

A Primer on Quantitative Approaches for Setting Recovery Targets and Identifying Critical Habitat of Species at Risk

M.S. Poesch ¹, J.M.R. Curtis ², and M.A. Koops ¹

¹ Fisheries and Oceans Canada
Great Lakes Laboratory for Fisheries and Aquatic Science
867 Lakeshore Road, Burlington, ON, L7R 4A6

² Fisheries and Oceans Canada
Pacific Biological Station
3190 Hammond Bay Road, Nanaimo, BC, V9T 6N7

2012

**Canadian Technical Report of
Fisheries and Aquatic Sciences 2983**



Fisheries and Oceans
Canada

Pêches et Océans
Canada

Canada

© Her Majesty the Queen in Right of Canada, 2012.

Published by:

Fisheries and Oceans Canada
200 Kent Street
Ottawa, Ontario
K1A 0E6

Correct citation for this publication:

Poesch, M.S., Curtis, J.M.R., and Koops, M.A. 2012. A primer on quantitative approaches for setting recovery targets and identifying critical habitat for species at risk. Can. Tech. Rep. Fish. Aquat. Sci. 2983: vii + 40 p.

TABLE OF CONTENTS

TABLE OF CONTENTS	iii
ABSTRACT	vii
1.0 BACKGROUND.....	1
2.0. INTRODUCTION.....	2
3.0. HABITAT SUITABILITY (HS) MODELS	3
3.1 CONCEPTUAL MODEL.....	4
3.2 STATISTICAL FORMULATION	5
3.3 CALIBRATION	8
3.4 VALIDATION.....	9
3.5 EVALUATING THE MODEL	10
3.6 EVALUATING ALTERNATIVE MODELLING APPROACHES.....	11
3.7 ADVANTAGES	12
3.8 LIMITATIONS	12
3.8.1 SAMPLE SIZE.....	12
3.8.2 SPATIAL SCALE.....	13
3.8.3 CHOICE OF MODELLING APPROACH	13
3.9 IMPLEMENTATION	13
4.0. INCIDENCE FUNCTION MODELS (IFM)	14
4.1 DATA REQUIREMENTS	14
4.1.1 EXTINCTION RATES	15
4.1.2 COLONIZATION RATES	15
4.1.3 DISPERSAL	16
4.2 DETERMINING POPULATION VIABILITY	16
4.3 ADVANTAGES	17
4.4 LIMITATIONS	18
4.5 IMPLEMENTATION	18
5.0. AREA PER INDIVIDUAL (API).....	20
5.1 DATA REQUIREMENTS	20
5.2 ADVANTAGES	21
5.3 LIMITATIONS	21
5.4 IMPLEMENTATION	21
6.0. MINIMUM AREA FOR POPULATION VIABILITY (MAPV).....	22
6.1 DATA REQUIREMENTS	22
6.2 ADVANTAGES	23
6.3 LIMITATIONS	23
6.4 IMPLEMENTATION	23

7.0. SPATIALLY EXPLICIT POPULATION MODELS	24
7.1 DATA REQUIREMENTS	24
7.2 ADVANTAGES	28
7.3 LIMITATIONS	28
7.4 IMPLEMENTATION	28
8.0. REFERENCES.....	29

TABLES

Table 1 – Quantitative methods used to build habitat suitability models for species at risk.....	7
Table 2 – Confusion matrix comparing model predictions to the actual data.	10
Table 3 – Examples of dispersal kernels available for developing incidence function models. All formulas give the proportion of individuals able to move from patches i and j , given distance (d_{ij}). Adopted from Chapman et al. (2006).	17
Table 4 – Software packages used for forecasting incidence function models to estimate population viability.	19

FIGURES

Figure 1 – Conceptual framework for developing a predictive habitat suitability model using two datasets, one for fitting the model (calibration, left side), and one for validating and evaluating the model (validation and evaluation, right side). Adapted from Guisan and Zimmerman 2000.	4
Figure 2 - A schematic of the key steps for carrying out a global sensitivity analysis of a habitat-based, spatially-explicit PVA model with RAMAS GIS (Akçakaya et al. 2005) and GRIP (Curtis & Naujokaitis-Lewis 2008a,b) version 2.0 (figure from Naujokaitis-Lewis and Curtis, unpublished).	26
Figure 3 - The relative importance of PVA model input parameters on extinction probability of 45 invertebrate and vertebrate species. Relative importance is measured simply as the Spearman correlation coefficient, with larger median coefficients corresponding to more influential parameters (Curtis, unpublished data).	27
Figure 4 - The minimum number of discrete populations (subpopulations) required to achieve three recovery targets (extinction probability < 0.2, 0.1, 0.05). Each black line represent the relative density of estimates for one of 45 vertebrate and invertebrate species, while the red line represents the expected number of populations across all species (Curtis, unpublished data).	28

ABSTRACT

This primer outlines common approaches for setting recovery targets and quantifying critical habitat for species at risk. Critical habitat is protected under the Species at Risk Act, and defined as the habitat necessary and sufficient for species survival or recovery. Key questions that need to be addressed when identifying critical habitat relate to the amount, quality and configuration of habitat needed to achieve a specified population or habitat-based recovery target. The process of delineating population and habitat based recovery targets is, by its nature, complex and quantitative approaches are often required. This primer introduces commonly used quantitative approaches for setting recovery targets and for identifying critical habitat, including: habitat suitability (HS) models; incidence function models (IFM); demographic based estimates of: area per individual (API), and minimum area for population viability (MAPV); and habitat based spatially explicit population viability analyses (PVA). In this primer, each modelling approach is introduced, data requirements are described, and advantages and limitations are discussed. Included in this primer, readers will find specific references to software or freely available code to allow them to implement these quantitative approaches themselves.

1.0 BACKGROUND

This manuscript presents several quantitative methods used in setting recovery targets and identifying critical habitat. It is assumed readers will be familiar with the following companion documents:

Grinnell, M.H. and Curtis, J.M.R. 2012. User manual for PatchImportance 1.0: Quantifying relative habitat patch importance based on metapopulation persistence and minimum abundance. Can. Tech. Rep. Fish. Aquat. Sci. 2977: vi + 41 p.

Grinnell, M.H. and Curtis, J.M.R. 2011. User manual for NetworkDistances 1.0: Calculating network-wise distances between habitat patches for spatially restricted species. Can. Tech. Rep. Fish. Aquat. Sci. 2960: iv + 29 p.

Vélez-Espino, L.A., R.G. Randall and M.A. Koops. 2010. Quantifying habitat requirements of four freshwater species at risk in Canada: Northern Madtom, Spotted Gar, Lake Chubsucker, and Pugnose Shiner. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/115.

Curtis, J.M.R. and I. Naujokaitis-Lewis. 2008. Sensitivity of population viability to spatial and non-spatial parameters using GRIP. *Ecological Applications* 18:1002-1013

Curtis, J.M.R. and I. Naujokaitis-Lewis. 2008. Source code for GRIP 1.0. *Ecological Archives* A013-033-S1

Vélez-Espino, L.A. and M. A. Koops. 2007. A quantitative approach to assessing allowable harm in species at risk: application to the Laurentian black redhorse (*Moxostoma duquesnei*). Canadian Science Advisory Secretariat 2007/051.

And previous technical reports:

de Kerckhove, D. T., K. E. Smokorowski, and R. G. Randall. 2008. A primer on fish habitat models. Can. Tech. Rep. Fish. Aq. Sci. 2817:iv + 58 p.

Minns, C.K. 2003. An area-per-individual (API) model for estimating critical habitat requirements in aquatic species-at-risk. DFO Can. Sci. Adv. Sec. Res. Doc. 2003/074.

2.0. INTRODUCTION

The *Species at Risk Act* - and other endangered species legislation around the world (e.g. *Endangered Species Act* of the United States) - requires the identification and protection of the habitat necessary to ensure the recovery and survival of the species at risk (i.e. critical habitat). Habitat impacts are the top threat to freshwater biodiversity (Jelks et al. 2008) and the one of the top threats to marine biodiversity (following exploitation; Hutchings and Reynolds 2004). Evidence suggests a clear link between protection of critical habitat and species recovery (Clark et al. 2002; Taylor et al. 2005; Hagen & Hodges 2006).

Understanding how to set appropriate recovery targets and identify critical habitat is an important theoretical and applied problem. Habitat units have unequal biological value, some are more important than others for maintain biodiversity (Economo 2011). Understanding which habitats are therefore critical for species survival and recovery is important for not only ensuring adequate protection of species at risk, but for prioritizing conservation efforts. Developing robust, and scientifically defensible, methods for setting recovery targets is fundamental for ensuring critical habitat is implemented correctly (Carroll et al. 1996; Ruckelshaus et al. 2002).

To accomplish the identification of critical habitat, models are needed which link population occurrence, persistence, and viability to habitat. These can then be used to set recovery goals and targets, and to identify critical habitat. The goal of this primer is to introduce quantitative methods commonly used throughout the process of setting recovery targets and identifying critical habitat, their advantages and limitations, with the hope that it may aid in implementing their use. The methods include: habitat suitability (HS) models, incidence function models (IFM), area per individual (API), minimum area for population viability (MAPV) and habitat based spatially explicit models of population viability (PVA). These methods have been used to: set recovery targets (e.g. minimum viable population (PVA, MAPV); identify suitable habitat (HS models); assess trade-offs and develop rules of thumb (e.g. PVA); evaluate our ability to achieve recovery targets under alternative scenarios (e.g. IFM, PVA).

3.0. HABITAT SUITABILITY (HS) MODELS

Habitat suitability models correlate species distribution or abundance to one or more ecological relevant variables. Habitat suitability models have become important tools in modelling changes in biodiversity (Guisan and Zimmermann 2000; Guisan and Thuiller 2005). Habitat suitability models have been used in a wide range of applications including predicting: changes in species distributions in relation to climate change (Thuiller 2004; Araújo et al. 2005), the establishment and spread of invasive species (Hartley et al. 2006), and localities with rare or endangered species (Rodriguez et al. 2007; Marmion et al. 2009; Franklin et al. 2009). With the increasing availability of remote-sensing data, and advances of geographic information systems, researchers are often no longer data limited, but are expanding the use of habitat suitability models to include more applications and broader spatial scales (Guisan and Zimmermann 2000; Marmion et al. 2009). The degree to which habitat suitability models can be reliably used to predict rare or uncertain events remains a fundamental challenge for the conservation of biodiversity (Zimmermann et al. 2007).

Building a habitat suitability model requires several steps. First, a conceptual model is needed to help shape the statistical model (i.e. purely empirical, purely mechanistic, purely theoretical, or some combination). Once a conceptual model is formulated, a study design can be implemented which helps identify the data requirements and the sampling regime. From there, a statistical formulation can be assigned and the model can be calibrated, validated and evaluated (Fig. 1). Each of these steps requires careful consideration as choices at each level may have substantial influence towards the resultant habitat suitability model (Dormann et al. 2008).

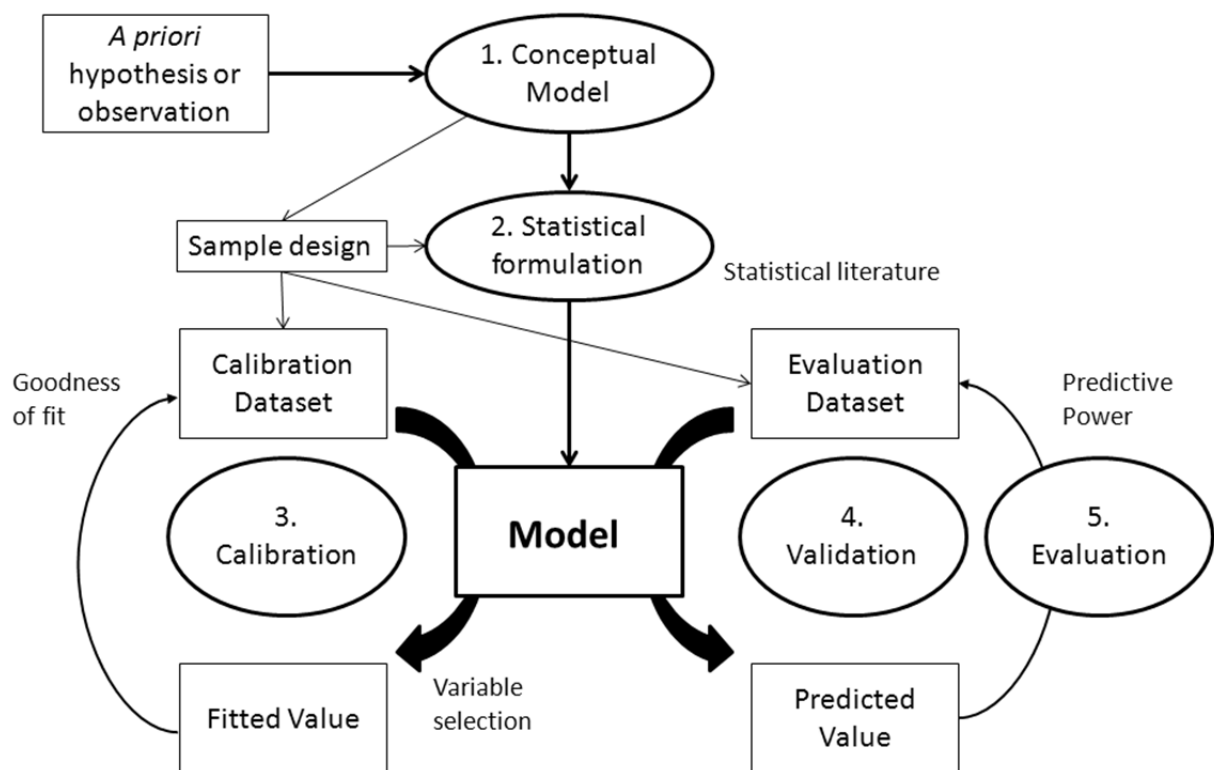


Figure 1 – Conceptual framework for developing a predictive habitat suitability model using two datasets, one for fitting the model (calibration, left side), and one for validating and evaluating the model (validation and evaluation, right side). Adapted from Guisan and Zimmerman 2000.

