



**HAL**  
open science

## Structural uncertainty in models projecting the consequences of habitat loss and fragmentation on biodiversity

Shawn J. Leroux, Cécile H. Albert, Anne-Sophie Lafuite, Bronwyn Rayfield, Shaopeng Wang, Dominique Gravel

► **To cite this version:**

Shawn J. Leroux, Cécile H. Albert, Anne-Sophie Lafuite, Bronwyn Rayfield, Shaopeng Wang, et al.. Structural uncertainty in models projecting the consequences of habitat loss and fragmentation on biodiversity. *Ecography*, 2017, 40 (1), pp.36-47. 10.1111/ecog.02542 . hal-01681549

**HAL Id: hal-01681549**

**<https://hal.science/hal-01681549>**

Submitted on 4 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Accepted Article

## Structural uncertainty in models projecting the consequences of habitat loss and fragmentation on biodiversity

Shawn J. Leroux<sup>1</sup>, Cécile H. Albert<sup>2</sup>, Anne-Sophie Lafuite<sup>3</sup>, Bronwyn Rayfield<sup>4</sup>, Shaopeng Wang<sup>5,6</sup>,

Dominique Gravel<sup>7</sup>

<sup>1</sup>Department of Biology, Memorial University of Newfoundland, St. John's, NL, Canada

<sup>2</sup>Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE), Aix Marseille Université, CNRS, IRD, Avignon Université, Marseille, France

<sup>3</sup>Centre for Biodiversity Theory and Modelling, Theoretical and Experimental Ecology Station, CNRS, Paul Sabatier University, Moulis, France

<sup>4</sup>Département des Sciences Naturelles, Université du Québec en Outaouais, Institut des sciences de la forêt tempérée, Ripon, QC, Canada

<sup>5</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

<sup>6</sup>Institute of Ecology, Friedrich Schiller University Jena, Jena, Germany

<sup>7</sup>Département de biologie, Université de Sherbrooke, Sherbrooke, QC, Canada

**Corresponding author:** Shawn J. Leroux, Department of Biology, Memorial University of Newfoundland, St. John's, NL, Canada. E-mail: sleroux@mun.ca

**Decision date:** 05-Sep-2016

---

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/ecog.02542].

## ABSTRACT

Ecological theory is essential to predict the effects of global changes such as habitat loss and fragmentation on biodiversity. The species-area relationship (SAR), metapopulation models (MEP) and species distribution models (SDM) are commonly used tools incorporating different ecological processes to explain biodiversity distribution and dynamics. Yet few studies have compared the outcomes of these disparate models and investigated their complementarity. Here we show that the processes underlying SAR (patch area), MEP (patch isolation) and SDM (environmental conditions) models can be compared with a common statistical framework. Our approach allows for species and community-level predictions under current and future landscape scenarios, facilitates multi-model comparison and provides the machinery for integrating multiple mechanisms into one model. We apply this framework to the distribution of eight focal vertebrate species in current and future projected landscapes where 10% of the landscape is lost to land-use change in southwestern, Quebec, Canada. Based on a model selection approach, we found that a model including patch area was the top ranked model for four of our focal species and models including patch isolation and environmental conditions were the top ranked models for two focal species each. Community-level predictions of models based on patch area, patch isolation and environmental conditions for both current and future landscapes showed high spatial overlap, however, patch area models always predicted a reduction of species richness per patch whereas both the patch isolation and environmental conditions models predicted an increase or decrease in species richness per patch following habitat loss and fragmentation. Our comparative tool will allow ecologists and conservation practitioners to relate structural uncertainty to key mechanisms underlying each model. Ultimately, this approach is one step in the direction of deriving robust predictions for the change and loss of biodiversity under global change, which is key for informing conservation plans.

**Key words:** Conservation, Generalized Linear Models, Global Change, Metapopulation model, Prediction, Species-Area Relationship, Species Distribution Model

## INTRODUCTION

Tools for predicting biodiversity change are essential for biodiversity conservation in the face of the global changes (Cardinale et al. 2012, Vellend et al. 2013, Dornelas et al. 2014, Ewers et al. This issue). Ecologists have derived from theory (Thuiller et al. 2013) a suite of models to predict the distribution, abundance, and diversity of species based on environmental conditions and ecological processes (hereafter biodiversity models). Some biodiversity models are commonly used for biodiversity conservation and land-use planning (e.g. species distribution models; Leroux et al. 2007) while others have seen limited applications (e.g. neutral theory). In most case studies, however, a single biodiversity model is applied to address a particular conservation problem (but see e.g. Moilanen and Hanski 1998, Keith et al. 2008, Dullinger et al. 2011, Fordham et al. 2013), therefore, we lack a good understanding of how the predictions from various models may differ, and the potential for multiple models to provide complementary insights on a given problem (Krosby et al. 2015, Brudvig et al. This issue).

It is not surprising that researchers and conservation practitioners usually apply a single biodiversity model to address a particular conservation problem. The study area, biodiversity element of interest, conservation goals and data availability may dictate the choice of model. For example, a species with very strict habitat requirements, such as the Glanville fritillary butterfly (*Melitaea cinxia*) in Finland occupies a landscape of habitat patches surrounded by inhospitable matrix and may be most amenable to metapopulation models (e.g. Moilanen and Hanski 1998). Given that the majority of respondents to a recent survey on quantitative training among early career ecologists were not comfortable with their understanding of mathematical models within their field (Barraquand et al. 2014), we surmise that researcher and practitioner experience with quantitative techniques may also explain the paucity of case studies applying multiple biodiversity models. But perhaps the most important reason for the lack of multi-model comparisons is that we do not have a framework for comparing models that are based on different assumptions (e.g. environmental filtering, species-area relationships), parameterized with different types of data (e.g. abundance, presence/absence, co-occurrence), and which provide disparate predictions (e.g. species-level, community-level).

Biodiversity models have a key role to play in exploring the causes and consequences of biodiversity change (Pereira et al. 2010, Thuiller et al. 2013). The primary driver of species extinctions and turnover is

Accepted Article

habitat loss and fragmentation (Newbold et al. 2015). Habitat loss and fragmentation results in the direct removal of species and their habitat but also the indirect loss of species by providing the pathways for resource extraction (e.g. fishing, forestry) and the arrival of non-native organisms. For example, Kaufman et al. (2009) report higher exploitation rates of native lake trout (*Salvelinus namaycush*) and higher prevalence of non-native smallmouth bass (*Micropterus dolomieu*) in lakes with good road access than lakes with poor road access in northeastern Ontario, Canada. Many biodiversity models have been used to predict the effects of habitat loss and fragmentation on biodiversity. For example, Cord and Rödder (2011) incorporate remotely sensed enhanced vegetation index data to improve predictions of anuran species distribution in fragmented habitats in Mexico. Developing a framework for comparing different biodiversity models will allow us to explore and potentially reduce structural uncertainty arising from model specific assumptions and to better manage model complexity.

Here we seek to fill this gap by showing how a common and simple statistical framework can be used for comparing the predictions from different biodiversity models. We begin by reviewing the history, formulation, predictions and applications in biodiversity conservation of three simple but widely applied biodiversity models; species-area relationships (SAR), metapopulation models (MEP), and species distribution models (SDM). Then we map the processes behind these three models to a common statistical framework and apply it to a systematic comparison with a case study of land-use planning in a fragmented landscape in southwestern Quebec, Canada. We end with perspectives for multi-model comparison and integration for real-world biodiversity conservation and land-use planning.

## REVIEW OF THREE SIMPLE BIODIVERSITY MODELS

We review species-area relationships, metapopulation models, and species distribution models (Table 1) because they are widely used in ecology and conservation and they are relatively simple with few parameters which makes them easy to parameterize in practice. Classic SAR makes predictions at the community-level whereas classic MEP and SDM models make predictions at the species-level.

### *Species-area relationships*

SARs are among the best-known and most documented observations in ecology (Rosenzweig 1995). SARs describe the dependence of species richness on area at various spatial scales, along environmental gradients and across landscape configurations. A power function  $S=cA^z$  was the first model proposed to predict the

Accepted Article

change of species richness ( $S$ ) with area ( $A$ ) across isolated ecosystems (Arrhenius 1921) where  $c$  and  $z$  are empirical constants. This formulation of SARs receives good support from ecological data and theory (Dengler 2009).

