

A spatially explicit method for evaluating accuracy of species distribution models

Mary Smulders¹, Trisalyn A. Nelson^{1*}, Dennis E. Jelinski², Scott E. Nielsen³ and Gordon B. Stenhouse⁴

¹Spatial Pattern Analysis and Research Laboratory, Department of Geography, University of Victoria, PO Box 3060, Victoria, BC, V8W 3R4, Canada, ²Laboratory for Landscape and Ecosystem Ecology, Department of Geography, University of Victoria, PO Box 3060, Victoria, BC, V8W 3R4, Canada, ³Department of Renewable Resources, University of Alberta, Edmonton, T6G 2H1, Canada, ⁴Foothills Model Forest Grizzly Bear Research Group, Box 6330, Hinton, Alberta, Canada

ABSTRACT

Aim Models predicting the spatial distribution of animals are increasingly used in wildlife management and conservation planning. There is growing recognition that common methods of evaluating species distribution model (SDM) accuracy, as a global overall value of predictive ability, could be enhanced by spatially evaluating the model thereby identifying local areas of relative predictive strength and weakness. Current methods of spatial SDM model assessment focus on applying local measures of spatial autocorrelation to SDM residuals, which require quantitative model outputs. However, SDM outputs are often probabilistic (relative probability of species occurrence) or categorical (species present or absent). The goal of this paper was to develop a new method, using a conditional randomization technique, which can be applied to directly spatially evaluate probabilistic and categorical SDMs.

Location Eastern slopes, Rocky Mountains, Alberta, Canada.

Methods We used predictions from seasonal grizzly bear (*Ursus arctos*) resource selection functions (RSF) models to demonstrate our spatial evaluation technique. Local test statistics computed from bear telemetry locations were used to identify areas where bears were located more frequently than predicted. We evaluated the spatial pattern of model inaccuracies using a measure of spatial autocorrelation, local Moran's *I*.

Results We found the model to have non-stationary patterns in accuracy, with clusters of inaccuracies located in central habitat areas. Model inaccuracies varied seasonally, with the summer model performing the best and the least error in areas with high RSF values. The landscape characteristics associated with model inaccuracies were examined, and possible factors contributing to RSF error were identified.

Main conclusions The presented method complements existing spatial approaches to model error assessment as it can be used with probabilistic and categorical model output, which is typical for SDMs. We recommend that SDM accuracy assessments be done spatially and resulting accuracy maps included in model metadata.

Keywords

Grizzly bear, local indicator of spatial autocorrelation (LISA), model evaluation, resource selection functions (RSF), spatial analysis, species distribution.

*Correspondence: Trisalyn A. Nelson, Spatial Pattern Analysis and Research Laboratory, Department of Geography, University of Victoria, PO Box 3060, Victoria, BC, V8W 3R4, Canada.
E-mail: trisalyn@uvic.ca

INTRODUCTION

Owing to their diverse application and influence in conservation decisions, understanding species distribution model (SDM) accuracy, and particularly the variation in their

accuracy across the study area, is critical. Mapped SDMs, which combine observed occurrences with geographical environmental data to predict species distribution, are a key conservation tool and are applied across a range of biogeographical scales (Olden *et al.*, 2002; Araùjo & Guisan, 2006;

Raes & ter Steege, 2007). Most methods for evaluating SDMs are aspatial and present accuracy as an overall single statistic or error matrix. Overall values, also known as global values or assessments (Zhang *et al.*, 2005; Barry & Elith, 2006), summarize model error for entire study areas and may combine accuracy of locations of accurate and poor prediction (Zhu, 2001). As study area size increases, SDM error is likely to vary spatially given heterogeneity expected over large spatial extents (McGwire & Fisher, 2001; Zhang *et al.*, 2005; Osborne *et al.*, 2007). There may be a spatial pattern to SDM error which could be used to improve model accuracy (e.g. McGwire & Fisher, 2001; Pontius & Schneider, 2001; Lobo & Martín-Piera, 2002; Barry & Elith, 2006; Lobo *et al.*, 2008).

Generally, and with SDMs, there is a growing recognition of the importance of spatially assessing model accuracy (McGwire & Fisher, 2001; Fielding, 2002; Zhang *et al.*, 2005; Barry & Elith, 2006). If there is no spatial component to model error, all geographical variables have been accounted for and correctly specified (Anselin, 1988). However, when model errors exhibit spatial patterning, characterizing the spatial nature of error can provide information on variables to be added or applied differently (Nicholls, 1989; Laffan, 1999). This is particularly true when spatial patterns of model error can be mapped using local spatial analysis techniques (e.g. Zhang *et al.*, 2005). By integrating mapped model error with other data, the conditions at locations poorly modelled can be assessed and relationships used to enhance model performance (Lobo & Martín-Piera, 2002; Barry & Elith, 2006).

Arguably, the most common approach to spatially assessing model error is to apply local measures of spatial autocorrelation to residuals (Anselin, 1988; Zhang *et al.*, 2005; Wulder *et al.*, 2007). For example, local Moran's *I* has been used by researchers in SDM creation to examine clusters of model residuals that deviate from predicted results (e.g. Lobo & Martín-Piera, 2002; Zhang *et al.*, 2005; Osborne *et al.*, 2007). The spatial arrangement of errors provides valuable details about the model because there is a difference in the importance of errors that are randomly distributed and those that are clustered (Lobo *et al.*, 2008). Clusters of model error may indicate the existence of unaccounted for spatially structured variables (Lobo *et al.*, 2008) or model inputs with spatially inaccuracy (Fortin & Edwards, 2001; Anselin, 2002). For example, clusters of error could occur at an incorrectly defined vegetation class polygon boundary because the spatial location of a polygon boundary is often uncertain (Fortin & Edwards, 2001). Others have used methods such as geographically weighted regression (Laffan, 1999) to identify locations where errors occur in closer proximity than expected for random error processes. Common methods for assessing the spatial characteristics of model error require quantitative data and have limited application to SDMs with categorical outputs. Furthermore, the spatial methods outlined earlier are difficult to use when validation procedures use observed data to assess probabilistic model output. Observations in locations with low probabilities do not necessarily indicate errors (Nelson & Boots, 2005).

A wide variety of approaches are used for the prediction of species distribution (Guisan & Zimmermann, 2000). Typically, SDMs compare species' presence and species' absence with environmental data to estimate various aspects of the geographical distribution of species (Lobo *et al.*, 2010). Methods used to compute SDMs based on presence/absence data include the following: generalized linear models (GLM), generalized additive models (GAM), classification and regression tree analysis and artificial neural networks (ANN) (Guisan & Zimmermann, 2000; Brotons *et al.*, 2004). The reliability of absences depends on the species characteristics (Hirzel *et al.*, 2001), abundance (Kéry *et al.*, 2009) and survey methods (MacKenzie & Royle, 2005). While researchers may determine plant absence with reasonable confidence, identifying absence sites for wildlife, especially highly mobile and cryptic animals, is troublesome (Boyce *et al.*, 2002; Johnson *et al.*, 2006).

