Title
Quantitative Ecology and the Conservation of Biodiversity: Species Richness, Abundance, and Extinction in Human-Altered Landscapes

Author
Kitzes, Justin

Publication Date
2012

Peer reviewed|Thesis/dissertation

Quantitative Ecology and the Conservation of Biodiversity: Species Richness, Abundance, and Extinction in Human-Altered Landscapes

by

Justin Adam Kitzes

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy, and Management

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Adina M. Merenlender, Co-chair
Professor John Harte, Co-chair
Professor William Z. Lidicker, Jr.

Fall 2012

Quantitative Ecology and the Conservation of Biodiversity:
Species Richness, Abundance, and Extinction in Human-Altered Landscapes

by

Justin Adam Kitzes

Abstract

Quantitative Ecology and the Conservation of Biodiversity:
Species Richness, Abundance, and Extinction in Human-Altered Landscapes

by

Justin Adam Kitzes

Doctor of Philosophy in Environmental Science, Policy, and Management

University of California, Berkeley

Professor Adina M. Merenlender, Co-chair

Professor John Harte, Co-chair

The goal of conservation biology is to understand and prevent the loss of biological diversity. Modern conservation science relies heavily on four major quantitative methods: reserve site selection algorithms, species distribution models, population viability analyses, and species-area relationships. These methods, however, have several longstanding and unresolved shortcomings, including extensive data requirements, long computation times, and important simplifying assumptions, that limit their ability to inform conservation decisions in many real landscapes. This dissertation develops new approaches in quantitative ecology that address these shortcomings through the use of simulation modeling, probability theory, machine learning, modern statistics, and economic input-output analysis.

Chapter 2 examines the optimal design of reserve networks for preventing extinction of terrestrial mammal species, demonstrating that the match between a species' body size and the spatial scale of a landscape can be used to determine which species will benefit from a clustered reserve network design. Chapter 3 derives two new macroecological metrics, similar to the species-area relationship, that provide probabilistic estimates of single-species extinction risks and community-wide extinction rates in landscapes undergoing habitat loss. Chapter 4 uses data from acoustic surveys to examine the road ecology of California bats, finding that total bat activity, and the activity of four common species, is decreased in the vicinity of large highways. Chapter 5 presents a "wildlife footprint" analysis that combines global bird and mammal range maps, data on the human appropriation of net primary productivity, and economic input-output tables to link specific human consumption activities to decreases in wild bird and mammal populations.

Contents

Contents		i
Acknowledgements		ii
Chapter 1	Introduction	1
Chapter 2	Extinction risk and tradeoffs in reserve site selection for species of different body sizes	9
Chapter 3	Beyond the species-area relationship: improved estimates of extinction risks and rates following habitat loss and climate change	33
Chapter 4	California bats avoid roads	56
Chapter 5	The global wildlife footprint: linking bird and mammal population losses to economic consumption	71
Chapter 6	Conclusion	96

Acknowledgements

I begin by thanking my advisors, Adina Merenlender and John Harte, who are the types of mentors, colleagues, and friends that any young researcher would be privileged to work with. I thank the faculty at the University of California, Berkeley, who have shared countless insights on many topics in ecology and conservation biology, especially my committee members William Lidicker, Jr., Claire Kremen, and Wayne Getz, and Steve Beissinger, Perry de Valpine, and Matthew Potts.

I thank my collaborators who contributed to the research presented here, especially David Armitage, Eric Berlow, Erin Conlisk, Karl-Heinz Erb, Shane Feirer, Katsunori Iha, Neo Martinez, Erica Newman, Cristoph Plutzar, William Rainey, and Mark Wilber. I gratefully acknowledge financial support from the Department of Environmental Science, Policy, and Management and the National Science Foundation Graduate Research Fellowship Program (Grant No. DGE 0946797).

For their friendship and advice over the years, I give special thanks to my fellow students Ted Grantham, Mary Matella, Sarah Reed, and Adena Rissman of the Merenlender Lab, Manisha Anantharaman, Danielle Christianson, Stacy Jackson, Amber Kerr, Chloe Lewis, Andy Rominger, and Adam Smith of the Harte Lab, and Mez Baker, Tim Bean, Sibyl Diver, Melissa Eitzel, Jason Hwan, Matt Luskin, and Houston Wilson.

Last, but certainly not least, I thank all of my family and friends for their love and support, especially my parents Bill and Sandi Kitzes, my aunt Susan Liebson, and my uncle Jeff Kitzes, who have all taken a special interest in my graduate career. The most thanks of all are due to my wife and my best friend Audrey Kitzes, who has believed in me every step of the way.

Chapter 1

Introduction

At the start of the twenty-first century, almost no corner of the globe remains untouched by the influence of humans [1–3]. Humans currently occupy 38% of the Earth's terrestrial surface for agricultural production [4], appropriate 24% of global net primary productivity [5], use 30% of the terrestrial renewable fresh water supply [6], and emit sufficient greenhouse gases to change the workings of the global climate system [7]. The magnitude of these and other impacts has led atmospheric scientist and Nobel Laureate Paul Crutzen to propose that the planet has entered a new geological epoch, the Anthropocene, that is defined by the presence of human-induced global change [8, 9].

The scope and scale of human influence on the planet has led unavoidably to negative consequences for non-human species, with biologists suggesting that the planet may be in the midst of a sixth great mass extinction event [10]. Currently, 25% of mammals, 13% of birds, and 41% of amphibians are considered threatened, and most major indices of the health of global biodiversity are in decline [11, 12]. Both because non-human organisms, species, and communities arguably have intrinsic value in their own right and because the services provided by natural ecosystems form the foundation for human societies [3, 13], the potential for extensive losses of biodiversity represents a major concern for humanity.

Growing concerns about biodiversity loss led to the emergence in the mid-1980s of the interdisciplinary field of conservation biology, which has been described as a "mission-oriented discipline comprising both pure and applied science" that embodies the "overarching normative assumption. . . that biodiversity is good and ought to be preserved" [14–16]. Although the basic principles of conservation biology are not new [17], the field has been distinguished by its strong foundation in quantitative theory and modeling, recognition of the inherent value of biodiversity, and incorporation of tools and perspectives from outside of the biological sciences [16].

Much previous academic research in conservation biology has focused on the development of quantitative methods that assist practitioners in prioritizing management actions to reduce biodiversity loss. Four such methods that have become particularly integral to the practice of conservation biology are reserve site selection algorithms, which use

optimization and decision science techniques to design networks of reserves to protect biodiversity [18–20], species distribution models, which use niche theory and statistical tools to predict the occurrence of a species across an unsurveyed landscape [21–23], population viability analyses, which use stochastic modeling and other methods from population biology to evaluate extinction risk for individual species [24, 25], and species-area relationships, which use empirical data and ecological theory to predict the number of species that will be lost as habitat area contracts [26–30].

While these methods have been widely successful in supporting the practice of conservation, their real-world applications remain limited by several important shortcomings. Reserve site selection and species distribution models, for example, require spatially-explicit species occurrence data that may not be available for all species and landscapes [23, 31]. Site selection algorithms are also often criticized for ignoring the long-term persistence of species [32–34], and many niche-based investigations of species distributions do not incorporate features of the built environment, such as roads or buildings, that may have a significant influence on patterns of species occurrence in human-modified landscapes. Population viability analyses require detailed data on the life history traits of specific species in specific landscapes [25], and the long computation times associated with these models may prohibit the evaluation of many hundreds or thousands of species. While the species-area relationship requires little empirical data, it ignores critical realities of population and community dynamics, including the possibility of extinction debt or small population effects [35, 36], and predicts only expected extinction rates without providing a measure of uncertainty around these predictions. Finally, these methods are not often integrated with quantitative economic or social models, which would allow for better evaluation of the human causes and consequences of biodiversity loss. While much recent research has been directed towards addressing these shortcomings [e.g., 37–47], many of these issues remain largely unresolved to the detriment of conservation practice.

In this dissertation, I develop new approaches in quantitative ecology that address many of the important shortcomings described above through the use of emerging mathematical and computational techniques, including simulation modeling, probability theory, machine learning, modern statistics, and economic input-output analysis. These approaches collectively provide improved methods for measuring and predicting species richness, abundance, and extinction in human-altered landscapes. Throughout, I focus on the principles of generality and reproducibility, with the goal of providing methods that can be applied by conservation practitioners to a wide variety of landscapes and taxa.

Existing reserve site selection methods are often criticized for focusing on present-day species representation without concern for long-term species persistence. The stochastic population models that can be used to evaluate species persistence, however, require extensive data on species' life history traits that are often not available. Additionally, two general reserve design principles that are often followed in an attempt to increase species persistence probabilities, protecting large habitat patches and clustering habitat

patches, are often in conflict in real landscapes. Chapter 2 combines site selection methods, stochastic population viability models, and empirical allometric relationships to examine how extinction risk for terrestrial mammals of different body sizes is affected by the tradeoff between patch area and clustering in reserve networks. The results demonstrate that the spatial scale at which the clustering of habitat patches reduces extinction risk is different for species of different body sizes, and that there is thus no globally optimal level of patch clustering that will benefit all species. The optimal level of clustering for any specific species, however, can be predicted by matching the spatial scale of the reserve network to a species' dispersal distance, providing a rough rule of thumb that can be used to choose a level of patch clustering to benefit any one species.

When detailed data on a biological community are unavailable, the species-area relationship is often used to predict the number of species that will be driven extinct as habitat area contracts. The species-area relationship, however, predicts only the expected number of extinctions without any measure of uncertainty around this prediction, assumes implicitly that a species with one individual remaining after habitat loss is protected, and does not provide extinction risk predictions for individual species. Chapter 3 uses probability theory and modern statistics to derive two new macroecological metrics, the extinction-area relationship and the probabilistic species-area relationship, that address these three limitations. A general framework for deriving these new metrics from two widely studied macroecological distributions, the species abundance distribution and the species-level spatial abundance distribution, is presented, and explicit equations for both metrics based on a recently proposed maximum entropy theory of ecology are provided. Examination of the behavior of these two metrics finds that extinction predictions are strongly influenced by the choice of the minimum abundance required for a species to be considered protected, a parameter that cannot be adjusted in the classic species-area relationship. In light of this finding, ecologists and conservation biologists may wish to use these new metrics in place of the classic species-area relationship when making extinction predictions in poorly studied landscapes.

The niche-based study of species' distributions often focuses on climatic and other abiotic factors that determine patterns of species occurrence. In human-altered landscapes, however, the built environment may significantly influence the locations at which species are found. The factors that influence the spatial distribution of bat species are particularly poorly understood. Combining passive acoustic recording technology, machine learning techniques, and modern statistics, Chapter 4 presents the results of an empirical study of the effects of roads on bat activity. Recordings of bat calls were collected near three highways bordering San Francisco Bay, processed using a statistical classifier to identify each call to the species level, and evaluated with generalized linear mixed models to examine the extent to which bat activity is reduced in the vicinity of large roads. Total bat activity across all species is found to be reduced by approximately one-half at sampled points adjacent to roads as compared to control points 300 m from roadsides. The four most frequently recorded species also individually show decreased activity near roads. The

magnitude of this road effect, however, is found to vary with temperature, with differences in activity between roadside and control points decreasing as daily maximum temperature increases. These results have implications both for the conservation of bat species and for the provision of insect control services.

Despite many decades of research evaluating the direct, on-the-ground causes of biodiversity loss, quantitative studies of the links between biodiversity loss and the indirect, upstream economic activities that drive this loss remain uncommon. Such studies have the potential to help better integrate biodiversity considerations into economic decision making and better educate the general public on the ways in which their personal consumption decisions lead to biodiversity loss. Chapter 5 combines global range maps of terrestrial bird and mammal species, maps of the human appropriation of net primary productivity, and economic input-output tables to create a global "wildlife footprint" that links 57 types of human consumption activities in 113 regions of the world to global decreases in wild bird and mammal populations. Current land use practices are found to decrease global wildlife populations by 20–25% relative to the baseline populations that would be present in the absence of human land uses, representing an ongoing loss of approximately 25 billion individual birds and 1.1 trillion individual mammals. The nations of Brazil, China, India, Indonesia, the Philippines, Russia, and the United States have the largest impacts on global wildlife populations under several methods of measurement. Consumer purchases from agricultural sectors lead to the majority of wildlife population losses, with one dollar spent by a final consumer in several agricultural sectors leading to losses on the order of 0.1–1 individual birds or mammals. These results provide a new perspective on the global drivers of biodiversity loss and may help to identify new, consumption-based points of leverage for decreasing pressures on global biodiversity.

Chapter 6 concludes and suggests directions for future research.

References

[1] Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography* 19:589–606.

[2] Sanderson E, Jaiteh M, Levy M, Redford K (2002) The human footprint and the last of the wild. *BioScience* 52:891–904.

[3] Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis* (Island Press, Washington, D.C.), p 137.

[4] Foley JA, et al. (2011) Solutions for a cultivated planet. *Nature* 478:337–342.

[5] Haberl H, et al. (2007) Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 104:12942–12947.

[6] Postel SL, Daily GC, Ehrlich PR (1996) Human appropriation of renewable fresh water. *Science* 271:785–788.

[7] IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* eds Solomon S, et al. (Cambridge University Press, Cambridge, United Kingdom and New York, NY), p 996.

[8] Crutzen PJ (2002) Geology of mankind. *Nature* 415:23.