3.1 CONCEPTUAL MODEL

The process of building a habitat suitability model begins with the formulation of a conceptual model. As nature is too complex and heterogeneous to be predicted accurately in every aspect of time and space, simplification using a conceptual model is needed (Guisan and Zimmerman 2000). Conceptual models can be empirical, mechanistic, or analytical. A conceptual model defines a simplified viewpoint from which the statistical formulation should be applied. For example, an empirical conceptual model is based on the assumption of maximizing the fit (or minimizing error) to real data. Conversely, a mechanistic conceptual model can be based on physiological, causal or process based mechanics. These models are often not judged based on model fit, but on the theoretical correctness of the model. Finally, an analytical conceptual model is based on predicting accurate responses to a simplified reality, such as predator prey dynamics as used in the Lotka-Volterra model (Guisan and Zimmerman 2000).

Habitat suitability models are typically classified as empirical models as they are often based on a *a priori* hypotheses or observations that require phenomenological (i.e.

“real”) data (Guisan and Zimmerman 2000). For example, researchers may have some knowledge regarding the risks or threats to current distributions of species at risk and their habitats - such as those outlined in species recovery plans - and can use those to model species distributions. The assumption is that empirical data represent the “true form” of nature, or the ideal habitat for the species. As such empirical habitat suitability models may be limited in their applicability to specific localities (i.e. study systems) or certain spatial scales. On the other hand, conceptual models are not necessarily mutually exclusive. Habitat suitability models can also be classified as mechanistic or analytical, given the nature of the study (Guisan and Zimmerman 2000). For example, a habitat suitability model outlining soil or temperature conditions of an endangered plant may be more mechanistic than empirical due to the physiological requirements of the species and the causative nature of the variables.

The differences among conceptual models have important repercussions for modelling habitat suitability for species at risk (Araujo et al. 2006). For example, a purely mechanistic conceptual model, as outlined above, may be well suited for predicting the physiological constraints of a species (i.e. fundamental niche), but have no utility for predicting actual localities where species are found (i.e. realized niche; Hutchinson 1953; MacArthur 1968). Alternatively, a purely empirical conceptual model may be well suited for predicting the realized niche of a species, while saying nothing in regards to the physiological tolerances of the species (Austin et al. 1990). As there are broad spatial differences between the delineation of the fundamental versus realized niche of a species (Soberón and Peterson 2005), the delineation of a conceptual model provides important assumptions as to how species habitats are to be treated. Despite the importance of defining a conceptual model, it remains one of the fundamental challenges for implementing habitat suitability models for species at risk (Araujo et al. 2006). Ultimately the choice of conceptual model will be dependent on the goals of the study, the spatial scale, and the precision needed (Guisan and Zimmerman 2000).

The formation of a conceptual model ideally leads to several study design considerations, such as: 1) the choice of spatial scale (Levin 1990), 2) the choice of explanatory variables for the model, and 3) sampling design considerations. Some of the study design considerations for species at risk include: selection of appropriate gear for capturing the species (Poos et al. 2007, 2012), the amount of sampling effort needed (Poos et al. 2007), and the selectivity of gear types, amongst others (Thompson 2004).

3.2 STATISTICAL FORMULATION

The next step in developing a habitat suitability model is statistical formulation. Statistical formulation is the process of selecting an appropriate statistical approach for predicting a response variable (Fig. 1). Numerous approaches are now available (Table 2). These methods can be defined as: i) probabilistic methods, where a probability distribution is assumed, ii) classification methods, which recursively partition data into categories, or iii) machine learning methods, which iteratively fit a distribution to the data (Olden et al. 2008). Each of these modelling approaches has been used to predict species occurrences and describe habitat suitability (Thuiller 2003; Elith et al. 2006; Thuiller et al. 2006). The modelling methods shown, represent a continuum of use from

logistic regression - the most prevalent and widespread statistical method for modelling binary data (Hosmer and Lemeshow 1989) - to newer learning-based methods, such as multivariate adaptive regression splines (Friedman 1991), artificial neural networks (Olden and Jackson 2001; Olden et al. 2004), random forest (Cutler et al. 2007) and boosted regression trees (Elith et al. 2008).

Perhaps at the forefront of decisions needed for modelling habitat suitability of species at risk is the choice of statistical method. The choice of statistical method has large consequences for not only how the data are handled, but often on the conservation decisions being made (Loiselle 2003; Johnson et al. 2004; Wilson et al. 2005; Rodriguez et al. 2007; Marrimon et al. 2009). For example, Pearson et al. (2006) showed that distribution changes of South African plant species varied from a 92% loss to a 322% gain depending on the statistical model used. Dorman et al (2008) demonstrated that of several uncertainties in modelling habitat suitability - including variable selection and co-linearity between variables – that the choice of statistical method had the largest impact. Previous analyses of species at risk echo these concerns over the importance of model type (Poos and Jackson, In Revision). As quantitative comparisons have demonstrated, the successes of statistical modelling approaches can be largely data dependent and there is no clear indication of the preeminence of any singular approach (Olden and Jackson 2002; Araujo and New 2007). As such, a comparison of model performance across various statistical methods is preferable to naïve selection of a single approach (Guisan and Zimmerman 2000; Olden and Jackson 2001; 2002).

Determining the most appropriate statistical method requires several steps. For many probabilistic models the choice of theoretical probability distribution is often needed (however see multivariate adaptive regression splines; Friedman 1991; or LOWESS (McCullagh and Nelder. 1989). The choice of theoretical probability distribution can be done by comparing the empirical distribution of data to that of a theoretical distribution, of which several tests are available depending on the data (e.g. goodness of fit test, Kolmogorow-Smirnov; Guisan and Zimmerman 2000). In some cases data may need to be transformed (e.g. log or arc-sin transformation) or standardized (e.g. z-score) to fit the model assumptions (Faraway 2006). The choice between choosing an appropriate data transformation or data standardization method to fit a given theoretical probability distribution is often not straight forward. In some cases it may be less preferable to transform data to meet model assumptions than it would be to use an alternative probability distribution. For example if count data are zero inflated, it may be preferable to use a zero inflated Poisson distribution rather than attempt to standardize data and use a Poisson distribution (Hall 2000; MacKenzie et al. 2002). Alternatively it may be better to choose an alternative statistical approach where probability distributions are not required (e.g. general additive models, multivariate adaptive regression splines; McCullagh and Nelder, 1989; Guisan et al. 2006). The question as to how much one should manipulate the data to fit the model is not straight forward (Olden and Jackson 2000), however guidelines are available elsewhere (see Hosmer and Lemeshow 1989; McCullagh and Nelder 1989; Olden and Jackson 2000; Hastie et al. 2001; Faraway 2006).

Table 1 – Quantitative methods used to build habitat suitability models for species at risk.

Type of Approach ¹	Probability Distribution	Response Variables	Statistical Method	Key References	R Code Available ²
Probabilistic	Gaussian	Environmental variables (e.g. substrate size, percent cover, biomass)	Multiple regression (weighted averaging, least squares regression, general linear models, general additive models, regression trees, LOWESS, generalized linear mixed models)	Vincent and Haworth (1983); Hastie and Tibshirani (1987); Faraway (2006); Nakagawa and Cuthill (2007); Bolker et al. (2009).	Faraway (2006); Crawley (2007); Bolker (2008); Milborrow (2011); Venables and Ripley (2011).
	Poisson, Negative binomial	individual species counts	Multiple regression (general linear models, general additive models)	McCullagh and Nelder (1989); Nakagawa and Cuthill (2007).	Same as above
	Binomial / Logit	presence / absence, relative abundance	Multiple regression (general linear models, general additive models, multivariate adaptive regression splines)	Hosmer and Lemeshow (1989); Peeters and Gardeniers (1998); Pampel (2000).	Same as above
	Zero inflated Poisson (ZIP), Zero inflated Negative Binomial (ZINB)	individual species counts	Zero inflated models (zero-inflated Poisson (ZIP), zero-inflated negative binomial (ZINB))	Hall (2000); MacKenzie et al. (2002); Tyre et al. (2003); Martin et al. (2005); Welsh et al. (2006).	Zeileis et al. 2009.
	Fitted by model	presence only	BIOMOD ³ , GARP ⁴ , MAXENT, GRASP	Stockwell and Peters (1999); Lehman et al. (2003); Thuiller (2003); Guisan and Thuiller (2005); Philips et al. (2006).	Bivand et al. 2008; Jurka 2011
Classification Based	Fitted by model	presence / absence, relative abundance	Regression trees (Classification and regression trees, Multivariate adaptive regression splines); Discriminant Analysis	Friedman (1991); Vayssières et al. (1996); Moisen et al. (2006).	Ripley 2009; Bocard et al. 2011; Milborrow 2011.
Learning Based Methods	Fitted by model	individual species counts, presence / absence, relative abundance	Artificial neural networks, boosted regression trees, random forest.	Bishop (1995); Lek et al. (1996); Ridgeway (1999); Friedman et al. (2000); Olden and Jackson (2001); Cutler et al. (2007); Elith et al. (2008); Olden et al. (2008).	Cutler et al. 2007; Ridgeway 2007; Elith et al. 2008.

Note: ¹ For simplicity multivariate or Bayesian methods are not described here, but can be found elsewhere (Legendre and Legendre 1998; Ellison 2005; King 2010), including with R code (Bocard et al. 2011). ² In many cases code is available in the R programming language, freely available at URL: <http://www.r-project.org/>. ³ BIOMOD is available online at URL: <http://r-forge.r-project.org/projects/biomod/>. ⁴ GARP can be analyzed using DesktopGARP available at: URL: <http://www.nhm.ku.edu/desktopgarp>.

3.3 CALIBRATION

The next phase in developing the habitat suitability model is dividing the data into two separate data sets: one for calibration, the other for validation (although this may not always be necessary; Guisan and Zimmerman 2000). Calibration refers to the adjustment of an initial model to improve the agreement between the model output and the data set (Rykiel 1996; Olden et al. 2002).

One fundamental component of model calibration is choosing the right number of explanatory variables. This is one of the most difficult tasks in developing a habitat suitability model (Guisan and Zimmerman 2000). If too many explanatory variables are used, the model may perform well in calibration (maximize fit), but may not be generalizable when validated and evaluated. If too few explanatory variables are used, the model may perform sub-optimally or exclude important habitat components. Often the choice of the number of explanatory variables is a choice between model fit and generality (Johnson and Omland 2004). A model that maximizes fit will have no consideration of model generality and favors more explanatory variables. Alternatively, a model that maximizes generality will consider only the number of explanatory variables that maximize prediction, even if this reduces model fit.

There are numerous approaches for maximizing model fit. For example, researchers can find numerous methods to evaluate their models based on how well they fit the data (i.e. goodness of fit, χ^2 , coefficient of determination, R^2 , or deviance, D^2 ; Guisan and Zimmerman 2000). However, model fit may not always be an accurate reflection of the data (see 3.4 Validation) and can become biased by the number of explanatory terms. In these cases, adjustment is needed (e.g. R^2_{adj} ; Quinn and Keough 2002).

To improve model generality, a method of reducing the number of explanatory variables is needed (i.e. variable selection method). Debate over variable selection method exists. Some studies have shown that the choice of variable selection method can bias the resultant model and that variable selection should be avoided (Whittingham et al. 2006). Other studies suggest that variable selection procedures, when compared to other choices in developing habitat suitability models, may actually have very small impact on resultant models (Maggini et al. 2006; Meynard and Quinn 2007; Dormann et al. 2008).

The two most popular methods for variable selection are forward or backward selection (Burnham and Anderson 2002). Forward and backward selection procedures attempt to maximize the generality of a series of competing models by choosing the best suite of explanatory variables. Variables are selected either starting from the simplest model and iteratively adding variables (forward selection), or starting with the most complex model and iteratively removing variables (backward selection; McCullagh and Nelder 1989). Other variable selection methods include Gibbs sampling (George and McCulloch 1993), lasso (Tibshirani 1996) and best subset (Quinn and Keough

2002). Despite the litany of potential methods for variable selection, there does not appear to be a 'best' method (Murtaugh 2009).