As a result of difficulties in obtaining species absence data (Hirzel *et al.*, 2002), the majority of species data currently collected consists of presence-only species records (Zaniewski *et al.*, 2002) and have spawned SDMs structured around presence-only or presence-available data. Available species locations are defined as locations in the proximity of the species that are not currently being used. The most common method used to compute presence-only SDMs is based on environmental envelopes (e.g. BIOCLIM; e.g. Busby, 1991; Beaumont *et al.*, 2005) that classify species presence locations, based on environmental conditions, in a multidimensional envelope within space (Pearce & Boyce, 2006) and assign a classification to all locations based on modelled presence or absence (i.e. species core or range (presence) regions and species absence regions – Beaumont *et al.*, 2005).

For wide-ranging animals, a slightly different modelling approach has been applied that contrasts species presence to species available locations to assess how well a habitat is 'used' (Pearce & Boyce, 2006). Often based on radiotelemetry data (Frair *et al.*, 2004), there are four main SDMs approaches based on presence-available data (Pearce & Boyce, 2006): (1) ecological niche factor analysis (ENFA – Hirzel *et al.*, 2004), (2) case-control logic regression analysis (Keating & Cherry, 2004), (3) logistic discrimination model, approximated by a logistic regression algorithm (Keating & Cherry, 2004) and (4) resource selection function models (RSF – Manly *et al.*, 2002). RSF models are of particular interest as have been frequently applied for characterizing species distribution and abundance (Boyce & McDonald, 1999; Nielsen, 2005), species interaction (Johnson *et al.*, 2000) and species diversity (Nielsen *et al.*, 2003).

Specifically, RSF models describe the relative probability of use for various resource units across a study area based on habitat characteristics (Pearce & Boyce, 2006). While RSF models can use presence/absence data, presence/available RSF models have been successful at modelling the relative probability of species occurrence for a variety of species (e.g. Noss *et al.*, 2002; Johnson *et al.*, 2004; Nielsen, 2005). With the growing popularity of presence/available species models (Thomas & Taylor, 2006) and the increasing availability of

spatial data (Beale *et al.*, 2010), advanced model evaluation techniques have become necessary (Hirzel *et al.*, 2006).

Though SDM researchers have identified the need to address spatial effects (i.e. Lobo & Martín-Piera, 2002), spatially explicit assessment of SDM model accuracy is not standard. In part, this dearth of spatial assessments is because many of the methods available for spatial analysis require quantitative data, especially when analysis is local or mappable. However, SDMs outputs are probabilistic and often partitioned into categories of presence and absence. For instance, SDM may group species probability of occurrence, based on a threshold of 0.5, into two categorical outputs: species presence and species absence (Wilson *et al.*, 2005). A local spatial SDM accuracy assessment method is needed that can be applied to probabilistic and categorical SDM outputs.

We demonstrate a spatially explicit approach to quantify and map spatial variability in predictive success of categorical SDMs, using a case study for grizzly bears in the eastern slopes of the Canadian Rockies. Our model validation method is based on a presence/available RSF model but can be applied to all types of SDMs. We mapped locations where observed grizzly bear use was greater than expected based on the RSF model. To do so, we applied a spatially local test statistic to observed data and compared results to an expectation, for the same statistic, generated from a randomization conditioned on the RSF model output. The spatial pattern of unexpected habitat use was quantified using local Moran's I , a measure of spatial autocorrelation that identifies clusters and outliers of locations poorly predicted by the RSF model. Clusters of poorly predicted areas indicate subregions of the study area where the RSF model has under-predicted bear occurrence and may reflect a region for further analysis. As such, environmental conditions at unexpected clusters of locations were evaluated to inform RSF model input variables and attribute relationships.

METHODS

Study area

The 38,705 km² study area is situated along the eastern slopes of the Canadian Rocky Mountains in west-central Alberta (53°25'N, 117°34'W, Fig. 1). The climate is typified by an average temperature range of 11.5°C in the summer to -6.0°C in the winter and an annual precipitation of 538 mm (Beckingham *et al.*, 1996). Local climate is strongly influenced by the elevation, which varies from 770 m to > 3500 m. Because of the short growing season and limited availability of high protein foods, grizzly bear populations occur at relatively low densities (about 4.79 bears per 1000 km²) (Boulanger *et al.*, 2005).

Bear location data

From 1999 to 2003, female grizzly bear telemetry location data were collected as part of the Foothills Research Institute Grizzly



Figure 1 Study area and surrounding towns. The inset map shows the position of the study area within Canada.

Bear Research Program. Bears were captured using leg snaring and immobilization techniques (Stenhouse & Munro, 2000; Cattet *et al.*, 2003). All capture efforts followed the Canadian Council of Animal Care protocols (Animal Use Protocol number 20010016). Bears were collared with either a Televilt (Lindesberg, Sweden) Simplex GPS radio collar or an Advanced Telemetry System (ATS, Isanti, MN, USA) GPS radio collar. The collars download six spatial locations per day, at 4-h intervals, with a spatial accuracy of approximately 10–20 m (Stenhouse & Munro, 2000; Nielsen *et al.*, 2004b; Linke *et al.*, 2005). Telemetry data from 1999 to 2002 were used to develop the RSF models, while data from 2003 were used for model evaluation. Therefore, we were testing the RSF model's ability to predict locations of bears in the following year with data not used in model creation. Only female bears with 50 or more telemetry locations per season were included, as recommended by Leban *et al.* (2001). In total, we used 3570 bear location points from nine bears to evaluate the RSF model (Table 1).

Resource selection function model

Seasonal presence/available RSF models were created by Nielsen (2005) using telemetry data from 1999 to 2002 to evaluate third-order (Johnson, 1980) habitat selection for adult female grizzly bears using logistic regression that estimates relative likelihoods with an exponential model. The spatial resolution of the RSFs was 30 m by 30 m. To account for seasonal variations in habitat use, three separate RSF models were estimated based on major changes in the phenology of vegetation (i.e. forage species) and habitat use patterns (Nielsen *et al.*, 2003, 2004a). The three model seasons are as follows: spring (hypophagia, 1 May to 15 June), summer (early

Table 1 Number of adult female bear radiotelemetry locations per season.

Bear	Bear radiotelemetry locations			Total
	Spring	Summer	Autumn	
g03	96	180	232	508
g10	99	134	142	375
g12	199	170	–	369
g23	106	145	119	370
g28	234	242	273	749
g40	163	181	216	560
g48	–	62	56	118
g60	103	–	–	103
g100	102	135	181	418
Total	1102	1249	1219	3570

hyperphagia, 16 June to 15 August) and autumn (late hyperphagia, 16 August to 15 October). Values from the seasonal RSF models were classified into ten ordinal RSF bins ranging from 1 (relatively low probability of female bear occurrence) to 10 (relatively high probability of female bear occurrence). Environmental variables, recorded at presence and available locations, were used to quantify habitat selection patterns for the remainder of the study area and included land cover type, distance to edge, forest age, compound topographic index, terrain ruggedness index (Nellemann & Cameron, 1996) and global solar radiation (Table 2). [Table 2 includes a brief justification for each covariate, but see Nielsen (2005) for further discussion.] Telemetry points used for evaluation were evenly dispersed throughout the seasons with 1102 occurring

in spring, 1249 occurring in summer and 1219 occurring in autumn.

