syracuse, New York 13210
3 Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, Montana 59812
4 Research and Innovation Centre, Environment and Natural Resources Area, Edmund Mach Foundation, 38010 San Michele all’Adige, Trento, Italy
5 Stelvio National Park, 23032 Bormio, Sondrio, Italy
The home-range concept: Are traditional estimators still relevant with modern telemetry technology?  

John G. Kie, Jason Matthiopoulos, John Fieberg, Roger A. Powell, Francesca Cagnacci, Michael S. Mitchell, Jean-Michel Gaillard, and Paul R. Moorcroft

Abstract

Recent advances in animal tracking and telemetry technology have allowed the collection of location data at an ever-increasing rate and accuracy, and these advances have been accompanied by the development of new methods of data analysis for portraying space use, home ranges and utilization distributions. New statistical approaches include data-intensive techniques such as kriging and nonlinear generalized regression models for habitat use. In addition, mechanistic home range models, derived from models of animal movement behaviour, promise to offer new insights into how home ranges emerge as the result of specific patterns of movements by individuals in response to their environment. Traditional methods, such as kernel density estimators are likely to remain popular, because of their ease of use. Large data sets make it possible to apply these methods over relatively short periods of time, such as weeks or months, and these estimates may be analyzed using mixed-effects models, offering another approach to studying temporal variation in space-use patterns. Although new technologies open new avenues in ecological research, our knowledge of why animals use space in the ways we observe will only advance by researchers using these new technologies and asking new and innovative questions about the empirical patterns they observe.
DNRSURVEY – MOVING-MAP SOFTWARE FOR AERIAL SURVEYS

Robert G. Wright, Brian S. Haroldson, and Chris Pouliot

SUMMARY OF FINDINGS

Advances in Global Position System (GPS), Geographic Information System (GIS), and computer technologies have enhanced our ability to navigate aerial wildlife surveys and capture observational data. We combined these technologies into a moving-map, aerial survey software program herein referred to as DNRSurvey, which allows users to display and record their position over digital aerial photos, navigate without reliance on ground features, and record animal locations directly to Environmental Systems Research Institute (ESRI, Inc., Redlands, California) shapefiles and Windows (Microsoft Corporation, Redmond, Washington) audio files. This program has improved the precision and efficiency of our aerial surveys and reduced data-entry transcription time and errors. Although originally designed for an aerial platform, DNRSurvey is equally applicable for vehicle-based observations.

INTRODUCTION

Aerial surveys are commonly used to estimate abundance of waterfowl, ungulates, and other large mammals. Navigation during these surveys, which began with map and compass, has improved with developments in technology, transitioning through land-based radio transmitters (e.g., long range navigation [LORAN]; Boer et al. 1989, Leptich et al. 1994) to global, satellite-based systems (e.g., GPS; Bobbe 1992, Leptich et al. 1994). Anthony and Stehn (1994) created a software program (GPSTRACK) which displayed real-time aircraft positions over pre-defined transects on a laptop computer and recorded locations of wildlife observations along transect lines. Within the last decade, advances in GPS, GIS, and computer hardware technologies have greatly enhanced our ability to navigate aerial surveys and capture observational data, independent of aircraft location. We combined these technologies into a moving-map, aerial survey software program referred to as DNRSurvey. Using a tablet computer connected to a GPS receiver, we are able to view and record our real-time position over digital air photos, navigate without reliance on ground features, and record animal observation data (e.g., location, count, age/sex, cover type) directly to ESRI shapefiles and Windows audio files. DNRSurvey is not a GIS, but a data collection tool that incorporates relevant GIS functionality. Use of this program has improved the precision and efficiency of our aerial surveys and reduced data-entry transcription time and errors.