One general pattern revealed by SAR studies is that, all else being equal, larger areas are expected to support more species. Many mechanisms have been proposed to explain this observation, including the null hypothesis of random placement and passive sampling (Connor and McCoy 1979), the area per se hypothesis, i.e. lower extinction probability with larger areas (MacArthur and Wilson 1967), and the habitat diversity hypothesis, i.e. larger areas having greater habitat diversity (Rosenzweig 1995).

Recent models have tried to improve the predictive power of SARs through the integration of processes shaping species richness in fragmented landscapes (Whittaker et al. 2005). Specifically, insights from metapopulation theory led to the species-fragmented area-relationship model (Hanski et al. 2013), while niche theory is at the core of the matrix-calibrated SAR (Koh and Ghazoul 2010, Tanentzap et al. 2012) and the countryside SAR (Mendenhall et al. 2014). SARs have been used extensively in ecology and conservation biology to predict species richness from sample-area curves (e.g. Tjørve and Turner 2009), optimally design nature reserves (e.g. Kukkala and Moilanen 2013), and predict species extinctions based on habitat loss and fragmentation (e.g. Brooks et al. 1999).

#### ***Metapopulation models***

Spatially explicit metapopulation models describe the presence and absence of a species across a network of habitat fragments (Hanski 1999). The widely used incidence function model (IFM) for metapopulations represents the occupancy of each habitat patch as a Markov chain with two states, occupied (species is present) and vacant (species is absent) with transition probabilities between the two states determined by local extinction and dispersal-mediated colonization from neighboring patches (Hanski 1999).

The original IFM assumed extinction rates decreased with patch area, and colonization rates increased with the connectivity and area of occupied patches (Hanski 1999). While SAR and MEP models both consider the effects of patch area on species, the core of MEP theory is based on dispersal limitations and patch connectivity (Levins 1969, Hanski 1999). Connectivity between patches can be estimated with simple structural measures, such as straight-line distances between neighboring patches, or more complex

functional measures derived from network theory, circuit theory, or individual-based models (Coulon et al. 2015).

Both metapopulation (Akçakaya et al. 2006) and habitat connectivity (Crooks and Sanjayan 2006) models have become increasingly important in conservation. In future conservation applications, it may be important to combine these models because the relationship between habitat connectivity and ecological dynamics has been recently shown to depend not only on the spatial configuration of habitat (Holland and Hastings 2008, Thompson et al. 2014) but also on demographic parameters of the species (e.g. extinction and colonization probabilities; Gilarranz and Bascompte 2012).

### *Species distribution models*

Species distribution models (SDM) correlate species distribution data (e.g. occurrence, abundance) with environmental or spatial covariates to describe the relationship between occupancy and the environment and to predict species distribution within or outside the sampled area (Elith and Leathwick 2009, Araújo and Peterson 2012). Generalized linear models, boosted regression trees, random forest, and general additive models and the combination of several algorithms (i.e., ensemble forecasting procedure; Araújo and New 2007) are the most common statistical approaches for SDMs.

Given most SDMs predict a continuous probability of species occurrence, there are a number of methods for converting continuous model predictions to binary range classifications and these methods make a number of simplifying assumptions (see review in Liu et al. 2005).

While the strength of SDMs is the ease with which they can be parameterized for broad sets of species over large spatial extents, SDMs have been widely criticized because they do not account for many ecological processes underlying the observed species distributions (e.g. dispersal, biotic interactions; Elith and Leathwick 2009). A new generation of SDMs that integrate process-based approaches have modified classical SDMs in several ways: i) by including physiology (e.g. Kearney & Porter 2009), species interactions (e.g. Trainor & Schmitz 2014), population dynamics (e.g. Leroux et al. 2013) and movement (e.g. Vasudev et al. 2015), ii) by formulating hybrid models that pair correlative models with additional ecological processes (e.g. population viability model; Haby et al. 2013), and iii) by using co-occurrence data as a covariate in joint SDMs (Pollock et al. 2014).

SDMs are commonly used for a variety conservation applications including spatial conservation prioritization and land-use management (e.g. protected areas design; Leroux et al. 2007, Meller et al. 2014) and restoration planning (Chetkiewicz and Boyce 2009). They are also used to improve the sampling of species distributions (Guisan et al. 2006), or to map the potential risk of species invasions (e.g. Gallien et al. 2012), disease (e.g. Kulkarni et al. 2010), and species responses to climate change (e.g. Kerr et al. 2015).

## **A STATISTICAL FRAMEWORK TO MODEL THE CONSEQUENCES OF HABITAT LOSS AND FRAGMENTATION ON BIODIVERSITY**

Here we describe a general and common statistical approach to compare predictions from the different biodiversity models described above. We aim to emphasize the differences among biodiversity models and therefore focus on the key process represented by each of them.

### *A general model for species distribution in fragmented landscapes*

Let us consider a general situation of isolated patches of favorable habitat in a matrix of inhospitable areas. It is convenient to express the immigration and extinction dynamics at a location at the species level. The time dynamics of species  $i$  being present at location  $x$ ,  $p_{ix}$ , could be described as:

$$\frac{dp_{ix}}{dt} = I_{ix}(1 - p_{ix}) - E_{ix}p_{ix} \quad (1)$$

where  $I_{ix}$  is the immigration rate of species  $i$  from the ‘mainland’ to the location  $x$ , and  $E_{ix}$  is the extinction rate of species  $i$  at location  $x$ . Thus, both rates are species and location specific. At equilibrium, immigration balances extinction, and we get the probability of finding species  $i$  in locality  $x$ :

$$\hat{p}_{ix} = \frac{I_{ix}}{I_{ix} + E_{ix}} \quad (2)$$

where  $\hat{p}_{ix}$  denotes the equilibrium occurrence probability. We could also express this quantity with a linear model, using a logit transformation:

$$\text{logit}(\hat{p}_{ix}) = \log\left(\frac{\hat{p}_{ix}}{1 - \hat{p}_{ix}}\right) = \log(I_{ix}) - \log(E_{ix}) \quad (3)$$

It is straightforward to get back to the community level and derive the expected species richness at location  $x$  ( $\hat{S}_x$ ), assuming that species are distributed independently of each other (Gravel et al. 2011) and summing species-specific occurrence probabilities:

$$\hat{S}_x = \sum_{i=1}^V \hat{p}_{ix} \quad (4)$$



where  $V$  is the total number of species in the regional pool. If species are not distributed independently (e.g. predator-prey) then approaches such as joint species distribution models (Pollock et al. 2014) may be used for community level predictions.

It is not straightforward to measure  $I$  and  $E$  directly in the field and even less so for measuring species and location specific values. It is nonetheless possible to derive likely estimates according to species distribution data and some constraints such as patch area, isolation and local environmental conditions.

The distribution of individual species in a locality is usually modeled statistically as a binomial process conditional on local site characteristics:

$$\text{logit}(p_{ix}) = \beta_{i0} + \beta_{i1}v_{1x} + \dots + \beta_{im}v_{mx} \quad (5)$$

where  $v_{mx}$  is the value of covariate  $m$  at location  $x$  and  $\beta_{im}$  is the species-specific coefficient for this covariate. This, for instance, is a common approach for SDMs (reviewed in Elith and Leathwick 2009). Such a modeling framework may include linear and non-linear (e.g. quadratic) terms for covariates,  $v_{mx}$ , as well as interactions between covariates. It is flexible enough to include a diversity of covariates, such as patch area or isolation, allowing the direct comparison of SDM with SAR or MEP theories. To date several studies have incorporated patch and landscape characteristics into SDMs. For example, Dullinger et al. (2011), Betts et al. (2014), and McCune (2016) combine patch (e.g. area) and landscape (e.g. connectivity) and climate covariates to predict the distribution of Alpine plants in the Calcareous Alps of Austria, birds in Oregon, USA, and rare plants in southern Ontario, Canada, respectively. Boulangeat et al. (2012) added propagule pressure as a covariate in their SDMs of plant species in the French Alps. It is easy to map the above model (Eq. 5) onto Eq. 3 in order to interpret the resulting parameters in terms of immigration and extinction dynamics, thus explicitly bridging the gap with theory through the set of equations introduced above (Eqs 1-2).