Supplementary landscape data

Additional landscape data, namely distance to water and elevation, were used to explore characteristics of locations where the observed frequency of bears was statistically unexpected given chance, conditioned on RSF model output. These variables were chosen because of their use in other grizzly bear RSF modelling projects (e.g. Nielsen *et al.*, 2002; Nielsen *et al.*, 2009; Ciarniello *et al.*, 2007). Distance to water was calculated using the straight-line Euclidean distance to the nearest water feature. Elevation was based on a 100- m digital elevation model for the province of Alberta.

Model evaluation

Locations where bear habitat use was statistically higher than predicted by the RSF, which we refer to as *unexpected locations* (i.e. under-predicted; false negative), were identified by comparing the observed spatial pattern of bear telemetry points to a reference distribution of spatial pattern. The reference distribution was generated by a randomization conditioned on the RSF (see Fig. 2). Five steps were used to identify unexpected locations. First, the study area extent was limited to locations available to grizzly bears (*Areal restrictions*). Second, the number of telemetry points for randomization to each RSF bin was calculated using a utilization function (*Calculating telemetry points within each RSF bin*). Third, we randomized telemetry points conditional on the RSF

Table 2 Variables used in the creation of the resource selection functions (RSF) models, as demonstrated by Nielsen (2005).

Variable	Description	Justification
Land cover	Composed of 10 classes; 6 forest classes (closed conifer, open conifer, mixed, deciduous, treed-bog, and regeneration forest), 3 open classes (alpine/herbaceous, non-vegetative, and open-bog/shrub), and 1 anthropogenic class	Proven to be an important variable in other grizzly bear habitat mapping (Nielsen <i>et al.</i> , 2002, 2003)
Distance to forest/non-forest edge	Distance in metres (standardized to 100 m increments) to a forest edge	Accounts for grizzly bears preference for edge habitats (Nielsen <i>et al.</i> , 2004a)
Forest age	Extracted from Alberta Vegetation Index (AVI) data and GIS maps of fire history from Foothills Model Forest (FMF; Hinton, Alberta)	Represents grizzly bear's selection for different forest ages, i.e. young regenerating stands (Nielsen <i>et al.</i> , 2002)
Compound topographic index	Represents soil moisture	Impacts the bear's use of microsites in clearcuts and locations of key food resources (Nielsen <i>et al.</i> , 2004a,c)
Terrain ruggedness	Generated by modifying Nellemann & Cameron's (1996) equation that takes into account changes in aspect and elevation	Used previously to model grizzly bear habitat selection and mortality risk (Nielsen <i>et al.</i> , 2004c)
Global solar radiation	Calculated on 3 days throughout the summer	Combined with slope–aspect relationships has been shown to influence habitat selection of sites within forest stands by grizzly bears (Nielsen <i>et al.</i> , 2002, 2003, 2004a)

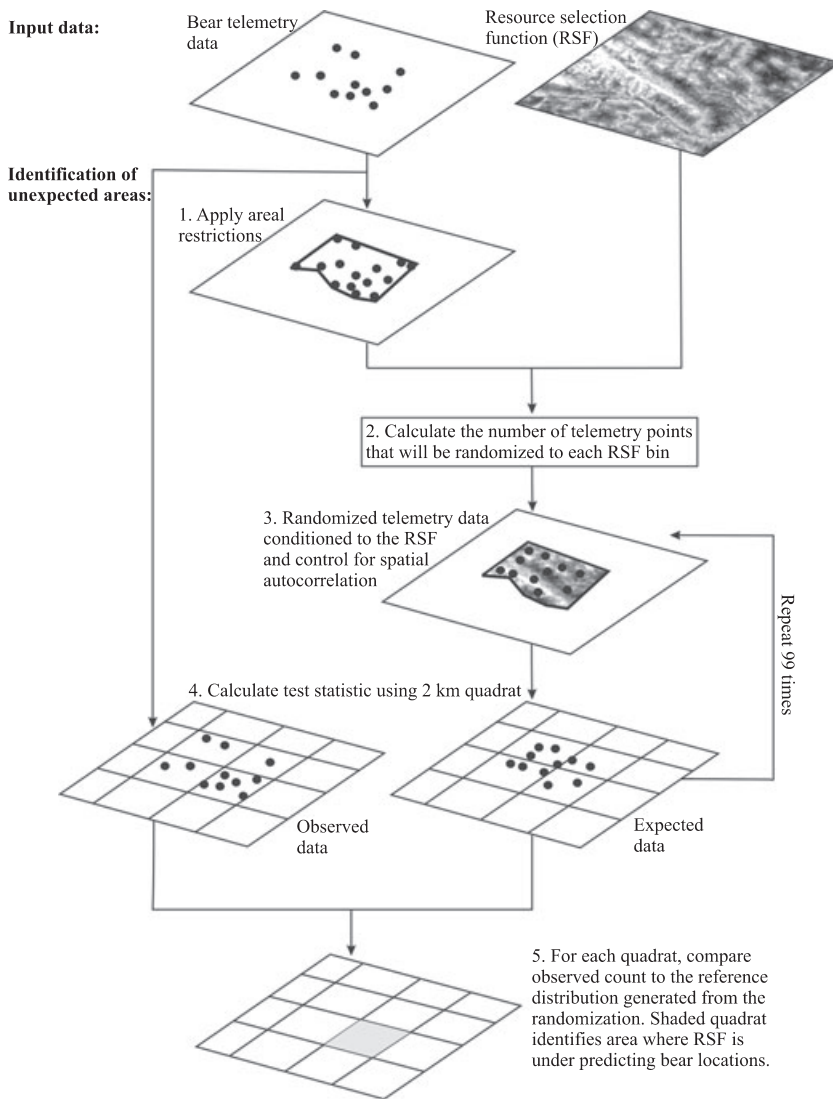


Figure 2 An overview of the methods used to identify unexpected locations.

and accounted for the spatial autocorrelation in the empirical data (*Conditional randomization*). The randomization procedure was repeated 99 times. Fourth, we applied a test statistic, quadrat counts, to the observed data and to all randomization outputs (*Test statistic*). The quadrat count was calculated for RSF bins categorized into groups of high, medium and low RSF values. Fifth, unexpected locations were identified by comparing the observed quadrat counts to the reference distribution of quadrat counts generated from the 99 randomizations (*Identifying unexpected locations*).

Areal restrictions

The spatial extent used for randomization can influence model accuracy assessment results (Lobo *et al.*, 2008). As the area from which available locations are drawn increases, the likelihood that these locations are more environmentally distinct from the species presence location also increases (Lobo *et al.*, 2008). As the RSF model we were evaluating was based on third-order (Johnson, 1980) habitat selection, the

spatial extent of each randomization was limited by fitting a minimum convex polygon (MCP) to each individual bear's telemetry data. MCPs have previously been used to define available locations for species (e.g. Frair *et al.*, 2004) and ensure that telemetry points are only assigned to areas deemed *available* to bears (Nielsen *et al.*, 2002; Buskirk & Millspaugh, 2006; Thomas & Taylor, 2006). We further restricted the study area by excluding water bodies and elevations above 2500 m (rock, snow, and glaciers), as higher elevations are not typically used by adult female bears (McLellan & Hovey, 2001).

