# Measuring and selecting scales of effect for landscape predictors in species–habitat models

Amanda E. Martin[1] and Lenore Fahrig

*Geomatics and Landscape Ecology Laboratory (GLEL), Ottawa-Carleton Institute of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6 Canada*

*Abstract.* Wildlife managers often use habitat models to determine species habitat requirements and to identify locations for conservation efforts, uses which depend on accurate specification of species–habitat relationships. Prior study suggests that model performance may be influenced by the way we measure environmental predictors. We hypothesized that species responses to landscape predictors are best represented by landscape composition-based measurements, rather than distance-based measurements. We also hypothesized that models using empirical data to select an appropriate scale of effect for each habitat predictor (multi-scale models) should perform better than models that assume a common scale of effect for all predictors (single-scale models). To test these hypotheses we constructed habitat models for three mammal species, *Mephitis mephitis*, *Mustela erminea*, and *Procyon lotor*, based on surveys conducted in 80 landscapes in southeastern Ontario, Canada. For each species we compared the performance of distance- and composition-based measurements, and we compared the performance of single- and multi-scale models. The composition-based measurement, measured at its empirically determined scale of effect, had greater explanatory power than the distance-based measurement of a given predictor more often than expected by chance, supporting our first hypothesis. Contrary to expectation, multi-scale models did not have better explanatory power or predictive performance relative to single-scale models. We identified and evaluated four potential mechanisms to explain this, and, depending on the species, we found that the best explanation was either that predictors have significant effects at a common scale or that, although the modeled effects were at multiple scales, they were of similar magnitude and direction at the scales modeled in single- and multi-scale models. Our results suggest that habitat modeling based on distance-based measurements could be improved by including composition-based measurements of landscape predictor variables, but that inclusion of predictor-specific scales of effect for composition-based measurements does not necessarily improve performance over models including composition-based measurements at a single scale. Conservation and wildlife management may be simplified when single-scale models perform as well as multi-scale models, as this suggests actions conducted at a single scale may address management objectives as well as actions taken at different scales for different landscape features.

Key words: *classification and regression tree; generalized additive model; generalized linear model; habitat suitability; landscape composition; landscape structure; mid-sized mammal; nearest-neighbor distance; occupancy model; receiver operating characteristic analysis; spatial scale; species distribution.*

## Introduction

Habitat modeling is commonly used in conservation planning and wildlife management. Habitat models are statistical frameworks that use empirical data to fit relationships between species occurrence or abundance and environmental predictor variables to define the direction and shape of the species response and make predictions of habitat suitability. These models have been used to identify landscape features and geographic locations for conservation efforts and to indicate the relative importance of different environmental condi-

tions (e.g., Peeters and Gardeniers 1998, Filipe et al. 2004, Balbontín 2005), which may be used to direct future research and prioritize among conservation objectives.

Although widespread in application, the ability of these models to accurately predict species habitat requirements and geographic distributions has been questioned (e.g., Beutel et al. 1999, Barry and Elith 2006). Models that fail to predict actual suitable habitat may underestimate species distributions and lead to lost opportunities for conservation; alternatively, models that predict suitable habitat where there is none may result in negligible outcomes for the focal species, as efforts are focused on unsuitable areas. Despite the known shortcomings, habitat models provide a useful tool, particularly where full empirical assessments of

[1] E-mail: Amanda.Martin@glel.carleton.ca

species distributions are not possible. In recognition of their utility, researchers have focused on understanding the factors influencing model performance, and improving study design and data analyses (e.g., MacKenzie et al. 2002, Segurado and Araújo 2004, Meynard and Quinn 2007).

One factor influencing the performance of habitat models is how the environmental predictor variables are measured. Measurement of landscape predictors in habitat models can generally be classified as either distance or composition based. Distance-based measurements measure the distance from the species sampling location to a given type of landscape feature; the most straightforward is the nearest neighbor distance, measured as the Euclidean distance to the nearest instance of that feature (e.g., distance to forest). Composition-based measurements are the amount of the landscape feature within a given area, commonly measured as its proportional cover, density, or count (e.g., proportional grass cover, housing density, or number of wetlands). Theoretical and empirical studies suggest that composition-based measurements of patch isolation are generally better predictors of dispersal success than are distance-based measurements of patch isolation (Moilanen and Nieminen 2002, Bender et al. 2003, Tischendorf et al. 2003). Given that species distributions are at least partly related to dispersal success, we hypothesized that composition-based habitat measurements should also predict species distributions better than distance-based habitat measurements.

When using composition-based measurements, the performance of habitat models depends on the selected scale, or size of the area within which the predictor values are assessed. Theory, supported by empirical research, suggests that species respond to their environmental context at characteristic scales (the "scale of effect"; Wiens 1989, Holland et al. 2004). If so, then misspecification of the scale may result in rejection of predictors that do influence the species, or predicted species–habitat relationships that are statistical artifacts. Despite this danger, the selection of scale appears to be arbitrary for most habitat models (Wheatley and Johnson 2009). Others rely on a priori selection of a scale based on some assumed link to the species biology (e.g., home range or territory size, dispersal distance; Rubbo and Kiesecker 2005, Müller et al. 2009). However, this still entails a risk of misspecification; for example, selecting the scale based on home range size when the true scale of effect is related to dispersal distance leads to underrepresentation of the landscape structure. Even if the correct biological link is assumed, the appropriate scale may be poorly estimated, because species traits may vary across the species distribution (e.g., Nilsen et al. 2005). Alternatively, the scale of effect may be estimated empirically, by assessing how the relationships between species occurrence (or abundance) and the predictor(s) change with scale.

Estimation of a scale of effect using empirical data is complicated when there are multiple landscape predictors because it is not clear whether the scale should be estimated independently for each predictor (hereafter referred to as "multi-scale" [MS] models) or if all landscape variables should be measured at a common scale ("single-scale" [SS] models). Literature to date suggests that MS models can outperform SS models (Graf et al. 2005, Boscolo and Metzger 2009, Kuhn et al. 2011), implying that scales of effect are predictor specific. This may occur if different environmental conditions are important at different life stages; for example, one landscape feature may limit dispersal success while another may influence foraging opportunities within the home range, leading to a larger scale of effect for the first predictor. In addition, the predictors themselves may influence their scale of effect; for example, scales of effect for predictors acting through predator–prey interactions may reflect the mobility of the interacting species.

Despite the support for MS models cited in the above paragraph, in some cases there is little difference between SS and MS models (Graf et al. 2005). If the SS habitat model for a species is as good as the MS model, management actions could be conducted at a single scale, greatly simplifying management. We propose the following four possible circumstances in which SS models should perform as well as MS models, and we outline methods by which they can be evaluated.

1) Species show little or no response to the candidate predictors. SS models are likely to perform similarly to MS models when the candidate predictors are poor predictors of species occurrence (or abundance), and the estimated scales of effect are due to chance. In this case, we should see a lack of significant effects of predictors across all scales (Fig. 1a).

2) Species response is not dependent on scale. Here, the species is affected by the environmental conditions similarly across all scales. For example, a species may avoid denning and foraging near roads, select a home range where road density is low, and have a low probability of surviving dispersal in regions where the chance of road mortality is high, resulting in negative responses to roads at all scales. In this case, we should detect significant effects of similar magnitude across all assessed scales (Fig. 1b).

3) Species response occurs at a common scale. If species occurrence (or abundance) is limited by a single life stage, one would expect the species to respond to environmental conditions at the scale relevant to space use during that stage. In this case, we should see a significant response to the predictors at the same (or similar) scale in SS and MS models. Since in reality species may exhibit a similar response over a range of scales, cases where SS and MS models estimate scales of effect within a range of adjacent, significant scales are considered functionally equivalent (Fig. 1c).
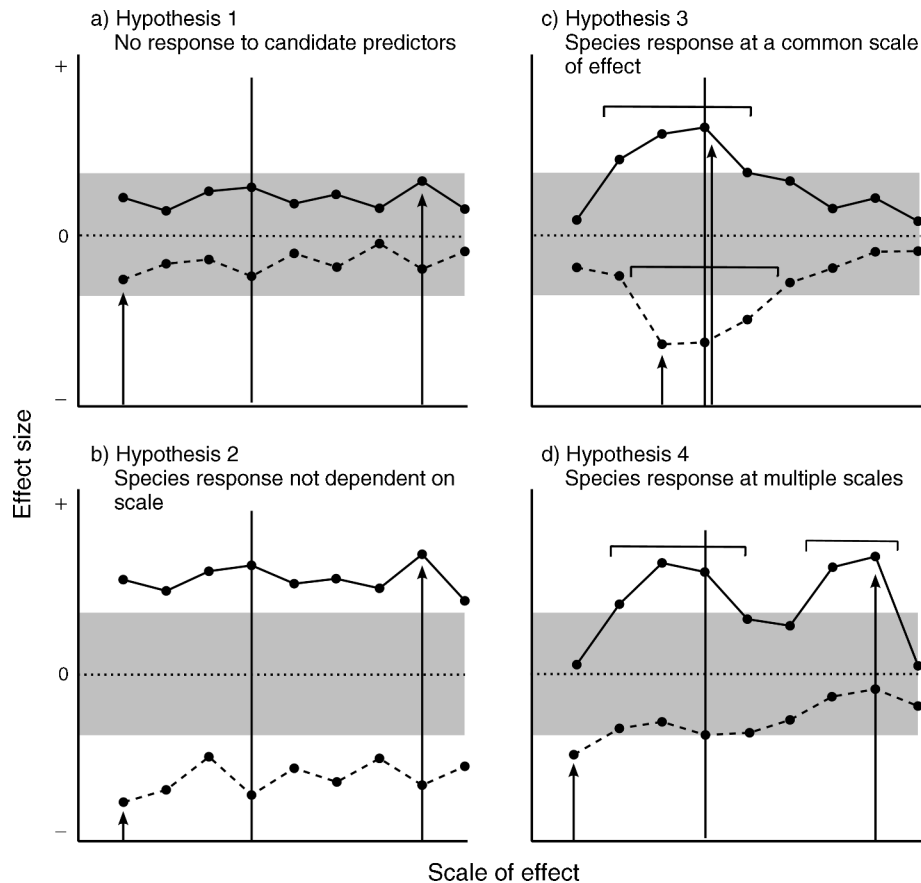
Fig. 1. Hypothetical examples of the changes in predictor variable effects with scale that would be consistent with each of the four mechanisms hypothesized to explain similarities in the performance of single- and multi-scale habitat models. The points represent assessed scales, and the lines show the predicted changes in the effect size for predictor variables with changes in scale for a predictor variable with a positive relationship to the species response (solid line), and a variable with a negative relationship (dashed line). Significant effects on the species response occur in the unshaded area, and nonsignificant effects occur in the gray-shaded area of the plot. Vertical lines indicate the apparent scales of effect for single-scale models, and arrows indicate the scales of effect for multi-scale models. Square brackets indicate the ranges of functionally equivalent scales, defined as adjacent scales having significant predictor effects.