Beginning with Eq. 5, we now turn to the interpretation of  $I_{ix}$  and  $E_{ix}$  from knowledge of patch area, patch isolation and environmental conditions, respectively, in order to link a single statistical approach with the different processes underlying the original SAR, MEP, and SDM models.

#### *Patch area: Application to species area relationship models*

It is commonly assumed that species extinction rates should decrease with increasing patch area (Hanski 1999). The underlying rationale is that larger patches sustain larger populations, thus reducing demographic

Accepted Article  
stochasticity (Lande et al. 2003) and the chance of random extinction despite favorable environmental conditions. Here, we assume  $E_{ix} = A^{-\beta_{i1}}$  and  $I_{ix}$  is a species-specific constant. According to Eq. 5, we could represent patch area dependent occurrence probabilities as:

$$\text{logit}(p_{ix}) = \beta_{i0} + \beta_{i1} \log(A) \quad (6)$$

where  $\beta_{i0}$  is a species-specific intercept related to the immigration rate. A SAR then can be computed by summing across all species, as described in Eq. 4.

#### *Patch isolation: Application to metapopulation models*

In the theory of island biogeography, MacArthur and Wilson (1967) hypothesized that  $I_{ix}$  scales inversely to the distance to the mainland. Later, Hanski (1999) developed spatially explicit metapopulation theory and generalized the approach to archipelagoes (or networks of patches), thus predicting that immigration should scale with the number of connected patches. More precisely, in a spatially explicit setting, the immigration rate could be described as:

$$I_{ix} = f\left(\sum_{y \in N_x} k_i(x, y) p_{iy}\right) \quad (7)$$

where  $f(\cdot)$  is some link function,  $N_x$  denotes the set of neighbor patches of  $x$ , and  $k_i(x, y)$  is a species-specific dispersal kernel describing the probability that a species colonizes patch  $x$  coming from patch  $y$ . Note that the summation is taken across all neighboring patches, assuming we know their occupancy  $p_{iy}$ . Unfortunately this information is often missing, in which case  $p_{iy}$  is simply considered as a constant (i.e. all neighboring patches have the same average occupancy) and  $I_{ix}$  depends simply on a local, species-specific measure of the connectivity of patch  $x$ :  $K_{ix} = \sum_{y=1}^N k_i(x, y)$ . Here we assume a power function for  $I_{ix} \propto K_{ix}^{\beta_{i1}}$  and  $E_{ix}$  is a species-specific constant. Using the same logic as Eq. 5, we would therefore model presence-absence from the connectivity matrix using the following statistical model:

$$\text{logit}(p_{ix}) = \beta_{i0} + \beta_{i1} \log(K_{ix}) \quad (8)$$

where  $\beta_{i0}$  is a species-specific intercept related to the extinction rate. This formulation of the MEP model focuses on dispersal limitations, which allows for quantitative comparison of predictions from models of patch area (i.e. SAR) and patch isolation (i.e. MEP).

#### *Environmental conditions: Application to species distribution models*

Local patch quality can also influence both species establishment success and extinction from a location. Establishment of a population is the end result of several processes taking place after first immigrants reach the location, from the initial dispersal event up to the development of a sustained population. Local environmental conditions could influence the germination success of seeds, the decision of an animal to stay at the location, or the growth of a population. Local conditions could also influence population size, and for the same argument as area, they could influence extinction rate. It is thus possible to express both immigration and the extinction rates as functions of local environmental conditions such that  $I_{ix} = f_i(v_{lx}, \dots, v_{mx})$  and  $E_{ix} = g_i(v_{lx}, \dots, v_{mx})$ . The corresponding statistical model of presence-absence could then be described by Eq. 5. Note that in this case, the absence of information about immigration and extinction dynamics prevents disentangling differential effects of the environment on each process. Such estimates could be obtained using state transition data.

*From theory to practice*

In practice, a species' occurrence probability does not necessarily follow the log-linear assumption described here. We took this approach as it maps naturally to generalized linear models, but emphasize that more flexible statistical models could be used. To compare the performance of patch area (SAR), patch isolation (MEP), or environmental conditions (SDM)-based models, we could compute model likelihoods and compare them by AIC, as long as we know the exact number of parameters. Goodness of fit for each model could be evaluated using statistics such as the pseudo  $R^2$ , or model predictive accuracy could be measured with the TSS (True Statistic Skill) and AUC (Area Under the receiver operating characteristic Curve, Allouche et al. 2006).

Overall, we show how we can build patch area, isolation and environmental conditions as covariates in a general statistical framework (Eq. 5) to produce models that map the underlying processes of SAR (Eq. 6), MEP (Eq. 8) and SDMs (Eq. 5) to a common currency. Although our SAR, MEP, and SDM formulations may differ from classic representations, we capture the essence of the classic models by incorporating key processes into our formulations.

**COMPARING PREDICTIONS OF BIODIVERSITY MODELS IN THE ST. LAWRENCE  
LOWLANDS SURROUNDING MONTREAL, CANADA**

We apply our statistical framework for comparing SAR, MEP, SDM models for focal species in a conservation case study in the fragmented, peri-urban landscape surrounding Montreal, Canada (centered on 45°400N, 73°150W). The landscape covers approximately 27 500 km<sup>2</sup>, the majority of which occurs in the fertile lowlands of the St. Lawrence River valley where small forest fragments are surrounded by agriculture and urban areas (Fig. 1, 2). We selected 8 vertebrate focal species: 5 birds and 3 amphibians, to represent the forest biodiversity in this region (Supplementary material Appendix 1). The focal species span a range of forest habitat preferences (i.e. composition and age ranging from relative generalists across all forest types to specialists preferring old, coniferous forest), patch size requirements (i.e. minimum patch area ranging from 0.5 to 5 ha) and dispersal abilities (i.e. median dispersal ranging from 16 to 40,889 m; Table 2, Supplementary material Appendix 1).

Following our framework, we fit statistical models of the presence/pseudo-absence of a focal species in a patch with a binomial error structure and logit link. We used a 1:1 presence:pseudo-absence sampling scheme with 100 different pseudo-absence data sets for each species. We compared the three model classes (patch area, patch isolation, environmental conditions) in terms of their ability to characterize present and predict future distributions of our 8 focal species. We hypothesized which model class would best fit each focal species' distribution based on dispersal and habitat preference traits. A detailed description of our study area, data, methods and hypotheses is provided in Supplementary material Appendix 1.

We used stepAIC to determine the most parsimonious set of covariates for each model class per set of pseudo-absences ( $n = 100$ ) per species ( $n = 8$ ). The three model classes included different covariates. As outlined in the statistical framework, the model with patch area is our representation of a SAR model ( $A_x$  in Eq. 6), the model with species-specific patch connectivity ( $k_{ix}$  in Eq. 7) is our representation of a MEP model, and the model with environmental conditions of soil drainage, soil deposit, % deciduous forest, % coniferous forest and maximum forest stand age is our representation of a SDM for all species. The presence of wetlands was included as an

additional covariate in the environmental conditions model for anuran species ( $v_{mx}$  in Eq. 5 Supplementary material Appendix 1).

Species-specific patch connectivity was defined as the sum of the dispersal probabilities between neighboring patches (see Eq. 7). We modeled dispersal probabilities as a negative exponential dispersal kernel with  $k_i(x, y) = 0.5 \times e^{-\text{dist}(x,y)/D50_i}$  where  $\text{dist}(x,y)$  is the Euclidean distance between the edges of patches  $x$  and  $y$  and  $D50_i$  is the median dispersal distance for species  $i$  (see Table 2). The set of neighboring patches in Eq. 7 (N) was defined as patches with adjacent polygons in a Voronoi tessellation of the landscape. Habitat patches neighboring the more continuous mountain areas were kept as sources, i.e. they were included in the habitat networks and the calculation of species-specific patch connectivity, but we did not predict the probability of occurrence of each focal species in these patches. To differentiate between the patch area and isolation effects processes, we only included isolation effects in our MEP model and patch area in our SAR model. For all three model classes, we tested linear and quadratic terms for each covariate but did not include covariate interaction terms.