Calculating the number of telemetry points expected in each RSF bin

To map locations where observed bear occurrence is unexpected, we first generated maps of expected bear occurrence as a baseline for comparison. A basic premise of an RSF model is that animal occurrence (per unit area use of habitats) should increase with increasing RSF bin value (Nielsen, 2005). The expected number of animal occurrences in each RSF bin

should, therefore, be a result of RSF bin value and its area. To determine how many bears are expected in each RSF bin, we followed Johnson *et al.* (2006) who evaluated a presence/available RSF model by comparing expected bin utilization values or the expected proportion of animal occurrences in each RSF bin, to observed bin animal frequencies. Accordingly, the utilization $U(x_i)$ value for each bin x_i is estimated as

$$U(x_i) = w(x_i)A(x_i) / \sum_j w(x_j)A(x_j), \quad (1)$$

where $w(x_i)$ is the raw midpoint value of the RSF bin x_i , $A(x_i)$ is the area of bin x_i , and x_j is the total number of RSF bins, $x_j = 1, 2, \dots, 10$ (Boyce & McDonald, 1999; Johnson *et al.*, 2006). The number of expected bear locations in each RSF bin (N_i) was calculated as:

$$N_i = N \times U(x_i) \quad (2)$$

where N is the number of observed telemetry points for the particular bear and $U(x_i)$ is utilization function calculated in equation (1). This equation relates a larger number of bear telemetry locations to high RSF bins and a smaller number to low RSF bins.

Conditional randomization

For each RSF bin, we randomized points conditioned on the RSF model and accounted for the spatial autocorrelation present in the empirical data. Wildlife location data are inherently clustered or spatially autocorrelated (Swihart & Slade, 1985; Otis & White, 1999; Cushman *et al.*, 2005). We replicated potential spatial autocorrelation to reduce the likelihood of Type 1 errors (see Fortin & Jacquez, 2000). Given that approximately half the quadrats contained telemetry points, we randomly selected half the quadrats and allocated telemetry data to these quadrats. Data were allocated following an inhomogeneous planar Poisson process (Getis & Boots, 1988) or a Poisson process that varies spatially, conditioned on the RSF model. The process of randomly selecting quadrats and conditionally allocating points was repeated 99 times.

Test statistic

We used quadrat counts as a test statistic and compared the number of telemetry points in each quadrat to the number of conditionally randomized points in the same quadrat. While a test statistic could be calculated for each RSF bin, to aid interpretation we aggregated bins into three equal area classes: low RSF (bins 1–4), medium RSF (bins 5–7) and high RSF (bins 8–10). We selected a quadrat size of 2 km by 2 km to ensure that there were an equal number of quadrats containing empirical telemetry points and quadrats remaining empty, as recommended by Getis & Boots (1988). If quadrats are too small, the likelihood of Type 1 errors will increase, as points may appear more or less clustered than random because of the quadrat size. If quadrats are too large, the likelihood of spatial stationarity within the quadrat is reduced and results will

reflect a mixture of ecological processes and limit interpretation. For each of the three RSF bin groups, the quadrat count was calculated for observed telemetry data and for the 99 randomizations.

Identifying unexpected locations

For each bear, statistically unexpected locations were detected in the three RSF bin groups by comparing the observed quadrat count to a reference distribution of count values at each quadrat generated from the 99 randomizations. An unexpected location occurs when the empirical test statistic (the observed quadrat count) has a higher value than that of the other 99 quadrat count permutations. We use a one-sided 99% ($\alpha = 0.01$) confidence interval based on the highest quadrat count in the 99 randomizations to identify such locations. Three unexpected locations categories were generated: (1) bear use greater than expected in the low RSF values; (2) bear use more than expected in the medium RSF values; and (3) bear use more than expected in the high RSF values.

While the RSF model was assessed using individual animals, results were summarized for all animals by dividing the number of times the quadrat was unexpected by the total number of bears intersecting the quadrat. Integrating individual bear results produced maps of quadrats having RSF values that are always, rarely, or never unexpected.

Exploring model errors (unexpected locations)

The spatial pattern of model error was quantified using local Moran's I that can be used to identify clusters and outliers of extreme values relative to the mean (Anselin, 1995; Nelson & Boots, 2008). Landscape characteristics (distance to water and elevation) underlying the statistically unexpected quadrats were assessed to demonstrate exploring variables contributing to errors of under-prediction. The relative frequency distribution for landscape characteristics within the statistically unexpected quadrats was compared to the distribution of values for all quadrats within the home range. Landscape characteristics were assessed for two grizzly bears for the low RSF values in the spring season. Spring was selected because the spring RSF model was the least accurate at predicting bear occurrence compared to the other seasonal RSF models. The two grizzly bears chosen had overlapping home ranges and the greatest number of statistically unexpected quadrats; thus, their habitat use was particularly problematic for the RSF model's predictive success.

RESULTS

Conditional randomization

Of the 2655 quadrats generated for individual bears, 1456 (55%) represented areas with more than one grizzly bear seasonal home range (Fig. 3). The RSF under-predicted bear occurrence at low RSF values in 152 quadrats (6%), medium

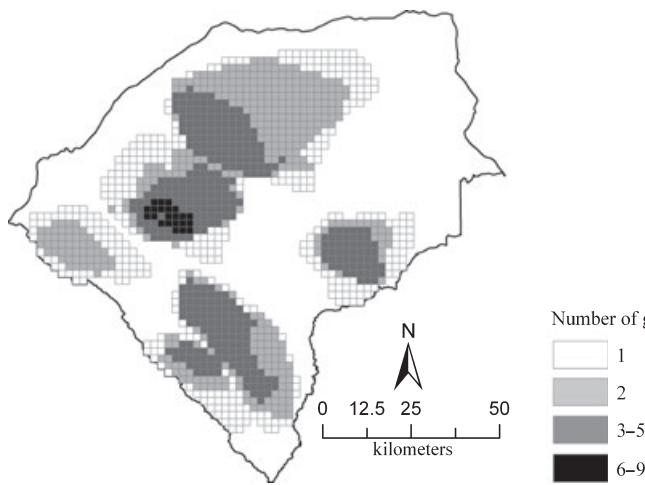


Figure 3 Quadrats used for validating the resource selection functions (RSF) model. Colour indicates the number of individual grizzly bear home ranges observed at each quadrat.

RSF values in 153 quadrats (6%) and high RSF values in 100 quadrats (4%) (Table 3). Seasonally, the spring RSF model contained the greatest proportion (5–9%) of unexpected quadrats (quadrats with more bear presence than expected), compared to 3–4% of quadrats in summer and 4–6% in autumn (Table 4).

Characterization of unexpected locations

The spatial pattern of unexpected quadrats for the RSF models is illustrated in Fig. 4. The unexpected quadrats were spatially clustered, especially in the northeast portion of the study area, while outliers tended to be located on home range peripheries. The frequency distribution of landscape characteristics underlying expected and unexpected quadrats for two grizzly bears in

low RSF values for spring can be seen in Fig. 5. The first bear had 9% of their quadrats identified as unexpected through the randomization process, while the second had 10% identified as unexpected. The cluster of the unexpected quadrats for these bears can be identified as extreme error in the east-central portion of Fig. 4. The frequency distribution of values for elevation for both bears was similar for the unexpected and the expected values (Fig. 6). In contrast, for distances to water, the unexpected quadrats in low RSF values occurred more frequently at distances between 10 and 14 km than the expected quadrats in low RSF values.