4) Species have a similar response at several scales. Species may display preference (or avoidance) for a landscape predictor at multiple scales (Chittaro 2004, Altmoos and Henle 2010), suggesting that the predictor may be relevant to space use at more than one life stage. The strongest scale of effect for a predictor may be different from the SS scale of effect, but if the predictor also has a strong effect at the SS scale (in the same direction), the SS model may be as good a predictor as the MS model. The data would support this hypothesis if the scales of effect for a predictor modeled in the SS and MS models, although not equivalent, produce effects of similar magnitude and direction at the SS and MS scales (Fig. 1d).

If the predictor has strong cross-scale correlations, it is possible that the species responds at a single scale, but the relationship is also well represented at other scales. In this case, one will be unable to interpret whether the

effect occurs at a single scale or at multiple scales. Conversely, if a predictor shows similar effects at multiple scales and cross-scale correlations in the predictor are low, one can say it is likely because the species responds at multiple scales.

The purpose of this study was to (1) compare the performance of distance- and composition-based measurements of landscape predictor variables, (2) compare the performance of SS and MS models, and (3) assess the hypothesized mechanisms for explaining cases where MS models do not outperform SS models. To do this, we constructed habitat models using remotely sensed landscape data and surveys of three terrestrial mid-sized mammals, *Mephitis mephitis* (striped skunk), *Mustela erminea* (ermine), and *Procyon lotor* (raccoon). For each species we assessed univariate models for candidate predictor variables, entered as both distance-based measurements, and as composition-based measurements across a range of spatial scales, using four statistical
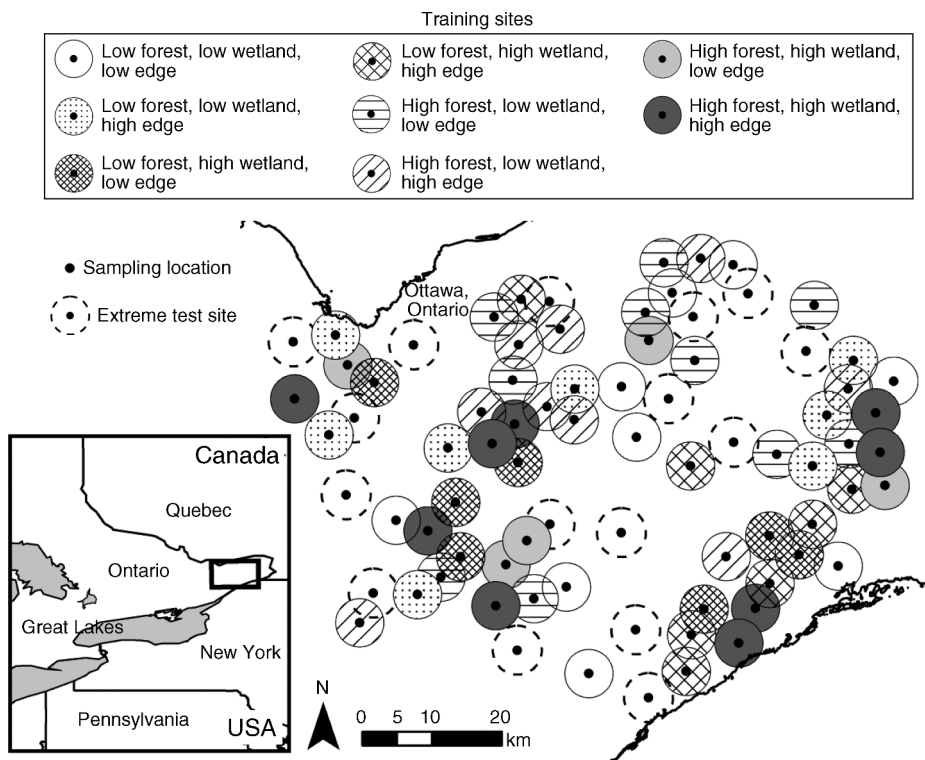
FIG. 2. Distribution of sampling sites in southeastern Ontario, Canada. Circles represent a 3.45 km radius. Internal patterns identify the eight combinations of low or high forest amount, wetland amount, and edge density for the 64 training sites. The 16 extreme test sites (circles with a dashed outline) have values of forest amount, wetland amount, and/or edge density outside the ranges for the training sites (circles with a solid outline), and were used to test extrapolation of the habitat models.

frameworks: occupancy models (OCC), generalized linear models (GLZ), generalized additive models (GAM), and classification and regression trees (CART). We tested the prediction that a composition-based measurement, measured at its appropriate scale, has greater explanatory power than its corresponding distance-based measurement. We then created SS and MS models for each statistical framework. For species where multi-scale models did not outperform single-scale models, we assessed our four hypothesized mechanisms (Fig. 1).

## MATERIALS AND METHODS

### Study area

We used footprint tracking stations to sample mid-sized mammals in southeastern Ontario, Canada, with all sites falling within ~60 km of the city of Ottawa (45°25′ N, 75°42′ W; Fig. 2). The study area lies within the Mixedwood Plains ecozone, which is characterized by deciduous and mixed deciduous–coniferous forests interspersed with alvars, grasslands, and wetlands (Wiken 1986) and has the highest human population density in Canada, with much of the landscape converted to agriculture and development. Although the smallest of Canada's ecozones, it is home to the largest number of imperiled species (Gibbs et al. 2009).

Mid-sized mammals we expected to sample included striped skunk, ermine, raccoon, *Sylvilagus floridanus*, *Marmota monax*, *Erethizon dorsatum*, *Mustela frenata*, and *Martes pennanti*.

Landscape features were described using remotely sensed data sets. Vector data sets represented the locations of roads (OMNR 2004) and streams (NRC 2002). We classified land cover at a resolution of 15 m, based on the Southern Ontario Land Resource Information System (SOLRIS; OMNR 2002), updated to include classifications of agricultural regions (J. Pasher et al., *unpublished manuscript*) and wetlands (OMNR 2009). We used high resolution (25 cm) orthorectified digital photos (OMNR 2008) to digitize expansion of the built-up class for the study region to account for urban sprawl since SOLRIS was created. These photos were also used to classify areas undefined in the amalgamated map and to digitize all buildings at least 10 m in width in a point vector data set, within a 3.45 km radius around all sampling locations.

### Site selection

We selected 80 survey sites; 80% (64) were selected to train the habitat models ("training sites"), and the remaining 20% to test the predictive performance of the models outside the data ranges of the training sites ("test sites"; Fig. 2). Hereafter, we use "site" to refer to the

sampling location plus its surrounding landscape, where the "sampling location" is the point where mammals were surveyed, and the "landscape" is the circular area centered on the sampling location within which the landscape predictors were measured.

Training sites were selected to minimize multicollinearity among predictors and nonindependence of sampling locations. We minimized multicollinearity among forest amount, wetland amount, and edge density. Forest and wetland amounts were their proportional covers within the landscape, and edge density was the total length of edges between all pairs of forest, wetland, grass, and crop fields, divided by the area of the landscape. We first ran a moving window analysis in FRAGSTATS (McGarigal et al. 2002) such that each pixel in our raster map of the study region became the center of a potential training site, and we calculated values for each of forest amount, wetland amount, and edge density at three circular window sizes (0.5, 1.0, and 2.0 km radius). We then defined ranges of "low" and "high" values for each variable, based on their distributions and correlations among them, and selected sites representing all eight combinations of low and high for each of the three variables at each of the three scales. To minimize correlations between the landscape predictors and conditions at the sampling locations, we distributed sites in each combination of the three landscape variables among five cover types at the sampling location (forest, wetland, built-up, grass, and crop) as evenly as possible. To avoid spatial dependence among sampling locations, all were separated by at least 4 km, since individuals were unlikely to be sampled at multiple sites if the distance between sites was greater than the home range width. Four kilometers exceeds the expected widths of home ranges for the target species (with a maximum mean home range of 6.55 km$^2$ for *M. pennanti*; Koen et al. 2007). This selection process for the training sites resulted in approximately eight sites per combination of the three landscape predictors, although this was not possible for some rare combinations (range from five sites with high forest, high wetland, and low edge to 10 sites with high forest, low wetland, and low edge).