SOFTWARE DEVELOPMENT

DNRSurvey was developed in Visual Basic (VB; Microsoft Corporation, Redmond, Washington) programming language and consists of 2 integrated components - Survey Editor (VB.NET) and MapView (VB 6.0). With Survey Editor, users create survey-specific data entry forms and shapefile attribute tables to record wildlife observations or other objects of interest (Figure 1). A variety of input controls, including textboxes, checkboxes, radio buttons, comboboxes, listboxes, and voice recording are available to customize data input (Figures 1 and 2). A spatial join feature allows attributes (e.g., public land survey features, plot number, acres) from another shapefile to be written to the survey shapefile when observations are recorded. Survey shapefile properties such as symbology and labeling can be pre-defined and a custom icon can be assigned to each survey form tool button (Figure 3).

MapView is the survey component of DNRSurvey and emulates a stripped-down version of an ESRI ArcMap data frame (Figure 4). It communicates with a GPS receiver via serial, USB or Bluetooth connection; displays a bread-crum trail of positional fixes; and pans the display window as needed. Background image and shapefile layers such as aerial photos, management unit boundaries, and survey plot boundaries are managed in a Table of Contents
window (Figure 4). Shapefile symbology and labeling can be customized and scale-dependent displays can be defined for all layers (Figure 5). Key functionalities, such as preset zoom scales, data backup, editing, and survey form activation are presented as toolbar buttons (Figure 4). Customized settings can be saved as a unique survey file (e.g., pa272_survey.lvs).

To begin collecting observations, users open the customized survey file, connect to the GPS receiver using the toolbar button, and select the data form tool button to make it active. The user records an observation by touching the screen where the object of interest is located and by completing the pop-up data form (Figure 4). Location coordinates and data form values are written directly to an output shapefile or audio file. Observations can be captured anywhere on the display or by accepting the current GPS position. The user edits an observation by selecting the Edit button and desired on-screen data point, and then by correcting erroneous data values in the pop-up data form. Pressing the data backup button copies all survey related data (e.g., observation shapefile, flight line shapefile, flight line text file) to a date/time-stamped working directory. The GPS coordinate properties (i.e., datum, coordinate system) are user-defined, but default to the North American Datum of 1983 (NAD83) and Universal Transverse Mercator (UTM) Zone 15N, respectively. In addition, the aircraft flight-line display and recording properties can be customized to meet the user’s needs (Figure 6).

DNRSurvey works on tablet computers running Windows XP and Windows 7 operating systems (Microsoft Corporation, Redmond, Washington). We recommend a minimum computer configuration which includes: 80 GB hard drive; 3 GB RAM; 1 GHz processor; 550 nit daylight-readable display; serial port and/or Bluetooth data link; and integrated keyboard. DNRSurvey is compatible with GPS receivers using Garmin (Garmin International, Inc., Olathe, Kansas) or National Marine Electronics Association (NMEA) output formats.

For cockpit deployment, we currently use a wireless configuration consisting of a Panasonic CF-19 Toughbook tablet computer (Panasonic Corporation, Secaucus, New Jersey) communicating with an fTech Solarius BT-25 SR Solar Bluetooth GPS receiver (fTech Corporation, Tainan, Taiwan). This configuration is convenient and enhances cockpit safety by eliminating loose cables. A Garmin GPSMAP196 mounted in the aircraft serves as a backup receiver. The computer battery lasts >3 hours and is replaced during each fuel stop. Battery life for the solar Bluetooth GPS is sufficient to last all day on a single charge.

We are currently working on additional enhancements and expect to complete software development by December 2011. Although originally designed for an aerial platform, DNRSurvey is equally applicable for vehicle-based observations and will be available at www.dnr.state.mn.us.

LITERATURE CITED

Figure 1. Survey Editor form building interface of DNRSurvey.

Figure 2. Data input values for combobox (species) and listbox (cover) controls are defined using lookup or user-defined tables via drop-down menus and tabs within the Survey Editor component of DNRSurvey.
Figure 3. Spatial join shapefiles, tool icons, and symbology and labeling properties are defined via drop-down menus and tabs within the Survey Editor component of DNRSurvey.

Figure 4. MapView interface component of DNRSurvey.
Figure 5. Symbology, labeling, and scale properties of background layers are defined via drop-down menus and tabs within the MapView component of DNRSurvey.

Figure 6. Datum and projection values, and flight-line display and recording properties are defined via drop-down menus and tabs within the MapView component of DNRSurvey.