[9] Steffen W, et al. (2011) The Anthropocene: from global change to planetary stewardship. *Ambio* 40:739–761.

[10] Barnosky AD, et al. (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471:51–57.

[11] Butchart SHM, et al. (2010) Global biodiversity: indicators of recent declines. *Science* 328:1164–1168.

[12] IUCN (2011) The IUCN Red List of Threatened Species. Version 2011.2.

[13] Callicott JB (2006) in *Principles of Conservation Biology*, eds Groom MJ, Meffe GK, Carroll CR (Sinauer Associates, Sunderland, MA), Third edition, pp 111–135.

[14] Noss R (1999) Is there a special conservation biology? *Ecography* 22:113–122.

[15] Soulé ME, Wilcox BA (1980) *Conservation Biology: An Evolutionary-Ecological Perspective* (Sinauer Associates, Sunderland, MA), p 395.

[16] Groom MJ, Meffe GK, Carroll CR (2006) *Principles of Conservation Biology* (Sinauer Associates, Sunderland, MA), Third edition, p 779.

[17] Meine C, Soule M, Noss RF (2006) "A mission-driven discipline": the growth of conservation biology. *Conservation Biology* 20:631–651.

[18] Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405:243–253.

[19] Ball I, Possingham H, Watts M (2009) in *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*, eds Moilanen A, Wilson K, Possingham H (Oxford University Press, Oxford, United Kingdom), pp 185–195.

[20] Sarkar S, et al. (2006) Biodiversity conservation planning rools: present status and challenges for the future. *Annual Review of Environment and Resources* 31:123–159.

[21] Guisan A, Zimmermann N (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147–186.

[22] Elith J, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.

[23] Franklin J (2010) *Mapping Species Distributions: Spatial Inference and Prediction* (Cambridge University Press, Cambridge, United Kingdom), p 338.

[24] Beissinger SR, McCullough DR (2002) *Population Viability Analysis* (University of Chicago Press, Chicago, IL), p 577.

[25] Akçakaya HR (2004) *Species Conservation and Management: Case Studies, Volume 1* (Oxford University Press, Oxford, United Kingdom), p 533.

[26] May RM, Lawton JH, Stork NE (1995) in *Extinction Rates*, eds Lawton JH, May RM (Oxford University Press, Oxford, United Kingdom), pp 1–24.

[27] Rosenzweig ML (1995) *Species Diversity in Space and Time* (Cambridge University Press, Cambridge, United Kingdom), p 436.

[28] Drakare S, Lennon JJ, Hillebrand H (2006) The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters* 9:215–227.

[29] Harte J, Smith AB, Storch D (2009) Biodiversity scales from plots to biomes with a universal species-area curve. *Ecology Letters* 12:789–797.

[30] He F, Hubbell SP (2011) Species-area relationships always overestimate extinction rates from habitat loss. *Nature* 473:368–371.

[31] Jetz W, McPherson JM, Guralnick RP (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution* 27:151–159.

[32] Williams PH, Araújo MB (2000) Using probability of persistence to identify important areas for biodiversity conservation. *Proceedings of the Royal Society B* 267:1959–1966.

[33] Cabeza M, Moilanen A (2001) Design of reserve networks and the persistence of biodiversity. *Trends in Ecology and Evolution* 16:242–248.

[34] Williams PH, Araújo MB (2002) Apples, oranges, and probabilities: integrating multiple factors into biodiversity conservation with consistency. *Environmental Modeling and Assessment* 7:139–151.

[35] Tilman D, May R, Lehman C, Nowak M (1994) Habitat destruction and the extinction debt. *Nature* 371:65–66.

[36] Smith AB (2010) Caution with curves: caveats for using the species-area relationship in conservation. *Biological Conservation* 143:555–564.

[37] Araújo MB, Williams PH, Fuller RJ (2002) Dynamics of extinction and the selection of nature reserves. *Proceedings of the Royal Society B* 269:1971–1980.

[38] Cabeza M, et al. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology* 41:252–262.

[39] Chave J, Wiegand K, Levin S (2002) Spatial and biological aspects of reserve design. *Environmental Modeling and Assessment* 7:115–122.

[40] Frank K, Wissel C (2002) A formula for the mean lifetime of metapopulations in heterogeneous landscapes. *The American Naturalist* 159:530–552.

[41] Moilanen A, Cabeza M (2002) Single-species dynamic site selection. *Ecological Applications* 12:913–926.

[42] Carroll C, Noss RF, Paquet PC, Schumaker NH (2003) Use of population viability analysis and reserve selection algorithms in regional conservation plans. *Ecological Applications* 13:1773–1789.

[43] McCarthy MA, Thompson CJ, Possingham HP (2005) Theory for designing nature reserves for single species. *The American Naturalist* 165:250–257.

[44] McCarthy MA, Thompson CJ, Williams NSG (2006) Logic for designing nature reserves for multiple species. *The American Naturalist* 167:717–727.

[45] Nicholson E, et al. (2006) A new method for conservation planning for the persistence of multiple species. *Ecology Letters* 9:1049–1060.

[46] Beier P, Brost B (2010) Use of land facets to plan for climate change: conserving the arenas, not the actors. *Conservation Biology* 24:701–710.

[47] Tanentzap AJ, Walker S, Theo Stephens RT, Lee WG (2012) A framework for predicting species extinction by linking population dynamics with habitat loss. *Conservation Letters*.

Chapter 2

Extinction risk and tradeoffs in reserve site selection for species of different body sizes

Abstract

As natural landscapes become fragmented, species' populations may become restricted to remnant habitat patches. Choosing which patches to protect often requires a tradeoff between maximizing local patch sizes to reduce local extinction risks and clustering patches to increase colonization rates. Here, I examine how these two considerations affect landscape-wide extinction risk for idealized terrestrial mammal species of different body sizes. Spatially-explicit stochastic, demographic metapopulation models, parameterized with empirical allometric relationships, are used to simulate extinction risk for 10 g to 100 kg species inhabiting patchy landscapes. An analysis of two-patch networks demonstrates that patch clustering decreases extinction risk only when inter-patch distances are within approximately 0.5 to 1.25 times a species' maximum observed dispersal distance. In an empirical, fragmented landscape in which a limited fraction of available habitat can be protected, this pattern accurately predicts that, relative to a network that maximizes mean patch area, a clustered patch network most decreases extinction risk for intermediate-sized species and increases extinction risk for the largest species. These results demonstrate that there is in principle no globally optimal level of patch clustering that minimizes extinction risk for all species and highlight rules of thumb for reserve network design based on the interaction of a species' body size and landscape scale.

Introduction

As global habitat loss and fragmentation continues, species' populations are becoming increasingly restricted to remnant habitat patches surrounded by a relatively inhospitable matrix [1, 2]. A major outstanding challenge in spatial ecology and conservation biology is to predict the extinction risk faced by these newly fragmented populations and to use this knowledge to prioritize the conservation of remaining patches to best minimize this risk.

As suggested by island biogeography and metapopulation theory [3–6], the two first-order factors that control diversity in patchy landscapes are local patch areas (primarily by influencing the probability of local patch extinction) and patch isolation (primarily by influencing the probability of the recolonization of a locally extinct patch). Land managers and policy makers with a fixed budget for conservation have thus long been advised to prioritize protection of the largest available habitat patches in a landscape and, presuming a relatively homogeneous matrix between patches, to simultaneously ensure that these patches are as close together as possible [7–9]. In any real landscape, however, it will generally be impossible to simultaneously maximize mean patch area and minimize inter-patch distance, barring the coincidental case in which the largest available habitat patches form the tightest cluster. Managers are thus frequently faced with a tradeoff between these two considerations, and the appropriate balance between the two should ideally be set by examining the risk of a species' extinction in each potential network.

Many methods have been proposed for estimating a species' extinction risk in patchy landscapes and selecting a network of patches to minimize this risk. Using detailed empirical data gathered for specific species and landscapes, spatially-explicit stochastic patch occupancy models [e.g., 10–13] and age or stage structured demographic models [e.g., 14–17] can be constructed to predict extinction risk in any patch network. These models lack generality, however, in that they are parameterized for a particular species in a particular landscape, and the extent of their broader applicability is often unclear. Conversely, more theoretical approaches have reached overarching conclusions about extinction risk as a function of species and network characteristics [e.g., 18–21]. These approaches, however, often do not consider the population dynamics of individual patches, are fully or partially spatially implicit, use aggregate constants to represent species' life history traits, and assume relatively simple statistical distributions for model parameters.

Methods of intermediate detail that balance the needs for generality and specificity remain relatively uncommon. A desirable method for patch selection should be broadly applicable and flexible enough to be used to explore the effects of different modeling assumptions, while also being detailed enough to directly incorporate key differences between species and landscapes.

One promising approach to achieving this balance involves the combination of empirical allometric relationships, known patterns in the relationship between body size and life history traits [22–24], and dynamic population models [25, 26]. Here, I construct and explore one such set of models for idealized terrestrial mammal species of different body sizes and use it to analyze the tradeoff between patch size and patch clustering in protected area networks. The models are spatially-explicit and account for decreases in reproductive rates, decreases in population densities, and increases in dispersal distances associated with increasing body size in mammals [27, 28]. These models are first used to examine extinction risk by body size in two-patch networks, across a range of patch areas and inter-patch distances, to identify general patterns in the relationship between extinction

risk, body size, and network characteristics. The analysis is then extended to an empirical landscape in northern California, where the patch networks that best minimize extinction risk for species of different body sizes are identified and compared.

Overall, the results demonstrate that, at any spatial scale, selecting clustered patch networks does not benefit all species equally and that the species that may benefit from clustering can be identified by matching dispersal distances to inter-patch distances. In principle, there is thus no globally optimal network design that best minimizes extinction risk for species of all body sizes, and the selection of priority patches for conservation should be made with the explicit knowledge of which species are advantaged and disadvantaged by any network design. These general findings, as well as the full allometry-based models themselves, could be broadly and rapidly extended to many taxonomic groups and landscapes and may be particularly useful for the practice of conservation planning in landscapes where little is known about resident species of conservation concern.

Methods

Population Models

To estimate extinction risk in patchy landscapes, spatially-explicit stochastic, demographic, single-stage, discrete-time metapopulation models are constructed for idealized terrestrial mammal species of different body sizes. Although these models describe a Markov process and could in principle be solved analytically [e.g., 29, 30], the number of possible states of the metapopulation grows as the product of the maximum population sizes of all patches, rendering a simulation approach necessary for the networks treated here. All analyses were carried out in R 2.13.2 [31].

The overarching goal is to create a model for each species that incorporates realistic landscape features, such as individual patch areas and locations, while relying, to the extent possible, only on life history parameters that can be predicted from body size. The models proposed here are based on five parameters for each species: maximum female offspring per year per adult female, y, mean female population density, s, maximum observed dispersal distance, x, and two parameters, θ and p_e, that characterize environmental stochasticity in the landscape. The first three of these can be estimated from a species' adult body mass, while the final two cannot be derived from first principles and must be assumed. A species' relative extinction risk in different patch networks, however, is relatively insensitive to all parameters except for dispersal distance (*Appendix*).

Beginning with a population in patch i at time t, $N(t)_i$, the population of i at time $t+1$, $N(t+1)_i$, is calculated by sequentially simulating birth and recruitment, death due to demographic stochasticity, and death due to environmental stochasticity.

$$N(t)_i^* = N(t)_i + \text{Poisson}[\Sigma^j z_{ij}\, yN(t)_j] \tag{2.1a}$$
$$N(t)_i^{**} = N(t)^* - \text{Binomial}[N(t)_i^*, d(t)_i] \tag{2.1b}$$
$$N(t+1)_i = \text{Binomial}[N(t)_i^{**}, \text{Beta}[\theta p_e, \theta(1-p_e)]] \tag{2.1c}$$

Equation (2.1a) models birth as a Poisson process with the mean number of recruits arriving in patch i given by the sum over all patches j, including $i = j$, of the product of the expected number of births in j, given by $yN(t)_j$, and the probability that a new born individual in patch j recruits to patch i, z_{ij}.

Calculating z_{ij} requires the choice and parameterization of a dispersal kernel for each species. Dispersal is modeled here using a radially symmetric bivariate Gaussian distribution, which, in a homogeneous landscape, can be interpreted phenomenologically as a diffusion process or mechanistically as the result of each individual undertaking a non-biased random walk from its place of birth [32, 33].

$$h_{ij} = \frac{1}{2\pi\sigma_h^2} \exp\left(-\frac{d_{ij}^2}{2\sigma_h^2}\right) \tag{2.2}$$

$$z_{ij} = \frac{h_{ij} A_i}{\sum^i h_{ij} A_i} \tag{2.3}$$

Equation (2.2) gives the probability that an individual born at the center of source patch j recruits to a point located at the center of destination patch i, h_{ij}. The product of h_{ij} and the area of patch i, A_i, gives an approximate probability that an individual born in patch j recruits to patch i. Equation (2.3) normalizes these probabilities for each source patch across all possible destination patches so that all newborn individuals recruit to a patch. Equation (2.2) represents only one reasonable choice for a dispersal kernel, and other "fat tailed" or superdiffusive dispersal kernels could be explored as well [34, 35].