The purpose of the 16 test sites was to test the performance of the habitat models built using the training site data when extrapolated to predictor values outside the ranges of the training site data. In selecting the training sites (see previous paragraph), the ranges of the predictor variables were restricted relative to their full potential ranges, because they were specifically selected to avoid multicollinearity. We therefore selected the 16 test sites to represent extreme values of the landscape predictors.

### Field data collection

We sampled mid-sized mammals using footprint tracking stations from 10 May to 27 August 2010, sampling each sampling location once per month (i.e., a
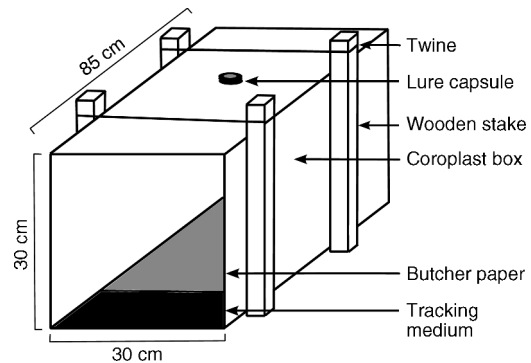


Fig. 3. Schematic of the enclosed tracking box used to sample mammals at footprint tracking stations.

total of four samples for each site). To sample, a single tracking station was erected, left active for one week, and removed from the sampling location at the end of that week. To minimize the effects of temporal variation, we placed sites into 10 groups, ensuring that sites with the same combination of forest, wetland, and edge were never in the same group. Each month the groups were randomly ordered such that five groups were sampled in the first two weeks of a month, and the remaining five groups in the second two weeks.

Mammals were surveyed using enclosed tracking boxes to exploit the tendency of many species to explore covered spaces (Foresman and Pearson 1998, Loukmas et al. 2003, Ebeling 2006, Rytwinski and Fahrig 2011). We constructed tracking boxes from coroplast, 85 × 30 × 30 cm (Fig. 3). We tied them with twine to four wooden stakes hammered into the ground. Tracking plates (29 × 81 cm) made from 0.16 cm thick polystyrene plastic were fitted to the bottom of the tracking boxes and partially covered with butcher paper, leaving ~12 cm of the plastic exposed at either end. The exposed plastic was painted with the tracking medium, a mixture of powdered carbon black and paraffin oil in a ratio of 1:3 by mass.

A lure was used to entice mammals into the vicinity of the tracking station. We chose a lure over bait to reduce the chance of the same individual revisiting the station (as there was no food reward associated with the station) and to maintain similar attractiveness over the week (which would be unlikely for baited stations). Kishel's Grub Essence (Kishel's Quality Animal Scents and Lures, Butler, Pennsylvania, USA) was placed in a housing capsule of galvanized plumbing supplies and suspended from the roof of the coroplast box, to avoid animals ingesting the lure.

Individuals moving through the station stepped in the tracking medium and transferred footprints to the butcher paper. Footprints were identified to species, using Elbroch (2003) and the Carleton University footprint library (Ottawa, Ontario, Canada). Sampling procedures followed guidelines from the Canadian

Council on Animal Care and were approved by the Carleton University Animal Care Committee.

### Data analysis

Models relating detection/non-detection to candidate predictors were created for the three species detected in at least 10% of the training sites. Statistical analyses were conducted in R (R Development Core Team 2011), unless otherwise specified.

We selected candidate predictors for each species based on literature review, and the species response was related to the candidate predictors using four statistical frameworks: OCC (MacKenzie et al. 2002), GLZ (McCullagh and Nelder 1989), GAM (Hastie and Tibshirani 1990), and CART (Breiman et al. 1984). We fit OCC models using the RMark R library (Laake et al. 2012) to interface with the MARK software package (White and Burnham 1999). This OCC model uses the detection history over four samples at each site and a multinomial maximum likelihood procedure with a logit link to model the probability of occurrence ($\psi$) and the probability of detection ($p$) as a function of the predictor(s). GLZ and GAM related the response to predictor(s) using the logit link and binomial distribution. For GAM, we used the gam R library (Hastie 2011) with a smoothing spline with four degrees of freedom. CART models were implemented in SPSS (2010) using the Gini index as a splitting criterion and pruning trees through cross-validation to select the smallest tree with a misclassification risk within one standard error of the minimum, where the misclassification risk estimates the proportion of data for which the tree predicts the incorrect class.

*Comparison of distance-based and composition-based measurements (Objective 1).*—To compare the explanatory power of distance-based and composition-based measurements, we determined whether the distance-based or the composition-based representation of each landscape predictor explained more of the variability in the species response, for each species and statistical framework (OCC, GLZ, GAM, and CART). Distance-based measurements were calculated in ArcGIS 10 (ESRI 2010), and composition-based measurements were calculated in the Geospatial Modeling Environment (*available online*).[2]

To make these comparisons we first had to identify the scale of effect for the composition-based measurement of each landscape predictor (for each species and statistical framework). Predictor values were calculated at multiple scales by systematically increasing the landscape radius (centered on the sample point) from 0.15 km to 3.45 km in increments of 0.15 km, for a total of 23 scales. To arrive at the 0.15 increment, for each predictor we initially calculated predictor values in increments of 0.03 km and calculated Spearman rank

---

correlations ($r_S$) between values at adjacent scales, increasing the increment by 0.03 km, and selected the largest increment for which $r_S > 0.70$ for all predictors. It is important to note that we are not attempting to pinpoint an exact scale of effect, but rather a general region of scales where the effect is strongest (e.g., a scale of effect of 1.50 km means the effect occurs within the range 1.35 to 1.65 km).

We estimated the scale of effect for each predictor by fitting univariate models at each scale and selecting the scale that maximized Nagelkerke $R^2$ ($R^2_N$, for OCC, GLZ, and GAM; Nagelkerke 1991) or minimized the misclassification risk (for CART). For OCC, we modeled $\psi$ as a function of the predictor, and a constant $p$. For CART models, in cases where the minimum misclassification risk could not distinguish among scales, we selected the scale that maximized the proportion of correct classifications, averaged across classes. If this still did not distinguish among scales, we selected the smallest scale from the set.

Once we had identified the scale of effect for the composition-based measurement of each predictor we calculated $R^2_N$ (or misclassification risk) for univariate models fitting the species response to the distance-based measurement of each predictor (for each species and statistical framework). For each statistical framework we tested the hypothesis that composition-based measurements had greater explanatory power than distance-based measurements using a $\chi^2$ test on the number of cases where each measurement type produced a higher $R^2_N$ (lower risk).

*Comparison of single-scale and multi-scale models (Objective 2).*—To compare SS and MS models, we first determined the "best" multivariate SS and MS models for each species and statistical framework, considering all candidate landscape predictors for the species plus the local cover type. For each landscape predictor we included the composition-based measurement at its empirically determined scale of effect (from Objective 1). Multicollinearity among landscape predictors for each species and model combination (SS and MS versions for OCC, GLZ, GAM, and CART) was assessed using the variance inflation factor (VIF) and pairwise $r_S$. For OCC, GLZ, and GAM we used multi-model inference based on Akaike Information Criterion for small sample sizes (AIC$_c$), using the RMark R library (Laake et al. 2012) for OCC and the MuMIn R library (Barton 2011) for GLZ and GAM. For GLZ and GAM we ran all subsets of predictors and retained all models that differed from the minimum AIC$_c$ ($\Delta$AIC$_c$) by $\leq 2$; if this included the null (intercept-only) model, we removed this framework from further consideration. For OCC, we did not estimate all possible models of the effects of all predictors on both the probability of occurrence, $\psi$, and the probability of detection, $p$, due to the prohibitive number of combinations (65 536). Instead, we first ran all subsets including predictors of $\psi$ only and selected a preliminary model (or set of

---

models) with $\Delta AIC_c \leq 2$. If the preliminary model set did not include the null model, we ran all possible combinations of the preliminary model set plus the predictors in these models as predictors of $p$, retaining all models with $\Delta AIC_c \leq 2$. We assessed residual autocorrelation (Global Moran's $I$) for the standardized residuals from all retained models in ArcGIS 10 (ESRI 2010). Models with significantly autocorrelated residuals were fit with an additional autocovariate parameter (Augustin et al. 1996). Variable selection in CART is intrinsic to the modeling process; thus, for CART, all variables included in the decision tree for the training data were included in the best model. The best SS model (or set of models) for each species and statistical framework was determined using the same approach that we used for the selection of the best MS model (or set of models), but for each model we determined the scale of effect for composition-based measurements based on the global multivariate model, and all composition-based measurements were entered at the same scale rather than at their individual scales of effect.