An alternative to the variable selection approach are multi-model comparisons (i.e. multi-model inference; Burnham and Anderson 2002). This approach attempts to find the most parsimonious model. Here, parsimony refers to the tradeoff between model fit (accuracy) and generality (Burnham and Anderson 2002). Akaike Information Criteria, and its affiliates (BIC, AIC_c), are the popular method of multi-model inference (but see consensus methods, Marimon et al. 2009). Multi-model inferences may be particularly useful for hypothesis testing (Johnson and Omland 2004), but may not be well suited for maximizing predictive performance.

It should be noted that for non-probabilistic methods (Table 1), the choice of modelling approach can be equivalent to the choice of variable selection method. This is due to the fact that most machine learning and classification based approaches iteratively solve for the most appropriate number of explanatory variables. For example, with classification and regression trees the model will recursively partition the data in dichotomous groups that maximize classification (Vayssières et al. 2000). Explanatory variables enter and re-enter the model several times to classify the response variable. This often leads to over-fitted models (i.e. too many terminal nodes; Chambers and Hastie 1993), which reduce model performance (Guisan and Zimmerman 2000). One disadvantage of machine learning approaches is that they can often be considered 'black boxes' for understanding how variables are selected and prioritized (Olden and Jackson 2002b; Olden et al. 2008). Caution is warranted in utilizing these methods appropriately for modelling species at risk (but see Hastie et al. 2001; Olden et al. 2008)

3.4 VALIDATION

Once the model has been calibrated, predictions can be produced. In many cases this is where researchers stop, with the belief that once a model has been constructed, and predictions produced, that the modelling process is complete (Olden et al. 2002). However, models have little merit without knowing how well those predictions differ from what was expected or from those based on chance. Model validation refers to methods that compare the predicted values of the calibrated model to an independent evaluation dataset (Fig. 1; Rykiel 1996). Unfortunately, despite the importance of model validation, it receives comparatively little attention relative to other modelling activities (Manel et al. 2001; Olden et al. 2002).

There are three main approaches used to validate the calibrated model: 1) jackknife, 2) bootstrapping, and 3) k -fold cross validation. The jackknife approach removes a single observation and constructs the model with the remaining $n-1$ observations. Model performance is calculated by repeating this procedure until all observations are removed (i.e. n times) and each observation can be predicted once (Olden et al. 2002). Bootstrapping randomly selects the number of observations to be including into the model and then re-samples the observations n times to obtain an estimate of model bias. Finally, k -fold cross validation splits the data into k subsamples, where a single subsample is used for validating the model, while the remaining $k-1$

subsamples are used to calibrate the model. The cross-validation process is repeated k times, where each validation subsample is used only once. The results of each of these validation methods are then aggregated (or averaged) into a single estimate of model performance. Bias can be assessed as the difference between the predicted values and the real parameter value in the evaluation dataset. When this difference is high, the model should be called into question (Guisan and Zimmerman 2000).

3.5 EVALUATING THE MODEL

Once the predicted values have been produced, they need to be assessed to determine model performance. Model evaluation refers to how well the model performs relative to some criteria, most often how well it predicts an independent evaluation dataset (although not necessarily so). In some cases, to evaluate a model, predictions to be transformed back to the scale of real observations to allow comparison (e.g. GLMs; Guisan and Zimmerman 2000). For continuous data, parameter estimates can be compared to the evaluation data using correlation coefficients (e.g. Pearson correlation) or non-parametric rank correlations (e.g. Kendall's t or Spearman's ρ ; Quinn and Keough 2002). For qualitative response variables, contingency tables may be used (e.g. χ^2) or Cohen's kappa (Cohen 1968). For presence/absence (i.e. binary) data, a decision is needed as to the threshold needed to define habitat suitability. In many cases researchers arbitrarily determine the threshold for habitat suitability (i.e. > 0.5 on a scale from 0 to 1). However, previous research has shown that the choice of thresholds can have a large influence on the results (Jiménez-Valverde and Lobo 2007; Lobo 2008), and this remains largely problematic for determining habitat suitability or predicting species at risk occurrences.

Thresholds for determining habitat suitability can be determined using a matrix comparing model predictions to actual data, known as a confusion matrix (Table 2). The confusion matrix can be used by researchers to compare the differences in correct classification (i.e. true positives or negatives) of either presence or absence at the expense of inflated error (i.e. false positives or negatives) at various model thresholds (Table 2). One could choose model thresholds which prioritize correct classification of species presence (true positive), at the expense of increased misclassification of species absences (false positives), or vice versa. Such a decision may be well suited for species at risk, where one would wish to emphasize correct predictions of the true occurrence of a species at risk, but would be willing to accept higher error rates (Loiselle et al. 2003). Another alternative is to use a balanced approach, with equal likelihood of misclassification of species presence and absence. This approach is more commonly used as it assumes equal consideration of modelling species presence and absence (Hartley et al. 2006).

Table 2 – Confusion matrix comparing model predictions to the actual data.

ACTUAL VALUE	PREDICTED VALUE	
	Species present	Species absent
	Species present	Species absent
	(true positives)	(false negatives)
	(false positives)	(true negatives)

3.6 EVALUATING ALTERNATIVE MODELLING APPROACHES

Modelling habitat suitability of species at risk is an iterative process (Fig. 1). Each step in developing a habitat suitability model requires careful consideration of each decision inherent in the model, and often re-analysis and comparison (Fig. 1). For example, researchers must decide on: 1) the conceptual model (e.g. empirical versus theoretical); 2) data type (e.g. abundance vs. presence / absence); 3) whether data should be standardized or transformed; 4) statistical formulation (e.g. probabilistic versus classification based); 5) variable selection method (e.g. forward or backward selection, AIC), 6) validation method (e.g. jackknife, bootstrap, k-fold cross validation); and for binary data, 7) the threshold for determining suitable from unsuitable habitat. As each of these decisions has been shown to alter model outputs (Dormann et al. 2008), criteria are needed for comparing between and among modelling approaches.

Models can be evaluated based on their overall performance, e.g. how well does the model fit the data (e.g. goodness of fit, see section 3.2). However for modelling species occurrences using binary (i.e. presence / absence) data, other modelling metrics are available including: model sensitivity (the ability of each model to correctly predict species presence); model specificity (the ability of each model to correctly predict species absence); and, overall classification (the ability of each model to correctly classify both species presence and absence). The use of these metrics provides an alternative means to evaluate each model, their comparative successes, and their associated errors, for both species presence and absence (Olden and Jackson 2002a). The distinction among modelling metrics is an important one as models for predicting species at risk are often complicated by zero inflated distributions (Cunningham and Lindenmayer 2006). In these cases, models may still perform well (i.e. high overall classification) due to the large number of species absences (i.e. high model specificity), while saying nothing regarding the habitats where species at risk are predicted to be found (i.e. model sensitivity).

One common approach for comparing among modelling approaches is to use a threshold independent method, such as a receiver operator characteristic curve (ROC; Manel 1999; Olden and Jackson 2002; Hartley et al. 2006). Similar to a confusion matrix (Table 2), ROC curves plot the rate of true positives (sensitivity) to false positives (1-specificity). The area under the ROC curve (commonly referred to as AUC) is considered a measure of model accuracy that is independent of a particular threshold (Fielding and Bell 1997). AUC varies from between 0.5 (the model is no better than chance) to an area of 1 (the model predicts perfectly; Fielding and Bell 1997). AUC is often used to compare between modelling approaches (e.g. Elith et al. 2006; Thuiller et al. 2006) or between competing sub-models, such as commonly used in machine learning methods (Olden et al. 2008). However, in some cases AUC may provide biased estimates and may not be appropriate for model comparisons (e.g. if omission and commission rates are unbalanced; see Lobo et al. 2008).

3.7 ADVANTAGES

Modelling existing and future habitats of species at risk has become an increasingly common application of habitat suitability models. One of the reasons for this popularity is that when faced with limited data on the distribution, abundance and dynamics of species at risk (Mace et al. 2005; Rodríguez 2007), habitat suitability models allow for the extrapolation of relatively few field samples to the entire potential range of a species. Also, as species at risk are, by their nature, rare and difficult to enumerate, habitat suitability models can provide a surrogate measure to prioritize conservation initiatives, even if sampling data are not available (Guisan et al. 2006). In these cases simple habitat suitability models can be produced using expert opinion or Geographic Information Systems (Guisan and Zimmerman 2000), which is a clear advantage for defining recovery targets for species at risk that are often data deficient.

Habitat suitability models have been used in a broad array of applications (Thuiller 2004; Araújo et al. 2005; Hartley et al. 2006; Rodríguez et al. 2007; Marmion et al. 2009; Franklin et al. 2009), and are flexible for use with many data types including presence only, presence absence, count data, ordinal data and continuous data (Table 1).

3.8 LIMITATIONS

Modelling the distributions of species at risk and their habitats is an activity filled with uncertainty. Species at risk are not only rare, but they are also difficult to detect, capture and enumerate, thereby complicating the evaluation of their habitat importance (Elith et al. 2002; Loiselle 2003; Heikkinen et al. 2006). Uncertainties in habitat suitability models for species at risk may arise during all stages of modelling including: determining an appropriate statistical formulation (Elith et al. 2006; Dormann et al. 2008), calibrating the model (e.g. deciding an appropriate theoretical probability or variable selection method); validating the model (e.g. determining validation procedure), and evaluating the model (e.g. determining thresholds for suitability). In addition, there are uncertainties with modelling endangered species due to missing data, the potential for spatial autocorrelation with model predictions (Fortin and Dale 2005; Dormann et al. 2007; Bivand 2008), and for determining future changes to habitat under climate change (Thuiller 2004; Araújo et al. 2006); to name a few (see Loiselle 2003; Araújo and Guisan 2005; Burgman et al. 2005; Barry and Elith 2006; Austin 2007).

3.8.1 SAMPLE SIZE

The application of habitat suitability models to identify critical habitat of species at risk remains controversial (Loiselle 2003; Burgman and Lindenmayer 2005; Araújo and Guisan 2006). One difficulty with using the standard sets of habitat suitability models used by the majority of ecologists, is that they are often compromised when analyzing data limited to a few sites and scales (Ellison and Agrawal 2005; Araújo and Guisan 2006). This condition of limited occurrences, in turn, produces data sets that have many complicating statistical issues, including zero inflated bias, increased co-linearity between variables and prevalence bias (Guisan et al. 2006; Dormann et al. 2008).

One often disregarded aspect of modelling habitat suitability for species at risk is the consideration of sample size requirements needed to determine habitat associations. Many traditional approaches, (e.g. general linear models; Table 1) require relatively large sample sizes, especially when attempting to identify habitat suitability (Pearce and Ferrier 2000; Stockwell and Peterson 2002). There are several reasons for this. First, given the highly complex nature of ecological niches (Hutchinson 1957), few sample locations are unlikely to accurately depict species habitats (Carroll and Pearson 1998; Wisz et al. 2008). Second, as sample size decreases the influence of outliers increases, thereby causing the potential for a few sample points to skew habitat relationships (Stockwell and Peterson 2002). Finally, interactions among habitat features may exponentially increase the number of sample locations needed to estimate habitat suitability (Wisz et al. 2008). For probabilistic regression, it is generally accepted that for every additional environmental variable added to the habitat suitability model, an additional ten sample sites are needed to adequately evaluate their contribution.

3.8.2 SPATIAL SCALE

Species-habitat relationships are strongly dependent on processes that operate at different spatial scales (Hunsaker and Levine 1995; Jackson et al. 2001). At broader spatial scales (i.e. macro scale: $10^3 - 10^5 \text{ km}^2$) climate, geomorphological or edaphic factors may limit species distribution or abundances (Tonn 1990). At more intermediate spatial scales (i.e. meso-scale: $10^2 - 10^3 \text{ km}^2$), abiotic conditions strongly alter resources and create heterogeneity within a given watershed or lake system (Utz and Hilderbrand 2011). At even finer spatial scales ($10^{-2} - 10^2 \text{ km}^2$), morphology (e.g. area, depth, structural complexity; Grossman and Freeman 1987; Poff et al. 2006) and biotic interactions help structure localized fauna. Therefore, to truly understand the habitat suitability of freshwater ecosystems, multi-scale approaches are needed, which remain relatively scarce in habitat suitability models (Jackson et al. 2001).

3.8.3 CHOICE OF MODELLING APPROACH

The selection of statistical method is a vitally important decision in developing a habitat suitability model (Elith et al. 2006; Aruajo and New 2007; Dormann et al. 2008; Marmion et al. 2009). Despite nearly two decades of comparative analysis, to date no preeminent approach exists. As such, it is recommended that habitat suitability modelling be treated as both an adaptive and iterative process. Developing habitat suitability models without proper validation or evaluation will undoubtedly lead to poor recovery potential for species at risk. One method to reduce uncertainty is to utilize a comparative approach incorporating all decisions inherent in modeling habitat suitability and testing for sensitivity (e.g. Dormann et al. 2008).