Equation (2.1b) models death due to demographic stochasticity as a binomial process in which the mean per capita death rate for patch i at time t, $d(t)_i$, is density dependent. A general functional form for $d(t)_i$ should be chosen so that when $N(t)_i^* = 0$ there are no deaths, when $N(t)_i^* = K_i$, the carrying capacity of patch i, the expected number of deaths is equal to the expected number of births, and $0 \leq d(t)_i \leq 1$. These criteria are met simply by Eq. (2.4).

$$d(t)_i = 1 - \frac{1}{1 + y[N(t)_i^*/K_i]} \tag{2.4}$$

K_i is calculated as the product of mean female population density for a species, s, and the area of patch i, multiplied conservatively by a factor of 10 to estimate a maximum population size for each patch.

Finally, Eq. (2.1c) models death due to exogenous environmental stochasticity. Phenomenologically, the desired expression should model simple "catastrophe" dynamics in which (a) the two most likely outcomes each year are no additional death or local patch extinction, with no additional death being the most likely outcome, and (b) the risk of local patch extinction is a function of population size, with smaller populations subject to the highest risk. These two goals are met flexibly by a beta-binomial distribution, which is parameterized here with a single shape parameter θ and a mean per capita yearly survival probability p_e [36]. In all simulations, $\theta = 1$, giving a strongly U-shaped distribution with peaks at 0 and $N(t)_i^{**}$. The parameter p_e is then varied to adjust the relative probabilities of the two extrema, with lower values of p_e leading to a higher probabilities of local patch extinction.

These models are not intended to capture the complexity of the many population processes that are present for any real species in any real landscape. I propose, however, that the behavior of these simple, general models can provide a starting point for exploring the extinction risk faced by different species in different types of landscapes.

Allometric Data

The population models described above are parameterized for idealized terrestrial mammal species. Terrestrial mammals are chosen as a focal species assemblage for this analysis as terrestrial mammals are often the target of conservation plans, have a wide range of body sizes, and have well-characterized allometric relationships. For a species of any adult body mass, m (g), log-log linear regressions using global data for terrestrial mammals [28] are used to calculate yearly reproductive output, y (maximum female young per adult female per year), as one-half the product of maximum litter size and the number of litters per year, and population density, s (females per hectare).

$$\log_{10} y = 0.72\ (\text{SE } 0.03) - 0.06\ (\text{SE } 0.01) \log_{10} m \tag{2.5}$$

$$\log_{10} s = 3.88\ (\text{SE } 0.07) - 0.74\ (\text{SE } 0.02) \log_{10} m \tag{2.6}$$

The maximum litter size, litters per year, and population density values for each species, as reported in [28], are the median values of all measures of central tendency, from all localities, reported in all surveyed literature.

Maximum observed dispersal distance, x (km), is estimated from a log-log linear regression for all species of terrestrial mammals presented by Sutherland et al. [27], here Eq. (2.7).

$$\log_{10} x = 6.46 \text{ (SE } 1.23) + 0.68 \text{ (SE } 0.08) \ \log_{10} m \tag{2.7}$$

The value of x is used to set the parameter σ_h, the variance of the Gaussian dispersal kernel used in the models, by calculating x to be the distance below which 99% of newborn individuals recruit. The relationship $\sigma_h = 0.329x$ was determined by simulation to give this probability. Further information on the definition and measurement of maximum observed dispersal distance can be found in [27].

Values of these parameters for species of various body sizes are given in Table (2.1).

Two-Patch Networks

The models are first applied to two-patch networks to explore changes in extinction risk with patch area and inter-patch distance. Models are parameterized for species with adult body masses of 0.01 kg, 0.1 kg, 1 kg, 10 kg, and 100 kg, representing the approximate range of body sizes of terrestrial mammals of western North America. The probability of extinction by year 100, based on 1,000 replicate simulations, is calculated for each species at combinations of patch area and inter-patch distance ranging from 1 to 200 hectares and 1 to 50 kilometers, respectively. Initial populations for each species in each patch are set to one-half of the carrying capacity of that patch. Environmental stochasticity is set with the parameter $p_e = 0.85$.

As the distance between patches decreases to zero, the dynamics of these two-patch networks do not approach the dynamics of a single patch, as the two patches experience uncorrelated environmental stochasticity and the home ranges of individual organisms cannot overlap both patches (see *Discussion*). In general, including positive correlations in dynamics between nearby patches to reflect region-wide stochastic processes would tend to increase extinction risks for species in clustered patch networks.

Multi-Patch Networks

The models are then applied to an empirical, multi-patch landscape from which a subset of patches with a fixed total area are to be selected for conservation. For each species, the goal is to select a subset of available patches to create a protected area network, with total area less than a specified threshold, that minimizes the extinction risk for that species. This approach is similar to the maximal covering problem in the reserve site selection literature [37, 38] but with the use of single-species extinction risk as the objective function instead of the number of species represented in the network.

Reserve networks are commonly selected with either exact global optimization [9, 37, 39, 40] or heuristic procedures [41]. While heuristics are not guaranteed to find globally optimal reserve network designs, many heuristics have been shown to produce near optimal solutions to other reserve design problems [41, 42]. At present, the long computation

times associated with stochastic simulation of these population models preclude the use of exact global optimization. Instead, a first-in greedy heuristic [43] is used to select reserve networks. This heuristic begins by selecting the single patch that most decreases extinction risk for the species (in these models, always the largest patch in the landscape) and adding at each step the remaining patch whose addition leads to the lowest network-wide extinction risk, until the maximum total area that can be protected is reached.

In a two-patch network, the probability of network-wide extinction can be used directly as a metric for comparing different networks. In networks containing patches that vary widely in area, however, the contribution of small patches to reducing extinction risk is dwarfed by the variance in extinction risk associated with the stochastic dynamics of large patches in the network. As such, the mean number of occupied patches in a candidate reserve network is used in place of network-wide extinction risk as the site selection criterion in the empirical, multi-patch landscape.

This procedure is used to select a reserve network that maximizes survival probability for species of 0.01 kg, 0.1 kg, 1 kg, 10 kg, and 100 kg adult body masses in a 57 patch fragmented landscape located east of Santa Rosa, California. This landscape is a low-elevation mixed oak savannah [44] with 57 remnant patches of natural habitat interspersed with urban and agricultural development [45]. Patch sizes range from approximately 5 hectares to 150 hectares, and pairwise inter-patch distances range from several hundred meters to approximately 20 kilometers.

All multi-patch simulations use parameters identical to the two-patch simulations described above, except for $p_e = 0.7$. Habitat quality within all patches is assumed to be identical, and the matrix between habitat patches is assumed to be homogeneous. To examine qualitative patterns in the configuration of the networks selected by each species, complete reserve networks containing up to 66% of the total available area are selected 10 times for each species.

A quantitative examination of the tradeoff between patch area and patch clustering is also conducted by calculating the extinction risk faced by each species in a reference network, consisting of the four largest available patches in the landscape, and a clustered network, a network of the same total area that includes three of these four patches and several smaller patches nearby (Fig. 2.3, inset). Extinction risks are calculated for a 100 year time horizon using 2,000 replicate simulations. For comparability between species, the value of p_e for each species that gives a 50% probability of extinction in the reference network in year 100 is identified, and the probability of extinction in the clustered network at this value of p_e is calculated for comparison.

Results

In two-patch networks, extinction risk varies greatly as a function of body size, patch area, and inter-patch distance (Fig. 2.1). Importantly, inter-patch distance does not affect the probability of survival for a species evenly across all distances. Rather, each species exhibits a unique and limited "characteristic range", defined here as the range of inter-patch distances within which inter-patch distance substantially influences survival probability, that is indicated by the steeply sloping isoclines of survival probability in Fig. (2.1). Only within this characteristic range do changes in inter-patch distance substantially increase or decrease a species' extinction risk, while outside of this range, inter-patch distance is largely irrelevant to extinction risk. Broadly speaking, this range is defined by a lower threshold below which a species can nearly equally disperse between patches at all distances and an upper threshold above which a species cannot disperse between patches regardless of distance.

Within the models presented here, any differences in extinction risk between species reflects only differences in yearly reproductive rate, population density, and maximum observed dispersal distance. The characteristic range of each species, however, is to a first approximation only a function of that species' maximum observed dispersal distance x (Fig. 2.1, *Appendix*). The characteristic range falls at distances of 0.14 to 0.35 km for a 0.01 kg species, 0.67 to 1.69 km for a 0.1 kg species, 3.23 to 8.08 km for a 1 kg species, 15.5 to 38.6 km for a 10 kg species, and 74.0 to 185 km for a 100 kg species (Tab. 2.1).

In a multi-patch landscape, clustered patch networks will generally have shorter inter-patch distances, more patches, and lower mean patch areas than a reference network selected to maximize mean patch area. In general, the first two of these factors have the potential to lead to lower extinction risks in clustered networks, while the third has the potential to lead to higher extinction risks in clustered networks. The analysis of the differences between a clustered network and a reference network that maximizes mean patch area is thus reminiscent of the long-standing single large or several small (SLOSS) debate in conservation biology [reviewed in 46], where here the relationship between inter-patch distance and dispersal is explicitly modeled.

Considering specifically the importance of patch recolonization and local patch extinction risk, the two-patch model results lead to an initial hypothesis regarding which species might experience reduced extinction risk in a multi-patch network that is clustered at any scale. Relative to the reference network that maximizes mean patch area, clustering may specifically benefit, via increasing recolonization rates, those species whose characteristic range matches the scale of the patch cluster. If, however, this clustering also results in a decrease in mean patch area, all species may potentially experience an additional "cost" associated with an increase in local patch extinction risk. As a first approximation, any given scale of patch clustering might thus be expected to most reduce network-wide extinction risk for species whose characteristic range matches the cluster scale.

This hypothesis is first tested qualitatively in a empirical, multi-patch landscape in which a reserve design heuristic is used to select a network that minimizes extinction risk for each species, given that the entire set of patches cannot be protected (Fig. 2.2). In this landscape, many pairwise inter-patch distances are on the order of 1 to 10 km, a spatial scale that overlaps most substantially with the characteristic ranges of the 0.1 kg species and 1 kg species (Tab. 2.1). For simplicity and generality, no specific differences in habitat within or between patches are assumed.

As predicted, the chosen reserve network for the 0.1 kg species qualitatively displays clustering at small spatial scales within this landscape, the 1 kg species network displays clustering across larger scales, and the networks selected by both larger and smaller bodied species are similar to the reference network. In a hypothetical landscape in which inter-patch distances are artificially increased by a factor of five, the pattern of patch clustering selected here by the 1 kg species is selected by the 10 kg species, also as predicted (*Appendix*).

Figure (2.3) compares, for each species, extinction risk in a reference network consisting of the four largest available patches and extinction risk in a clustered network of the same total area consisting of three of these patches and a set of smaller nearby patches (similar to the network selected by the 1 kg species). At a species-specific level of p_e chosen such that each species has a 50% probability of survival to year 100 in the reference network, probabilities of survival to year 100 are 6%, 33% 49%, and 20% greater in the clustered network for the 0.01 to 10 kg species, and 20% lower in the clustered network for the 100 kg species. As expected from the qualitative results, the species that most benefits from a clustered network design is the intermediate-sized 1 kg species, with the benefits of clustering decreasing for successively larger or smaller body sizes.

Discussion

The two-patch analysis finds that each idealized terrestrial mammal species of a given body size has a unique "characteristic range" within which changes in inter-patch distance influences extinction risk, but above and below which extinction risk is insensitive to distance. This characteristic range is found from approximately 0.5 to 1.25 times a species' maximum observed dispersal distance and varies greatly by species.

In any real landscape, inter-patch distances will fall within the characteristic range of only a subset of all species inhabiting that landscape. These species may experience lower extinction risk in clustered patch networks due to increased recolonization rates, while species whose characteristic range does not match the scale of inter-patch distances may experience higher extinction risk due to the inevitable decrease in mean patch area associated with selecting clustered patches. These patterns are confirmed in the empirical multi-patch analysis, where a clustered network most lowers extinction risk for medium bodied species and increases extinction risk for the largest species.

These findings violate two assumptions that may often be made by conservation planners and managers when selecting patches for protection, which are that a clustered network of patches (a) most benefits the largest bodied species in the landscape (which has the highest probability of local extinction and hence supposedly will most benefit from increased recolonization rates), and (b) will not substantially harm the probability of survival for any species in the landscape. These intuitions may be behind the widespread application of many methods that arbitrarily increase patch clustering in reserve network designs, including the boundary length modifier in the popular conservation planning software program MARXAN [47], the selection philosophy embodied in the worst-out heuristic used in the software program Zonation [48], and attempts to incorporate spatial adjacency into exact reserve network optimization models [e.g., 40, 49–51].

In the empirical landscape examined here, these assumptions are clearly violated, as the largest bodied species exhibits substantially higher extinction risk in a clustered patch network, while other species exhibit lower extinction risks. These results clearly demonstrate that the decision to select a clustered or non-clustered network necessarily represents an implicit decision to favor the protection of some species over others. There is thus, in principle, no globally optimal level of patch clustering for all species.

Given these findings, conservation planners are left with three reasonable options when attempting to design a reserve network to decrease extinction risk for multiple species of different body sizes. First, network design decisions may be made for the largest bodied species present in the landscape, which will generally have the highest absolute extinction risk in any network. Second, if the basic body size structure of the community is known, an explicit decision may be made to cluster patches or not to cluster patches based on the characteristic ranges of the species present in the landscape and the overall scale of inter-patch distances. In the empirical landscape examined here, for example, if species of 100 kg are not present or not in need of conservation, a clustered network may be preferred simply based on the knowledge of approximate dispersal distance for medium sized mammals and the tens-of-kilometers scale of the network. Third, a detailed simulation such as the one presented here could be conducted for any real landscape to help inform conservation planning decisions.