To compare the best SS model (or set of models) to the best MS model (or set of models) for each species and statistical framework, we compared explanatory power, predictive performance, and predictive performance during model extrapolation. Explanatory power was the $R_N^2$ for the training data for OCC, GLZ, and GAM, using model-weighted average $R_N^2$ ($\overline{R_N^2}$) when more than one model was retained (Boscolo and Metzger 2009), and using the misclassification risk for the training data for CART. Predictive performance was the area under the receiver operating characteristic (ROC) curve (AUC; Hanley and McNeil 1982), calculated using the PresenceAbsence R library (Freeman 2007). For each of the eight model combinations (SS and MS versions of OCC, GLZ, GAM, and CART) for each species, we used random sampling cross-validation, dividing the 64 training sites into two subsets 1000 times, where 80% of sites were used to estimate the parameters and make predictions for the remaining 20% of sites. For OCC, GLZ, and GAM we used model-weighted average predictions from the retained models. Predictive performance during extrapolation was estimated by using the models parameterized with the training sites to make predictions for the extreme test sites and comparing AUC between SS and MS models for a given species and framework using the bootstrap method in the pROC R library, with 5000 samples (Robin et al. 2011).

*Comparison of hypotheses for cases where the multiscale model does not outperform the single-scale model (Objective 3).*—We determined which of the hypothesized mechanisms were most likely driving any similarities in MS and SS model performance. For this we first selected the modeling framework with the highest predictive performance for a given species: GLZ for striped skunk and GAM for ermine and raccoon (see *Results*). We then examined each landscape predictor

common to the best SS model (or model set) and best MS model (or model set), looking for patterns consistent with each of the hypothesized mechanisms, as in Fig. 1. We limited comparisons to the predictors common to the best SS and MS models, as the hypothesized mechanisms (with the exception of hypothesis 1) depend on responses to predictor variables in both the SS and MS models. For GLZ we used the standardized regression coefficient ($\beta$) and its significance at all assessed scales from univariate models for each composition-based landscape predictor (see Objective 1 above) to estimate change in effect with scale. For GAM we used the range of the smoothed function as a measure of the effect size and its significance at all scales from the univariate models. Similarity in the direction and magnitude of relationships for a given predictor at the SS and MS scales was calculated as the $r_S$ between the predicted probabilities of occurrence from the univariate model with the predictor measured at the SS scale and the predicted probabilities from the univariate model with the predictor measured at the MS scale; positive correlations of greater than 0.70 were interpreted as high similarity in the predictor effects at the SS and MS scales. If a predictor was significant at only some (ranges of) scales and produced effects of similar magnitude and direction at the SS and MS scales (i.e., consistent with hypothesis 4), we calculated the $r_S$ between the predictor variable measured at the SS scale and at the MS scale; $r_S$ values less than 0.70 were interpreted as support for a conclusion of a species response at multiple scales.

## RESULTS

Three target species were detected at >10% of training sites: striped skunk (30%), ermine (19%), and raccoon (61%). Forest, wetland, grass, crop, housing, road, and edge were assessed for skunk (Storm 1972, Larivière et al. 1999, Bixler and Gittleman 2000, Larivière and Messier 2000) and raccoon (Pedlar et al. 1997, Prange et al. 2003, Henner et al. 2004, Barding and Nelson 2008). For ermine we assessed forest, wetland, grass, crop, housing, road, and stream (Debrot and Mermod 1983, Murphy and Dowding 1994, Alterio et al. 1998, Lisgo 1999, Sidorovich and Solovej 2007). Distance- and composition-based measurements were made for each candidate landscape predictor (Table 1). For all species, local cover type (forest, wetland, built-up, grass, or crop) at the sampling location was included as a categorical candidate predictor, to address the effects of local conditions.

### Comparison of distance-based and composition-based measurements (Objective 1)

Composition-based measurement of a predictor variable, when measured at the appropriate scale of effect, generally had greater explanatory power than the distance-based measurement of the same predictor. Models with composition-based measurements outper-

TABLE 1. Distance- and composition-based measurements of landscape predictor variables used to model species–habitat relationships for striped skunk, ermine, and raccoon, with units of measurement in parentheses.

| Feature type | Predictor | Distance-based measurement | Composition-based measurement |
|---|---|---|---|
| Land cover | forest, wetland, grass, crop | Euclidean distance from the sampling location to the edge of the nearest raster cell classified as the given cover type (km) | proportional cover of the given cover type within the landscape |
| Linear feature | road, edge,† stream | Euclidean distance from the sampling location to the nearest point along the linear feature (km) | total length of the feature divided by the area of the landscape (km/km$^2$) |
| Point feature | housing | Euclidean distance from the sampling location to the nearest instance of the feature (km) | total number of points divided by the area of the landscape (per km$^2$) |

† Includes edges between all pairs of forest, wetland, grass, and crop fields.

formed models with distance-based measurements significantly more often than expected by chance for all statistical frameworks (OCC, $\chi^2 = 5.76$, df $= 1$, $P = 0.016$; GLZ, $\chi^2 = 3.86$, df $= 1$, $P = 0.050$; GAM, $\chi^2 = 17.19$, df $= 1$, $P < 0.001$; CART, $\chi^2 = 13.76$, df $= 1$, $P < 0.001$); overall, out of 84 cases (three species × four statistical frameworks × seven landscape predictors), the model with the composition-based measurement made at the empirically determined scale of effect had a greater $R^2_N$ (or lower misclassification risk) than the analogous model with the distance-based measurement 70 times (Appendix A).

### Comparison of single-scale and multi-scale models (Objective 2)

Comparisons of the best multivariate SS and MS habitat models for each species and statistical framework showed no improvement in model performance for MS relative to SS models (see Appendix B for details of the models selected during the model selection process). After removing the raccoon SS and MS GLZ and OCC models (as the most supported model set included the null model), we found that MS models did not increase the explanatory power relative to SS models. Out of 10 cases (three species × four statistical frameworks, minus the two removed raccoon model frameworks) MS models produced larger $\overline{R^2_N}$ (or smaller misclassification risk) five times and smaller $\overline{R^2_N}$ (or larger risk) five times (Fig. 4).

MS models also failed to improve predictive performance relative to SS models. In the random sampling cross-validation trials, comparisons between SS and MS models for a given species and statistical framework showed the median AUC for MS models was generally greater than that for SS models, with the exception of the skunk GAM, ermine CART, and raccoon CART, where the MS model had a lower median AUC (Fig. 5). Despite the generally greater median AUC for MS models, model performance varied widely among trials, and overlap in the distribution of AUC values between SS and MS models suggests there was little difference in performance, and that the performance of SS relative to MS models depended on the specific split of the data set.

The conclusion that MS models did not improve performance relative to SS models was supported by the ROC analysis for models extrapolated to the test sites (Fig. 6). During extrapolation, SS models predicted the probability of occurrence of the species as well as, or better than, MS models. For all three species, there were no differences in the AUC of SS and MS models across statistical frameworks (all $P > 0.05$; Appendix C), with the exception of the comparison of the SS and MS CART for ermine, where the SS model had a larger AUC than the MS model ($D = 1.96$, $P = 0.050$).

We note that there was no need for addition of an autocovariate parameter or removal of highly correlated predictors for any of the models presented above. Moran's $I$ tests for residual autocorrelation in the best SS and best MS OCC, GLM, and GAM were nonsignificant for all species ($P > 0.05$; Appendix D). VIF for landscape predictors for each species and model combination (SS and MS versions for OCC, GLZ, GAM, and CART) suggested multicollinearity was within acceptable levels (VIF $< 10$), although in some cases (particularly for SS models), VIF indicated moderate collinearity (Appendix D). Although overall multicollinarity was within acceptable levels, pairwise comparisons among landscape predictors for each species and model combination revealed moderate to high collinearity between housing and road densities in most species and model combinations, with high collinearity ($r_S > 0.70$) for the raccoon MS OCC and GLZ models (Appendix D).

### Comparison of hypotheses for cases where the multi-scale model does not outperform the single-scale model (Objective 3)

Since the MS models did not outperform the SS models for any species, we further evaluated the three species to determine which (if any) of the hypothesized mechanisms were driving similarities in model performance, using the statistical framework with the greatest predictive performance during cross-validation and limiting analyses to the predictors common to the best SS model (or set of models) and the best MS model (or set of models). Although performance varied widely
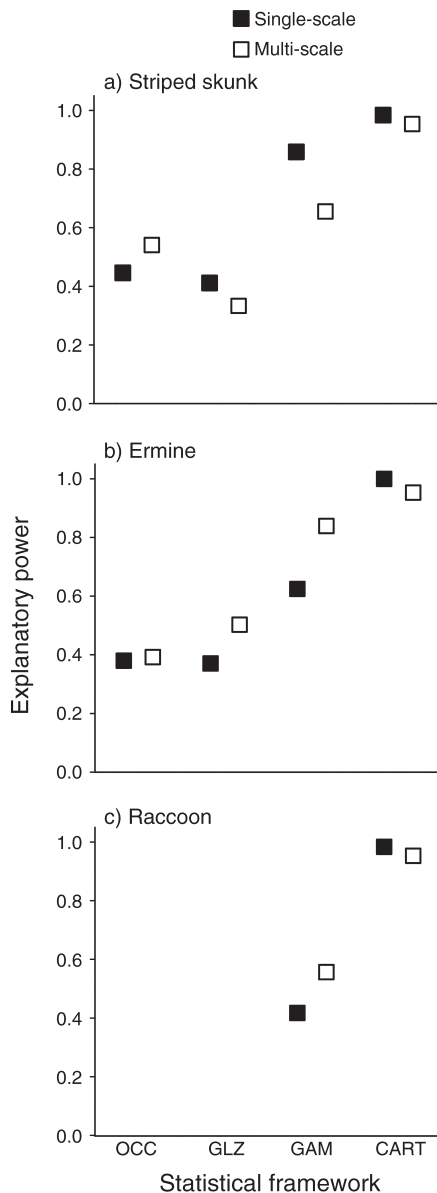
FIG. 4. Explanatory power of single- and multi-scale habitat models predicting probability of occurrence for *Mephitis mephitis* (striped skunk), *Mustela erminea* (ermine), and *Procyon lotor* (raccoon), measured by Nagelkerke $R^2$ ($R_N^2$) for occupancy (OCC), generalized linear (GLZ), and generalized additive (GAM) models, and by 1 − misclassification risk for classification and regression tree (CART) models. Greater explanatory power is indicated by larger values. For cases where more than one model was retained ($\Delta AIC_c \leq 2$), we used model-weighted average values, weighting the $R_N^2$ by the Akaike weight for the model.