### **Spatial predictions of species occurrences under current and future land-use scenarios**

We randomly selected a single pseudo-absence iteration per species (results were similar when randomly taking another pseudo-absence iteration) to predict the probability of occurrence of the focal species across our full landscape based on patch area, patch isolation, and environmental condition models. Following our statistical framework (Eq. 4), we summed the species-specific probability of occurrences for each model to create a predictive “species richness” map and multiplied Eq. 4 by 1000 (Distler et al. 2015). We only included species for which the best model reached both a TSS > 0.3 and AUC > 0.7 for each model class (Thuiller et al. 2014). We calculated the spatial coherence of the patch area, patch isolation and environmental condition model predictions on species richness at time  $t$  with Spearman correlations ( $\rho$ ) for the

entire set of patches. In addition, we calculated Jaccard coefficients of the model predictions on species richness at time  $t$  for the top 10% richest forest patches by area.

We used the same set of patch area, patch isolation and environmental condition models parameterized under current conditions to predict the probability of occurrence of our focal species into the future (see Supplementary material Appendix 1 for description of future landscape scenario) and used the same metrics as above to compare future “species richness” maps.

All analyses were conducted with the statistical software R 3.0.2 (R Core Team 2015) and the species distribution modeling library biomod2 (Thuiller et al. 2009).

## RESULTS

### Patch area, patch isolation and environmental condition model selection

The patch area model was the most frequent top ranking model for four species (*P. cinereus* (97%), *R. sylvatica* (66%), *S. aurocapilla* (78%), and *S. varia* (81%)), the patch isolation model was the most frequent top ranking model for two species (*D. pileatus* (71%) and *S. canadensis* (81%)) and the environmental conditions model was the most frequent top ranking model for two species (*B. americanus* (52%) and *S. minor* (55%)). Five of the eight species had high model selection uncertainty with at least two model classes ranking as top model in  $\geq 20\%$  of pseudo-absence iterations (Table 3). The top models for all species had a mean predictive accuracy that was good to very good (TSS  $> 0.3$  and ROC  $> 0.7$ ; Thuiller et al. 2014), except for the top models for *B. americanus* and *D. pileatus* which had a mean predictive accuracy that was poor (TSS  $> 0.2$  and ROC  $> 0.6$ ) (Supplementary material Appendix 2). Given the poor fit of

models for *B. americanus* and *D. pileatus*, we did not include these two species in our calculations of predicted species richness (i.e. Eq. 4). The top-ranked model did not consistently match our hypotheses based on species traits (Table 2, Appendix 1).

### **Comparisons of spatial predictions for stacked species occurrences in current landscape**

All predictions of species richness for the current landscape are positively correlated and show high spatial overlap, but there is nonetheless substantial variability in patch-specific predictions among models (Fig. 3). The patch area and environmental condition model predictions were most similar based on Spearman's  $\rho$ , particularly for small patches ( $\rho = 0.6$ , Fig 3) whereas the patch isolation and environmental condition model predictions were least similar based on Spearman's  $\rho$  ( $\rho = 0.31$ , Fig 3). The spatial overlap in the 10% most species rich patches by area based on the Jaccard coefficient show slightly different patterns (Fig. 3). When considering only these richest patches, patch area and patch isolation models showed the highest spatial overlap (Jaccard coefficient = 0.82) and the patch area and environmental condition models showed the lowest spatial overlap (Jaccard coefficient = 0.59).

### **Comparisons of spatial predictions for stacked species occurrences under future land-use scenario**

Our 2050 landscape was composed of less habitat area (-10% of total area), more forest patches (9,585 patches representing a 30 % increase), but patches were smaller on average (median=18,900 m<sup>2</sup> representing a 55 % decrease; Fig. 2). Similar to the current landscape, all model predictions for the future landscape were positively correlated (Fig. 3). The patch area and environmental condition model predictions

Accepted Article  
were most similar whereas the patch isolation and environmental condition model predictions were least similar based on Spearman's  $\rho$  ( $\rho = 0.46$  and  $0.29$ , respectively, Fig 3). When considering only the 10% richest patches by area, patch area and patch isolation models showed the highest spatial overlap (Jaccard coefficient =  $0.82$ ) and both the patch area and environmental condition models and the patch isolation and environmental condition models had lower spatial overlap (both with Jaccard coefficient =  $0.62$ ).

There was considerable spatial variability of change in species richness within model class predictions (Supplementary material Appendix 3, Fig. A2). While patch area models, because of their definition, always predicted a reduction in species richness due to habitat loss, patch isolation and environmental condition models gave more contrasted results (Fig. 4). On the one hand, some patches had an apparent increase in species richness according to the patch isolation model, which is an artifact of an increase in the number of close neighboring patches due to fragmentation following habitat loss. On the other hand, the repositioning of patches following habitat loss lead to changes in average patch condition and mostly reduced species richness predictions according to the environmental condition model (Fig. 4). The greatest change in species richness following habitat loss was predicted by the patch isolation and patch area models (Fig. 4, Supplementary material Appendix 3, Fig. A2).

#### **PERSPECTIVES ON THE USE AND INTEGRATION OF SIMPLE BIODIVERSITY MODELS IN CONSERVATION**

Mathematical models have played a fundamental role in the development of both ecological theory (Caswell 1988, Codling and Dumbrell 2012) and a toolbox to predict the impacts of human activities on biodiversity



(e.g. Pereira et al. 2010). While there is an increasing number of studies incorporating environmental conditions and landscape predictors in the same model (e.g. Calabrese et al. 2014, Betts et al. 2014, McCune 2016), we are aware of only a few studies that compare the predictions of different theoretical models for a given landscape (see Moilanen and Hanski 1998, Keith et al. 2008, Dullinger et al. 2011, Fordham et al. 2013, Brudvig et al. This issue). Here, we apply a common and flexible statistical framework to compare the predictions of simple biodiversity models and apply it to investigate present and future biodiversity in a real, fragmented landscape. Our framework enabled us to uncover complementarities among model predictions and also limitations of incorporating a single mechanism for predicting the distribution of species richness; patch area (SAR), patch isolation (MEP) and environmental conditions (SDM).

It is common practice in ecology and conservation biology to compare the performance of different formulations *within* a class of biodiversity models. For example, DeCamargo and Currie (2014) compare the predictions and fit of three types of SAR models for bird communities in Ontario, Canada; the classic power function SAR model, a polynomial SAR model and the endemic-area model. Likewise, Meynard and Quinn (2007) compare the performance of four different SDMs on virtual species data; generalized additive models, logistic regression, classification trees, and the genetic algorithm for rule-set production (GARP). While comparing predictions within a model class is valuable, it is also limited because a particular model class assumes specific drivers of biodiversity distribution and inference from within model class comparisons are constrained to the adopted model structure. For example, SAR models assume that species richness is related to area and therefore residual variation that may be explained by other mechanisms (e.g. patch isolation, environmental conditions) will not be uncovered when comparing predictions of different SAR formulations.

Comparisons among model classes are rare because such comparisons require a flexible framework that can be applied to diverse models. We provide one solution to fill this gap with a generalized linear model that can be parameterized with presence/pseudo-absence data, which may allow ecologists and conservation biologists to evaluate predictions from different biodiversity models in order to improve predictions of the effects of global changes on biodiversity. This framework is equivalent to developing an SDM with patch, landscape and environmental covariates. However, the above framework allows ecologists to build on previous studies that integrate multiple classes of covariates into SDMs in a few ways. First, the approach we present here is based on theory and is not purely phenomenological. SAR, MEP and SDM have different

Accepted Article

histories and different user groups and the framework we outline allows us to relate each theory to a specific statistical model. The roadmap we provide is an explicit attempt to show how these different theories can be integrated into a common and simple statistical framework. As such, it allows us to test and compare hypotheses and reveal structural uncertainty in model selection (i.e. has the modeler selected the right model/process?). Second, our approach allows comparison of these three biodiversity models at species (Eq. 5) and community levels (Eq. 4). This is critical for disentangling how different species respond to different drivers and for comparing predictions from different models. Finally, the long-term trajectories of biodiversity in fragmented landscapes may differ (e.g. Collins et al. This issue, Ewers et al. This issue) and a process-based approach at species and community levels may yield more robust predictions for future scenarios and long-term land-use planning (Leroux et al. 2013, Thuiller et al. 2013).