Several limitations of the modeling framework and specific results presented here are important to note. First, results for species of each body size should be interpreted in light of the substantial variation around the allometric regressions used to parameterize the models used here (*Appendix*). In empirical applications, ranges of parameters could be used in place of central tendencies.

Second, the modeling framework presented here embodies a metapopulation perspective in which the only movement between patches occurs as a result of natal dispersal. Additional adult movements between patches, including those that would allow an individual to combine multiple patches into a single home range, are not incorporated. While this approximation may be reasonable when the matrix between patches is relatively inhospitable,

it will become less accurate for vagile species inhabiting patches that are very close together and separated by very permeable matrix habitat.

Third, these models rely on substantial simplifications of the dynamics of species' dispersal in complex landscapes. Dispersal can be considered to be a product of both a species' life history traits and landscape features, and the models presented here are constructed for landscapes in which the matrix between patches is homogeneous and no specific features such as corridors or impermeable barriers are present between patches [52]. Additionally, behavioral interactions between dispersing individuals and patch edges, which may be species-specific, are not included [53]. More detailed knowledge of landscape features and species' traits, where available, could be incorporated in this modeling framework by using this knowledge to increase or decrease dispersal probabilities between pairs of patches.

Fourth, like many general models of population dynamics in patchy landscapes [e.g., 13, 19, 46], these models do not include correlation in environmental stochasticity between patches, which, to the extent that populations physically closer together in space tend to have correlated dynamics, would likely increase absolute predicted extinction risk for clustered network designs, as nearby sites would be less likely to serve as a source of colonists in the event of a local extinction.

Fifth, all models were simulated for a 100 year time horizon and do not include temporal autocorrelation in patch dynamics. Longer time horizons and temporal correlations will likely increase the probability of local extinction, and, in contrast to the above, may increase the value of recolonization from nearby patches and hence the value of clustered designs.

Finally, empirical verification of extinction models such as these is critically important to testing their utility for real world conservation. I am not aware, however, of any data sets that record species persistence in fragmented landscapes over the century-long time scales simulated here. As such, the results presented here should be interpreted as indicative estimates of general patterns and should be used in empirical planning applications only when better data and models are lacking.

Three additional extensions to these models could aid their application to empirical problems in conservation planning. First, the framework presented here could be applied to other taxa with well-characterized allometric relationships, such as birds. Second, modifications could be made to this modeling framework in the event that specific knowledge of a species' life history traits or the characteristics of a specific landscape are known. Reproductive output and population density, for example, could be modified on a patch-by-patch basis to reflect variation in habitat quality, and information on the matrix between patches could be used to modify inter-patch dispersal probabilities. Third, the extinction-based approach described here could be combined with a representation-based reserve selection approach, such as that implemented in the popular conservation software packages MARXAN [47] or Zonation [48], to select networks that would represent the greatest number of species in a reserve network at a future time horizon, rather than at present.

Conclusion

The allometry-based models presented here provide a novel approach for estimating the extinction risk faced by species of different body sizes in any network of habitat patches. These models have the potential to inform the practice of conservation planning in fragmented landscapes, where managers are likely to face a tradeoff between selecting a protected area network that maximizes mean local patch area and a network that minimizes inter-patch distances.

The results presented here for idealized terrestrial mammal species suggest that a species may have a substantially lower extinction risk in a clustered network of patches when inter-patch distances are approximately 0.5 to 1.25 times the species' maximum observed dispersal distance. Other species, however, may experience elevated extinction risks in this clustered patch network due to the likely associated decrease in mean patch area. The decision to select any particular design for a protected area network must thus be made carefully in light of the unavoidable tradeoff between the extinction risks of different species. Beyond these general findings, these models also provide a quantitative site-selection method that can be used to provide a first-pass design for reserve networks based explicitly on species' extinction risk, even in the absence of detailed empirical data on the species and landscape in question.

Acknowledgements

I thank Steve Beissinger, John Harte, William Lidicker, Jr., Adina Merenlender, and Adam Smith for helpful comments on drafts of this chapter.

References

[1] Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34:487–515.

[2] Lindenmayer DB, Fischer J (2006) *Habitat Fragmentation and Landscape Change* (Island Press, Washington, D.C.), p 352.

[3] MacArthur RH, Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17:373–387.

[4] MacArthur RH, Wilson EO (1967) *The Theory of Island Biogeography* (Princeton University Press, Princeton, NJ).

[5] Hanski I (1997) in *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, eds Tilman D, Kareiva P (Princeton University Press, Princeton, NJ), pp 21–45.

[6] Hanski I (1999) *Metapopulation Ecology* (Oxford University Press, Oxford, United Kingdom), p 313.

[7] Diamond J (1975) The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation* 7:129–146.

[8] Nekola JC, White PS (2002) Conservation, the two pillars of ecological explanation, and the paradigm of distance. *Natural Areas Journal* 22:305–310.

[9] Williams JC, ReVelle CS, Levin SA (2004) Using mathematical optimization models to design nature reserves. *Frontiers in Ecology and the Environment* 2:98–105.

[10] Moilanen A, Cabeza M (2002) Single-species dynamic site selection. *Ecological Applications* 12:913–926.

[11] Cabeza M, Moilanen A (2003) Site-selection algorithms and habitat loss. *Conservation Biology* 17:1402–1413.

[12] Drechsler M, Frank K, Hanski I, O'Hara RB, Wissel C (2003) Ranking metapopulation extinction risk: from patterns in data to conservation management decisions. *Ecological Applications* 13:990–998.

[13] Nicholson E, et al. (2006) A new method for conservation planning for the persistence of multiple species. *Ecology Letters* 9:1049–1060.

[14] Beissinger SR, McCullough DR (2002) *Population Viability Analysis* (University of Chicago Press, Chicago, IL), p 577.

[15] Carroll C, Noss RF, Paquet PC, Schumaker NH (2003) Use of population viability analysis and reserve selection algorithms in regional conservation plans. *Ecological Applications* 13:1773–1789.

[16] Akçakaya HR (2004) *Species Conservation and Management: Case Studies, Volume 1* (Oxford University Press, Oxford, United Kingdom), p 533.

[17] Newbold SC, Siikamäki J (2009) Prioritizing conservation activities using reserve site selection methods and population viability analysis. *Ecological Applications* 19:1774–1790.

[18] Frank K (2005) Metapopulation persistence in heterogeneous landscapes: lessons about the effect of stochasticity. *The American Naturalist* 165:374–388.

[19] McCarthy MA, Thompson CJ, Possingham HP (2005) Theory for designing nature reserves for single species. *The American Naturalist* 165:250–257.

[20] McCarthy MA, Thompson CJ, Williams NSG (2006) Logic for designing nature reserves for multiple species. *The American Naturalist* 167:717–727.

[21] Drechsler M, Johst K (2010) Rapid viability analysis for metapopulations in dynamic habitat networks. *Proceedings of the Royal Society B* 277:1889–1897.

[22] Blueweiss L, et al. (1978) Relationships between body size and some life history parameters. *Oecologia* 37:257–272.

[23] Damuth J (1981) Population density and body size in mammals. *Nature* 290:699–700.

[24] Silva M, Brown JH, Downing JA (1997) Differences in population density and energy use between birds and mammals: a macroecological perspective. *Journal of Animal Ecology* 66:327–340.

[25] Belovsky GE (1987) in *Viable Populations for Conservation*, ed Soule ME (Cambridge University Press, Cambridge, United Kingdom), p 204.

[26] Wilson KA, et al. (2010) Conserving biodiversity in production landscapes. *Ecological Applications* 20:1721–1732.

[27] Sutherland GD, Harestad AS, Price K, Lertzman KP (2000) Scaling of natal dispersal distances in terrestrial birds and mammals. *Conservation Ecology* 4:16.

[28] Jones KE, et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90:2648.

[29] Day JR, Possingham HP (1995) A stochastic metapopulation model with variability in patch size and position. *Theoretical Population Biology* 48:333–360.

[30] Frank K, Wissel C (1998) Spatial aspects of metapopulation survival - from model results to rules of thumb for landscape management. *Landscape Ecology* 13:363–379.

[31] R Core Team (2012) R: A Language and Environment for Statistical Computing.

[32] Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38:196–218.

[33] Codling EA, Plank MJ, Benhamou S (2008) Random walk models in biology. *Journal of the Royal Society Interface* 5:813–834.

[34] Clark J, Fastie C, Hurtt G, Jackson S, Johnson C (1998) Reid's paradox of rapid plant migration. *BioScience* 48:13–24.

[35] Getz WM, Saltz D (2008) A framework for generating and analyzing movement paths on ecological landscapes. *Proceedings of the National Academy of Sciences of the United States of America* 105:19066–19071.

[36] Bolker BM (2008) *Ecological Models and Data in R* (Princeton University Press, Princeton, NJ), p 408.

[37] Önal H (2003) First-best, second-best, and heuristic solutions in conservation reserve site selection. *Biological Conservation* 115:55–62.

[38] Williams JC, ReVelle CS, Levin SA (2005) Spatial attributes and reserve design models: A review. *Environmental Modeling and Assessment* 10:163–181.

[39] Rodrigues AS, Gaston KJ (2002) Optimisation in reserve selection procedures - why not? *Biological Conservation* 107:123–129.

[40] Groeneveld RA (2010) Species-specific spatial characteristics in reserve site selection. *Ecological Economics* 69:2307–2314.

[41] Csuti B, et al. (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80:83–97.

[42] Pressey RL, Possingham HP, Day JR (1997) Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biological Conservation* 80:207–219.

[43] Church RL, Stoms DM, Davis FW (1996) Reserve selection as a maximal covering location problem. *Biological Conservation* 76:105–112.

[44] Clark HW (1937) Association types in the north coast ranges of California. *Ecology* 18:214–230.

[45] Merenlender AM, Brooks C, Shabazian D, Gao S, Johnston R (2005) Forecasting exurban development to evaluate the influence of land-use policies on wildland and farmland conservation. *Journal of Conservation Planning* 1:40–57.

[46] Ovaskainen O (2002) Long-term persistence of species and the SLOSS problem. *Journal of Theoretical Biology* 218:419–433.

[47] Ball I, Possingham H, Watts M (2009) in *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*, eds Moilanen A, Wilson K, Possingham H (Oxford University Press, Oxford, United Kingdom), pp 185–195.

[48] Moilanen A (2007) Landscape Zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation* 134:571–579.

[49] Önal H, Briers RA (2005) Designing a conservation reserve network with minimal fragmentation: A linear integer programming approach. *Environmental Modeling and Assessment* 10:193–202.

[50] Marianov V, ReVelle C, Snyder S (2008) Selecting compact habitat reserves for species with differential habitat size needs. *Computers & Operations Research* 35:475–487.

[51] Williams JC (2008) Optimal reserve site selection with distance requirements. *Computers & Operations Research* 35:488–498.

[52] Hilty JA, Lidicker Jr. W, Merenlender AM (2006) *Corridor Ecology: The Science And Practice of Linking Landscapes for Biodiversity Conservation* (Island Press), p 323.

[53] Lidicker Jr. W (1999) Responses of mammals to habitat edges: an overview. *Landscape Ecology* 14:333–343.

Body mass (kg)	y	s	x	CR
0.01	2.26	6.91	0.28	0.14 – 0.35
0.1	1.96	1.26	1.35	0.67 – 1.69
1	1.70	0.23	6.46	3.23 – 8.08
10	1.47	0.04	31.0	15.5 – 38.6
100	1.28	0.01	148	74.0 – 185

Table 2.1: Life history parameters, including yearly reproductive output, y (maximum female young per adult female per year), population density, s (females per hectare), and maximum observed dispersal distance, x (km) for idealized terrestrial mammals of different body sizes (see *Methods*). Each species' characteristic range, CR (km), is calculated as 0.5 to 1.25 times x (see *Results*).