Table 3 Quadrats where the number of observed bears is statistically higher than expected based on random process conditioned to the resource selection functions (RSF).

RSF value	Number of quadrats (%)
Low	152 (6)
Medium	153 (6)
High	100 (4)

Table 4 Number of quadrats for each season where the number of observed bears is statistically higher than expected based on random process conditioned to the resource selection functions (RSF).

RSF value	Number of quadrats*		
	Spring (%)	Summer (%)	Autumn (%)
Low	72 (8)	44 (4)	36 (5)
Medium	74 (9)	39 (3)	40 (6)
High	41 (5)	29 (3)	30 (4)

*Total number of quadrats: spring = 896, summer = 1056, autumn = 730.

DISCUSSION

Despite the wide-scale use of SDMs and recognition of the importance of spatial accuracy assessments, majority of SDM evaluation methods are largely aspatial (Zhang *et al.*, 2005). Given that most animal habitats have varying degrees of heterogeneity, and thus are spatially varied, this represents a limitation of such models. We introduce a new spatially explicit approach to evaluating species distributions and demonstrate the approach on an RSF model.

Spatially explicit model evaluation techniques typically require quantitative model output (e.g. Anselin, 1988; Zhang *et al.*, 2005). The primary benefit of our approach is that it enables assessment of model error for quantitative and categorical model output, which is typical for SDMs. Our approach is also useful when observed data are used for assessment of probabilistic model output, because observations at locations of low probabilities are not necessarily errors. A second benefit is that the conditional randomization approach may be used at multiple spatial scales. By varying the grain size, or even shape, of the local test statistics and the extent from which species locations are allocated, researchers could examine the impact of scale and determine at which scale the model is most accurate (Jelinski & Wu, 1996). For example, if we were interested in population-level grizzly bear habitat use, we could increase the grain size of the test statistic quadrats and change the extent to a landscape scale, as opposed to the individual

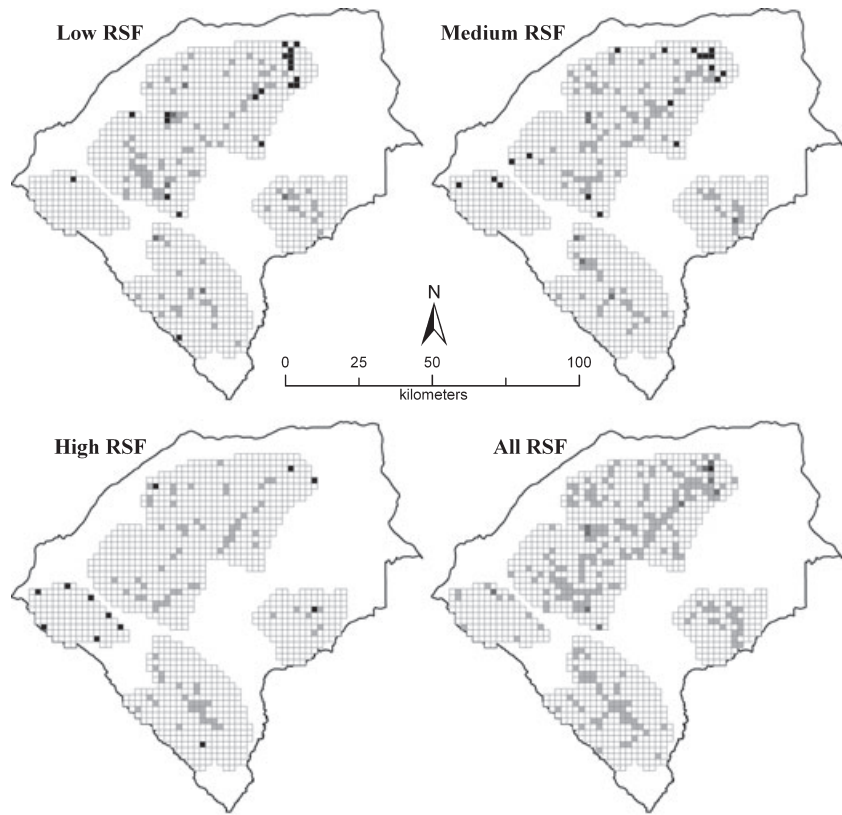


Figure 4 Quadrats with unexpected resource selection functions (RSF) values integrated for all individual bears and seasons. Colour represents the proportion of results indicating unexpected RSF value; calculated as the number of times an individual bear analysis determines values are unexpected divided by the total number of individual bears analyses performed.

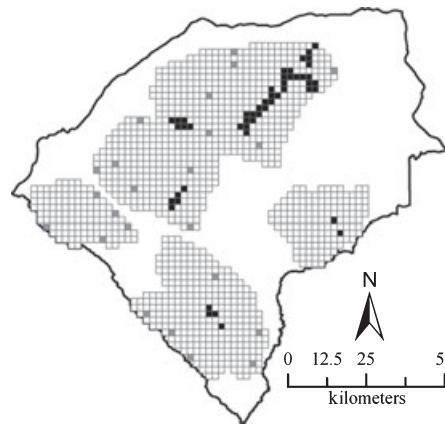
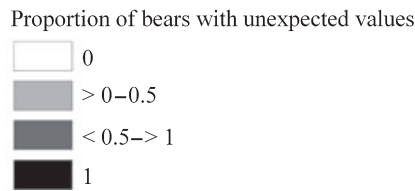
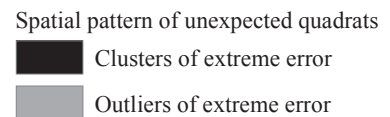


Figure 5 Spatial pattern of statistically unexpected quadrats identified by local Moran's *I*.



bear home ranges. A third benefit is that unlike analysing model residuals, our method is not based on the statistical model used to create the SDM and, as a result, can be applied to SDMs constructed with presence/available, presence-only and presence/absences species locations. Finally, the conditional randomization method enables identification of areas where modelled values are significantly different than expected

based on a null hypothesis. While the most common null hypothesis is that processes are random, other null hypotheses could be tested.

We demonstrated our conditional randomization method for assessing local, spatial SDM accuracy using female grizzly bear presence locations. The bear locations were conditioned on the RSF model, based on the RSF bin utilization value

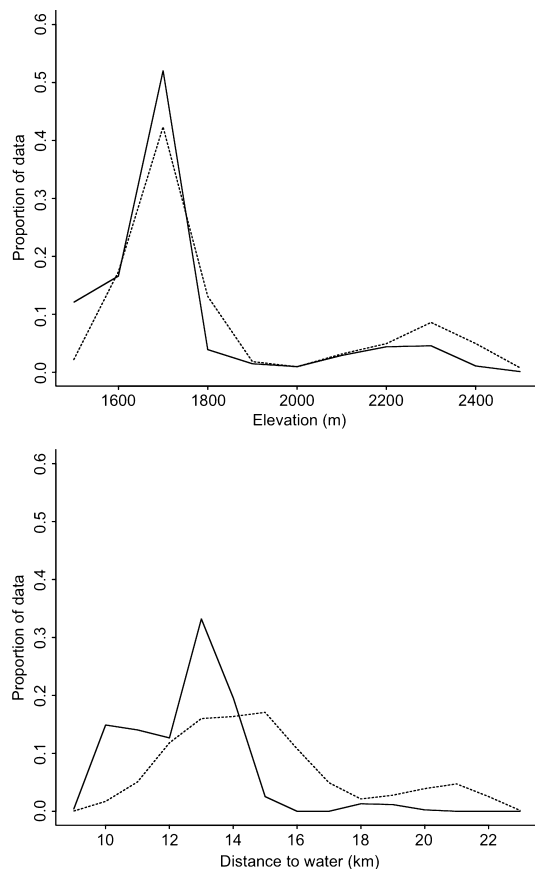


Figure 6 Distribution of elevation and distance to water for two bears in low resource selection functions (RSF) values during spring: unexpected areas are indicated by the solid lines; all locations within the bear's home ranges are indicated by the dotted lines.