The application of this framework to the case study revealed that patch area is an important driver of the distribution of many of the 8 focal vertebrate species, but that connectivity and environmental conditions were also important for at least four focal species (Table 3). Clearly, adoption of a single model class for all species would have overlooked key correlates of the distribution of some of the focal species. Haddad et al (This issue) also find evidence that patch area, isolation and configuration are key determinants of species richness in two experimental systems. Taken together, these results provide evidence in contrast to the recently proposed Habitat Amount Hypothesis (Fahrig 2013). Determining which model was best for each species a priori, based on traits alone, was not reliable (only matched the top ranked model in Table 3 for 50% of the species, Appendix 1) which underscores the need to consider multiple models. However, given the small number of focal species we consider, the overall low quality of these data (very incomplete set of presence/absence data with, on average 3% of the patches as occupied), the poor sampling of smaller patches, and the fact that the data come from various sampling protocols, we prefer to consider this as an illustration of model comparison and limit our biological interpretation of our results.

Habitat loss and fragmentation alter the distribution of patch area, connectedness and environmental conditions (Fahrig 2003). Consequently, a model that focuses on a single of these components of land-use change may not capture its full effects on biodiversity (Brudvig et al. This issue, Haddad et al. This issue). Disentangling the relative impacts of patch area and isolation on biodiversity has been a driver of fragmentation research for decades (Haddad et al. 2015). By comparing predictions from multiple models that

Accepted Article

incorporate the effects of different fragmentation processes we can better understand the relative and combined importance of these processes on species distribution and richness under land-use change. In the case study, patch area models, by definition, always predicted a patch-level decrease in species richness, whereas patch isolation and environmental condition models predicted either an increase or a decrease in patch-level species richness following habitat loss and fragmentation (Fig. 4). The predicted magnitude of change in any single patch was greatest for patch isolation and patch area models (Fig. 4). Over all patches, patch area and environmental condition models predicted an 11% decline and the patch isolation model predicted a 7% decline in species richness following habitat loss and fragmentation.

Application of these simple models to predict change in biodiversity is critical for conservation policy and practice (Resasco et al. This issue). For example, the Millennium Ecosystem Assessment (2005) applied SARs to predict species extinction rates following projected habitat loss. The choice of this model class for such predictions, however, presupposes key drivers of species extinction rates under global change. The framework we outline allows us to better capture how structural uncertainty arises from different underlying processes, which may help to explain why some model predictions are not consistent with empirical observations.

Some have argued for the need to integrate biodiversity models for better predictions of the effects of global change on biodiversity (Moilanen and Hanski 1998, Keith et al. 2008, Dullinger et al. 2011, Fordham et al. 2013, Brudvig et al. This issue). For example, Fordham et al. (2013) provide an overview of the key insights gained by integrating metapopulation dynamics into SDMs, Koh & Ghazoul (2010) incorporate aspects of patch quality in their matrix-calibrated SAR model and Talluto et al. (2016) develop a methodological approach to integrate models across spatial scales in order to reduce error propagation arising from mechanistic SDMs. Based on the statistical framework we adopt, an integrated model may take the form of Eq. 5 where the species occurrence probability is based on covariates describing local site characteristics including patch area, patch isolation and environmental conditions. Such an integration must balance the cost of added model complexity with the benefit of improved model fit and predictions (Merow et al. 2014).

## CONCLUSION

We provide a roadmap based on ecological theory for comparing and potentially integrating different classes of biodiversity models. The common and simple statistical framework we apply allows researchers and conservation practitioners to understand trade-offs between model complexity and model predictive accuracy and therefore make better informed modelling decisions with repercussions on predictions of biodiversity responses to global changes. Future development of this framework could involve applications to other biodiversity models (e.g. neutral model) and evaluation of integrated models. We hope that this approach encourages empiricists to collect and share biodiversity data that would allow more cross-model class comparison and integration. These cross-model class comparisons will facilitate feedbacks between theory and data which is essential for improving predictions for biodiversity change under global change.

#### **ACKNOWLEDGEMENTS**

SJL and DG were supported by Natural Science and Engineering Research Council of Canada Discovery Grants, CHA was supported by a Marie Curie International Outgoing Fellowship under the 7th European Community Framework Program (DYVERSE project, no. 272284), BR was funded by the OURANOS ÉCOBIOCC project, SW gratefully acknowledges the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118). ASL was supported by the TULIP Laboratory of Excellence (ANR-10-LABX-41) and the Midi-Pyrenees Region. This work contributes to the Labex OT-Med (no. ANR-11-LABX-0061) funded by the French government through the A\*MIDEX project (no. ANR-11-IDEX-0001-02). We thank A. Gonzalez, M. Dumitru, and B. Holt for their thoughtful help on earlier stages of this work.

## REFERENCES

- Akcakaya, H. R. et al. 2006. The role of metapopulations in conservation. - In: *Key Topics in Conservation Biology*. MacDonald, D., Service, K. (eds), Blackwell publishing, pp. 64–84.
- Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS): Assessing the accuracy of distribution models. – *J Appl Ecol* 43: 1223–1232.
- Araújo, M. and New, M. 2007. Ensemble forecasting of species distributions. - *Trends Ecol Evol* 22: 42–47.
- Araújo, M. B. and Peterson, A. T. 2012. Uses and misuses of bioclimatic envelope modeling. - *Ecology* 93: 1527–1539.
- Arrhenius, O. 1921. Species and Area. - *J Ecol* 9: 95.
- Barraquand, F. et al. 2014. Lack of quantitative training among early-career ecologists: a survey of the problem and potential solutions. - *PeerJ* 2: e285.
- Betts, M. G. et al. 2014. A species-centered approach for uncovering generalities in organism responses to habitat loss and fragmentation.– *Ecography* 37: 517-527.
- Boulangéat, I. et al. 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. – *Ecol Lett* 15: 584-593.
- Brooks, T. M. et al. 1999. Threat from deforestation to montane and lowland birds and mammals in insular South-east Asia. - *J Anim Ecol* 68: 1061–1078.
- Brudvig, L. et al. This issue. Evaluating conceptual models of landscape change. - *Ecography*
- Calabrese, J. M. et al. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecol Biogeogr* 23: 99-112.
- Cardinale, B. J. et al. 2012. Biodiversity loss and its impact on humanity. - *Nature* 486: 59–67.
- Caswell, H. 1988. Theory and models in ecology: a different perspective. – *Ecol Model* 43: 33–44.
- Chetkiewicz, C.-L. B. and Boyce, M. S. 2009. Use of resource selection functions to identify conservation corridors. – *J Appl Ecol* 46: 1036–1047.
- Codling, E. A. and Dumbrell, A. J. 2012. Mathematical and theoretical ecology: linking models with ecological processes. - *Interf Focus* 2: 144–149.
- Collins, C. et al. This issue. Fragmentation affects plant community composition over time. -*Ecography*.

- Connor, E. F. and McCoy, E. D. The statistics and biology of the species-area relationship. - *Am Nat* 113: 791–833.
- Cord, A. and Rödder, D. 2011. Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? - *Ecol Appl* 21: 3285–3298.
- Crooks and Sanjayan. 2006. *Connectivity Conservation*. - Cambridge University Press.
- Coulon, A. et al. 2015. A stochastic movement simulator improves estimates of landscape connectivity. - *Ecology* 96: 2203–2213.
- De Camargo, R. X. and Currie, D. J. 2015. An empirical investigation of why species–area relationships overestimate species losses. - *Ecology* 96: 1253–1263.
- Dengler, J. 2009. Which function describes the species-area relationship best? A review and empirical evaluation. - *J Biogeogr* 36: 728–744.
- Distler, T. et al. 2015. Stacked species distribution models and macroecological models provide congruent projections of avian species richness under climate change. - *J Biogeogr* 42: 976–988.
- Dornelas, M. et al. 2014. Assemblage time series reveal biodiversity change but not systematic loss. - *Science* 344: 296–299.
- Dullinger, S. et al. 2011. Patch configuration affects alpine plant distribution. – *Ecography* 34: 576–587.
- Elith, J. and Leathwick, J. R. 2009. Species distribution models: Ecological explanation and prediction across space and time. - *Annu Rev Ecol Evol Syst* 40: 677–697.
- Ewers, R. M. et al. This issue. Predicted trajectories of tree community change in Amazonian rainforest fragments. – *Ecography*.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. - *Annu Rev Ecol Evol Syst* 34: 487–515.
- Fahrig, L. 2013. Rethinking patch size and isolation effects: the habitat amount hypothesis. – *J Biogeogr* 40: 1649–1663.
- Fordham, D. A. et al. 2013. Tools for integrating range change, extinction risk and climate change information into conservation management. - *Ecography* 36: 956–964.
- Gallien, L. et al. 2012. Invasive species distribution models - how violating the equilibrium assumption can create new insights: Beyond the equilibrium assumption of SDMs. - *Global Ecol Biogeogr* 21: 1126–1136.