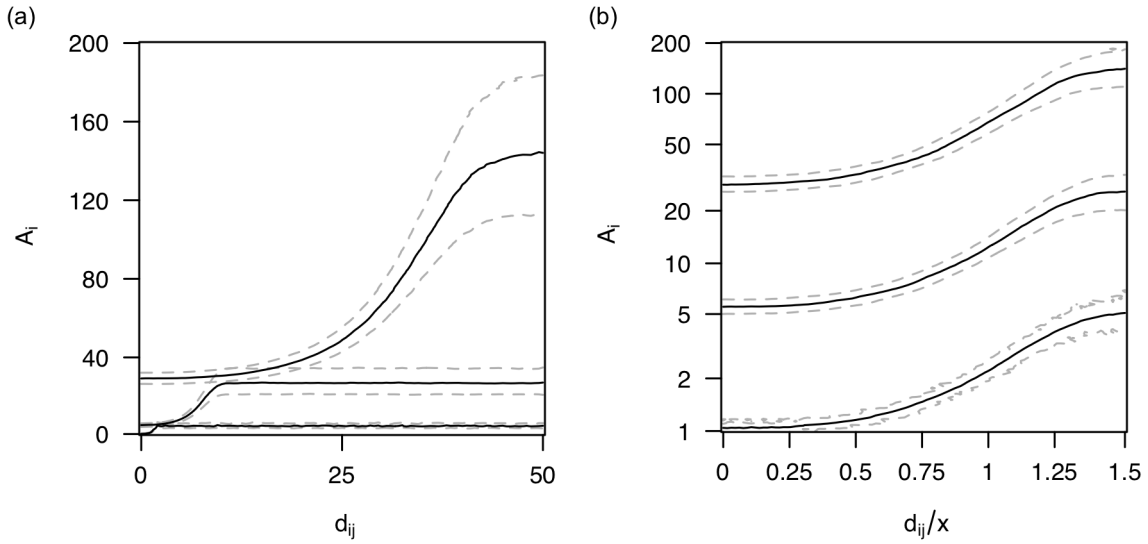

Figure 2.1: Isoclines of survival probability for species in two-patch networks with different patch areas, A_i (hectares), and inter-patch distances, d_{ij} (km). Solid lines, from bottom to top, show combinations of A_i and d_{ij} that give a 35% probability of survival to year 100 for species with adult body masses of 0.1 kg, 1 kg, and 10 kg, with dashed lines showing 25% and 45% probabilities. (a) Each species has a unique characteristic range, indicated by steeply sloping portions of the isoclines, within which inter-patch distance substantially affects extinction risk. (b) As in (a), but with the ratio of inter-patch distance to maximum observed dispersal distance on the x-axis and a log scale for area. The characteristic range for all three species falls at approximately 0.5 to 1.25 times the species' maximum observed dispersal distance.

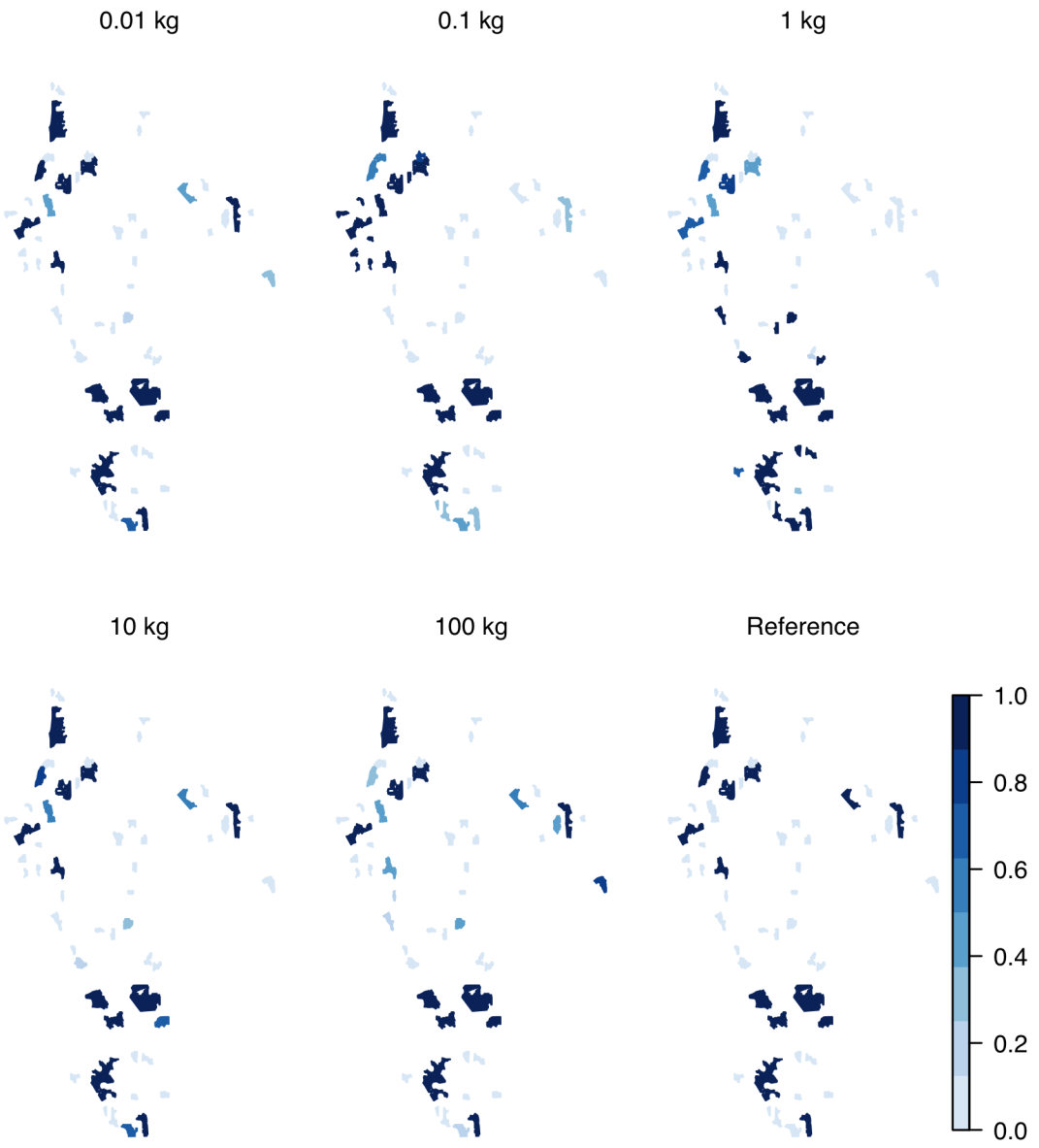

Figure 2.2: Reserve networks selected from a 57 patch empirical landscape for species of different body sizes, and a reference network that maximizes mean patch area. Darker colors indicate patches that were selected the most frequently by replicate reserve network design simulations. The characteristics of each species' reserve network are predicted by the match of inter-patch distances to that species' characteristic range.

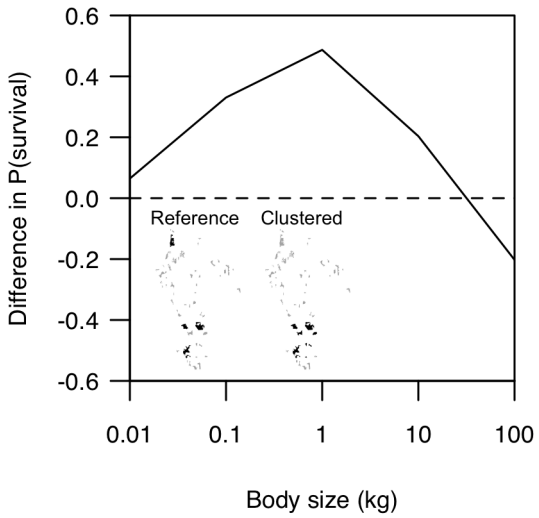

Figure 2.3: Difference in probability of survival to year 100 between a clustered reserve network and a reference network for species of different body sizes. p_e for each species is chosen to give a 50% probability of survival in the reference network and ranges from 0.58 to 0.77. Medium-bodied species have substantially higher survival probabilities in the clustered network, while the largest-bodied species has a lower survival probability in the clustered network.

Appendix

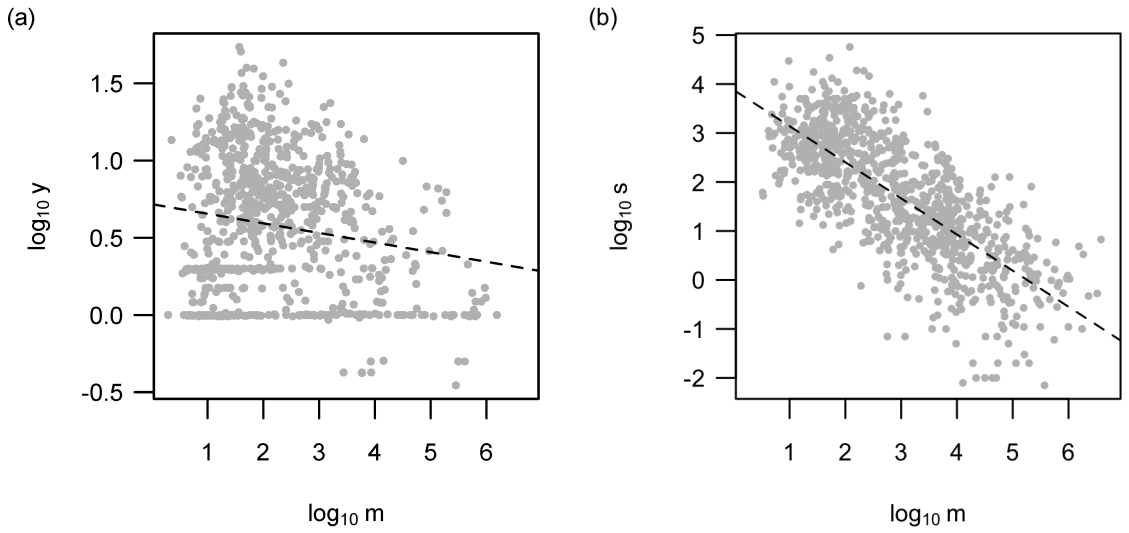

Figure 2.A1: Allometric relationships for terrestrial mammals based on data from Jones et al. 2009 (Ecological Archives E090-184-D1). (a) Maximum yearly reproductive output, y (female young per adult female per year), calculated as one-half the product of litter size and litters per year, as a function of body mass, m (g), for terrestrial mammals ($n = 813$). The shape of the log-log relationship between y and m is arguably not linear, and a reasonable alternative would be to set $\log_{10} y$ for all species to a global mean near 0.5. As characteristic range is relatively insensitive to y, this change would affect absolute probabilities of extinction for each species, but not preferred reserve networks (see Fig. 2.A2). (b) Population density, s (females per hectare), as a function of body mass, m (g), for terrestrial mammals ($n = 927$).

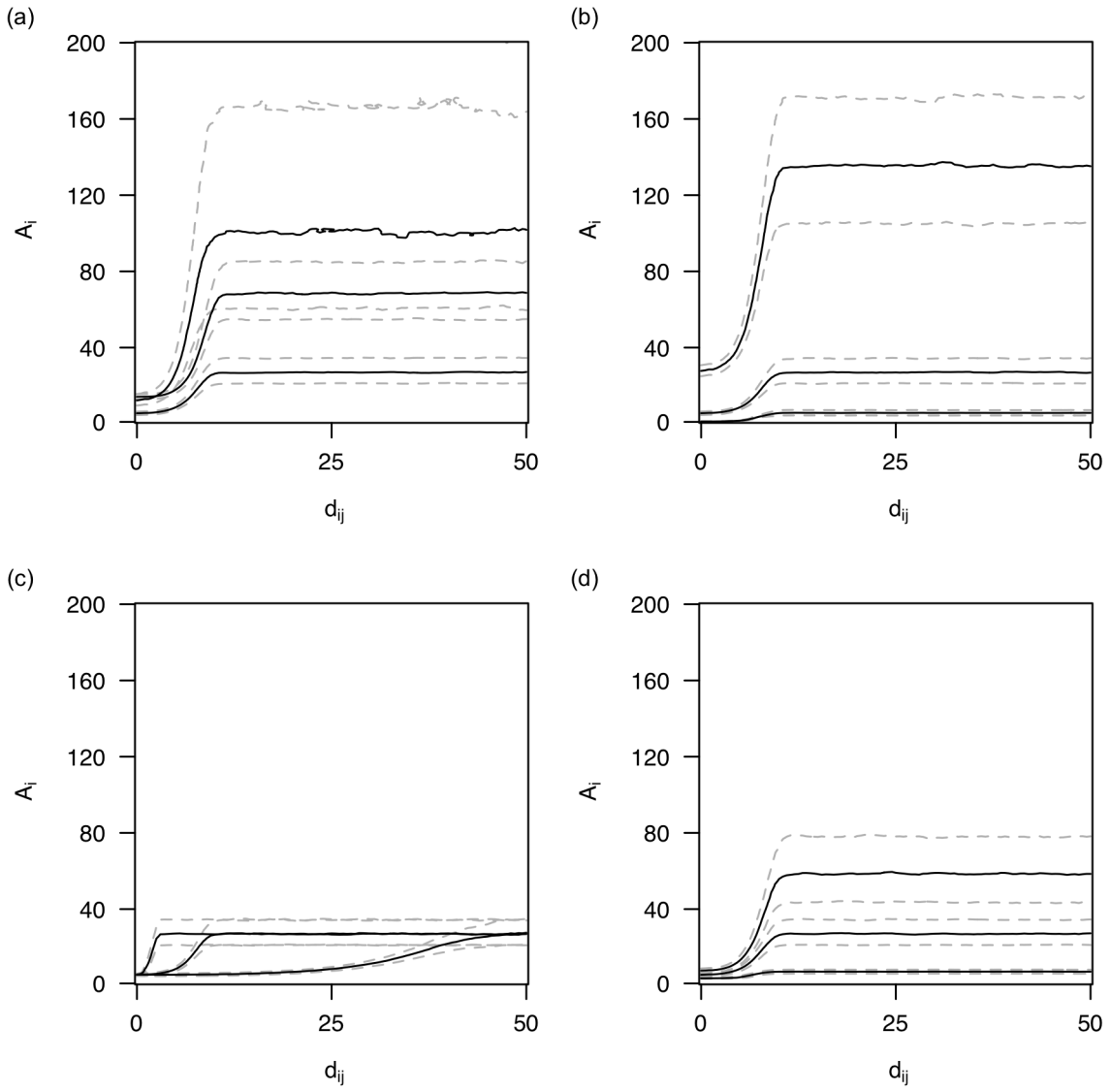

Figure 2.A2: Influence of model parameters on extinction risk for a 1 kg species in two-patch landscapes. Each sub-figure shows isoclines of 35% probability of survival to year 100 in solid lines, with 25% and 45% isoclines in surrounding dashed lines, for three different levels of a model parameter. (a) From bottom to top, yearly reproductive output y equal to 1, 5, and 0.2 times y drawn from regression on body mass. (b) From bottom to top, population density s equal to 5, 1, and 0.2 times s drawn from regression on body mass. (c) From left to right, maximum dispersal distance x equal to 0.2, 1, and 5 times x drawn from regression on body mass. (d) From bottom to top, environmental stochasticity parameter p_e equal to 0.95, 0.85, and 0.8. While all parameters influence absolute extinction risk (the height and shape of the isoclines), only x substantially influences the location of the characteristic range, the range of inter-patch distances with steeply sloping isoclines within which inter-patch distance substantially affects extinction risk.