($U(x_i)$; the number of bear presence locations that is predicted to be in each RSF bin according to resource preference and area bin). Boyce *et al.* (2002) and Johnson *et al.* (2006) both used the utilization formula combined with k -fold cross-validation to evaluate spatially explicit RSF model predictions. Boyce *et al.* (2002) compared the area-adjusted utilization value for each RSF bin [$U(x_i)$] to the RSF bin rank using Spearman rank correlation. Instead of relying on bin ranking, Johnson *et al.* (2006) calculated the expected number of animal locations in each RSF bin (N_i) from the utilization value [$U(x_i)$] and compared this value to the observed number using linear regression and chi-square tests. Our method differs from that proposed by Boyce *et al.* (2002) and Johnson *et al.* (2006) because we compared the expected number of bear telemetry locations in each RSF bin (N_i) to a randomization conditioned on the RSF model and summarized with a test statistic. As a result, unlike the previous authors who obtained overall, global accuracy measures, we were able locally and spatially evaluate the RSF model.

Using conditional randomization and a grizzly bear presence data set, we demonstrate that the RSF models predict bear occurrence well, as only approximately 5% of the study area

was identified as under-predicting bear occurrence. High RSF values were predicted most correctly in comparison with low and medium RSF values. However, there was model error in all RSF values, indicating that the model was under-predicting bear occurrence at all levels of habitat selection.

Although RSF models performed best in high RSF values, unexpected locations identified in these high values may indicate the need for further scaling of the RSF model output. The high RSF values may, in fact, not be ranked high enough and may be too inclusive. As RSF values are categorical, the actual difference in range of raw RSF values varies between each categorical bin. For example, the difference in raw bear selection values between bins 1 and 2 is smaller than the difference between bins 9 and 10. The number of RSF bins and the range of raw RSF values included in each bin can be arbitrary, and the sensitivity of RSF model results to the RSF bin designation has been noted by other researchers (e.g. Boyce *et al.*, 2002; Hirzel *et al.*, 2006). The RSF input attribute relationships could also be adjusted to better distinguish differences between the high and medium RSF bins. Researchers have noted that similar issues of convoluted habitat preferences result from uneven distribution and proportion of presence/absence sampling (see Real *et al.*, 2006). Irrespective of these sampling issues, identifying locations in high RSF values where bear occurrence is greater than predicted is particularly important when making local management decisions. Locations where predicted use is high and actual use even higher may not be caused by scaling or sampling, but may actually represent important areas for habitat conservation.

The accuracy of the grizzly bear RSF models varied seasonally with the summer RSF model performing best and spring RSF model performing poorest (i.e. rank of performance: summer > autumn > spring). This result contrasts with the finding of Nielsen (2005) who, using Somer's D and Spearman's rank correlation, found that the autumn RSF model performed best and the summer RSF model performed poorest (i.e. performance rank: autumn > spring > summer). The disparity between our spatial and Nielsen's (2005) aspatial accuracy assessment techniques relates to Nielsen's (2005) rank method as it is insensitive to the amount of variation between RSF bin values. By incorporating expected habitat, via conditional randomization, we used probabilistic comparisons that determine whether bear habitat use was greater than expected based on chance conditioning.

Using local spatial statistics, we identified clusters and outliers of significant error in unexpected locations. Unexpected clusters were located in core areas that coincided with increased use by bears. The exception to this is the linear cluster of unexpected quadrats in the northeast portion of the study area, which resulted from a single bear. Relative to other bears, this particular bear's seasonal home range had an exceptionally large spatial extent that was more than twice as large as other bears' ranges (1302 km² compared to an average of 505 km²), and its movements were fundamentally different than that for the other bears. Hence, we consider this bear's movements to be a statistical outlier. Using individual-based

analyses, we were able to identify this cluster as resulting from a bear's behaviour and not as an area of high inaccuracy in the RSF model.

Contrasting to the clusters of error, outliers of error occurred on the peripheries of home ranges, usually in locations of few bears. Outliers are locations where low accuracy is surrounded by high accuracy. In cases where model spatial resolution is fine, outliers may have minimal impact. As model spatial resolution becomes coarser, outliers represent a greater area, and errors may warrant further investigation because model relationships may be unique at these locations.

When the observed spatial pattern of habitat use is highly variable, we recommend considering randomization algorithms that account for empirical variation in the spatial pattern of telemetry locations (e.g. the frequency distribution of the observed number of telemetry points within each quadrat). During randomization, the number of telemetry points to be allocated to each quadrat may be randomly extracted from the frequency distribution and thus will more appropriately mirror the observed bear behaviour. In some instances, it may also be necessary to vary quadrat size when the spatial extent of the home range is inconsistent but the number of telemetry points is similar.

The methods we have outlined can be used to both validate a model and identify adjustments to improve model performance. By identifying locations of model inaccuracy, one can recognize and map errors. Furthermore, one can examine landscape characteristics at these locations that may be causing their deviation from the model's prediction. We demonstrated this examination of characteristics with two variables, elevation and distance to water, both of which were not included in the creation of the RSF model, but were shown to be relevant in other grizzly bear RSF models (e.g. Nielsen *et al.*, 2002, 2009; Ciarniello *et al.*, 2007). While elevation was similar in both expected and unexpected areas, the spring RSF model was under-predicting distances to water between 10 and 14 km. In other words, there were more bear occurrences than expected in low RSF values within this distance to water. Once locations of model inaccuracy are identified, there is an opportunity to spatially adjust model parameters to incorporate observed relationships at poorly predicted locations. For example, bear occurrence may be higher than predicted in some low RSF bins because of the proximity of these locations to high resource abundance. Spatially explicit parameters could be incorporated into the creation of the RSF model by adding a spatial neighbourhood attribute that considers the relationship between the RSF value of a cell and its neighbours. It may be beneficial for a cell with a low RSF value surrounded by high RSF values to receive a higher value in the final RSF model.

While we demonstrated our conditional randomization method using an RSF model, it is applicable to all SDMs, including those constructed from presence/absence data. For example, our method could be used to assess the spatial accuracy of a plant SDM built on species presence/absence data. Relevant areal restrictions may be applied to limit the randomization extent (see Lobo *et al.*, 2008 for relative

occurrence area discussion). The plant presence/absence data could be randomized conditioned on the SDM model such that presence locations are randomized to species occurrence areas and absences randomized to areas where the species is not found. A quadrat could be used as the test statistic to summarize the observed plant presence and absence in comparison with the randomized 'expected' results. Local Moran's *I* could then be applied separately for the presence and the absence data to identify clusters of inaccuracy in model presence and absence predictions. Lastly, environment conditions at the error clusters may be analysed to ascertain which environmental variables could be adjusted or added to better determine species presence and/or absence.