3.9 IMPLEMENTATION

Many software packages exist that allow researchers to develop habitat suitability models (e.g. SPSS, JMP), including software that is freely available (see Table 1 for examples).

4.0. INCIDENCE FUNCTION MODELS (IFM)

Many species with spatially structured populations are in decline, and population viability models provide a quantitative simulation-based evaluation of species viability to inform management decisions (Frank and Wissel 1998; Akçakaya 2000). Incidence function models determine population viability based on the connectivity of sites (i.e. patches) within a metapopulation (Hanski 1999), and the dispersal ability of species to reach those sites once they become extirpated (i.e. rescue effect).

Incidence function models are spatially explicit (they consider patch location) and provide a realistic evaluation of patch dynamics (Hanski 1999; March 2008). By quantifying patch dynamics, incidence function models can be used to better understand the importance of ecological processes such as species specific dispersal, patch quality and landscape influences (Moilanen and Hanski 1998), and to inform conservation management through evaluation of the minimum amount of habitat or population size needed to maintain viability (Hanski 1999; Robert 2009).

4.1 DATA REQUIREMENTS

Incidence function models use a time-continuous Markov-chain model (Hanski 1999). Each patch (i) is assumed to be in one of two states, vacant ($x_i = 0$) or occupied ($x_i = 1$). Changes in these states can occur from a patch becoming vacant due to local extinction ($x_i: 1 \rightarrow 0$) or correlated extinction (i.e., regional stochasticity) from another patch ($x_j, x_i: 1 \rightarrow 0$). Alternatively a vacant patch can become occupied ($x_i: 0 \rightarrow 1$) via colonization from another patch (j). The state of the whole metapopulation (x_1, \dots, x_n) is given by a vector of states x_i of these individual patches. Differences in the state of each patch are determined by the balance between colonization and extinction rates (see below), with some level of stochasticity (Lande 1993; Moilanen 1999). Assuming constant but patch-specific probabilities of colonization (C_i) and extinction (E_i) in unit time, the simplest form of the incidence of the species in patch i is given by:

$$J_i = \frac{C_i}{C_i + E_i}$$

or if rescue is considered:

$$J_i = \frac{C_i}{C_i + E_i - C_i E_i}$$

4.1.1 EXTINCTION RATES

Extinction rates can be quantified in many ways (Hanski 1999). The simplest form for determining extinction rate (E_i) is using the area of the patch (A_i), given by:

$$E_i = \frac{e}{A_i^x}, \text{ for } A \geq e^{1/x}$$

where e defines the extinction probability of a patch of unit size, and x defines the scaling of the extinction risk with patch area (Hanski 1998; Moilanen 2004). This model assumes that probability of extinction generally depends on population size, which can be extrapolated using species-area relationships. The relationship between extinction and species-area has been demonstrated on both empirical and theoretical grounds (Lande 1993; Foley 1994; Hanski 1994; 1999; Hanski et al. 1996).

4.1.2 COLONIZATION RATES

Colonization rate (C_i) is the rate at which individuals move from one patch (i) to colonize another patch (j). Colonization rate depends on three factors: 1) the mean number of migrants (M_i) per year, 2) the probability of a migrant starting from one patch reaching the other patch, based on the distance between the patches (d_{ij}) and the dispersal ability of the species (α , see dispersal kernel); and 3) patch Area (A_j , Hanski et al. 1996; Hanski 1999).

Given these factors, colonization rate can be quantified as:

$$C_i = \frac{M_i^2}{M_i^2 + y^2}$$

where:

$$M_i = \beta S = \beta \sum_j p_j e^{-\alpha d_{ij}} A_j$$

and p_j is 0/1 indicator of site occupancy, and e , β , and y are parameters that are fit using probabilistic regression (Hanski et al. 1996; ter Braak et al. 1998; Oksanen 2004).

The fit of the probabilistic regression can be shown arithmetically. For example, given the equations for colonization and extinction (above), the incidence function can be reduced to:

$$J_i = \left[1 + \frac{ey}{S_i^2 A_i^x} \right]^2$$

based on connectivity (S_i) and patch size A_i . This can be parameterized as a linear model for the log-odds of incidence:

$$\text{logit}(J_i) = \beta_0 + 2 \log S + \beta_1 \log A$$

which defines a generalized linear model with logistic link function (ter Braak et al. 1998; Oksanen 2004). Examples of this derivation and of ecological data are available in ter Braak et al. (1996).

4.1.3 DISPERSAL

Dispersal (α) is parameterized in incidence function models using a dispersal kernel. Dispersal kernels fit empirical estimates of species movement to a mathematical function. They have been used extensively in metapopulation models and assume that patch accessibility is dependent on distance (Hansson 1991; Hanski et al. 1996; Moilanen 2004). One advantage of dispersal kernels is that they allow researchers the ability to extrapolate relationships in patch occupancy, often using species life-history characteristics, without the need for labour-intensive field studies (Moilanen 2004; Heinz et al. 2005). For example, body size allometry can be used to estimate dispersal ability (see Minns 1995) and incorporated into incidence function models without the need for expensive field studies (e.g. mark-recapture). Such measures are often useful when dispersal data are not available, as is often the case with species at risk (Velez-Espino and Koops 2008; Velez-Espino and Koops 2009a; Velez-Espino and Koops 2009b).

Choosing an appropriate dispersal kernel can often be challenging (Heinz et al. 2005; 2006). First, numerous dispersal kernels exist (see Table 3). Second, studies have shown that the quantification of dispersal can alter estimates of metapopulation viability (Ovaskainen and Hanski 20004; Heinz et al. 2005; 2006; Revilla and Wiegand 2008), sometimes drastically so (Poos and Jackson 2012). Most models of metapopulation viability use a negative exponential decay dispersal kernel (Hanski 1994; Vos et al. 2001; Frank and Wissel 2002), although it may be sensitive in some instances (Poos and Jackson 2012).

4.2 DETERMINING POPULATION VIABILITY

Incidence function models can be used to estimate population viability by adding stochasticity into the model and building forward projections (see Table 3). Each of these modelling approaches deal with adding stochasticity into the model differently, and in some cases, colonization potential as well (e.g. R and SPOMSIM versus META-X; Table 4).

Table 3 – Examples of dispersal kernels available for developing incidence function models. All formulas give the proportion of individuals able to move from patches i and j , given distance (d_{ij}). Adopted from Chapman et al. (2006).

Kernel	Formula	Key References
Negative exponential	$\exp(-\sigma_1 d_{ij})$	Hanski and Gilpin (1997) ; Hanski (1999); Hanski et al. (2000)
Gaussian	$\exp(-d_{ij}/2\sigma^2)$	Turchin (1998)
Inverse power function	$\begin{cases} 1 & \text{if } i = j \\ \min(1, d_{ij}^{-\alpha_2}) & \text{if } i \neq j \end{cases}$	Chapman et al. (2006)
Extended negative exponential decay	$\exp(-\alpha_3 d_{ij}^{\beta_2})$	Taylor (1978)
Fat tail	$\frac{1}{1 + \alpha_4 d_{ij}^{\beta_2}}$	Moilanen (2004)
Sigmoidal function	$1 - e^{-a * e^{(-\beta * d_{ij})}}$	Heinz et al. (2006)

Note: σ indicates variance. Parameters α & β are from fitting the dispersal kernel to empirical data.

4.3 ADVANTAGES

Incidence function models are a crucial tool for the management of species at risk (Ackakaya 2000). Incidence function models have been used extensively to model the viability of spatially structured populations (Hanski 1999; Moilanen 1999) including several endangered species of butterfly (Hanski 1999; Grimm and Storch 2000). As incidence function models are among the simplest form of population viability analyses (Akçakaya and Sjögren-Gulve 2000), they do not require demographic or stage data, but only occupancy, colonization and extinction rates, which can be easily estimated from empirical data (Hanski 1994,1999; Moilanen 1999, 2004; Grimm et al. 2004). This is a clear advantage over other modelling approaches (e.g. demographic PVA, API, MAPV, or other spatially explicit models, Sections 5.0-7.0), especially when data is limited, as the case with many species at risk.

Incidence function models have become relatively commonplace in conservation biology. Modifications have been made to incidence function model to provide more realism, such as consideration of patch quality (Moilanen and Hanski 1998; Ovaskainen and Hanski 2002), improved dispersal metrics (Ovaskainen 2004; Heinz et al. 2005), incorporation of transition states (Thomas and Hanski 2004), and rigorous parameter estimation techniques (Moilanen 1999; Dreschler et al. 2003). Quantitative comparisons

of incidence function modelling approaches have shown them to be comparable to other spatially explicit models (Kindvall et al. 2000; Keeling 2002).

4.4 LIMITATIONS

There is an ongoing debate whether incidence function models are useful in cases where the assumptions of classic (i.e. Levin's type) metapopulations are not met (see Levins 1969; Harrison 1994; Baguette 2004; Hanski 2004; Shreeve et al. 2004). For example, empirical studies have demonstrated large temporal variation in patch dynamics (i.e. colonization and extinction rates) can lead to sensitivity of incidence function models (Crone et al. 2001; Thomas et al. 2002). In addition, there are few empirical examples of metapopulations that meet the assumption of a constant pulse of extinction-colonization (e.g. pool frog; Sjogren-Gulve 1991; Glanville Fritillary butterfly; Hanski et al. 1994); while the vast majority do not (Harrison 1994; Baguette et al 2004). However, contrary to such criticisms, incidence function models have been shown to be appropriate for use from a range of spatially structured populations: from classic metapopulations to species found in fragmented landscapes with patchy distributions (Ovaskainen and Hanski 2004). For example, using an incidence function model with individual based background, Ovaskainen and Hanski (2004) provide a unifying framework for incorporating metapopulation dynamics into incidence function models.

Studying metapopulations in stream settings is challenging and has limitations which should be noted for developing incidence function models. Gotelli and Taylor (1999) showed that stream fish may not fit Levin's type metapopulation models as migration may cause asynchrony in upstream versus downstream movement. Fagan (2002) demonstrated how a dendritic network can provide additional isolation of patches not encountered in terrestrial landscapes. As streams represent linear, stepping stone metapopulations, most software packages (Table 4) are not well suited for modelling these systems, and custom approaches are needed (e.g. Grinnell et al. 2011).

Understanding how species specific dispersal and patch specific qualities can alter incidence function models is important for developing robust models (Heinz et al. 2006; Poos and Jackson 2012). Species specific dispersal has been shown to alter incidence function models, and in turn rates of population viability (Poos and Jackson 2012). For example, Heinz et al. (2005; 2006) used simulated data to demonstrate that viability of patches in a metapopulation can change based on the way dispersal was modeled. Better integration of species specific behaviour is needed into incidence function models (Tischendorf 2001; Vos et al. 2001; Baguette and vanDyck 2007). In addition, patch specific processes can alter metapopulation dynamics in many ways (Roitberg and Mangel 1997; Hanski and Moilanen 1998; Schtickzelle et al. 2006). Patch quality can affect both the probabilities of colonization and extinction of an empty patch (Hanski and Moilanen 1998).

4.5 IMPLEMENTATION

Statistical packages - including freely available code – and ecological examples for developing incidence function models can be found in Table 4.

Table 4 – Software packages used for forecasting incidence function models to estimate population viability.

Software	Ecological Example	Key Reference
SPOMSIM	Avian conservation (Beissinger et al. 2006) ; marmot metapopulations (Ozgul et al. 2006)	Moilanen (2004)
R ¹	Stream fishes (Poos and Jackson 2012)	Oksanen (2004)
PATCH	Regional conservation planning, e.g. Grizzly bears, wolves, wolverines (Carroll et al. 2003; 2004)	Schumaker (1998)
HEXSIM	Ord's kangaroo mouse (Heinrichs et al. 2010)	Schumaker (2012)
META-X	Butterfly metapopulations (McIntrie et al. 2007)	Grimm et al. (2004)
ALEX	Reintroduction of the Greater Bilby (Southgate and Possingham 1995)	Possingham and Alex (2004)
Conefor Sensinode 2.2	Capercaillie in Catalonia, Spain (Pascual-Hortal and Saura 2008)	Saura and Torne (2009)

Note: See also ter Braak et al. (1998) and Baguette (2004) for a review of ecological examples. ¹ R statistical package is freely available at URL: <http://www.r-project.org/>

5.0. AREA PER INDIVIDUAL (API)

Space use is related to body size; larger individuals need and use more space. This is the basic ecological concept underlying the area per individual (API) approach. Ecology has identified body size as a driver in many ecological processes (e.g., Schmidt-Nielsen 1984, Peters 1986, Hildrew et al. 2007), including the number of individuals that can occupy a given space. For example, Cyr et al. (1997) provide an allometry between body size and density (number of individuals per unit area) for aquatic organisms, and similar relationships have been demonstrated for fishes (e.g., Boudreau and Dickie 1989, Randall et al. 1995). Area per individual is the inverse of density.