MS GAM producing the greatest median AUC, although the SS GAM did not perform better than the MS GLZ or the SS and MS OCC (Fig. 5).

For striped skunk, performance was most consistent with the hypothesis that the species has a similar response to the environmental context at several scales (hypothesis 4); however, interpretation of this as a
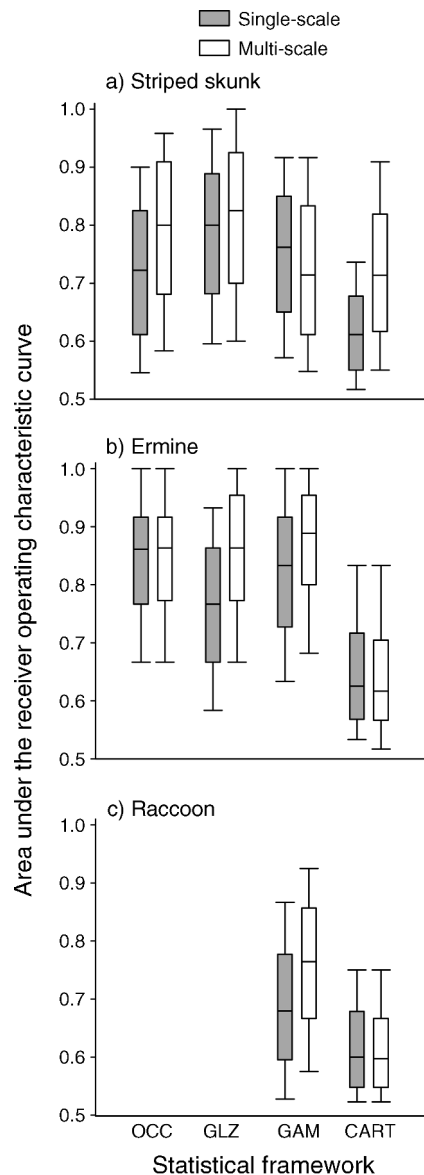


FIG. 5. Predictive performance of single- and multi-scale habitat models predicting the probability of occurrence for striped skunk, ermine, and raccoon under random sampling cross-validation (1000 iterations), measured by the area under the receiver operating characteristic curve. Relationships were modeled using occupancy (OCC), generalized linear (GLZ), generalized additive (GAM), and classification and regression tree (CART) models. In the boxplots, the horizontal line shows the median, the box represents the 25th and 75th percentiles, the top whisker ranges from the 75th to the 90th percentile, and the bottom whisker ranges from the 25th to the 10th percentiles.

during cross-validation trials, SS and MS GLZ produced greater median AUC values relative to SS and MS OCC, GAM, and CART for striped skunk, and SS and MS GAM produced greater median AUC values relative to SS and MS CART for raccoon (Fig. 5). For the ermine we selected the GAM framework based on the
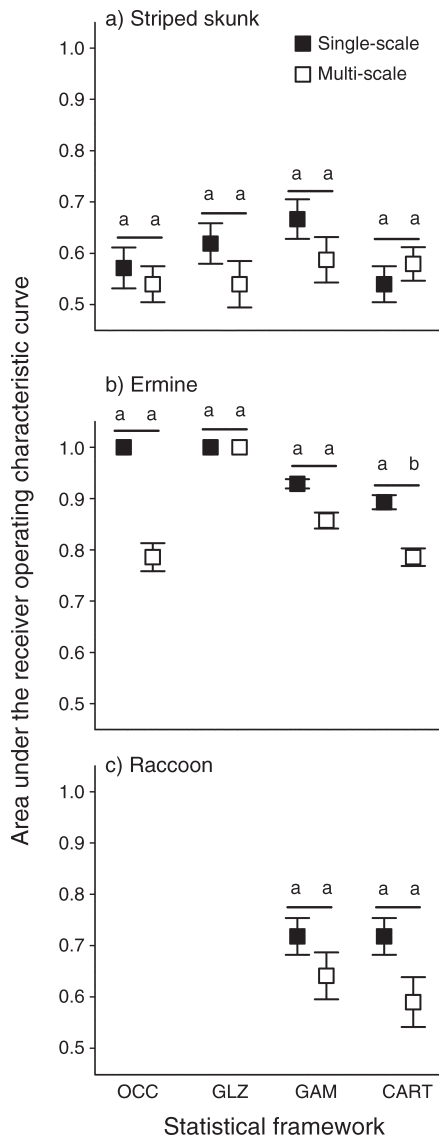
Fig. 6. Predictive performance of single- and multi-scale models predicting probability of occurrence for striped skunk, ermine, and raccoon and for the extreme test data set ($n = 16$), measured by the area under the receiver operating characteristic curve ($\pm$SE). Relationships were modeled using occupancy (OCC), generalized linear (GLZ), generalized additive (GAM), and classification and regression tree (CART) models. Lowercase letters denote differences for pairwise comparisons of single- and multi-scale models ($\alpha = 0.05$).

response at multiple scales was confounded by high cross-scale correlations in the predictor at the SS and MS scales. Hypothesis 4 predicted significant effects at only some scales (or ranges of scales), lack of functionally equivalent scales of effect in SS and MS models, and effects of similar magnitude and direction at the SS and MS scales. Our support for hypothesis 4 was based on the species response to proportional forest, wetland, and crop covers and edge density, which were

included in both the SS and MS model set (Table 2). Changes in $\beta$ with scale suggest the species responded to forest cover at two ranges of scales, peaking at 0.45 km (the scale of effect for MS models) with a second, smaller peak at 1.65 km (the scale of effect for SS models; Fig. 7), and forest cover had effects of similar magnitude and direction at these scales, as indicated by the positive correlation between the probabilities of occurrence predicted by forest cover at the SS and MS scales ($r_S = 0.74$). Although proportional wetland cover, crop cover, and edge density did not result in significant responses at the scale of effect modeled in SS models, the similarities in the magnitude and direction of effects for these predictors in the SS and MS models likely enhanced the similarity in model performance driven by forest cover; there were high positive correlations between the probabilities of occurrence predicted by wetland cover at the SS and MS scales ($r_S = 0.87$), crop cover at the SS and MS scales ($r_S = 0.81$), and edge density at the SS and MS scales ($r_S = 0.77$). However, interpretation of this as a response at multiple scales was confounded by high cross-scale correlations in the predictor at the SS and MS scales for forest, wetland, and crop covers and edge density, with $r_S$ values equal to that of the correlations between predictions stated in the previous sentences. Proportional grass cover and housing density had nonsignificant effects at all assessed scales (all $P > 0.10$); however, this does not support hypothesis 1 in this case (i.e., that the species does not respond to the predictors), since both SS and MS models did include predictors with significant effects.

For ermine, the most likely explanation for the similarity in SS and MS model performance was hypothesis 3 (i.e., the species responds to the environmental context at a common scale of effect). This hypothesis predicted significant effects over a limited range of scales and a functionally equivalent scale of effect in SS and MS models. Support for this hypothesis was based on the species response to housing density. Housing density was the only predictor included in both the SS and MS GAM models (Table 2), and the same scale of effect (1.20 km) was found for housing density in the SS and MS models (Fig. 7).

For raccoon, the most likely explanation for the similarity in SS and MS model performance also appears to be hypothesis 3. Support for this hypothesis was based on the species response to proportional crop cover, which, along with local cover type and housing density, were the only predictors included in both the SS and MS GAM models (Table 2). The same scale of effect (1.50 km) was found for crop cover in the SS and MS models (Fig. 7). Although housing density had nonsignificant effects at all assessed scales (all $P > 0.10$), this provides little support for hypothesis 1 in this case (i.e., that the species does not respond to the predictors), since both SS and MS models did include predictors with significant effects; however, we note that model performance, while within the reasonable range for the

TABLE 2. Summary of predictor variables selected (shown with an X) through multi-model inference for single-scale (SS) and multi-scale (MS) generalized linear models (for striped skunk *Mephitis mephitis*) and generalized additive models (for ermine *Mustela erminea* and raccoon *Procyon lotor*), where $n$ indicates the number of models included in model averaging ($\Delta AIC_c \leq 2$).

| Species and model | $n$ | Local cover | Forest | Wetland | Grass | Crop | Housing | Road | Edge | Stream |
|---|---|---|---|---|---|---|---|---|---|---|
| **Striped skunk** | | | | | | | | | | |
| SS | 6 | | X | X | X | X | X | X | X | |
| MS | 10 | | X | X | X | X | X | | X | |
| **Ermine** | | | | | | | | | | |
| SS | 2 | | | | | X | X | X | | |
| MS | 1 | | | X | | | X | | | X |
| **Raccoon** | | | | | | | | | | |
| SS | 2 | X | | | | X | X | | | |
| MS | 4 | X | X | | X | X | X | | X | |

MS model (AUC > 0.70; Swets 1988), was poor for the SS model.