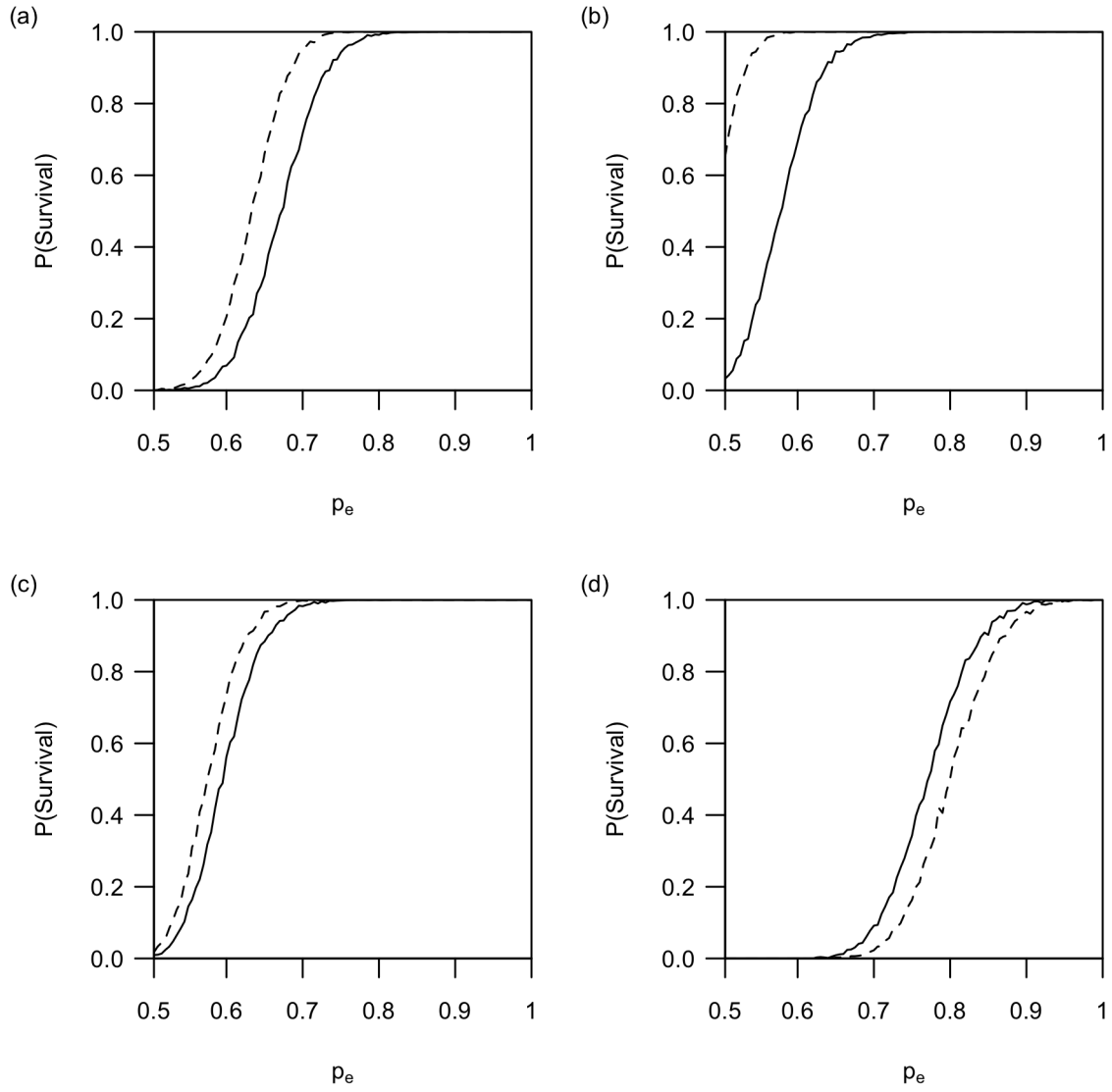

Figure 2.A3: Probability of survival to year 100 for species of (a) 0.1 kg, (b) 1 kg, and (c) 10 kg, and (d) 100 kg body masses in a reference network (solid line) consisting of the four largest patches in the empirical landscape and a clustered network (dashed line) (see Fig. 2.3, inset). The three smaller bodied species have a higher probability of survival in the clustered network across all values of p_e, while the 100 kg species has a higher probability of survival in the reference network for all p_e. The preferred reserve network design for each species is thus independent of the chosen level of environmental stochasticity. Results for the 0.01 kg species are not shown (survival probabilities in the two networks are very similar for all levels of p_e for this species).

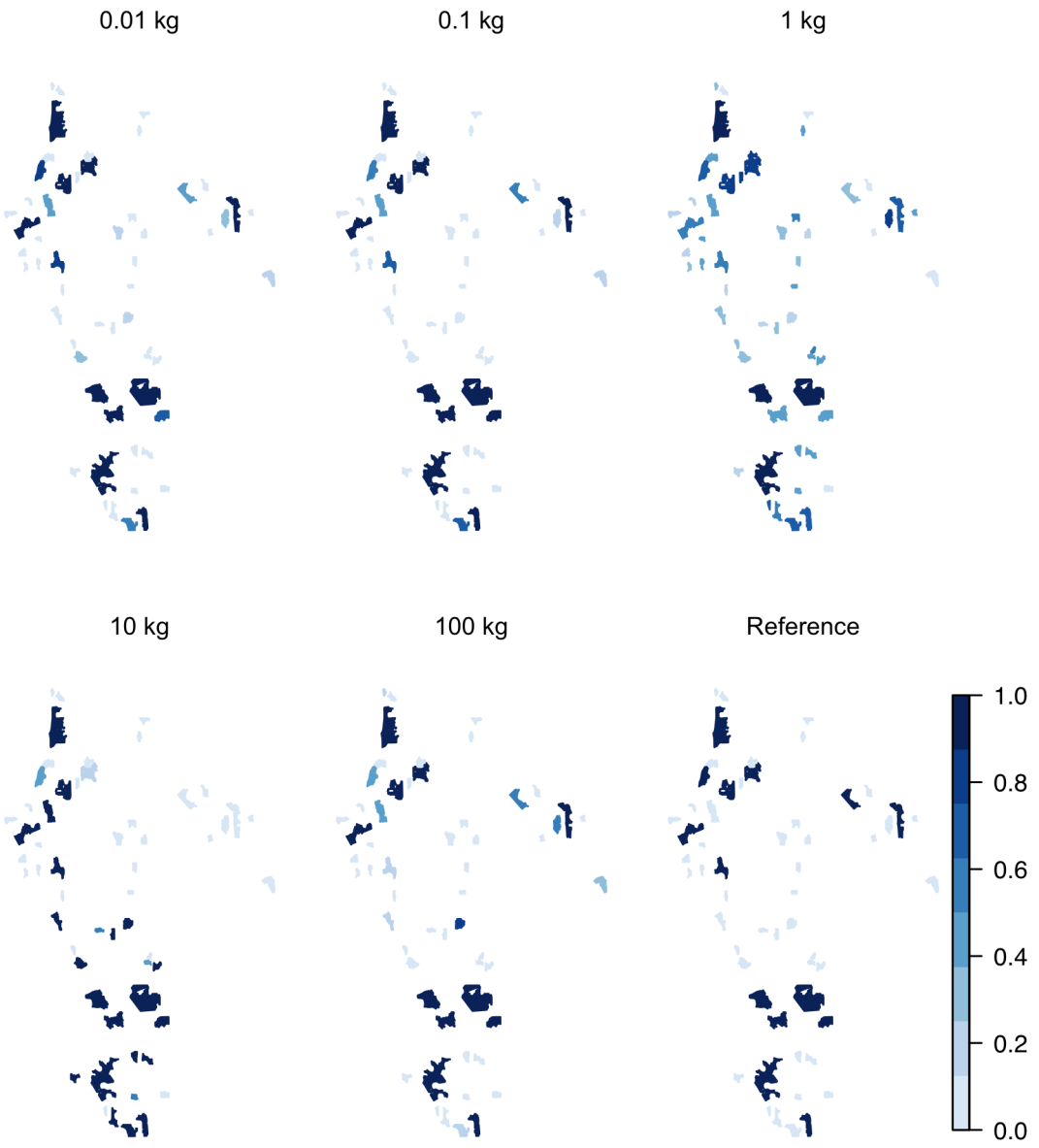

Figure 2.A4: As in Fig. 2.2, but with all inter-patch distances increased artificially by a factor of five (distances not shown to scale). As predicted by the match between characteristic range and inter-patch distances, the reserve network selected for the 10 kg species now demonstrates a clustered design similar to that of the reserve network formerly selected for the 1 kg species, and the network selected for the 1 kg species is now more similar to the reference network.

Chapter 3

Beyond the species-area relationship: improved estimates of extinction risks and rates following habitat loss and climate change

Abstract

The species-area relationship is often used to predict extinction rates following habitat loss when detailed species-level data are not available. The species-area relationship, however, only predicts the expected number of extinctions without an associated measure of uncertainty, assumes a species is protected if only one individual of a species remains, and does not address extinction risk for individual species, all of which may limit its utility for informing decisions in real landscapes. Here, I use two well-studied ecological distributions, the species abundance and species-level spatial abundance distributions, to derive two new extinction metrics that address these shortcomings: the extinction-area relationship, the probability of a single species experiencing extinction following habitat loss or range contraction, and the probabilistic species-area relationship, the probability of any number of species, from zero to the original complement, remaining in a landscape following habitat loss. I present explicit equations for these metrics based on predictions from a recently proposed maximum entropy theory of ecology and provide methods for parameterizing both metrics using types of limited empirical data that are commonly available in real landscapes. Examination of the behavior of the two metrics demonstrates that extinction predictions are strongly influenced by the choice of the minimum abundance required for a species to be considered protected, a parameter that cannot be adjusted in the classic species-area relationship. Given the importance of this parameter and of quantifying the uncertainty of extinction predictions, ecologists and conservation biologists may wish to consider using the two new metrics presented here in place of the classic species-area relationship when estimating extinction risks and rates in poorly studied landscapes.

Introduction

Habitat loss and climate change represent two of the most significant ongoing threats to global biodiversity [1]. Many approaches have been developed to assess the expected levels of species loss within a landscape due to these threats, including niche-based species distribution models [2, 3], dynamic population viability analyses [4–6], and static conservation planning algorithms [7–12]. These approaches, however, all share in common the need for empirical data such as maps of species' presence and absence across a landscape or detailed knowledge of species' life history traits.

In the absence of such data, extinction estimates are often based on applications of the species-area relationship, a metric that predicts the scaling of species richness with habitat area [13, 14]. Although many mathematical forms of the species-area relationship have been proposed [15–17], a power law, $S = cA^z$, with a slope z near 0.25 is often used for predicting large-scale losses of biodiversity [18–20] and, at smaller spatial scales, as an input into optimization methods used to design networks of protected areas to minimize biodiversity loss [21–25].

The utility of the species-area relationship for estimating extinction rates, however, has been called into question. Many critiques have centered on the accuracy of the assumptions used to select particular forms of the relationship [26–28], with several specifically noting that the assumption of a power law form for the species-area relationship is not often valid [14, 29]. However, several additional limitations inherent to the species-area relationship itself may also limit its usefulness for real-world conservation applications.

First, and perhaps most importantly, the classic species-area relationship provides an estimate only of the expected number of species lost following a reduction in available habitat, without information on uncertainty in this estimate. In practice, an estimate of 50% species loss, for example, may have very different meaning if losses of 49% to 51% are nearly guaranteed versus if losses from 10% to 90% are all plausible. A more useful metric would thus estimate not only the expected level of species loss following habitat loss but also a probability of any level of loss, ranging from zero to all species initially present. Such a measure would allow the predictions of the species-area relationship to be used in a risk assessment framework, similar to population viability analysis [4, 6], in which managers could set an acceptable level of risk of loss for a given number of species. For example, a manager or policy maker may wish to protect an amount of habitat such that there is no more than a 95% chance of losing 50% or more of all species initially present in a landscape.

Second, by its definition, the species-area relationship counts a species as present in an area if at least one individual of that species is found in that area [27, 30]. In practice, however, a single individual may be insufficient to allow for the long-term persistence of a species, and it may be desirable to set a higher population threshold above which a species is considered protected. The species-area relationship itself does not allow for this adjustment.