Individuals need space to complete life processes such as finding sufficient food for energy demands (with larger individuals having greater energy requirements), refuge from predators, and reproduction. The API approach assumes that when an individual has less space (due to limited habitat availability or increased density) there is a cost, and the cost is realized through reduced survival. Essentially, the API approach is a density-dependent mechanism consistent with the concept of a carrying capacity. An area of habitat can support a limited number of individuals and this capacity diminishes as the size of individuals increases.

The API approach assumes a basal survival (s_b) when there are no habitat limitations or density-dependent effects. Under these conditions, the habitat available to each individual is equal to or greater than the API requirement. As the area available per individual declines below the API, due to either reductions in available habitat or increases in the number of individuals using the same habitat area, survival declines. The simplest assumption is a linear decline (Fig. 1). This scalar to survival can be added to a population model to evaluate the population-level effects of habitat limitations (e.g., Minns et al. 1996, Minns 2003, Vélez-Espino and Koops 2009).

5.1 DATA REQUIREMENTS

In addition to a population model, the API approach requires (i) an estimate of API, (ii) an estimate of habitat available for each life stage, and (iii) an estimate of abundance or density in each habitat. The habitat and abundance (or density) estimates, while not trivial to acquire, are not unique to this approach and will not be dealt with further here. There are a couple of approaches to estimating API. The preferred approach is an estimate based on observed densities in healthy populations of the species. Data from an area comparable to the available habitat (e.g., type of habitat and ecosystem, quality, etc.) can provide the most directly applicable estimate of API. In the absence of species-specific information, an allometry can be used to estimate density based on body size. API is estimated as the inverse of density. An estimate of API for each life stage should be calculated as body size differences among life stages will change habitat requirements and density-dependent effects.

5.2 ADVANTAGES

The API approach is based on the known, proven ecological principle that body size determines space use with larger individuals using and requiring more habitat than smaller individuals. There are minimal data requirements which can be estimated (e.g., using allometries) for species with very limited data. The API approach can also be expanded to include habitats of differing quality through the addition of a scalar (Minns 2003).

5.3 LIMITATIONS

The API approach, as currently applied, assumes that density directly affects survival. While there are ecological mechanisms where density can affect growth or reproduction, there has not yet been an attempt to frame these effects within an API approach. The API approach assumes that all individuals are equal. If multiple stages use the same habitat, the API approach does not account for the differences in competitive ability between individuals of different life stages (e.g., juveniles versus adults) if they share the same habitat. The API approach is not spatially-explicit. It deals with the total amount of habitat and assumes that the configuration of that habitat (e.g., one large patch versus many small patches) does not affect the individuals or population. This assumes that there are no edge effects and no costs to moving among patches.

5.4 IMPLEMENTATION

Code for developing API is currently not available. However, examples of demographic models are available from Morris and Doak (2002), which will allow users the ability to implement API.

6.0. MINIMUM AREA FOR POPULATION VIABILITY (MAPV)

The minimum area for population viability (MAPV) concept is an approach to quantifying habitat needs for population viability based on the minimum viable population (MVP) concept and the API approach (Vélez-Espino et al. 2010). As a generalization, it posits that the amount of habitat needed to support a recovered population is a product of the number of individuals in the recovered population and the amount of habitat each individual needs to support their life processes. Specifically, it has been applied in cases where the recovery target is set for demographic sustainability by identifying the MVP and habitat needs are calculated based on an API, or:

$$(1) \quad MAPV = MVP \times API$$

On its own, equation 1 provides an estimate of habitat area needed to support a specified number of adults (the MVP). To ensure long-term sustainability of the population, habitat is needed for all life stages. This can be quantified by first calculating the number of juveniles and young of the year (YOY) needed to support this number of adults. The stable stage distribution from a population model at equilibrium (i.e., with a population growth rate of $\lambda = 1$) allows the calculation of these numbers. Equation 1 can then be repeated for each life stage and then summed across all life stages to get an estimate of the amount of habitat needed to sustain the whole population at MVP (Fig. 2). Note that the API for each life stage will differ since adults are larger than juveniles which are larger than YOY.

If a life stage occupies multiple habitats (e.g., a summer and a winter habitat), then MAPV calculations need to be applied to each of these habitats. Exceptions to the calculation of area requirements based on API will include habitats with functions that are not related to the productivity of those habitats. For example, migration corridors need to be included in critical habitat identification but would not be based on API considerations. Instead, the size of migration corridors needs to maintain connectivity among habitats and allow passage of individuals at the appropriate time. Spawning habitat area may also be different than an area estimate based on adult API. Minns (2003) proposed calculating spawning habitat based on the API calculated from egg diameter. However, this may under-estimate spawning habitat needs of large-bodied fishes (e.g., Lake Sturgeon, Randall 2008) or fishes with eggs that drift (e.g., Western Silvery Minnow).

6.1 DATA REQUIREMENTS

The MAPV approach requires a recovery target (e.g., an estimate of MVP), an estimate of API for each life stage (see Section 5.0), and an estimate of the stable stage distribution. Calculation of the stable stage distribution requires a population model.

6.2 ADVANTAGES

The MAPV approach is based on the API approach and inherits its advantages. It has minimal data requirements and can be applied to species with limited data.

6.3 LIMITATIONS

The MAPV estimate is a minimum and should be considered a first-order approximation of the amount of habitat need to support a population. Assuming that the API estimates are correct, the MAPV will perch the population on the edge of density-dependent effects. This will make the population very sensitive to increases in density (e.g., due to higher abundances) or decreases in habitat (which may not be human-induced). Simulations by Young and Koops (2012) showed that a population with MAPV habitat available will, on average, maintain a population at the MVP, but that the population will often fluctuate below MVP. Vélez-Espino et al. (2010) recommended adding a buffer zone around the MAPV amount of habitat. Young and Koops (2012) found that a habitat area one and a half times larger than the MAPV would ensure that the population was above MVP 95% of the time.

Like the API approach, the MAPV approach is not spatially-explicit, but deals with the total aggregate of habitat area. As such, it does not capture edge effects or the costs associated with transiting among patches. There is no consideration given to the configuration of habitat, which has been shown to influence the effectiveness of protected areas.

6.4 IMPLEMENTATION

Code for developing MAPV is currently not available. However, examples of demographic models are available from Morris and Doak (2002), which will allow users the ability to implement MAPV.

7.0. SPATIALLY EXPLICIT POPULATION MODELS

Population viability analysis (PVA) was originally used to simulate population dynamics, quantify extinction risk or estimate time to extinction while accounting for uncertainty. During the past 30 years, PVA applications have diversified to include estimating minimum viable population (MVP) sizes (e.g. Reed et al. 2003), predicting population dynamics under alternative resource management scenarios (e.g. Curtis & Vincent 2008), identifying critical habitat (e.g. Curtis and Naujokaitis-Lewis 2008a) and using sensitivity analysis to prioritize research needs and recovery activities for species at risk (Naujokaitis-Lewis et al. 2009). The structure of PVA models and the kinds of predictions generated by these models vary widely and according to study objectives, data availability, species biology, and the expertise and capacity of modelers to build them. In this section, we focus on a class of complex, data-intensive Monte-Carlo simulation models, namely habitat-based, spatially-explicit PVA models, and discuss its potential applications for setting population and habitat-based recovery targets and identifying critical habitat.

7.1 DATA REQUIREMENTS

Habitat-based, spatially-explicit models typically integrate a habitat suitability sub-model and a population dynamics sub-model with functions that link the two. RAMAS GIS is one of the most commonly-used software platforms for developing these models (Naujokaitis-Lewis 2009). The RAMAS GIS user manual (Akçakaya et al. 2005) provides an excellent primer for those wishing to learn more about how to develop and apply these simulation models. Here, we provide an overview of the approach, its potential applications, and tools that have been developed for use with RAMAS software for species at risk recovery planning and critical habitat identification, including the freely available program GRIP (Curtis & Naujokaitis-Lewis 2008a,b) version 2.0 (Fig. 2), written in the statistical programming language R (R Development Core Team 2011).

When developing a habitat-based, spatially-explicit PVA model with RAMAS software, the first step is to develop a habitat suitability model (see Section 3). Once developed, the habitat suitability model provides the basis for defining the spatial structure of discrete patches of suitable habitat with RAMAS Spatial, by specifying a habitat suitability threshold (above which populations are self-sustaining) and a neighbourhood distance, which reflects the daily movement patterns of individuals (Akçakaya et al. 2005). In RAMAS software, discrete patches of suitable habitat are treated synonymously as discrete populations (if occupied) linked by varying rates of dispersal. The dynamics of each population are typically modeled with a stage-based transition matrix whereby fecundities or survival rates or both are varied at each time step according to functions including those governing density-dependence, spatial autocorrelation, exploitation, landscape dynamics, climate change, and stochastic processes including the frequency and impacts of catastrophes. Population dynamics can also vary according to dispersal functions, and functions that link initial or global conditions that vary according to habitat suitability. As one example, the carrying capacity, K , of discrete patches of suitable habitat could be specified as a function of total habitat suitability or total patch area. Once the population sub-model is specified

and initialized, RAMAS Metapop can be used to simulate the dynamics of spatially-structured populations in the landscape defined with RAMAS Spatial. For more details, please refer to Akçakaya et al. (2005).

One of the key limitations of applying habitat-based, spatially-explicit PVA models, regardless of the objectives, is dealing with a number of important sources of uncertainty. Uncertainties in PVA models can derive from natural variation, errors in parameter estimation, incorrectly specified model structures, vague terminology, and spatial inaccuracies. Thus best practices for such PVA models includes carrying out a systematic and comprehensive sensitivity analysis to evaluate the influence of input parameters and model structure on predictions (Naujokaitis-Lewis et al. 2009). Global sensitivity analyses, whereby all parameters of interest are varied simultaneously, are particularly useful for ranking the relative influence of parameters on model predictions, identifying, assessing interactions among parameters, and prioritizing alternative management actions. Global sensitivity analyses, however, are rare in PVAs likely due to the computational requirements of varying habitat and demographic parameters simultaneously (Naujokaitis-Lewis et al. 2009).

GRIP 2.0 is a tool that can be used to automate global sensitivity analyses of habitat-based, spatially-explicit PVA models, based on GRIP 1.0 (Curtis & Naujokaitis-Lewis 2008b) a program designed to automate sensitivity analyses of spatially-implicit PVA models created with RAMAS Metapop software (Akçakaya et al. 2005). GRIP 2.0 uses the baseline map of discrete patches of suitable habitat defined with RAMAS Spatial as a template for creating a specified number of alternative realizations of landscapes (Figure 3). For each alternative landscape, a unique population model is developed by randomly varying the parameters in the population dynamics model as in GRIP 1 and described in Curtis and Naujokaitis-Lewis (2008a,b). Once the replicate landscapes and corresponding population dynamics models are created, each with a unique set of input parameters, GRIP 2.0 submits the corresponding input files for RAMAS to run the simulations, and then collates the model output for further analysis. The GRIP scripts can also be customized easily to vary parameters and functions systematically to create a series of alternative management scenarios for comparison under uncertainty (e.g. comparison of alternative minimum size limits for fisheries management, Curtis & Vincent 2008).