- Gilarranz, L. and Bascompte, J. 2012. Spatial network structure and metapopulation persistence. – *J Theor Biol* 297: 11-16.
- Gravel, D. et al. 2011. Trophic theory of island biogeography. - *Ecol Lett* 14: 1010–1016.
- Guisan, A. et al. 2006. Using niche-based models to improve the sampling of rare species. - *Conserv Biol* 20: 501–511.
- Haby, N. A. et al. 2013. Scale dependency of metapopulation models used to predict climate change impacts on small mammals. - *Ecography* 36: 832–841.
- Haddad, N. M. et al. 2015. Habitat fragmentation and its lasting impact on Earth’s ecosystems. - *Science Advances* 1: e1500052–e1500052.
- Haddad, N. M. et al. This issue. Experimental evidence does not support the Habitat Amount Hypothesis. – *Ecography*.
- Hanski, I. 1999. *Metapopulation ecology*. - Oxford University Press.
- Hanski, I. et al. 2013. Species-fragmented area relationship. – *P Natl Acad Sci USA* 110: 12715–12720.
- Holland, M. D. and Hastings, A. 2008. Strong effect of dispersal network structure on ecological dynamics. - *Nature* 456: 792–794.
- Kaufman, S. D. et al. 2009. Impacts of road access on lake trout (*Salvelinus namaycush*) populations: regional scale effects of overexploitation and the introduction of smallmouth bass (*Micropterus dolomieu*). - *Can J Fish Aquat Sci* 66: 212–223.
- Kearney, M. and Porter, W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species’ ranges. - *Ecol Lett* 12: 334–350.
- Keith, D. A. et al. 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. - *Biol Lett* 4: 560–563.
- Kerr, J. T. et al. 2015. Climate change impacts on bumblebees converge across continents. - *Science* 349: 177–180.
- Koh, L. P. and Ghazoul, J. 2010. A matrix-calibrated species-area model for predicting biodiversity losses due to land-use change: Matrix-calibrated species-area model. - *Conserv Biol* 24: 994–1001.
- Krosby, M. et al. 2015. Focal species and landscape “naturalness” corridor models offer complementary approaches for connectivity conservation planning. - *Landscape Ecol* 30: 2121–2132.

- Kukkala, A. S. and Moilanen, A. 2013. Core concepts of spatial prioritisation in systematic conservation planning: Concepts of systematic conservation planning. - *Biol Rev* 88: 443–464.
- Kulkarni, M. A. et al. 2010. High resolution niche models of malaria vectors in northern Tanzania: A new capacity to predict malaria risk? - *PLoS ONE* 5: e9396.
- Lande, R. et al. 2003. Stochastic population dynamics in ecology and conservation. - Oxford University Press.
- Leroux, S. J. et al. 2007. Accounting for system dynamics in reserve design. - *Ecol Appl* 17: 1954–1966.
- Leroux, S. J. et al. 2013. Mechanistic models for the spatial spread of species under climate change. - *Ecol Appl* 23: 815–828.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. - *B Entomol Soc Am* 15: 237-240.
- Liu, C. et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. - *Ecography* 28: 385–393.
- MacArthur, R. H. and Wilson, E. O. 1967. *The Theory of Island Biogeography*. - Princeton University Press.
- McCune, J. L. 2016. Species distribution models predict rare species occurrences despite significant effects of landscape context. *J Appl Ecol*. Doi: 10.1111/1365-2664.12702.
- Meller, L. et al. 2014. Ensemble distribution models in conservation prioritization: from consensus predictions to consensus reserve networks. - *Divers Distrib* 20: 309–321.
- Mendenhall, C.D. et al. 2014. Predicting biodiversity change and averting collapse in agricultural landscapes. - *Nature* 509:213-217.
- Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? - *Ecography* 37: 1267–1281.
- Meynard, C. N. and Quinn, J. F. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species: Comparison of species-distribution models. - *J Biogeogr* 34: 1455–1469.
- Millennium Ecosystem Assessment 2005. *Ecosystems and Human Well-being: Biodiversity Synthesis*. World Resources Institute, Washington, DC.
- Moilanen, A. and Hanski, I. 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. - *Ecology* 79: 2503–2515.



- Newbold, T. et al. 2015. Global effects of land use on local terrestrial biodiversity. - *Nature* 520: 45–50.
- Pereira, H. M. et al. 2010. Scenarios for global biodiversity in the 21st century. - *Science* 330: 1496–1501.
- Pollock, L. J. et al. 2014. Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). - *Methods Ecol Evol* 5: 397–406.
- R Development Core Team. 2015. R: a language and environment for statistical computing, v.3.0.2. R Foundation for Statistical Computing, Vienna, Austria.
- Resasco, J. et al. This issue. Theory, experiments, and the conservation of fragmented landscapes. – *Ecography*.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. - Cambridge University Press.
- Talluto, M. et al. 2016. Cross-scale integration of knowledge for predicting species ranges : a metamodeling framework. *Global Ecol Biogeogr* 25 : 238-249.
- Tanentzap, A. J. et al. 2012. A framework for predicting species extinction by linking population dynamics with habitat loss. - *Conserv Lett*: 149-156.
- Thompson, P. et al. 2014. Robustness of the spatial insurance effects of biodiversity to habitat loss. - *Evol Ecol Res* 16: 445–460.
- Thuiller, W. et al. 2009. BIOMOD - a platform for ensemble forecasting of species distributions. - *Ecography* 32: 369–373.
- Thuiller, W. et al. 2013. A road map for integrating eco-evolutionary processes into biodiversity models. - *Ecol Lett* 16: 94–105.
- Thuiller, W. et al. 2014. Does probability of occurrence relate to population dynamics? - *Ecography* 37: 1155–1166.
- Tjørve, E. and Turner, W. R. 2009. The importance of samples and isolates for species-area relationships. - *Ecography* 32: 391–400.
- Trainor, A. M. and Schmitz, O. J. 2014. Infusing considerations of trophic dependencies into species distribution modelling. - *Ecol Lett* 17: 1507–1517.
- Vasudev, D. et al. 2015. From dispersal constraints to landscape connectivity: lessons from species distribution modeling. - *Ecography* 38: 967–978.

Vellend, M. et al. 2013. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. - *P Natl Acad Sci USA* 110: 19456–19459.

Whittaker, R. J. et al. 2005. Conservation biogeography: assessment and prospect. - *Divers Distrib* 11: 3–23.

Supplementary material (Appendix EXXXXX at <[www.oikosoffice.lu.se/appendix](http://www.oikosoffice.lu.se/appendix)>). Appendix 1–3.

## TABLE LEGENDS

Table 1. Summary of the three biodiversity models we considered in this study; SAR, MEP and SDM. The summary includes information on the goals, history and recent developments, key assumptions and mechanisms, predictions and software for each model class.