Finally, the species-area relationship provides only an aggregate prediction of extinction rates across all species without any information about differences in extinction risk for individual species. The species-area relationship thus does not allow, for example, the extinction of a rare species to be weighted differently, or analyzed apart, from that of a common species. If available, species-level metrics similar to the species-area relationship could be used for the assessment of extinction risk for individual species of conservation concern when insufficient data are available for a fully informed population model. Such metrics would be particularly useful for evaluating the impacts of climate change, as they would allow for an estimate of a species' extinction risk following a projected level of range contraction [19].

To address these three shortcomings, I here present a probabilistic framework for deriving two new ecological metrics that are similar to the species-area relationship but provide richer estimates of extinction risks and rates. The first metric, which applies at the species level, is the extinction-area relationship, $\varepsilon(a)$, which gives the probability of an individual species experiencing extinction following a given level of habitat loss or range contraction. The second metric, which applies at the community level, is the probabilistic species-area relationship, $\rho(S)$, which gives the probability of a given number of species being protected in a landscape, from zero to the initial number present, following habitat loss.

These two new metrics can be derived from two well-studied ecological distributions, the species abundance distribution and the species-level spatial abundance distribution. Using predictions for these two underlying distributions from a recently proposed maximum entropy theory of ecology [30, 31], I present equations for these metrics and demonstrate how both metrics can be parameterized using limited empirical data of the type often available in real landscapes. Examination of the behavior of these two metrics demonstrates that the choice of the minimum abundance required for a species to be considered protected has a significant influence on extinction predictions. In light of these results, ecologists and conservation biologists may wish to consider using these two new metrics in place of the classic species-area relationship when estimating extinction risks and rates in poorly studied landscapes.

Methods

Two underlying distributions

Consider a single landscape or species' range of original area A_0 that contracts to area A following habitat loss or climate change, and define a as the fraction of habitat that remains, $a = A/A_0$. Define n_0 for any species as that species' original abundance in area A_0, and S_0 as the number of species originally present in A_0. Finally, define a minimum critical abundance, n_c, as the minimum population size that must remain after habitat loss for a species to be considered protected. Biologically, n_c can be considered to be the critical population size below which deterministic Allee effects or demographic stochasticity may

be expected to produce to a very high probability of extinction. The parameter n_c is thus philosophically equivalent to the quasi-extinction thresholds or minimum viable population sizes that are associated with demographic extinction analyses [4, 6, 32–34].

Within this context, the overall goal of this analysis is to derive explicit equations for two new extinction metrics, the extinction-area relationship, $\varepsilon(a)$, and the probabilistic species-area relationship, $\rho(s)$. These metrics are derived using a static sampling approach that, similar to previous derivations of the classic species-area relationship [35, 36], is based on two underlying probability distributions: the species abundance distribution, $\phi(n)$, and the species-level spatial abundance distribution, $\Pi(n)$. This statistical approach does not explicitly incorporate many of the features of empirical landscapes and communities, including habitat heterogeneity, population dynamics, and other ecological mechanisms, but instead represents these to the extent that they combine to produce particular forms of these two underlying distributions.

The first underlying distribution is the species abundance distribution, the probability that a randomly selected species from a community has abundance n, $\phi(n|\boldsymbol{\theta}_\phi)$, where $\boldsymbol{\theta}_\phi$ represents parameters describing the shape of the distribution. Many forms for the species abundance distribution have been described [35, 37, 38], with perhaps the two most common being the logseries distribution [39] and the lognormal distribution [40, 41].

The second underlying distribution is the species-level spatial abundance distribution, the probability that a randomly selected area A within a large landscape of area A_0 will contain n individuals of a given species that had n_0 individuals present in A_0, $\Pi(n|a, n_0, \boldsymbol{\theta}_\Pi)$, where $\boldsymbol{\theta}_\delta$ represents additional parameters describing the shape of the distribution. An associated cumulative distribution function, $\Pi(n \leq n_c|a, n_0, \boldsymbol{\theta}_\Pi)$, giving the probability that a randomly selected area A will contain n_c or fewer individuals of a species, can be calculated as the sum of $\Pi(n)$ from zero to n_c. Many distributions have been proposed to describe $\Pi(n)$ [30, 36, 42–44], including most prominently the Poisson or binomial distributions, corresponding to the hypothesis that individuals of a species are randomly placed in space [45] and the negative binomial distribution, which allows for variable levels of species' aggregation [35, 46–48].

In all analysis that follows, two key assumptions are made about the application of these two distributions to species in real landscapes. First, the abundances of all species present in a landscape are assumed to represent random draws from the species abundance distribution. Second, the spatial distributions of species are assumed to be independent of each other, so that the presence or absence of one species in an area A does not affect the presence or absence of other species in that area. While this latter assumption is known to be incorrect in many real landscapes, many major theories in spatial macroecology share these assumptions and yet have been widely successful in reproducing empirically realistic patterns of species abundance and distributions [30, 49]. As successful prediction of spatial patterns, not mechanistic understanding of ecological processes, is the ultimate goal of the present work, these two assumptions are used in subsequent derivations. Extreme violations

of these assumptions, however, may lead real landscapes to deviate substantially from predictions presented here.

Extinction-area relationship

The extinction-area relationship, $\varepsilon(a|n_c, n_0, \boldsymbol{\theta})$, gives the probability that an individual species experiences extinction when its habitat or range shrinks to a fractional area $a = A/A_0$, where A_0 is the original area inhabited by the species and A is the area remaining after habitat loss or range contraction. In addition to a, this metric also depends on n_c, the critical minimum abundance at or below which a species is considered extinct, n_0, the original abundance of the species in A_0, and potentially a set of additional parameters $\boldsymbol{\theta}$. Like the classic species-area relationship, $\varepsilon(a)$ is not a probability distribution, as the sum of extinction probability across all a need not sum to one.

In terms of the underlying distributions, $\varepsilon(a|n_c, n_0, \boldsymbol{\theta})$ is given simply by the cumulative species-level spatial abundance distribution, $\Pi(n \leq n_c|a, n_0, \boldsymbol{\theta}_\Pi)$.

$$\varepsilon(a|n_c, n_0, \boldsymbol{\theta}_\Pi) = \Pi(n \leq n_c|a, n_0, \boldsymbol{\theta}_\Pi) \tag{3.1}$$

The value of n_c need not be constant across species but could instead be a function of n_0 or additional species-specific parameters. For example, if a certain fraction f of a species' original population must remain following habitat loss for that species to be protected, n_c may be set equal to fn_0.

If n_c, n_0, and the parameters $\boldsymbol{\theta}_\Pi$ are known, Eq. (3.1) may be used directly to calculate the extinction-area relationship. Frequently, however, n_0 and $\boldsymbol{\theta}_\Pi$ may not be known, but repeated samples of a species' abundance, $\mathbf{n}_m$, in several randomly selected small plots in A_0, each of area A_m, may be available. Using these samples, maximum likelihood estimates and confidence intervals for the parameters n_0 and $\boldsymbol{\theta}_\Pi$ may be obtained using Eq. (3.2).

$$\mathcal{L}(n_0, \boldsymbol{\theta}_\Pi|\mathbf{n}_m) = \prod_{i=0}^{k} \Pi(n_i|a_m, n_0, \boldsymbol{\theta}_\Pi) \tag{3.2}$$

With the resulting maximum likelihood estimates and confidence intervals for the parameters n_0 and $\boldsymbol{\theta}_\Pi$, a maximum likelihood estimate and confidence interval for $\varepsilon(a)$ can be obtained by translating associated values of n_0 and $\boldsymbol{\theta}_\Pi$ into values of $\varepsilon(a)$ using Eq. (3.1) with the known values of a_m and n_c [50].

Probabilistic species-area relationship

The probabilistic species-area relationship, $\rho(S|S_0, a, \boldsymbol{\theta})$, gives the probability of S species being protected in a landscape after habitat loss reduces available area to $a = A/A_0$. In

addition to a, $\rho(S)$ depends on S_0, the number of species originally present in area A_0 prior to habitat loss, and potentially on additional parameters $\boldsymbol{\theta}$. $\rho(n)$ is a true probability distribution, with the probability of all values of S from 0 to S_0 summing to one. The expected value of $\rho(S)$ for any value of a gives the expected number of surviving species after habitat loss, similar to the prediction of the classic species-area relationship.

In the event that the species abundance distribution for species present in area A_0, $\phi(n_0)$, is known or can be inferred from theory, the probability that a single randomly selected species found in A_0 will be protected after habitat loss can be calculated by combining the species abundance distribution with the cumulative species-level spatial abundance distribution.

$$g(a, n_c, \boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\Pi) = \sum_{n_0=1}^{\infty} [1 - \Pi(n \leq n_c | a, n_0, \boldsymbol{\theta}_\Pi)]\, \phi(n_0 | \boldsymbol{\theta}_\phi) \tag{3.3}$$

The value of $g(a, n_c, \boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\Pi)$ gives the probability that a randomly selected species from S_0 is found in area A. This probability depends on the fraction of original area remaining after habitat loss, a, the minimum critical abundance at or below which a species is considered protected, n_c, and any additional parameters $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\Pi$ needed to describe the shapes of the species abundance and species-level spatial abundance distributions. If the shape of the distribution $\Pi(n \leq n_c)$ varies with species abundance, n_0, as would occur, for example, if more or less abundant species had systematically more or less aggregated spatial distributions [42], the parameters $\boldsymbol{\theta}_\Pi$ may themselves be a function of n_0 or other variables.

If S_0 is known, the abundance of each species in A_0 can be considered to be a random draw from $\phi(n_0)$, and interspecific spatial patterns are independent (see *Methods: Two underlying distributions*), the distribution $\rho(S)$ is the outcome of S_0 Bernoulli trials, each with probability g of success, and $\rho(S)$ is thus a binomial distribution.

$$\rho(S | S_0, a, n_c, \boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\Pi) = \rho(S | S_0, g) = \binom{S_0}{S} g^S (1 - g)^{S_0 - S} \tag{3.4}$$

The mean of $\rho(S)$ is $S_0 g$ and the variance of $\rho(S)$ is $S_0 g(1 - g)$.

For conservation applications, it is reasonable to assume that the values of S_0 and a will be known, as these data are also required to use the classic species-area relationship for extinction prediction. In the event that the parameters $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\Pi$ are also known, in which case the species abundance and species-level spatial abundance distributions are specified exactly, Eqs. (3.3) and (3.4) can be used directly to calculate $\rho(S)$.

Although the requirements of prior knowledge of $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\Pi$ may appear limiting, under the maximum entropy theory predictions for $\phi(n_0)$ and $\Pi(n)$, $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\Pi$ represent only a single unknown parameter μ_ϕ. In the special case in which only one parameter is unknown, this parameter can be inferred conveniently from repeated measurements of the number of species found in small measurement plots of area A_m. A maximum likelihood estimate of μ_ϕ, along with confidence intervals, can be calculated by first calculating a likelihood profile for the parameter g_m, the probability that a randomly selected species from S_0 is found in area a_m, using Eq. (3.4).

$$\mathcal{L}(g_\mathrm{m}|\mathbf{S}_\mathrm{m}) = \prod_{i=0}^{k} \rho(S_i|S_0, g_\mathrm{m}) \tag{3.5}$$

Because, in this scenario, any value of g is uniquely specified by a, n_c, and the single unknown μ_ϕ, a likelihood profile for g_m can be translated directly into a likelihood profile for μ_ϕ using Eq. (3.3) with $a = a_\mathrm{m}$ and $n_\mathrm{c} = 0$. A maximum likelihood estimate for $\rho(S)$ at any area a can then be obtained by evaluating Eqs. (3.3) and (3.4) with the maximum likelihood estimate of μ_ϕ. A confidence interval for the distribution $\rho(S)$ can also be calculated by evaluating $\rho(S)$ at all values of μ_ϕ within a known confidence interval calculated from the likelihood profile. In using this procedure, the measurement area A_m may be different from, and in most practical cases will be substantially smaller than, areas A and A_0.

In the special case that the exact abundance of every individual species in A_0 is known instead of the distribution $\phi(n_0)$, a different procedure for estimating $\rho(S)$ may be used. In this case, the probability of each individual species being protected in an area A can be calculated directly using Eq. (3.1). Presuming statistical independence of all species, $\rho(S)$ will then be described by a Poisson-binomial distribution, a generalization of the binomial distribution that allows for variable probabilities of success in each Bernoulli trial [51, 52]. As these data are very unlikely to be available in practice, this calculation method is not pursued further.

Maximum entropy theory of ecology

Although the framework described above can be used with any choice of distributions for $\phi(n_0)$ and $\Pi(n)$, this study uses the forms of these distributions that are predicted by a recently proposed maximum entropy theory of ecology [30, 31]. In this theory, probability distributions such as $\phi(n_0)$ and $\Pi(n)$ are derived by maximizing the information entropy of the distribution in question, subject to a set of known constraints. This procedure provides the "least-biased" distribution possible, or the distribution that assumes the least information, given the constraints [53]. The maximum entropy theory predictions of the $\phi(n_0)$ and $\Pi(n)$ distributions have been extensively tested and found to fit a wide range of empirical data sets [29–31, 54].