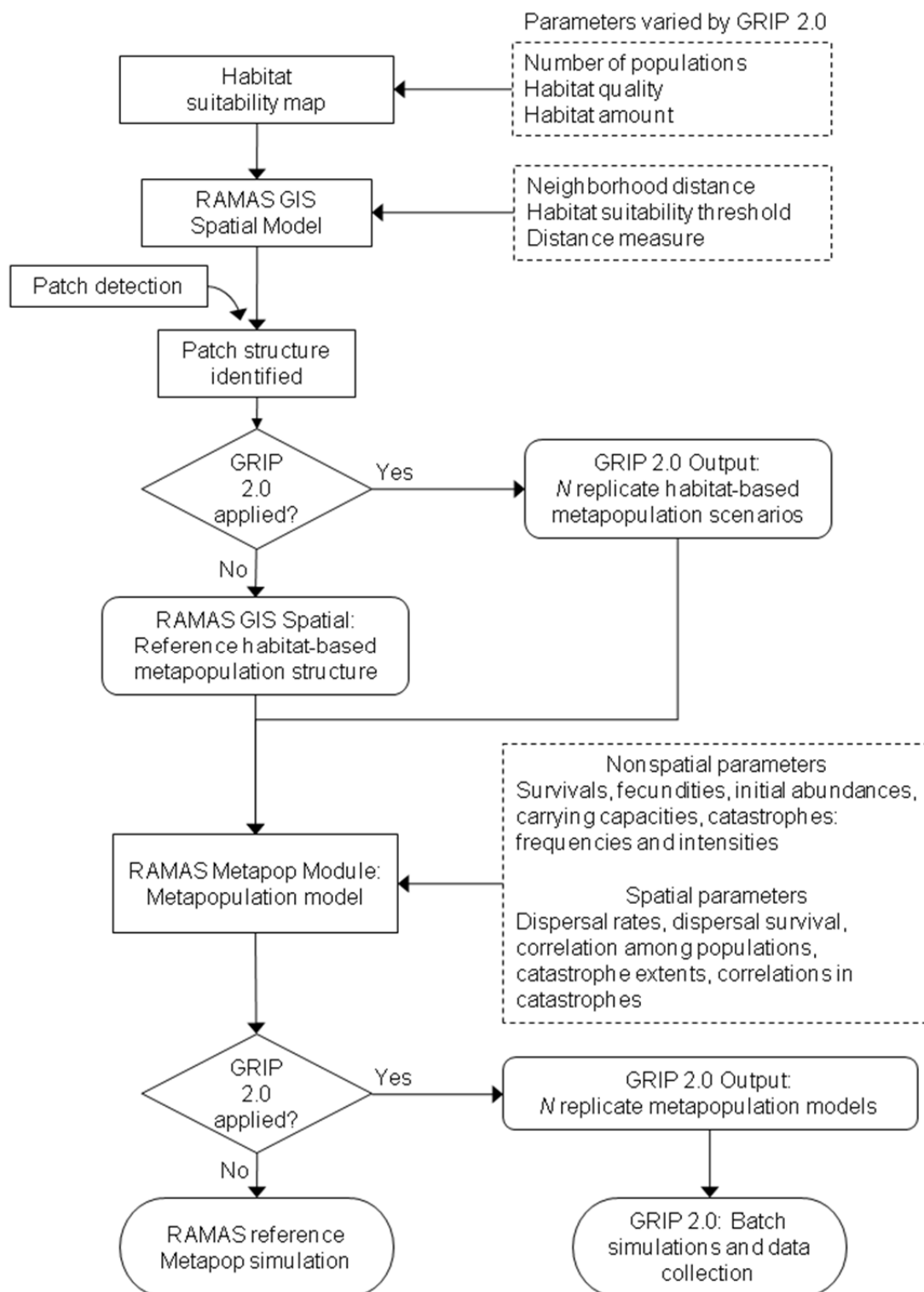


Figure 2 - A schematic of the key steps for carrying out a global sensitivity analysis of a habitat-based, spatially-explicit PVA model with RAMAS GIS (Akçakaya et al. 2005) and GRIP (Curtis & Naujokaitis-Lewis 2008a,b) version 2.0 (figure from Naujokaitis-Lewis and Curtis, unpublished).

Output from a global sensitivity analysis of a habitat-based, spatially-explicit PVA model can be used to rank the relative importance of model parameters (see Curtis & Naujokaitis-Lewis 2008a) and to set recovery targets, given specified recovery goals (Fig. 3). Influential parameters can help prioritize research (e.g. if parameter values are uncertain) or help identify potential management actions that could influence population viability. For instance, if adult survival rate is highly influential of model predictions (e.g. extinction probability or metapopulation abundance) but is unknown or uncertain due to limited data, that parameter could be flagged as a high priority for further research. By contrast if the same influential survival rate was well known for a population, the parameter could be used to prioritize management strategies that enhance adult survival. Figure 3 summarizes the relative importance of input parameters from PVA models of 45 vertebrate and invertebrate species. Across these species, the carrying capacity, initial abundance, and number of discrete populations have the strongest influence on population viability.

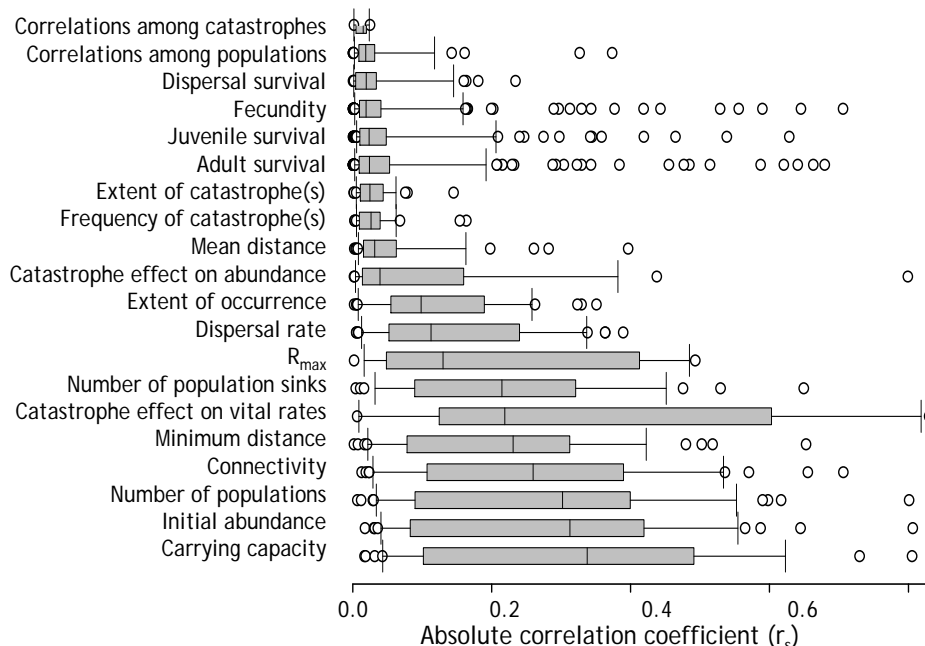


Figure 3 - The relative importance of PVA model input parameters on extinction probability of 45 invertebrate and vertebrate species. Relative importance is measured simply as the Spearman correlation coefficient, with larger median coefficients corresponding to more influential parameters (Curtis, unpublished data).

Using the same simulation output from a global sensitivity analysis, one can estimate the habitat carrying capacity, number of populations and configuration of those populations required to meet specified recovery goals under uncertainty. Figure 4, for example, provides estimates of the minimum number of discrete populations required to meet three recovery goals for the same 45 species: ensure the probability of extinction remains at or below 0.2, 0.1, or 0.05. Such simulation results and analyses, whether based on a single species or pooled among many, can be used to inform decisions on the amount and configuration of habitat required for species survival and recovery.

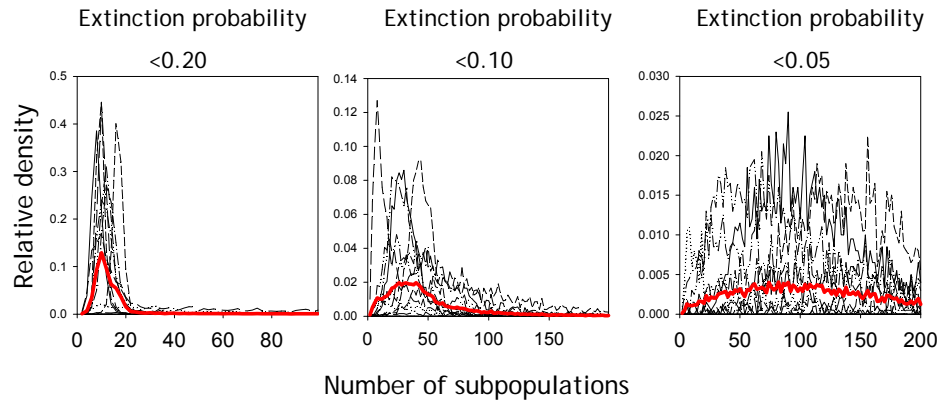


Figure 4 - The minimum number of discrete populations (subpopulations) required to achieve three recovery targets (extinction probability < 0.2, 0.1, 0.05). Each black line represent the relative density of estimates for one of 45 vertebrate and invertebrate species, while the red line represents the expected number of populations across all species (Curtis, unpublished data).

7.2 ADVANTAGES

The key advantage of coupling global sensitivity analysis with habitat-based, spatially-explicit PVA models is the ability to evaluate the influence of parameter values, model structure, and alternative landscape and population management scenarios on model predictions of interest, in a transparent, repeatable manner while accounting for uncertainties. Specifically, these models allow users to evaluate the relative influence of habitat quality, habitat quantify, and habitat configuration on population dynamics. Although rarely attempted, PVA model predictions can also be validated and evaluated against empirical data.

7.3 LIMITATIONS

The key limitations of habitat-based, spatially-explicit PVA models include 1) data requirements, 2) many sources of uncertainty, and 3) a capacity for estimating parameters, programming, running simulations and analyzing multivariate data.

7.4 IMPLEMENTATION

Code for implementing global sensitivity analyses using GRIP 1.0 is available for the R programming language (R Development Team 2011) from Curtis & Naujokaitis-Lewis (2008b). Code for GRIP 2.0 is currently awaiting publication and will be available later in 2012. Code is also available for determining the importance of habitat patches (Grinnell and Curtis 2012) and for calculating network distances between habitat patches in spatially restricted species (Grinnell and Curtis 2011).

8.0. REFERENCES

- Akçakaya, H. R. 2000. Viability analyses with habitat-based metapopulation models *Population Ecology* **42**:45-53.
- Akçakaya, H. R. 2005. RAMAS GIS: Linking Landscape Data with Population Viability Analysis (ver. 5.0). Applied Biomathematics, Setauket, New York.
- Akçakaya, H. R. and P. Sjögren-Gulve. 2000. Population viability analysis in conservation planning: an overview. *Ecological Bulletins* **48**:9-21.
- Araújo, M. B. and A. Guisán. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* **33**:1677-1688.
- Araujo, M. B. and M. New. 2007. Ensemble forecasting of species distributions. *TRENDS in Ecology and Evolution* **22**:42-47.
- Araújo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005. Validation of species–climate impact models under climate change *Global Change Biology* **11**:1504-1513.
- Austin, M. P., A. O. Nicholls, and C. R. Margules. 1990. Measurement of the realized qualitative niche: environmental niches of five Eucalyptus species. *Ecological Monographs* **60**:161-177.
- Baguette, M. 2004. The classical metapopulation theory and the real, natural world: a critical appraisal *Basic and Applied Ecology* **5**:213-224.
- Baguette, M. and H. VanDyck. 2007. Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology* **22**:1117-1129.
- Beissinger, S. R., J. R. Walters, D. G. Catanzaro, K. G. Smith, J. B. Dunning, S. M. Haig, B. R. Noon, and B. M. Stith. 2006. Modeling approaches in Avian conservation and the role of field biologists. *Ornithological Monographs* **59**:1-56.
- Bivand, R. S., P. Edzer, and V. Gomez-Rubio. 2008. *Applied spatial data analysis with R*. Springer, New York.
- Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, New York, U.S.A.
- Boersma, P. D., P. Kareiva, W. F. Fagan, J. A. Clark, and J. M. Hoekstra. 2001. How Good Are Endangered Species Recovery Plans? *BioScience* **51**:643-649.
- Bolker, B. M. 2008. *Ecological models and data in R*. Princeton University Press, Princeton, NJ.
- Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution* **24**:127-135.
- Borcard, D., F. Gillet, and P. Legendre. 2011. *Numerical ecology in R*. Springer, New York.
- Boudreau, P. R. and L. M. Dickie. 1989. Biological model of fisheries production based on physiological and ecological scalings of body size. *Canadian Journal of Fisheries and Aquatic Sciences* **46**:614-623.
- Brown, J. H., D. W. Mehlman, and G. C. Stevens. 1995. Spatial variation in abundance. *Ecology* **76**:2028-2043.
- Burgman, M., D. B. Lindenmayer, and J. Eltith. 2005. Managing landscapes for conservation under uncertainty. *Ecology* **86**:2007-2017.

- Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.
- Carroll, R., C. Augspurger, A. Dobson, J. Franklin, G. Orians, W. Reid, R. Tracy, D. Wilcove, and J. Wilson. 1996. Strengthening the use of science in achieving the goals of the Endangered Species Act: an assessment by the Ecological Society of America. *Ecological Applications* **6**:1-11.
- Carroll, C., R. F. Noss, P. C. Paquet, and N. H. Schumaker. 2004. Extinction debt of protected areas in developing landscapes. *Conservation Biology* **18**:1110-1120.
- Carroll, S. S. and D. L. Pearson. 1998. The effects of scale and sample size on the accuracy of spatial predictions of tiger beetle (Cicindelidae) species richness. *Ecography* **21**:401-414.
- Chambers, J. M. and T. J. Hastie. 1993. Statistical Models. Chapman and Hall, London.
- Clark, J. A. and E. Harvey. 2002. Assessing multi-species recovery plans under the Endangered Species Act. *Ecological Applications* **12**:655-662.
- Cohen, J. 1968. Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* **70**:213-220.
- Crawley, M. J. 2007. The R book. John Wiley and Sons Ltd., West Sussex, England.
- Crone, E. E., D. Doak, and J. Pokki. 2001. Ecological influences on the dynamics of a field vole metapopulation. *Ecology* **82**:831-843.
- Culter, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. *Ecology* **88**:2783-2792.
- Cunningham, R. B. and D. B. Lindenmayer. 2005. Modeling count data of rare species: Some statistical issues. *Ecology* **86**:1135-1142.
- Curtis, J.M.R. and I.R. Naujokaitis-Lewis. 2008a. Sensitivity of population viability to spatial and nonspatial parameters using GRIP. *Ecological Applications* **18**:1002-1013.
- Curtis, J. M. R. and I. Naujokaitis-Lewis. 2008b. Source code for GRIP 1.0. *Ecological Archives* A013-033-S1.
- Curtis, J.M.R. and A.C.J. Vincent. 2008. Use of Population Viability Analysis to Evaluate CITES Trade-Management Options for Threatened Marine Fishes. *Conservation Biology* **22**:1225-1232.
- Cyr, H., J. A. Downing, and R. H. Peters. 1997. Density-body size relationships in local aquatic communities. *Oikos* **79**:333-346.
- Dormann, C. F., J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Kissling, I. Kühn, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**:609-628.
- Dormann, C. F., O. Purschke, J. R. G. Márquez, S. Lautenbach, and B. Schröder. 2008. Components of uncertainty in species distribution analysis: A case study of the great grey shrike. *Ecology* **89**:3371-3386.
- Dreschler, M., K. Frank, I. Hanski, R. B. O'Hara, and C. Wissel. 2003. Ranking metapopulation extinction risk: From patterns in data to conservation management decisions. *Ecological Applications* **13**:990-998.