	<b>Species-area relationships (SARs)</b>	<b>Metapopulation models (MEPs)</b>	<b>Species distribution models (SDMs)</b>
<b>Goal</b>	SARs describe the dependence of species richness on area at various spatial scales	MEPs model the presence and absence of a species across a network of discrete and dissimilar habitat patches	SDMs describe the relationship between species and its environment and predict species distribution within or outside the sampled area
<b>History and recent development</b>	A power function $S=cA^z$ was first proposed to model the change of species richness (S) with area (A) across isolated ecosystems (Arrhenius 1921). Although other forms of SARs have been proposed, the power model is the most commonly used form. Recent developments include the species-fragmented area relationship, matrix-calibrated SAR and the countryside SAR.	Spatially implicit MEPs were first proposed in 1960s. They were then developed to incorporate landscape configuration and heterogeneity in spatially explicit MEPs (by Hanski and colleagues). Recent efforts pair MEPs with network models of habitat connectivity, to better account for indirect dispersal pathways along stepping stone patches and least-cost links.	The history of SDMs is firmly rooted in ecological studies of biodiversity along environmental gradients, the concept of ecological niche and in classic studies of species' responses to their environment. Recent efforts aim towards mechanistic SDMs, which link SDMs with process-based approaches, e.g. population viability model, physiology, species co-occurrences.

Key assumptions or mechanisms	The mechanisms of SARs can be scale dependent. At the small extents, they depend on species abundance and spatial distribution; at intermediate extents, habitat heterogeneity, and self-similar species distributions are dominant factors; at very large extents, evolutionary history is the main driver.	In a patchy landscape, population persistence is governed by the processes of local extinctions and re-colonization. Local extinction rate is a function of patch area and colonization rate is a function of the connectivity of a patch, which is related to its area and distance to other occupied patches.	i) The current distribution of a species depicts its suitable habitat, ii) species distribution are determined by the explanatory variables included in the models iii) species are at pseudo-equilibrium with environmental variables, and iv) the relationship between species distribution and explanatory variables does not change in time.
Prediction	Species richness increases with area, which is usually described by: $S=cA^z$ . The exponent $z$ can be scale dependent; for instance, SARs can exhibit tri-stage patterns from local to continental or global scales	Metpopulation persistence decreases with local extinction rates, and increases with colonization rates and the metapopulation capacity, which is determined by landscape structure.	SDMs predict a continuous probability of species occurrence. Several methods can be used to convert such continuous model predictions to binary range.
Software	Any software with a statistical toolbox (e.g. R, Matlab, etc.)	R packages: <i>grainscape</i> , <i>MetaLandSim</i> , <i>igraph</i> and stand alone software <i>Graphab</i>	R packages: <i>dismo</i> and <i>biomod2</i> .

Table 2. Traits of focal species used in the case study. Traits characterize dispersal ability and habitat preference: minimum habitat patch area (MA); median dispersal distance based on natal dispersal ( $D_{50}$ ); preference for deciduous (D), mixed (M) or coniferous (C) forest; preferred forest type (Type); sensitivity to soil drainage (SD); and sensitivity to distance to wetlands (DW). Grey shading indicates a preference or sensitivity to feature.









Species	MA (ha)	$D_{50}$ (m)	Preferred habitat					
			Forest				SD	DW
			D	M	C	Type		
 <i>Bufo americanus</i>	0.5	2,795				any		
 <i>Dryocopus pileatus</i>	1	8,187				dense & old		
 <i>Plethodon cinereus</i>	0.27	16				dense		
 <i>Rana sylvatica</i>	0.5	564				dense		
 <i>Scolopax minor</i>	2	34,317				low height		
 <i>Seiurus aurocapilla</i>	5	1,286				dense & old		
 <i>Sitta canadensis</i>	3	1,827				dense & old		
 <i>Strix varia</i>	1	40,889				old		

Table 3. Summary of patch area, patch isolation and environment condition model class comparison for our eight focal species. We report the percentage of iterations for each model class (n = 100 sets of pseudo-absence) where it was ranked as the top model based on  $\Delta AIC$ . The best supported model for each species is shaded in grey. We report the fit statistics (TSS, AUC) for each model in Supplementary material Appendix 2, Table A1).

Species	Model classes		
	Patch area	Patch isolation	Environmental conditions
<i>Bufo americanus</i>	47	1	52
<i>Dryocopus pileatus</i>	6	71	23
<i>Plethodon cinereus</i>	97	0	3
<i>Rana sylvatica</i>	66	0	34
<i>Scolopax minor</i>	45	0	55
<i>Seiurus aurocapilla</i>	78	0	22
<i>Sitta canadensis</i>	18	81	1
<i>Strix varia</i>	81	11	8

## FIGURE LEGENDS

Figure 1. Location of the study area in and around the St. Lawrence River near Montreal, Quebec, Canada.

Red lines outline the boundary of the St. Lawrence lowlands ecoregion. Map is presented with North America Datum 1983 and Transverse Mercator projection.

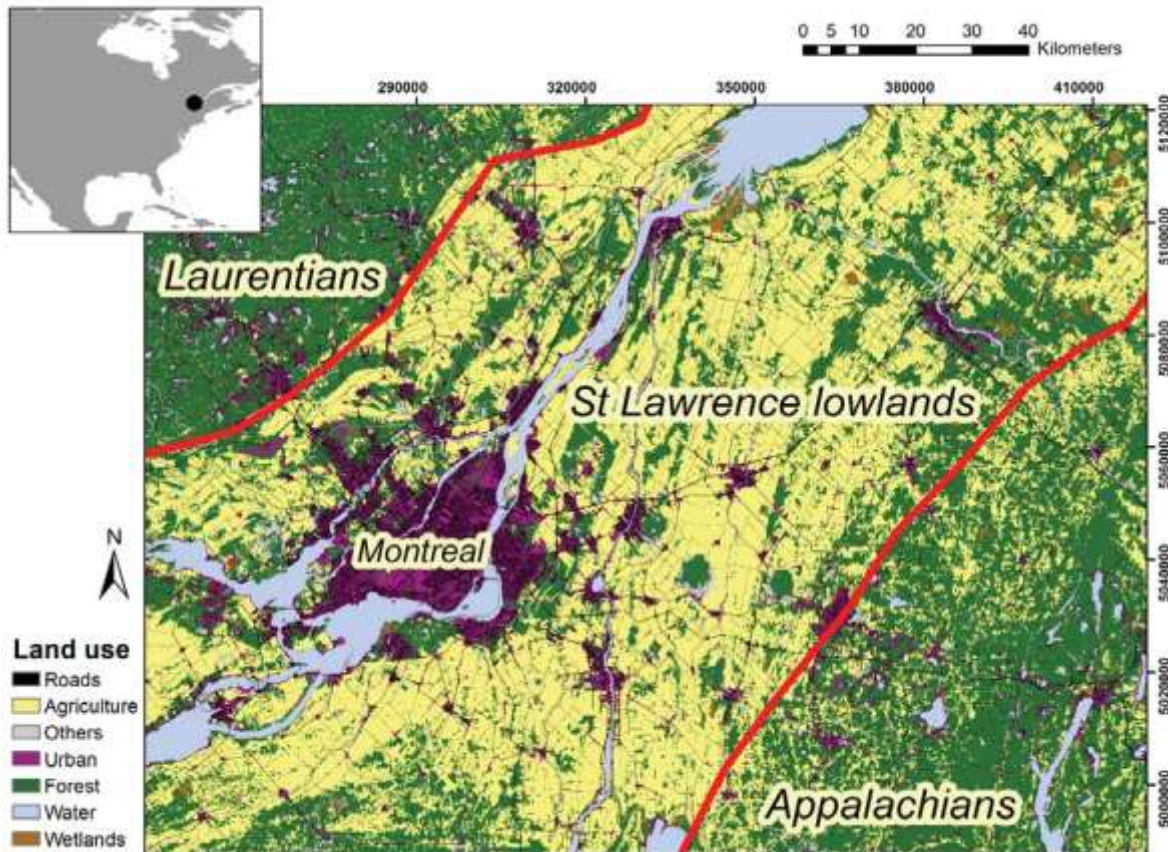




Figure 2. Forest patch network within the St Lawrence lowlands used in the analysis. Black patches were part of the 10% of the landscape that was removed with our “Business As Usual” landscape change scenario for 2050 (see supplementary material Appendix 1). The landscape change scenario for 2050 created a more fragmented landscape with a 30% increase in the number patches, a 55% decrease in the median size of patches and patches with more irregular edges (mean perimeter:area ratio: +29%).