CONCLUSIONS

The risk in using single or summary statistics to validate the performance of a model across a study area has been cautioned by some researchers (e.g. McGwire & Fisher, 2001), because underlying processes may vary geographically, particularly in large, diverse areas (Dale & Fortin, 2002). Our findings support this notion of spatial variability in model accuracy. With the growth of large datasets, assessing model spatial accuracy is becoming increasingly critical (McGwire & Fisher, 2001). Information on geographical variations in model accuracy provided in the metadata may be important for making local land use decisions. The conditional randomization method presented here has several advantages. The main advantage over other spatial analysis approaches is that it is designed for use with categorical and probabilistic model output. It enables local and mappable evaluation of accuracy, is independent of thresholds (unlike confusion matrix-based techniques), and can be applied to both presence/absence and presence/availability models. Furthermore, when integrated with additional data, accuracy assessment results can provide information on how to improve future model output.

ACKNOWLEDGEMENTS

Funding was primarily provided by the Foothills Research Institute Grizzly Bear Project. Many thanks to all those associated with funding, data collection and data preparation. Research was also funded by NSERC. Thanks to Karen Laberee for editorial support and two reviewers for their comments on a previous version of this paper.

REFERENCES

- Anselin, L. (1988) *Spatial econometrics: methods and models*. Kluwer Academic, Dordrecht.
- Anselin, L. (1995) Local indicators of spatial association. *Geographical Analysis*, **27**, 93–115.
- Anselin, L. (2002) Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, **27**, 247–267.

- Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modeling. *Journal of Biogeography*, **33**, 1677–1688.
- Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**, 413–423.
- Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecological Letters*, **13**, 246–264.
- Beaumont, L.J., Huges, L. & Poulsen, M. (2005) Predicting species distribution: use of climate parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, **186**, 250–269.
- Beckingham, J.D., Corns, I.G.W. & Archibald, J.H. (1996) *Field guild to ecosites of west-central Alberta*. Natural Resources of Canada, Canadian Forest Service, Northwest Region. Northern Forest Centre, Edmonton, Alberta, Canada.
- Boulanger, J., Stenhouse, G.B., Proctor, M., Himmer, S., Paetkau, D. & Cranston, J. (2005) *2004 Population inventory and density estimates for the Alberta 3B and 4B Grizzly Bear Management Area*. Alberta Sustainable Resources Development, Fish and Wildlife Division, Hinton, Alberta, Canada.
- Boyce, M.S. & McDonald, L.L. (1999) Relating populations to habitats using resource selection functions. *Trends in Ecology and Evolution*, **14**, 268–272.
- Boyce, M.S., Vernier, P.R., Nielsen, S.E. & Schmieglow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, **157**, 281–300.
- Brottons, L., Thuiller, W., Araujo, M.B. & Hirzel, A.H. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.
- Busby, J.R. (1991) BIOCLIM: a bioclimatic analysis and prediction system. *Plant Protection Quarterly*, **6**, 8–9.
- Buskirk, S.W. & Millsbaugh, J.J. (2006) Metrics for studies of resource selection. *Journal of Wildlife Management*, **70**, 358–366.
- Cattet, M.R.L., Christison, K., Caulkett, N.A. & Stenhouse, G.B. (2003) Physiologic responses of grizzly bears to different methods of capture. *Journal of Wildlife Diseases*, **39**, 649–654.
- Ciarniello, L.M., Boyce, M.S., Heard, D.C. & Seip, D.R. (2007) Components of grizzly bear habitat selection: density, habitat, roads and mortality risk. *Journal of Wildlife Management*, **71**, 1446–1457.
- Cushman, S.A., Chase, M. & Griffin, C. (2005) Elephants in space and time. *Oikos*, **109**, 331–341.
- Dale, M.R.T. & Fortin, M.-J. (2002) Spatial autocorrelation and statistical tests in ecology. *Ecoscience*, **9**, 162–167.
- Fielding, A.H. (2002) What are the appropriate characteristics of an accuracy measure? *Predicting species occurrence* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 271–280. Island Press, Washington.
- Fortin, M.-J. & Edwards, G. (2001) Delineation and analysis of vegetation boundaries. *Spatial uncertainty in ecology: implications for remote sensing and GIS applications* (ed. by C.T. Hunsaker, M.F. Goodchild, M.A. Friedl and T.J. Case), pp. 158–174. Springer, New York.
- Fortin, M.-J. & Jacquez, G. (2000) Randomization tests and spatially autocorrelated data. *Bulletin of the Ecological Society of America*, **81**, 201–205.
- Frair, J.L., Nielsen, S.E., Merrill, E.H., Lele, S.R., Boyce, M.S., Munro, R.H.M., Stenhouse, G.B. & Beyer, H.L. (2004) Removing GPS collar bias in habitat selection studies. *Journal of Applied Ecology*, **41**, 201–212.
- Getis, A. & Boots, B. (1988) *Point Pattern Analysis Scientific Geography Series*, vol 8. Sage Publications, Newbury Park.
- Gusian, A. & Zimmerman, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hirzel, A.H., Helfer, V. & Métral, F. (2001) Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, **145**, 111–121.
- Hirzel, A.H., Hausser, J., Chessel, D. & Perrin, N. (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, **83**, 2027–2036.
- Hirzel, A.H., Posse, B., Oggier, P.-A., Crettenand, Y., Glenz, C. & Arlettaz, R. (2004) Ecological requirements of reintroduced species and the implications for release policy: the case of the bearded vulture. *Journal of Applied Ecology*, **41**, 1103–1116.
- 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*, **199**, 142–152.
- Jelinski, D.E. & Wu, J. (1996) The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, **11**, 129–140.
- Johnson, D.H. (1980) The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, **61**, 65–71.
- Johnson, B.K., Kern, J.W., Wisdom, M.J., Findholt, S.L. & Kie, J.G. (2000) Resource selection and spatial separation of mule deer and elk during spring. *Journal of Wildlife Management*, **64**, 685–697.
- Johnson, C.J., Seip, D.R. & Boyce, M.S. (2004) A quantitative approach to conservation planning: using resource selection functions to map the distribution of mountain caribou at multiple spatial scales. *Journal of Applied Ecology*, **41**, 238–251.
- Johnson, C.J., Nielsen, S.E., Merrill, E.H., McDonald, T.L. & Boyce, M.S. (2006) Resource selection functions based on use-availability data: theoretical motivation and evaluation methods. *Journal of Wildlife Management*, **70**, 347–357.
- Keating, K.A. & Cherry, S. (2004) Use and interpretation of logistic regression in habitat selection studies. *Journal of Wildlife Management*, **68**, 774–789.
- Kéry, M., Royle, J.A., Plattner, M. & Dorazio, R.M. (2009) Species richness and occupancy estimation in communities subject to temporary emigration. *Ecology*, **90**, 1279–1290.
- Laffan, S.W. (1999) *Spatially assessing model error using geographical weighted regression*. *GeoComputation 99*. Available at: http://www.geocomputation.org/1999/086/gc_086.htm.