- Economo, E. P. 2011. Biodiversity conservation in metacommunity networks: linking pattern and persistence. *American Naturalist* **177**:167-180.
- Efron, B. 1986. How biased is the apparent error rate of a prediction rule? *Journal of the American Statistical Association* **81**:461-470.
- Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* **77**:802-813.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**:129-151.
- Ellison, A. M. 2004. Bayesian inference in ecology. *Ecology Letters* **7**:509-520.
- Fagan, W. F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology* **83**:3243-3249.
- Faraway, J. J. 2006. *Extending the linear model with R*. Chapman & Hall, Boca Raton, FL.
- Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence absence models. *Environmental Conservation* **24**:38-49.
- Foley, P. 1994. Predicting extinction times from environmental stochasticity and carrying capacity. *Conservation Biology* **8**:124-137.
- Fortin, M. J. and M. R. T. Dale. 2005. *Spatial analysis – a guide for ecologists*. Cambridge University Press, London.
- Frank, K. and C. Wissel. 1998. Spatial aspects of metapopulation survival: from model results to rules of thumb for landscape management. *Landscape Ecology* **13**:363-379.
- Franklin, J., K. E. Wejnert, S. A. Hathaway, C. J. Rochester, and R. N. Fisher. 2009. Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. *Diversity and Distributions* **15**:167-177.
- Friedman, J. 1991. Multivariate adaptive regression splines. *Annals of Statistics* **19**:1-141.
- Friedman, J., T. Hastie, and R. Tibshirani. 2000. Additive logistic regression: A statistical view of boosting. *The Annals of Statistics* **28**:337-374.
- George, E. I. and R. E. McCulloch. 1993. Variable selection via Gibbs sampling. *Journal of the American Statistical Association* **88**:881-889.
- Gotelli, N. J. and C. M. Taylor. 1999. Testing metapopulation models with stream-fish assemblages. *Evolutionary Ecology Research* **1**:835-845.
- Grimm, V. and I. Storch. 2000. Minimum viable population size of capercaillie *Tetrao urogallus*: Results from a stochastic model. *Wildlife Biology* **5**:219-225.
- Grimm, V., H. Lorek, J. Finke, F. Koester, M. Malachinski, M. Sonnenschein, A. Moilanen, I. Storch, A. Singer, C. Wissel, and K. Frank. 2004. META-X: generic software for metapopulation viability analysis. *Biodiversity and Conservation* **13**:165-188.

- Grinnell, M.H. and Curtis, J.M.R. 2011. User manual for NetworkDistances 1.0: Calculating network-wise distances between habitat patches for spatially restricted species. Can. Tech. Rep. Fish. Aquat. Sci. 2960: iv + 29 p.
- Grinnell, M. H. and J. M. R. Curtis. 2012. User manual for PatchImportance 1.0: Quantifying relative habitat patch importance based on metapopulation persistence and minimum abundance. Can. Tech. Rep. Fish. Aquat. Sci. 2977:vi + 41 p.
- Grossman, G. D. and M. C. Freeman. 1987. Microhabitat use in a stream fish assemblage. *Journal of Zoology* **212**:151-176.
- Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. E. Zimmermann. 2006. Using niche-based distribution models to improve the sampling of rare species. *Conservation Biology* **20**:501-511.
- Guisan, A. and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**:993-1009.
- Guisan, A. and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**:147-186.
- Hagen, A. N. and K. E. Hodges. 2006. Resolving critical habitat designation failures: reconciling law, policy, and biology. *Conservation Biology* **20**:399-407.
- Hall, D. B. 2000. Zero-inflated Poisson binomial regression with random effects: a case study. *Biometrics* **56**:1030-1039.
- Hastie, T. and R. Tibshirani. 1987. Generalized additive models: some applications. *Journal of the American Statistical Association* **82**:371-386.
- Hanski, I. 1994. Patch-occupancy in fragmented landscapes. *Trends in Ecology and Evolution* **9**:131-135.
- Hanski, I. 1999. *Metapopulation Ecology*. Oxford University Press, New York.
- Hanski, I. 2004. Metapopulation theory, its use and misuse. *Basic and Applied Ecology* **5**:225-229.
- Hanski, I., J. Alho, and A. Moilanen. 2000. Estimating the parameters of survival and migration of individuals in metapopulations. *Ecology* **81**:239-251.
- Hanski, I. and M. E. Gilpin. 1997. *Metapopulation biology: Ecology, genetics, and evolution*. Academic Press, San Diego, California.
- Hanski, I., A. Moilanen, T. Pakkala, and M. Kuussaari. 1996. The quantitative incidence function model and persistence of an endangered butterfly metapopulation. *Conservation Biology* **10**:578-590.
- Hansson, L. 1991. Dispersal and connectivity in metapopulations. *Biological Journal of the Linnean Society* **42**:89-103.
- Harrison, S. 1991. Local extinction in a metapopulation context. *Biological Journal of the Linnean Society* **42**:73-888.
- Hartley, S., R. Harris, and P. J. Lester. 2006. Quantifying uncertainty in the potential distribution of an invasive species: climate and the Argentine ant. *Ecology Letters* **9**:1068-1079.
- Hastie, T., R. Tibshirani, and J. H. Friedman. 2001. *The elements of statistical learning: data mining, inference, and prediction*. Springer, Berlin, Germany.
- Heinrichs, J. A., D. A. Bender, D. L. Gummer, and N. H. Schumaker. 2010. Assessing critical habitat: Evaluating the relative contribution of habitats to population persistence. *Biological Conservation* **143**:2229-2237.

- Heinz, S. K., L. Conradt, C. Wissel, and K. Frank. 2005. Dispersal in fragmented landscapes: Deriving a practical formula for patch accessibility. *Landscape Ecology* **20**:83-99.
- Heinz, S. K., C. Wissel, and K. Frank. 2006. The viability of metapopulations: individual dispersal behaviour matters. *Landscape Ecology* **21**:77-89.
- Hildrew, A. G., D. G. Raffaelli, and R. Edmonds-Brown. 2007. *Body size: The structure and function of aquatic ecosystems*. Cambridge University Press, Cambridge.
- Hill, J. K., C. D. Thomas, and O. T. Lewis. 1996. Effects of habitat patch size and isolation on dispersal by *Hesperia comma* butterflies: implications for metapopulation structure. *Journal of Animal Ecology* **65**:725-735.
- Hosmer, D. W. and S. Lemeshow. 1989. *Applied Logistic Regression*. John Wiley and Sons, New York.
- Hunsaker, C. T. and D. A. Levine. 1995. Hierarchical approaches to the study of water quality in rivers. *BioScience* **45**:193-203.
- Hutchings, J. A. and J. D. Reynolds. 2004. Marine fish population collapses: consequences for recovery and extinction risk. *BioScience* **54**:297-309.
- Hutchinson, G. E. 1953. The concept of pattern in ecology. *Proceedings of the National Academy of Sciences USA* **105**:1-12.
- Jackson, D. A. and H. H. Harvey. 1989. Biogeographic associations in fish assemblages: local versus regional processes. *Ecology* **70**:1472-1484.
- Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities: the roles of biotic, abiotic and spatial factors? *Canadian Journal of Fisheries and Aquatic Sciences* **58**:157-170.
- Jelks, H. L., S. J. Walsh, N. M. Burkhead, S. Contreras-Balderas, E. Diaz-Pardo, D. A. Hendrickson, J. Lyons, N. E. Mandrak, F. McCormick, J. S. Nelson, S. P. Platania, B. A. Porter, C. B. Renaud, J. J. Schmitter-Soto, E. B. Taylor, and M. L. Warren. 2008. Conservation status of imperiled North American freshwater and diadromous fishes *Fisheries* **33**:372-386.
- Jiménez-Valverde, A. and J. M. Lobo. 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica* **31**:361-369.
- Johnson, J. B. and K. S. Omland. 2004. Model selection in ecology and evolution. *TRENDS in Ecology and Evolution* **19**:101-108.
- Jurka, T. P. 2011. Maxent v.1.3.1: low-memory multinomial logistic regression with support for text classification. URL: <http://cran.rproject.org/web/packages/maxent/index.html>
- Keeling, M. J. 2002. Using individual-based simulations to test the Levins metapopulation paradigm. *Journal of Animal Ecology* **71**:270-279.
- Kindvall, O. 2000. Comparative precision of three spatially realistic simulation models of metapopulation dynamics. *Ecological Bulletins* **48**:101-110.
- King, R. 2010. *Bayesian analysis for population ecology*. Chapman and Hall/CRC, Boca Raton, FL.
- Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* **142**:911.

- Lehmann, A., J. McOverton, and J. R. Leathwick. 2003. GRASP: generalized regression analysis and spatial prediction. *Ecological Modelling* **160**:165-183.
- Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, and S. Aulagnier. 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* **90**:39-52.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* **73**:1943-1967.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.* **71**:237-240.
- Lobo, J. M. 2008. More complex distribution models or more representative data? *Biodiversity Informatics* **5**:14-19.
- Lobo, J. M., A. Jiménez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**:145-151.
- Loiselle, B. A., C. A. Howell, C. H. Graham, J. M. Goerck, T. Brooks, K. G. Smith, and P. H. Williams. 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* **17**:1591-1600.
- MacArthur, R. H. 1968. The theory of the niche. Pages 159-176 *in* R. C. Lewontin, editor. *Population biology and evolution*. Syracuse University Press, Syracuse, New York.
- Mace, G. M., H. P. Possingham, and N. Leader-Williams. 2005. Prioritizing choices in conservation. Pages 17-34 *in* D. MacDonald and K. Services, editors. *Key topics in conservation biology*. Blackwell Publishers, Oxford.
- MacNally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between - and reconciliation of - 'predictive' and 'explanatory' models. *Biodiversity and Conservation* **9**:655-671.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**:2248-2255.
- Maggini, R., A. Lehmann, N. E. Zimmermann, and A. Guisan. 2006. Improving generalized regression analysis for the spatial prediction of forest communities. *Journal of Biogeography* **33**:1729-1749.
- Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**:921-931.
- Marmion, M., M. Parviainen, M. Luoto, R. K. Heikkinen, and W. Thuiller. 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* **15**:59-69.
- Marsh, D. 2008. Metapopulation viability analysis for amphibians *Animal Conservation* **11**:463-465.
- Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters* **8**:1235-1246.
- McCullagh, P. and J. A. Nelder. 1989. *Generalised linear models*. 2nd edition. Chapman and Hall, London.

- McIntrie, E. J. B., C. B. Schlotz, and E. E. Crone. 2007. Designing a network for butterfly habitat restoration: where individuals, populations and landscapes interact. *Journal of Applied Ecology* **44**:725-736.
- Meynard, C. N. and J. F. Quinn. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography* **34**:1455-1469.
- Milborrow, S. 2011. earth v.3.2-1: Multivariate adaptive regression spline models. URL: <http://cran.r-project.org/web/packages/earth/index.html>
- Minns, C. K. 1995. Allometry of home range size in lake and river fishes. *Canadian Journal of Fisheries and Aquatic Sciences* **52**:1499-1508.
- Minns, C. K. 2003. An area-per-individual (API) model for estimating critical habitat requirements in aquatic species-at-risk. DFO Can. Sci. Adv. Sec. Res. Doc. 2003/074.
- Minns, C. K., R. G. Randall, J. E. Moore, and V. W. Cairns. 1986. A model simulating the impact of habitat supply limits on northern pike, *Esox Lucius*, in Hamilton Harbour, Lake Ontario. *Canadian Journal of Fisheries and Aquatic Sciences* **53**:20-34.
- Moilanen, A. 1999. Patch occupancy models of metapopulation dynamics: Efficient parameter estimation using implicit statistical inference. *Ecology* **80**:1031-1043.
- Moilanen, A. 2004. SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. *Ecological Modelling* **179**:533-550.
- Moilanen, A. and I. Hanski. 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. *Ecology* **79**:2503-2515.
- Moisen, G. G., E. A. Freeman, J. A. Blackard, T. S. Frescino, N. E. Zimmermann, and T. C. Edwards. 2006. Predicting tree species presence and basal area in Utah: A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling* **199**:176-187.
- Morris, W. F. and D. F. Doak. 2002. Quantitative conservation biology: Theory and practice of population viability analysis. Sinauer Associates Inc., Sutherland, MA.
- Murtaugh, P. A. 2009. Performance of several variable-selection methods applied to real ecological data. *Ecology Letters* **12**:1061-1068.
- Nakagawa, S. and I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews* **82**:591-605.
- Naujokaitis-Lewis, I. R., J. M. R. Curtis, P. Arcese, and J. S. Rosenfeld. 2009. Sensitivity analyses of spatial population viability analysis models for species at risk and habitat conservation planning. *Conservation Biology* **23**:225-229.
- North, A. and O. Ovaskainen. 2007. Interactions between dispersal, competition and landscape heterogeneity. *Oikos* **116**:1106-1119.
- Oksanen, J. 2004. Incidence Function Model in R. URL: <http://cc.oulu.fi/~jarioksa/opetus/openmeta/metafit.pdf>
- Ovaskainen, O. and I. Hanski. 2004. From individual behavior to metapopulation dynamics: Unifying the patchy population and classic metapopulation models. *American Naturalist* **164**:364-377.
- Olden, J. D. and D. A. Jackson. 2000. Torturing data for the sake of generality: how valid are regression models. *Ecoscience* **7**:501-510.