## DISCUSSION

### Comparison of distance-based and composition-based measurements (Objective 1)

We predicted that the composition-based measurement of a given landscape predictor, measured as the proportional cover or density at the appropriate scale of effect, would have greater explanatory power than a distance-based measurement of the same predictor. This prediction was largely based on published comparisons of distance- and composition-based measurements of patch isolation (Moilanen and Nieminen 2002, Bender et al. 2003, Tischendorf et al. 2003), under the assumption that species distributions are, at least in part, related to dispersal success. Our results support this prediction. We suggest that habitat modeling based on distance-based measurements could be improved by including composition-based measurements of landscape predictor variables.

The higher performance of composition- over distance-based measurements of landscape predictors is intuitive, as distance-based measurements only account for the proximity to attractive or avoided landscape features, and not the amount of the feature in the landscape. The amount of a given landscape feature, and the corresponding abundance of resources or deleterious agents associated with that feature, are likely to be important for population persistence in a given location. For example, as feeding may temporarily deplete local food availability, increasing foraging area is likely to promote persistence as individuals can switch feeding grounds when food availability is low and spread feeding pressure over a greater area, allowing resources to regenerate in depleted regions (Watts 1998). Measuring distance to alternate feeding grounds would be a poor indicator of their availability since nearby feeding areas could cover only a small area, and thus, not significantly supplement the local population, while extensive feeding areas farther away could provide important food supplementation. Similarly, increasing road density may decrease the probability of persistence in an area, as the number of crossings during daily movements is likely to increase with roads in the landscape, increasing the probability of road mortality (Beringer et al. 1990, Gibbs and Shriver 2005). Measuring distance to the nearest road may be a poor indicator of this effect; if there is only one nearby road, the probability of mortality may be lower than if there are several roads somewhat farther away from the sampling location.

In this study we used a relatively simple distance-based measurement (nearest neighbor distance), selected because of its frequent use in habitat models (e.g., Clark et al. 1993, Balbontín 2005, Rainho and Palmeirim 2011). Despite its simplicity, we feel it likely that the relative performance of distance- and composition-based measurements would be similar for more complex distance-based metrics. In a simulation study to determine which isolation metrics best correlated with immigration rates, a more complex distance-based metric (Voronoi polygon area) did not improve performance over the nearest neighbor distance (Bender et al. 2003). Similarly, mensurative studies have found no consistent improvement in performance between the nearest neighbor distance and more complex distance-based metrics of isolation, including area-weighted nearest neighbor distance, Voronoi polygon area, summed Euclidean distance to all other patches, and the summed exponentially decreasing function of distance between patches (Moilanen and Nieminen 2002, Matter et al. 2005).

We note that, for some statistical frameworks, the species response may be best represented at a larger scale than was assessed in this study (i.e., at a scale greater than 3.45 km). In particular, for predictors with a maximum $R_N^2$ (or minimum risk) at 3.45 km, such as proportional wetland and crop cover for skunk GLZ, proportional wetland cover for ermine GAM, and edge density for raccoon GAM, we can only infer that the species response to the composition-based measurement
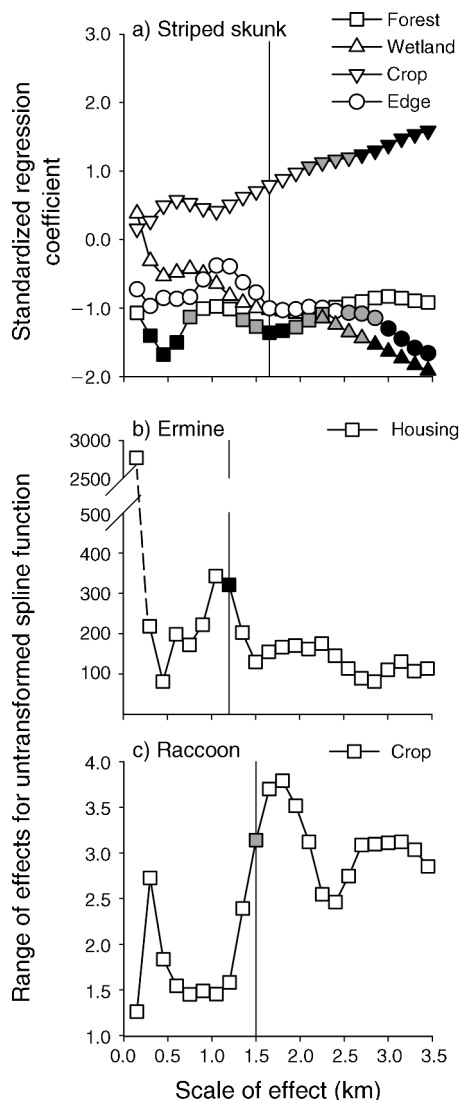
FIG. 7. Variation in predictor effects with scale, where scale is measured as the radius of the circular landscape (km). Values are shown for univariate models for the predictors common to the best single- and multi-scale models for the generalized linear modeling framework (for striped skunk) and the generalized additive modeling framework (for ermine and raccoon). For striped skunk, predictor effects are represented by the standardized regression coefficients; for ermine and raccoon, predictor effects are represented by the range of the untransformed spline function. Only predictors which had significant effects at at least one of the assessed scales ($\alpha = 0.10$) are shown. Symbol fills indicate significance at $\alpha = 0.05$ (black), at $\alpha = 0.10$ (gray), or nonsignificant (white). Vertical lines identify the scales selected for single-scale models.

was best represented at 3.45 km or larger. Ideally, analyses would have continued until a point at which a maximum could be reached; however, analyses of larger scales were not possible, due to limitations in the available land cover data and high overlap of landscapes. Nevertheless, inclusion of additional scales would likely not have altered the results of our

comparison of distance- and composition-based measurements of landscape predictors. If a predictor represented by the composition-based measurement at 3.45 km would have been better represented at a larger scale, and had a greater $R_N^2$ (or lower risk) than the distance-based measurement at 3.45 km, the larger scale would also perform better than the distance-based measurement, as this larger scale would necessarily have a greater $R_N^2$ (or lower risk) than at 3.45 km. Conversely, if this predictor had a lower $R_N^2$ (or greater risk) at 3.45 km relative to the distance-based measurement, the distance-based measurement may erroneously be considered to be a better representation of the species response to that predictor. However, in our study there was only one case where the distance-based predictor outperformed a predictor with a measured scale of effect at 3.45 km (road for striped skunk, under the OCC framework; Appendix A). Therefore, although for some predictors the most appropriate scale of effect may have been larger than what was assessed in this study, this limitation is unlikely to have altered the outcome of this analysis.

Although composition-based measurements of landscape predictors generally outperformed distance-based measurements, there may be benefits to consideration of both distance- and composition-based measurements of a predictor, and selection of the most appropriate method of measurement based on a comparison of model fit. In this study there were predictors for each species for which the distance-based measurement explained more of the variability in response than the composition-based measurement. Furthermore, we must be cautious in extrapolating the results of a study of three species to all species; although most research to date supports the use of composition- over distance-based measurements, results of a mensurative study of immigration in two invertebrate species suggests that the relative performance of distance- and composition-based measurements may vary by species (Matter et al. 2005).

### Comparison of single-scale and multi-scale models (Objective 2)

Our prediction that MS habitat models would have greater explanatory power and predictive performance than SS models was based on the assumption that there are differences in the scales of effect among landscape predictors, and that the failure to account for predictor-specific scales of effect in SS models results in poorer model performance than MS models. Instead, the results of our analyses suggest that MS models did not improve performance relative to SS models for the three species we studied (striped skunk, ermine, and raccoon). We found no difference in explanatory power, with SS outperforming MS in half of the models across species and statistical frameworks. We note that the slightly higher median AUC values for MS models that we generally observed during cross-validation were consis-

tent with the results of previous studies, where ROC analyses suggested we could expect at most a 0.05 increase in the AUC for MS models over SS models (Boscolo and Metzger 2009, Graf et al. 2005, Kuhn et al. 2011). In previous studies, such differences were interpreted as support for MS over SS models; however, the wide variability between cross-validation trials in our study suggests we should not interpret our results in this manner. In addition, we conducted a novel comparison of the predictive performance during model extrapolation between SS and MS models, which supported the conclusion that MS were not superior to SS models for our three study species.