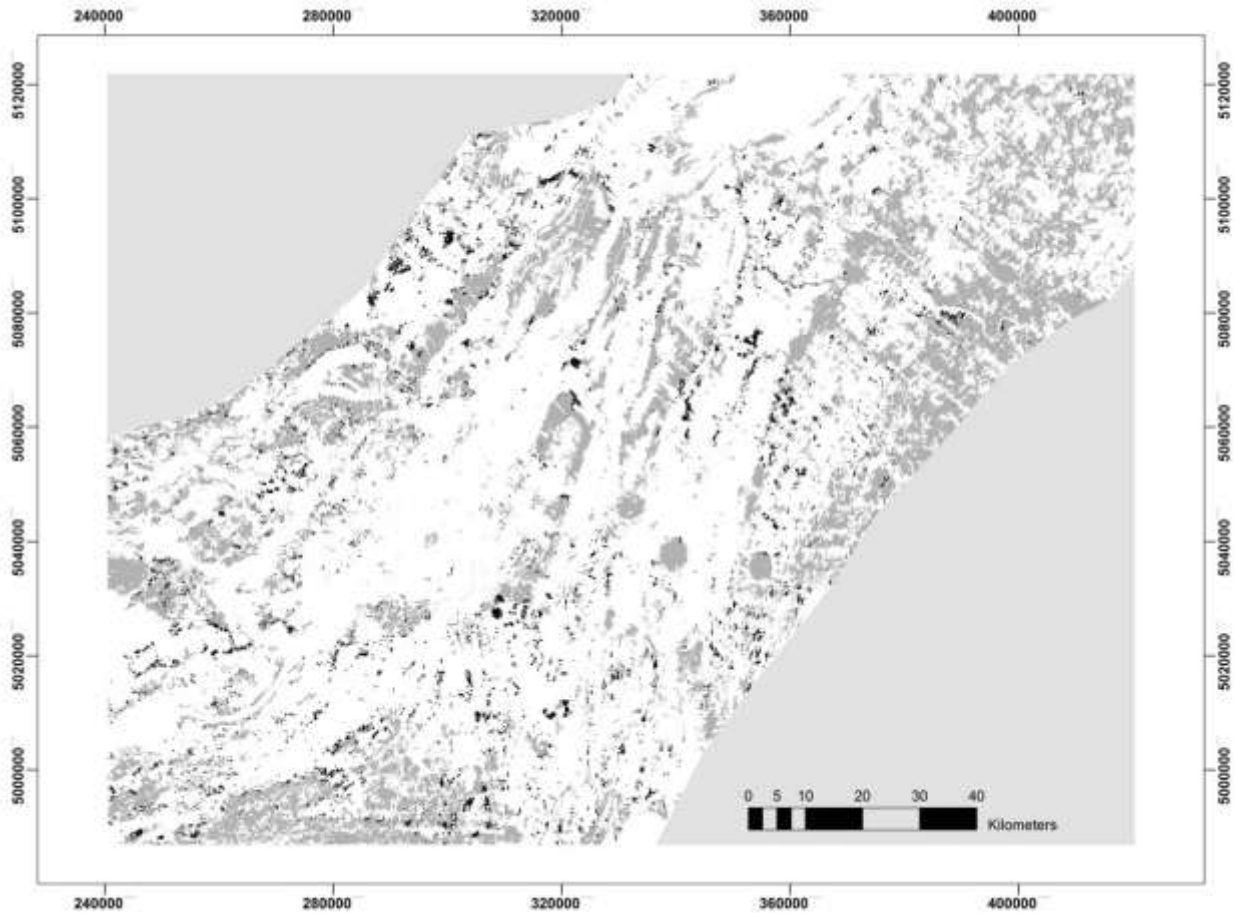




Figure 3. Pairwise correlations between model predictions for the current (top row) and future (bottom row) landscapes. Spearman  $\rho$  is reported for all patches and the Jaccard Coefficient for the 10% most species rich patches by area as predicted by each model. Black dots are the patches that make up the top 10% area of richest patches (along the x-axis). The axes are the predicted species richness (i.e. Eq. 4) multiplied by 1000 for each model. The two species (*B. americanus*, *D. pileatus*) with poor model fit were excluded from the species richness calculation.

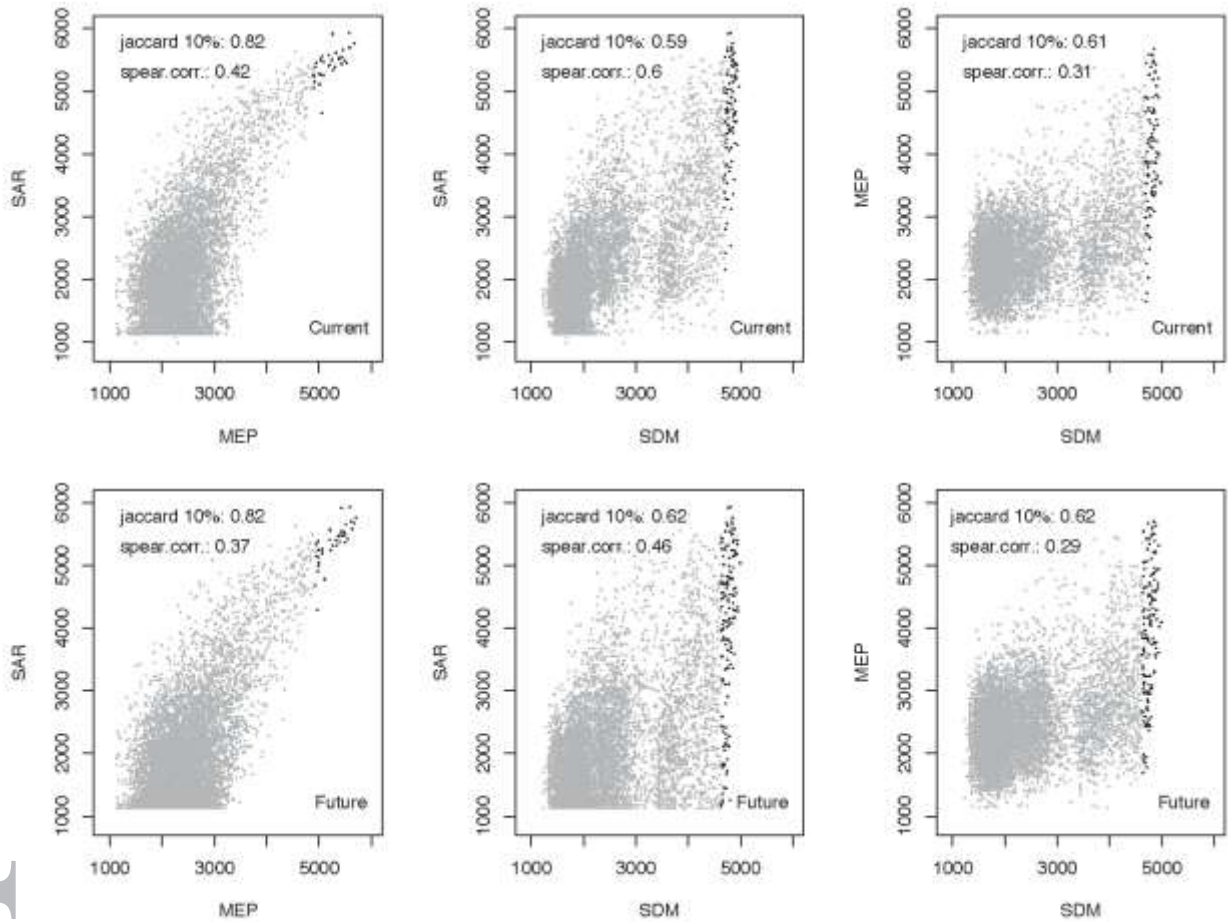


Figure 4. Predictions of current and future species richness (Eq. 4) multiplied by 1000 for the patch area, patch isolation, and environmental condition models. The two species (*B. americanus*, *D. pileatus*) with poor model fit were excluded from the species richness calculation. Each point represents the predictions (current and future) for the centroid location of a current patch. Future predictions of species richness = 0 are for patches where the centroid location was lost with habitat loss.