- Leban, F.A., Wisdom, M.J., Garton, E.O., Johnson, B.K. & Kie, J.G. (2001) Effect of sample size on the performance of resource selection analysis. *Radio-tracking and animal populations* (ed. by J.J. Millsbaugh and J.M. Marzluff), pp. 291–307. Academic Press, New York.
- Linke, J., Franklin, S.E., Huettmann, F. & Stenhouse, G.B. (2005) Seismic cutlines, changing landscape metrics and grizzly bear landscape use in Alberta. *Landscape Ecology*, **20**, 811–826.
- Lobo, J.M. & Martín-Piera, F. (2002) Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conservation Biology*, **16**, 158–173.
- Lobo, J.M., Jimenez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.
- Lobo, J.M., Jimenez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.
- MacKenzie, D.M. & Royle, J.A. (2005) Designing efficient occupancy studies: general advice and tips on allocation of survey effort. *Journal of Applied Ecology*, **42**, 1105–1114.
- 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*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- McGwire, K.C. & Fisher, P. (2001) Spatial variable thematic accuracy: beyond the confusion matrix. *Spatial uncertainty in ecology: implications for remote sensing and GIS applications* (ed. by C.T. Hunsaker, M.F. Goodchild, M.A. Friedl and T.J. Case), pp. 308–329. Springer, New York.
- McLellan, B.N. & Hovey, F.W. (2001) Habitats selected by grizzly bears in a multiple use landscape. *The Journal of Wildlife Management*, **65**, 92–99.
- Nellemann, C. & Cameron, R.D. (1996) Effects of petroleum development on terrain preferences of calving caribou. *Arctic*, **49**, 23–28.
- Nelson, T.A. & Boots, B. (2005) Identifying insect infestation hot spots: an approach using conditional spatial randomization. *Journal of Geographical Systems*, **7**, 291–311.
- Nelson, T.A. & Boots, B. (2008) Detecting spatial hot spots in landscape ecology. *Ecography*, **31**, 556–566.
- Nicholls, A.O. (1989) How to make biological surveys go further with generalized linear models. *Biological Conservation*, **50**, 51–75.
- Nielsen, S.E. (2005) *Habitat ecology, conservation and projected population viability of grizzly bears (Ursus arctos L.) in west-central Alberta, Canada*. Dissertation. Department of Biological Studies, University of Alberta, Edmonton, Alberta, Canada.
- Nielsen, S.E., Boyce, M.S., Stenhouse, G.B. & Munro, R.H.M. (2002) Modeling grizzly bear habitats in the yellowhead ecosystem of Alberta: taking autocorrelation seriously. *Ursus*, **13**, 45–56.
- Nielsen, S.E., Boyce, M.S., Stenhouse, G.B. & Munro, R.H.M. (2003) Development and testing of phenologically driven grizzly bear habitat models. *Ecoscience*, **10**, 1–10.
- Nielsen, S.E., Boyce, M.S. & Stenhouse, G.B. (2004a) Grizzly bear and forest I. Selection of clearcuts by grizzly bears in west-central Alberta, Canada. *Forest Ecology and Management*, **199**, 51–65.
- Nielsen, S.E., Herrero, S., Boyce, M.S., Mace, R.D., Benn, B., Gibeau, M.L. & Jevons, S. (2004b) Modelling the spatial distribution of human-caused grizzly bear mortalities in the Central Rockies ecosystem of Canada. *Biological Conservation*, **120**, 101–113.
- Nielsen, S.E., Munro, R.H.M., Bainbridge, E.L., Stenhouse, G.B. & Boyce, M.S. (2004c) Grizzly bears and forestry II. Distribution of grizzly bear foods in clearcuts of west-central Alberta, Canada. *Forest Ecology and Management*, **199**, 67–82.
- Nielsen, S.E., Cranston, J. & Stenhouse, G.B. (2009) Identification of priority areas for grizzly bear conservation and recovery in Alberta, Canada. *Journal of Conservation Planning*, **5**, 38–60.
- Noss, R.F., Carroll, C., Vance-Borland, K. & Wuerthner, G. (2002) A multicriteria assessment of the irreplaceability and vulnerability of sites in the Greater Yellowstone Ecosystem. *Conservation Biology*, **16**, 895–908.
- Olden, J.D., Jackson, D.A. & Peres-Neto, P.P. (2002) Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society*, **131**, 329–336.
- Osborne, P.E., Foody, G.M. & Suarez-Seoane, S. (2007) Non-stationarity and local approaches to modeling the distribution of wildlife. *Diversity and Distribution*, **13**, 313–323.
- Otis, D.L. & White, G.C. (1999) Autocorrelation of location estimates and the analysis of radiotracking data. *The Journal of Wildlife Management*, **63**, 1039–1044.
- Pearce, J.L. & Boyce, M.S. (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, **43**, 405–412.
- Pontius, R.G. & Schneider, L.C. (2001) Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, **85**, 239–248.
- Raes, N. & ter Steege, H. (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727–736.
- Real, R., Barbosa, A.M. & Vargas, J.M. (2006) Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, **13**, 237–245.
- Stenhouse, G.B. & Munro, R.H.M. (2000) *Foothills Model Forest Grizzly Bear Research Program: 1999 Annual Report*. pp. 107. Hinton, Alberta.
- Swihart, R.K. & Slade, N.A. (1985) Testing for independence of observations in animal movements. *Ecology*, **66**, 1176–1184.
- Thomas, D.L. & Taylor, E.J. (2006) Study designs and tests for comparing resource use and availability II. *Journal of Wildlife Management*, **70**, 324–336.

- Wilson, K.A., Westphal, M.I., Possingham, H.P. & Elith, J. (2005) Sensitivity of conservation planning to different approaches to using predictive species distribution data. *Biological Conservation*, **122**, 99–112.
- Wulder, M.A., White, J.C., Coops, N.C., Nelson, T. & Boots, B. (2007) Using local spatial autocorrelation to compare outputs from a forest growth model. *Ecological Modelling*, **209**, 264–276.
- Zaniewski, A.E., Lehmann, A. & Overton, J.M. (2002) Predicting species spatial distribution using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, **157**, 261–280.
- Zhang, L., Gove, J.G. & Heath, L.S. (2005) Spatial residual analysis of six modeling techniques. *Ecological Modelling*, **186**, 154–177.
- Zhu, A.-X. (2001) Modeling spatial variation of classification accuracy under fuzzy logic. *Spatial uncertainty in ecology: implications for remote sensing and GIS applications* (ed. by C.T. Hunsaker, M.F. Goodchild, M.A. Friedl and T.J. Case), pp. 330–349. Springer, New York.

BIOSKETCH

Mary Smulders is a research assistant in the Spatial Pattern Analysis and Research (SPAR) Laboratory, Geography Department, University of Victoria, Canada and a GIS Analyst for Living With Lions, Kenya. Her primary interest is developing and employing spatial analysis techniques to study animal, mainly carnivore, movement and habitat use.

Author contributions: All aspects of the project were lead by M.S. with assistance from the other authors. T.A.N. also advised on methods, while D.E.J. and S.E.N. contributed to biological interpretations and applications. G.B.S. provided the data.

Editor: David Richardson