### Comparison of hypotheses for cases where the multi-scale model does not outperform the single-scale model (Objective 3)

To our knowledge, this study is the first attempt to explain cases where MS models do not outperform SS models. Our results were most consistent with two of the four hypothesized mechanisms: for striped skunk, the hypothesis that the species shows a similar response to the environmental context at several scales (hypothesis 4; Fig. 1d) and for ermine and raccoon, the hypothesis that the species response to the environmental context occurs at a common scale of effect (hypothesis 3; Fig. 1c). Evidence for these hypotheses over hypothesis 2 (i.e., that the species response is not dependent on scale) is consistent with the literature, which suggests that responses across scales may be rare relative to a response at one or a few scales (Chittaro 2004, Altmoos and Henle 2010), and these scales are most likely determined by the scale(s) of the species' interactions with the landscape.

It is important to note that, for the striped skunk, interpretation of the measured similar response to the environmental context at several scales could not be definitively attributed to a species preference (or avoidance) for a given landscape predictor at multiple scales, due to the high cross-scale correlations in the predictor values at the scales of effect modeled in SS and MS models. While the results may have been observed because the predictors were relevant to space use at more than one life stage for the skunk, with predictors acting at a common scale of effect (i.e., the SS scale of effect), the results are also consistent with an interpretation that the species does not respond to predictors at the SS scale of effect, but rather, is well represented at this scale due to similarities in the landscape structure at the SS and MS scales of effect.

It is possible that the ecological characteristics of the study species may influence whether SS models perform similarly to MS models. For example, in this study, all three species are considered to be habitat generalists. It may be that habitat specialists have strong responses at multiple scales that cannot be adequately represented at a single scale, by virtue of their more stringent habitat requirements. However, we note that the three bird species modeled by Boscolo and Metzger (2009) were considered to be forest specialists, and the small improvements in MS over SS performance for these species were consistent with our study. Another situation where MS models may outperform SS models may be for species that require complementary habitat, as certain environmental factors may have more localized effects around a given habitat, while others may act at a larger scale, such as those influencing the ability of the species to move between required resource patches. These potential relationships between the ecological characteristics of a species and the relative performance of SS and MS models are highly speculative at this point, as they have not yet been examined, and are thus an area for future research.

An alternative explanation for the similar performance of SS and MS models may be that our method for finding the scales of effect in MS models may not have accurately identified the scales of effect, resulting in reduced model performance. While selection of the scales for SS models was based on multivariate models including all candidate landscape predictors, we found the scales for MS models by fitting univariate models for each predictor at each scale, which does not take into account variability explained by the other predictors when fitting parameter values. It was not logistically possible to model all combinations of all parameters at all scales. Our approach of assessing seven landscape predictor variables at 23 scales for a given species and statistical framework using univariate models required 161 models; to assess all combinations of the seven predictors at 23 scales would require 3 404 825 447 models. We note that our approach is similar to the approach taken by other researchers, and that consideration of all possible combinations of predictors at all scales (Boscolo and Metzger 2005) does not appear to appreciably improve performance of MS relative to SS models over the results where univariate models were used to select the scales of effect (Graf et al. 2005). Limitation of the assessed scales to 3.45 km or less may also have reduced the performance of MS models. For some MS models, the species response may be best represented at a larger scale than was assessed in this study (i.e., at a scale greater than 3.45 km), while for SS models, the selected scale of effect was less than 3.45 km for all three species. If the species response to a predictor was best represented at a scale greater than 3.45 km, inclusion of the predictor at a scale of 3.45 km may have resulted in reduced model performance. Therefore, we expect the outcome of these analyses would be similar for a case where scales were selected by comparing models of all combinations of all parameters at all scales, but cannot discount the possibility that assessment of scales greater than 3.45 km would have improved the performance of MS relative to SS models.

Overall, our results suggest that MS models may not improve model performance relative to SS models, and that this may be expected for species which show a

strong response to certain elements of the environmental context at a common spatial scale or similar responses at several scales, although interpretation in the latter case was confounded by high cross-scale correlations in this study. This is not to suggest that MS modeling should be abandoned, particularly when there is sufficient prior knowledge to hypothesize that the species may respond to different environmental factors at different spatial scales. Rather, these results suggest that we should not discount the value of SS models and assessing the relative importance of predictors in a habitat model, as focus on a single scale may be sufficient if the factor or factors driving the species response to the environmental context act at a single scale. SS models may be particularly useful when models are used to inform conservation and management actions, as implementation may be greatly simplified if different actions can be implemented at the same scale.

Our evaluation of the hypothesized mechanisms that may drive similarities in model performance used only the results from the statistical framework producing the greatest predictive performance, as measured by the AUC, as higher predictive performance is expected to indicate a better fit to the assumptions of the statistical framework. The lack of superior performance of OCC relative to the other statistical frameworks was unexpected, as prior study suggests that OCC generally perform better than models that do not account for detectability (Rota et al. 2011). Occupancy models have been gaining popularity in response to concerns that traditional methods fail to account for biases due to imperfect detection and potential effects of the environmental variables of interest on detection, leading to biased assessments of predictor effects (e.g., Gu and Swihard 2004). There are several reasons why OCC performed more poorly than expected. First, our data may have violated the assumptions of OCC. For example, violation of the assumption of "closure" (i.e., no change in occupancy during the study period) has been shown to lead to overestimates of the probability of occurrence (Rota et al. 2009). In addition, the quality of OCC estimates tend to decline with smaller sample sizes, fewer replicates, and lower $p$ (Guillera-Arroita et al. 2010), conditions that apply to some degree to our study. Furthermore, in some cases, fitting a model that more closely reflects the ecological process may be more important than explicitly accounting for possible sampling bias; that is, using more flexible models such as GAM may produce superior models relative to OCC if the underlying relationships are nonlinear. While OCC modeling may be extended to the GAM framework, to our knowledge the algorithms to do so have not been made available in statistical software. In accordance with previous recommendations (Beutel et al. 1999, Segurado and Araújo 2004), we suggest using multiple statistical frameworks, as done here, to ensure modeled relationships show the best possible fit to the empirical data.

## Conclusions

To our knowledge, this is the first habitat modeling study to explicitly examine the relative performance of distance- and composition-based measurements of landscape predictors, and to evaluate mechanisms that could explain cases where MS models do not outperform SS models. Understanding the effects of different methods of predictor variable measurement and scale selection on model performance should help improve habitat modeling, and thus improve their usefulness for conservation planning and wildlife management. Our results suggest that priority should be placed on composition-based measurements, where the scale of measurement is selected based on empirical data. However, for all three species, there was at least one predictor for which the distance-based measurement was superior, suggesting it may often be beneficial to include both methods of measurement and select the appropriate measurement based on model fit. Our results further suggest that MS models may not improve model performance relative to SS models, and that this may be expected for species which show a strong response to certain elements of the environmental context at a common spatial scale or similar responses at several scales. For these species management actions that involve landscape features may be simplified, as actions can be taken at a common scale.

## Literature Cited

Alterio, N., H. Moller, and H. Ratz. 1998. Movements and habitat use of feral house cats, *Felis catus*, stoats *Mustela erminea* and ferrets *Mustela furo*, in grassland surrounding yellow-eyed penguin *Megadyptes antipodes* breeding areas in spring. Biological Conservation 83:187–194.

Altmoos, M., and K. Henle. 2010. Relevance of multiple spatial scales in habitat models: a case study with amphibians and grasshoppers. Acta Oecologica 36:548–560.

Augustin, N. H., M. A. Mugglestone, and S. T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology 33:339–347.

Balbontín, J. 2005. Identifying suitable habitat for dispersal in Bonelli's eagle: an important issue in halting its decline in Europe. Biological Conservation 126:74–83.

Barding, E. E., and T. A. Nelson. 2008. Raccoons use habitat edges in northern Illinois. American Midland Naturalist 159:394–402.

Barry, S., and J. Elith. 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43:413–423.

Barton, K. 2011. MuMIn: multi-model inference. R package version 0.13.21. R Foundation for Statistical Computing, Vienna, Austria. http://CRAN.R-project.org/package=MUMIn

Bender, D. J., L. Tischendorf, and L. Fahrig. 2003. Using patch isolation metrics to predict animal movement in binary landscapes. Landscape Ecology 18:17–39.

Beringer, J. J., S. G. Seibert, and M. R. Pleton. 1990. Incidence of road crossing by black bears on Pisgah National Forest, North Carolina. Pages 85–92 in L. M. Darling and W. R. Archibald, editors. Bears: their biology and management. Volume 8: A selection of papers from the Eighth International Conference on Bear Research and Management, Victoria, British Columbia, Canada, February 1989. International Association of Bear Research and Management, Victoria, British Columbia, Canada.

Beutel, T. S., R. J. S. Beeton, and G. S. Baxter. 1999. Building better wildlife-habitat models. Ecography 22:219–223.

Bixler, A., and J. L. Gittleman. 2000. Variation in home range and use of habitat in the striped skunk (Mephitis mephitis). Journal of Zoology 251:525–533.

Boscolo, D., and J. P. Metzger. 2009. Is bird incidence in Atlantic forest fragments influenced by landscape patterns at multiple scales? Landscape Ecology 24:907–918.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Wadsworth International Group, Belmont, California, USA.

Chittaro, P. M. 2004. Fish-habitat associations across multiple spatial scales. Coral Reefs 23:235–244.