- Olden, J. D. and D. A. Jackson. 2001. Fish habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. *Transactions of the American Fisheries Society* **130**:878-897.
- Olden, J. D. and D. A. Jackson. 2002. A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* **47**:1976-1995.
- Olden, J. D. and D. A. Jackson. 2002b. Illuminating the 'black box': Understanding variable contributions in artificial neural networks. *Ecological Modelling* **154**:135-150.
- Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society* **131**:329-336.
- Olden, J. D., J. J. Lawler, and N. L. Poff. 2008. Machine learning methods without tears: A primer for ecologists. *Quarterly Review of Biology* **83**:171-193.
- OMNR. 2008. Southern Ontario Land Resource Information System (SOLRIS) Phase 2 – Data Specifications. Ontario Ministry of Natural Resources, Science and Information Branch, Peterborough, Ontario.
- Ovaskainen, O. and I. Hanski. 2004. From individual behavior to metapopulation dynamics: Unifying the patchy population and classic metapopulation models. *American Naturalist* **164**:364-377.
- Ozgul, A., K. B. Armitage, D. T. Blumstein, D. H. vanVueren, and M. K. Oli. 2006. Effects of patch quality and network structure on patch occupancy dynamics of a yellow-bellied marmot metapopulation. *Journal of Animal Ecology* **75**:191-202.
- Pascual-Hortal, L. and S. Saura. 2008. Integrating landscape connectivity in broad-scale forest planning through a new graph-based habitat availability methodology: application to capercaillie (*Tetrao urogallus*) in Catalonia (NE Spain) *European Journal of Forest Research* **127**:23-31
- Pampel, F. C. 2000. Logistic regression: A primer. Sage Publications, London.
- Pearce, J. and S. Ferrier. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling* **128**:127-147.
- Pearson, R. G., W. Thuiller, M. B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson, and D. C. Lees. 2006. Model-based uncertainty in species' range prediction. *Journal of Biogeography* **33**:1704-1711.
- Peeters, E. and J. P. Gardeniers. 1998. Logistic regression as a tool for defining habitat requirements of two common gammarids. *Freshwater Biology* **39**:605-615.
- Peters, R. H. 1986. The ecological implications of body size. Cambridge University Press, Cambridge.
- Phillips, S. J., D. R. Anderson, and R. E. Shapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**:231-259.
- Poff, N. L., B. P. Bledsoe, and C. O. Cuhaciyan. 2006. Hydrologic variation with land use across the contiguous United States: Geomorphic and ecological consequences for stream ecosystems. *Geomorphology* **79**:264-285.
- Pollock, J. F. 2006. Detecting population declines over large areas with presence-absence, time-to-encounter, and count survey methods *Conservation Biology* **20**:882-893.

- Poos, M. S. and D. A. Jackson. In Revision. Conservation by consensus: reducing uncertainties in modeling the distribution of endangered species using habitat-based ensemble models. *Ecography*.
- Poos, M. S., N. E. Mandrak, and R. L. McLaughlin. 2007. The effectiveness of two common sampling methods for sampling imperiled freshwater fishes. *Journal of Fish Biology* **70**:691-708.
- Poos, M. S., N. E. Mandrak, and R. L. McLaughlin. 2008. A practical framework for selecting among single species, multi-species and ecosystem-based recovery plans. *Canadian Journal of Fisheries and Aquatic Sciences* **65**:2656-2666.
- Poos, M. S., D. Lawrie, C. Tu, D. A. Jackson, and N. E. Mandrak. 2012. Developing local and regional population estimates for an endangered freshwater minnow, the reidside dace (*Clinostomus elongatus*), in Canada. *Aquatic Conservation* **22**:47-57.
- Possingham, H. and I. Davies. 2003. ALEX: A model for the viability analysis of spatially structured population s. *Biological Conservation* **73**:143-150.
- Quinn, G. P. and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge, UK.
- R Development Core Team. 2011. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (ISBN: 3-900051-07-0).
- Randall, R. G., J. R. M. Kelso, and C. K. Minns. 1995. Fish production in freshwaters: are rivers more productive than lakes? . *Canadian Journal of Fisheries and Aquatic Sciences* **52**:631-643.
- Randall, R. G. 2008. Narrative description and quantification of the habitat requirements of Lake Sturgeon, *Acipenser fulvescens* in the Great Lakes and upper St. Lawrence River. DFO Can. Sci. Advis. Sec. Res. Doc. 2008/015.
- Revilla, E. and T. Wiegand. 2008. Individual movement behavior, matrix heterogeneity, and the dynamics of spatially structured populations. *PNAS* **105**:19120-19125.
- Ridgeway, G. 1999. The state of boosting. *Computing Sciences and Statistics* **31**:172-181.
- Ridgeway, G. 2007. General boosted models: a guide to the gbm package. URL: <http://cran.r-project.org/web/packages/gbm/vignettes/gbm.pdf>
- Ripley, B. 2009. nnet: Feed-forward neural networks and multinomial log-linear models. URL: <http://cran.r-project.org/web/packages/nnet/index.html>
- Ripley, B. 2011. Package 'tree' v. 1.0-29. For R programming language. URL: <http://cran.r-project.org/web/packages/tree/index.html>
- Robert, A. 2009. The effects of spatially correlated perturbations and habitat configuration on metapopulation persistence. *Oikos* **118**:1590-1600.
- Rodríguez, J. P., L. Brotons, J. Bustamante, and J. Seoane. 2007. The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions* **13**:243–251.
- Roitberg, B. D. and M. Mangel. 1997. Individuals on the landscape: Behavior can mitigate landscape differences among habitats. *Oikos* **80**:234-240.
- Ruckelshaus, M. H., P. Levin, J. B. Johnson, and P. M. Kareiva. 2002. The Pacific Salmon wars: what science brings to the challenge of recovering species. *Annual Review of Ecology and Systematics* **33**:665-706.

- Rykiel, E. J. 1996. Testing ecological models: the meaning of validation. *Ecological Modelling* **90**:229-244.
- Saura, S. and J. Torne. 2009. Conefor Sensinode 2.2: A software package for quantifying the importance of habitat patches for landscape connectivity. *Environmental Modelling and Software* **24**:135-139.
- Schmidt-Nielsen, K. 1984. *Scaling: why is animal size so important?* Cambridge University Press, Cambridge.
- Shreeve, T. G., R. L. H. Dennis, and H. VanDyck. 2004. Resources, habitats and metapopulations - whither reality? *Oikos* **106**:404-408.
- Schtickzelle, N., G. Mennechez, and M. Baguette. 2006 Dispersal depression with habitat fragmentation in the bog fritillary butterfly. *Ecology* **87**:1057-1065.
- Schumaker, N. H. 1998. A users guide to the PATCH model EPA/600/R-98/135. U.S. Environmental Protection Agency, Environmental Research Laboratory, Corvallis, Oregon.
- Schumaker, N. H. 2008. HexSim (Version 1.3.6.9). US Environmental Protection Agency, Environmental Research Laboratory, Corvallis, Oregon.
- Sjogren-Gulve, P. 1991. Extinctions and isolation gradients in metapopulations: the case of the pool frog (*Rana lessonae*). *Biological Journal of the Linnean Society* **42**:135-147.
- Soberón, J. and A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species distributional areas. *Biodiversity Informatics* **2**:1-10.
- Southgate, R. and H. Possingham. 1995. Modelling the reintroduction of the greater bilby *Macrotis lagotis* using the metapopulation model Analysis of the Likelihood of Extinction (ALEX). *Biological Conservation* **73**:151-160.
- Sponseller, R. A., E. F. Benfield, and H. M. Valett. 2001. Relationships between land use, spatial scale and stream macroinvertebrate communities. *Freshwater Biology* **46**:1409-1424.
- Stockwell, D. and D. Peters. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information* **13**:143-158.
- Stockwell, D. R. B. and A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* **148**:1-13.
- Taylor, M. T., K. F. Suckling, and J. J. Rachlinski. 2005. The effectiveness of the Endangered Species Act: a quantitative analysis. *BioScience* **55**:360-368.
- ter Braak, C. J. F., I. Hanski, and J. Verboom. 1998. The incidence function approach to modelling of metapopulation dynamics. Pages 167-188 in J. Bascompte and R. V. Solé, editors. *Modeling spatio-temporal dynamics in ecology*. Springer, New York.
- Thomas, C. D., R. J. Wilson, and O. T. Lewis. 2002. Short-term studies underestimate 30-generation changes in a butterfly metapopulation. *Proceedings of the Royal Society of London B* **269**:563-569.
- Thompson, W. L. 2004. *Sampling rare or elusive species*. Island Press, Washington D.C.

- Thuiller, W. 2003. BIOMOD: optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology* **9**:1353-1362.
- Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology* **10**:2020-2027.
- Thuiller, W. F., G. Midgley, M. Rougeti, and R. Cowling. 2006. Predicting patterns of plant species richness in megadiverse South Africa. *Ecography* **29**:733-744.
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B* **58**:267-288.
- Tischendorf, L. 2001. Can landscape indices predict ecological processes consistently? *Landscape Ecology* **16**:235-254.
- Tonn, W. M. 1990. Climate change and fish communities: A conceptual framework. *Transactions of the American Fisheries Society* **119**:337-352.
- Turchin, P. B. 1998. Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates, Sunderland, MA.
- Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Paris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys by estimating false negative error rates in presence-absence data. *Ecological Applications* **13**:1790-1801.
- Utz, R. M. and R. H. Hilderbrand. 2011. Interregional variation in urbanization-induced geomorphic change and macroinvertebrate habitat colonization in headwater streams. *Journal of the North American Benthological Society* **30**:25-37.
- Vayssières, M. P., R. E. Plant, and B. H. Allen-Diaz. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science* **11**:679-694.
- Vélez-Espino, L. A. and M. A. Koops. 2008. Recovery potential assessment of reddsidedace (*Clinostomus elongatus*) in Canada. Canadian Science Advisory Secretariat, Ottawa, ON.
- Vélez-Espino, L. A. and M. A. Koops. 2009a. Recovery potential assessment for lake sturgeon in Canadian designatable units. *North American Journal of Fisheries Management* **29**:1065-1075.
- Vélez-Espino, L. A. and M. A. Koops. 2009b. Quantifying allowable harm in species at risk: application to the Laurentian black redhorse (*Moxostoma duquesnei*). *Aquatic Conservation: Marine and Freshwater Ecosystems* **19**:676-688.
- Vélez-Espino, L. A., R. G. Randall, and M. A. Koops. 2010. Quantifying habitat requirements of four freshwater species at risk in Canada: Northern Madtom, Spotted Gar, Lake Chubsucker, and Pugnose Shiner. *DFO Can. Sci. Advis. Sec. Res. Doc.* 2009/115.
- Venables, W.N. and Ripley, B.R. 2011. MASS: Support Functions and Datasets for Venables and Ripley's MASS, in, R package version 7.3-11. URL: <http://cran.r-project.org/web/packages/MASS/index.html>
- Vincent, P. J. and J. M. Haworth. 1983. Poisson regression models of species abundances. *Journal of Biogeography* **10**:153-160.
- Vos, C. C., J. Verboom, P. F. M. Opdam, and C. J. F. ter Braak. 2001. Toward ecologically scaled landscape indices. *American Naturalist* **157**:24-41.

- Welsh, A. H., R. B. Cunningham, C. F. Donnelly, and D. B. Lindenmayer. 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling* **88**:297-308.
- Whittingham, M. J., P. A. Stephens, R. B. Bradbury, and R. P. Freckleton. 2006. Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology* **75**:1182-1189.
- Wilson, K. A., M. I. Westphal, H. P. Possingham, and J. Elith. 2005. Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation* **122**:99-112.
- Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, and A. Guisán. 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* **14**:763-773.
- Zeileis, A., C. Kleiber, and S. Jackman. 2009. Regression models for count data in R. URL: <http://cran.r-project.org/web/packages/pscl/vignettes/countreg.pdf>