Clark, J. D., J. E. Dunn, and K. G. Smith. 1993. A multivariate model of female black bear habitat use for a geographic information system. Journal of Wildlife Management 57:519–526.

Debrot, S., and C. Mermod. 1983. The spatial and temporal distribution pattern of the stoat (Mustela erminea L.). Oecologia 59:69–73.

Ebeling, C. A. 2006. Comparison of detection methods for three sympatric skunk species. Dissertation. Angelo State University, San Angelo, Texas, USA.

Elbroch, M. 2003. Mammal tracks and sign: a guide to North American species. Stackpole Books, Mechanicsburg, Pennsylvania, USA.

ESRI. 2010. ArcInfo 10 Geographic Information System. Environmental Systems Research Institute, Redlands, California, USA.

Filipe, A. F., T. A. Marques, S. Seabra, P. Tiago, F. Ribeiro, L. Moreira Da Costa, I. G. Cowx, and M. J. Collares-Pereira. 2004. Selection of priority areas for fish conservation in Guadiana River basin, Iberian Peninsula. Conservation Biology 18:189–200.

Foresman, K. R., and D. E. Pearson. 1998. Comparison of proposed survey procedures for detection of forest carnivores. Journal of Wildlife Management 62:1217–1226.

Freeman, E. 2007. PresenceAbsence: presence-absence model evaluation. R package version 1.1.5. R Foundation for Statistical Computing, Vienna, Austria. http://CRAN.R-project.org/package=PresenceAbsence

Gibbs, J. P., and W. G. Shriver. 2005. Can road mortality limit populations of pool-breeding amphibians? Wetlands Ecology and Management 13:281–289.

Gibbs, K. E., R. L. Mackey, and D. J. Currie. 2009. Human land use, agriculture, pesticides and losses of imperiled species. Diversity and Distributions 15:242–253.

Graf, R. F., K. Bollman, W. Suter, and H. Bugmann. 2005. The importance of spatial scale in habitat models: capercaillie in the Swiss Alps. Landscape Ecology 20:703–717.

Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. Biological Conservation 116:195–203.

Guillera-Arroita, G., M. S. Ridout, and B. J. T. Morgan. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1:131–139.

Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology 143:29–36.

Hastie, T. 2011. gam: generalized additive models. R package version 1.04. R Foundation for Statistical Computing, Vienna, Austria. http://CRAN.R-project.org/package=gam

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. Chapman and Hall, New York, New York, USA.

Henner, C. M., M. J. Chamberlain, B. D. Leopold, and L. W. Burger, Jr. 2004. A multi-resolution assessment of raccoon den selection. Journal of Wildlife Management 68:179–187.

Holland, J. D., D. G. Bert, and L. Fahrig. 2004. Determining the spatial scale of species' response to habitat. BioScience 54:227–233.

Koen, E. L., J. Bowman, C. S. Findlay, and L. Zheng. 2007. Home range and population density of fishers in eastern Ontario. Journal of Wildlife Management 71:1484–1493.

Kuhn, A., J. Copeland, J. Cooley, H. Vogel, K. Taylor, D. Nacci, and P. August. 2011. Modeling habitat associations for the common loon (Gavia immer) at multiple scales in northeastern North America. Avian Conservation and Ecology 6:4.

Laake, J., E. Rakimberdiev, B. Collier, J. Rotella, and A. Paul. 2012. RMark: an alternative approach to building linear models in MARK. R package version 2.1.0. R Foundation for Statistical Computing, Vienna, Austria. http://www.phidot.org/software/mark/rmark/

Larivière, S., and F. Messier. 2000. Habitat selection and use of edges by striped skunks in the Canadian prairies. Canadian Journal of Zoology 78:366–372.

Larivière, S., L. R. Walton, and F. Messier. 1999. Selection by striped skunks (Mephitis mephitis) of farmsteads and buildings as denning sites. American Midland Naturalist 142:96–101.

Lisgo, K. A. 1999. Ecology of the short-tailed weasel (Mustela erminea) in the mixedwood boreal forest of Alberta. Dissertation. University of British Columbia, Vancouver, British Columbia, Canada.

Loukmas, J. J., D. T. Mayack, and M. E. Richmond. 2003. Track plate enclosures: box designs affecting attractiveness to riparian mammals. American Midland Naturalist 149:219–224.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

Matter, S. F., T. Roslin, and J. Roland. 2005. Predicting immigration of two species in contrasting landscapes: effects of scale, patch size and isolation. Oikos 111:359–367.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, New York, New York, USA.

McGarigal, K., S. A. Cushman, M. C. Neel, and E. Ene. 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. Version 3.3. University of Massachusetts, Amherst, Massachusetts, USA. http://www.umass.edu/landeco/research/fragstats/fragstats.html

Meynard, C. N., and J. F. Quinn. 2007. Predicting species distributions: a critical comparison of the most common statistical frameworks using artificial species. Journal of Biogeography 34:1455–1469.

Moilanen, A., and M. Nieminen. 2002. Simple connectivity measures in spatial ecology. Ecology 83:1131–1145.

Müller, D., B. Schröder, and J. Müller. 2009. Modelling habitat selection of the cryptic hazel grouse Bonasa bonasia in a montane forest. Journal of Ornithology 150:717–732.

Murphy, E. C., and J. E. Dowding. 1994. Range and diet of stoats (Mustela erminea) in a New Zealand beech forest. New Zealand Journal of Ecology 18:11–18.

Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. Biometrika 78:691–692.

Nilsen, E. B., I. Herfindal, and J. D. C. Linnell. 2005. Can intraspecific variation in carnivore home-range size be explained using remote-sensing estimates of environmental productivity? Ecoscience 12:68–75.

NRC. 2002. National topographic data base. Natural Resources Canada, Earth Sciences Sector, Sherbrooke, Quebec, Canada.

OMNR. 2002. Southern Ontario land resource information system (2000–2002). Ontario Ministry of Natural Resources, Ottawa, Ontario, Canada.

OMNR. 2004. Ontario road network net element. Ontario Ministry of Natural Resources, Ottawa, Ontario, Canada.

OMNR. 2008. Digital raster acquisition project eastern Ontario orthoimagery. Ontario Ministry of Natural Resources, Ottawa, Ontario, Canada.

OMNR. 2009. Wetland unit. Ontario Ministry of Natural Resources, Ottawa, Ontario, Canada.

Pedlar, J. H., L. Fahrig, and H. G. Merriam. 1997. Raccoon habitat use at two spatial scales. Journal of Wildlife Management 61:102–112.

Peeters, E. T. H. M., and J. J. P. Gardeniers. 1998. Logistic regression as a tool for defining habitat requirements of two common gammarids. Freshwater Biology 39:605–615.

Prange, S., S. D. Gehrt, and E. P. Wiggers. 2003. Demographic factors contributing to high raccoon densities in urban landscapes. Journal of Wildlife Management 67:324–333.

R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rainho, A., and J. M. Palmeirim. 2011. The importance of distance to resources in the spatial modelling of bat foraging habitat. PLoS ONE 6:19227.

Robin, X., N. Turck, J. C. Sanchez, and M. Müller. 2011. pROC: display and analyze ROC curves. R package version 1.4.3. R Foundation for Statistical Computing, Vienna, Austria. http://CRAN.R-project.org/package=pROC

Rota, C. T., R. J. Fletcher, Jr., R. M. Dorazio, and M. G. Betts. 2009. Occupancy estimation and the closure assumption. Journal of Applied Ecology 46:1173–1181.

Rota, C. T., R. J. Fletcher, Jr., J. M. Evans, and R. L. Hutto. 2011. Does accounting for imperfect detection improve species distribution models? Ecography 34:659–670.

Rubbo, M. J., and J. M. Kiesecker. 2005. Amphibian breeding distribution in an urbanized landscape. Conservation Biology 19:504–511.

Rytwinski, T., and L. Fahrig. 2011. Reproductive rate and body size predict road impacts on mammal abundance. Ecological Applications 21:589–600.

Segurado, P., and M. B. Araújo. 2004. An evaluation of methods for modelling species distributions. Journal of Biogeography 31:1555–1568.

Sidorovich, V. E., and I. A. Solovej. 2007. The stoat *Mustela erminea* population decline in northern Belarus and its consequences for weasels *Mustela nivalis*. New Zealand Journal of Zoology 34:9–23.

SPSS. 2010. SPSS statistics. Version 19. IBM Corporation, Armonk, New York, USA.

Storm, G. L. 1972. Daytime retreats and movements of skunks on farmlands in Illinois. Journal of Wildlife Management 36:31–45.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. Science 240:1285–1293.

Tischendorf, L., D. J. Bender, and L. Fahrig. 2003. Evaluation of patch isolation metrics in mosaic landscapes for specialist vs. generalist dispersers. Landscape Ecology 18:41–50.

Watts, D. P. 1998. Long-term habitat use by mountain gorillas (*Gorilla gorilla beringei*). 2. Reuse of foraging areas in relation to resource abundance, quality, and depletion. International Journal of Primatology 19:681–702.

Wheatley, M., and C. Johnson. 2009. Factors limiting our understanding of ecological scale. Ecological Complexity 6:150–159.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46 Supplement:120–138. http://www.phidot.org/software/mark/

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.

Wiken, E. B. 1986. Terrestrial ecozones of Canada. Environment Canada, Lands Directorate, Ottawa, Ontario, Canada.