The maximum entropy theory of ecology predicts that the species abundance distribution takes the form of a logseries distribution, which depends on only one parameter, p. The parameter p can be related to the mean of the distribution μ_ϕ, equal to the average number of individuals per species [39].

$$\phi(n_0|\mu_\phi) = \frac{-1}{\log(1-p)}\frac{p^{n_0}}{n_0} \tag{3.6}$$

$$\mu_\phi = \frac{-1}{\log(1-p)}\frac{p}{(1-p)} \tag{3.7}$$

Although the original maximum entropy theory of ecology [30, 31] proposes an upper-truncated variant of the logseries distribution, the classic logseries distribution is used here for mathematical simplicity, recognizing that as the total number of individual organisms across all species becomes large relative to the mean number of individuals per species, the difference between the truncated and non-truncated distributions becomes negligible.

The maximum entropy theory of ecology predicts that the species-level spatial abundance distribution takes the form of an upper-truncated geometric distribution, which depends only on the parameters a and n_0 [30, 31].

$$\Pi(n|a, n_0) = \frac{(1-q)q^n}{1-q^{n_0+1}} \text{ for } 0 \leq n \leq n_0, \text{ otherwise } 0 \tag{3.8}$$

$$an_0 = \frac{q}{1-q} - \frac{(n_0+1)q^{n_0+1}}{1-q^{n_0+1}} \tag{3.9}$$

The cumulative distribution function $\Pi(n \leq n_c)$ is the sum from zero to n_c of $\Pi(n)$.

$$\Pi(n \leq n_c|a, n_0) = \frac{q^{n+1}-1}{q^{n_0+1}-1} \text{ for } 0 \leq n \leq n_0 \tag{3.10}$$

Although, as previously noted, any two distributions may be used to represent $\phi(n_0)$ and $\Pi(n)$, the use of Eqs. (3.6) and (3.8) for subsequent analysis is well supported by extensive empirical testing [29–31, 54]. These equations arguably represent effective "null models" for extinction prediction in the absence of data suggesting that other distributions would be more appropriate. Due to the implicit equations needed to calculate the parameters p and q from μ_ϕ, a, and n_0, equations for the extinction-area relationship and probabilistic species-area relationship that are based on Eqs. (3.6) and (3.8) do not have simple closed form solutions and must be evaluated numerically.

Results

Extinction-area relationship

Using predictions of the maximum entropy theory of ecology, the extinction-area relationship for a species with known initial abundance n_0 is fully specified by n_0, a, the fraction of original area remaining after habitat loss, and n_c, the critical minimum abundance at or below which a species is considered extinct.

$$\varepsilon(a|n_c, n_0) = \frac{q^{n_c+1} - 1}{q^{n_0+1} - 1} \tag{3.11}$$

The parameter q in Eq. (3.11) can be calculated from a and n_0 using Eq. (3.9).

Figure (3.1) shows the extinction-area relationship for species with different initial abundances and for assumptions of different minimum critical abundances. For all initial abundances, extinction risk initially falls rapidly as a increases from zero. With $n_c = 0$, implying that the presence of a single individual is sufficient for a species to be protected, species with an initial abundance of greater than approximately 1000 individuals in A_0 are very likely to be protected at all but the highest levels of habitat loss (Fig. 3.1a).

Increasing n_c, however, substantially raises a species' probability of extinction. Using a critical abundance of $n_c = 10$ instead of $n_c = 0$, for example, raises the extinction risk for a species with $n_0 = 100$ to approximately that of a species with $n_0 = 10$ (Fig. 3.1b). With $n_c = 0$, only 1% of A_0 must be protected for a species with $n_0 = 100$ to have more than a 50% probability of survival. However, at $n_c = 10$, over 15% of A_0 must be protected to achieve this same probability of protection.

If n_0 for a species is not known, but samples of the species' abundance, $\mathbf{n}_m$, at area A_m are available, a maximum likelihood estimate of n_0 can be calculated by maximizing Eq. (3.12) with respect to n_0.

$$\mathcal{L}(n_0|\mathbf{n}_m) = \prod_{i=0}^{k} \frac{(1-q)q^{n_i}}{1 - q^{n_0+1}} \tag{3.12}$$

A 95% confidence interval for n_0, based on the critical value of a one-tailed Chi-squared test with one degree of freedom, can be found by locating values of n_0 above and below the maximized likelihood, $\hat{\mathcal{L}}$, such that $-\log(\mathcal{L}) = -\log(\hat{\mathcal{L}}) + 1.92$ [50]. The extinction-area relationship for the maximum likelihood estimate of n_0, and for lower and upper bounds of the confidence interval for n_0, can then be calculated using Eq. (3.11).

Figure (3.2) gives examples of two maximum likelihood estimates of $\varepsilon(a)$, with confidence intervals, for simulated samples consisting of 5 and 50 plots in which a species' abundance

is measured. A measurement area $a_{\mathrm{m}} = 0.001$, an initial abundance $n_0 = 1000$, and a critical abundance $n_{\mathrm{c}} = 0$ are used for these illustrations.

Probabilistic species-area relationship

Using predictions of the maximum entropy theory of ecology, the probability of S species remaining in a landscape after habitat loss can be calculated from knowledge of four parameters: a, the fraction of habitat remaining after habitat loss, n_{c}, the critical abundance at or below which a species is considered extinct, S_0, the total number of species in the original landscape, and μ_ϕ, the mean number of individuals per species in the original landscape.

In the event that all four of these parameters are known or can be reasonably assumed, the probability of S species remaining in A is given by Eqs. (3.13) and (3.14), with p and q calculated from Eqs. (3.7) and (3.9).

$$\rho(S|S_0, g) = \binom{S_0}{S} g^S (1-g)^{S_0 - S} \tag{3.13}$$

$$g(a, n_{\mathrm{c}}, \mu_\phi) = \sum_{n_0=1}^{\infty} \left[\left(1 - \frac{q^{n_{\mathrm{c}}+1} - 1}{q^{n_0+1} - 1} \right) \left(\frac{-p^{n_0}}{n_0 \log(1-p)} \right) \right] \tag{3.14}$$

The summation in Eq. (3.14) has no apparent closed form solution and must be evaluated numerically. However, if $n_{\mathrm{c}} = 0$ and $A \ll A_0$, the first term in the summation is well-approximated by $an_0/(an_0 + 1)$ [30], leading to the simplified Eq. (3.15).

$$g(a, n_{\mathrm{c}}, \mu_\phi) = -\frac{a}{\log(1-p)} \sum_{n_0=1}^{\infty} \frac{p^{n_0}}{an_0 + 1} \tag{3.15}$$

Figure (3.3) shows the probabilistic species-area relationship for different levels of a and n_{c}. In all calculations, $S_0 = 100$ and $\mu_\phi = 1{,}000$. As expected, the number of species protected decreases as a decreases, with the distribution $\rho(S)$ taking the form of a binomial distribution. The variance of this distribution is maximal at an expected value of S equal to one half S_0. As with the extinction-area relationship, the parameter n_{c} has a significant influence on the distribution $\rho(S)$. At $a = 0.1$, for example, changing n_{c} from zero to ten has approximately the same effect on extinction rates as decreasing the remaining habitat area A by an order of magnitude.

In applications of the probabilistic species-area relationship, the values of S_0, a, and n_{c} are likely to be known or can be assumed. The parameter μ_ϕ, the mean of the species abundance distribution $\phi(n_0)$, however, will likely need to be inferred from empirical data.

Given knowledge of S_0, and recognizing that μ_ϕ can be estimated by N_0/S_0, where N_0 is the total number of individuals of all species present in A_0, a simple method for inferring μ_ϕ rests on the empirical observation that the number of individuals of all species in a community generally scales linearly with area [30, 38]. As such, the total number of individuals N_m in several small plots of area A_m may be measured and averaged, giving $\bar{N}_\mathrm{m}$, with which an estimate of μ_ϕ can be obtained from Eq. (3.16).

$$\hat{\mu}_\phi = \frac{N_0}{S_0} = \frac{\bar{N}_\mathrm{m} A_0}{A_\mathrm{m} S_0} \tag{3.16}$$

While this approach ignores uncertainty in the measurement of N_m, this estimate may be sufficiently accurate for many landscapes in which the total abundance of all species is believed to scale linearly with area.

For some landscapes and communities, however, a count of individual organisms may be difficult to obtain, but counts of the number of species found in many small census plots of area A_m, $\mathbf{S}_\mathrm{m}$, may be available. In this event, a maximum likelihood estimate for the parameter g_m, the probability that a randomly selected species from S_0 is found in an area A_m, can be calculated using Eq. (3.17).

$$\mathcal{L}(g_\mathrm{m}|\mathbf{S}_\mathrm{m}) = \prod_{i=0}^{k} \binom{S_0}{S_i} g_\mathrm{m}^{S_i} (1 - g_\mathrm{m})^{S_0 - S_i} \tag{3.17}$$

Presuming $A_\mathrm{m} \ll A_0$, Eqs. (3.15) and (3.7) can be used with $a_\mathrm{m} = A_\mathrm{m}/A_0$ and $n_\mathrm{c} = 0$ to translate each value of g_m into a corresponding value of μ_ϕ, giving $\mathcal{L}(\mu_\phi|\mathbf{S}_\mathrm{m})$. Maximizing this likelihood function gives a central estimate of μ_ϕ, given the sampled data. Similar to the extinction-area relationship, a 95% confidence interval for μ_ϕ can be found by locating values of μ_ϕ above and below the maximum likelihood estimate such that $-\log(\mathcal{L}) = -\log(\hat{\mathcal{L}}) + 1.92$, where $\hat{\mathcal{L}}$ is the maximized likelihood value. A confidence interval for each value of S can then be found by calculating $\rho(S)$ for many values of μ_ϕ within its confidence interval and finding the maximum and minimum values of $\rho(S)$ for each S.

Figure (3.4) gives an example of three confidence intervals estimated for $\rho(S)$ at three different levels of a, given a sample of 10 small measurement plots in which S is counted. In all examples, $a_\mathrm{m} = 0.00001$, $S_0 = 100$, and $\mu_\phi = 10{,}000$.

Discussion

To address several shortcomings inherent in applying the classic species-area relationship to extinction prediction, this chapter has proposed two new metrics for estimating extinction

risks and rates, the extinction-area relationship and the probabilistic species-area relationship. The extinction-area relationship provides estimates of single-species extinction risk as a function of the level of habitat contraction, a result that may be particularly useful when applied in concert with species distribution models that predict changes in species' ranges following climate change. The probabilistic species-area relationship provides the probability of any number of species being protected, from zero to the original number present, following habitat loss, a distribution that allows predictions of extinction rates to be incorporated into a risk management framework for setting habitat protection targets.

Both of these extinction metrics can be derived from two well-studied underlying ecological distributions, the species abundance distribution and species-level spatial abundance distribution. By using the forms of these underlying distributions predicted by the maximum entropy theory of ecology, explicit equations for both extinction metrics can be derived. Using likelihood-based inference methods, both metrics can be parameterized using limited empirical data of the type that is often available in real landscapes.

The ability to specify a minimum critical abundance at or below which a species is considered extinct, n_c, is a particularly important contribution of the two proposed metrics. The results presented here show that even a relatively small increase in n_c, from 0 to 10, for example, can lead to substantial changes in the probability of a species' extinction, even for relatively abundant species. If a very low probability of extinction is desired for a given species, even a small increase in n_c may require dramatically more area to be protected. These findings suggest that the appropriate level for the minimum critical abundance parameter should be considered carefully in future applications of ecological metrics to extinction estimation.

Although the specific results of this analysis are based on the predictions of the maximum entropy theory of ecology, the general framework presented here can and should be used to derive the extinction-area relationship and probabilistic species-area relationship from other choices of underlying distributions, as different underlying distributions could potentially lead to substantially different extinction predictions. For example, the logseries species abundance distribution predicts a relatively large number of rare species (the mode of the logseries is always at one individual per species), whereas other "humped" distributions such as the lognormal [40, 41] or zero-sum multinomial [38, 55] predict relatively less rarity for any value of μ_ϕ. As extinction risk decreases rapidly as a species' abundance increases, calculations of the probabilistic species-area relationship using the logseries species abundance distribution will tend to estimate higher extinction rates than estimates made using these other distributions.

With regard to the species-level spatial abundance distribution, more aggregated spatial distributions will tend to increase the probability of extinction when n_c is small. Consider, for example, the case in which an original landscape of area A_0 is divided into 10 plots, only one of which will remain after habitat loss ($a = 0.1$), and $n_c = 0$, such that every plot in which a species is present can protect that species. If a species' distribution is

minimally aggregated, such that it appears in every plot, the probability of extinction will be zero, while if a species is maximally aggregated, such that all individuals of that species appear in a single plot, the probability of extinction will be 90%. In this case, increasing aggregation clearly increases the probability of extinction. Compared to the truncated geometric distribution used here, a binomial, or random placement, spatial abundance distribution will thus lead to lower extinction risk, as will a negative binomial spatial abundance distribution with a clustering parameter $k > 1$ [35, 42, 47, 48]. Conversely, negative binomial spatial abundance distributions with $0 < k < 1$ will lead to higher predicted extinction risks.

This positive relationship between aggregation and extinction risk, however, applies only when n_c is small. Consider now the same landscape, but with $n_c = n_0$, such that all individuals of a species must remain after habitat loss for a species to be protected. In this case, maximal aggregation, in which all individuals are found in one of the 10 plots, gives the species the same 90% probability of extinction, but if a species is less aggregated, such that individuals are spread out across multiple plots, then extinction is certain. The choice of n_c thus has the potential to affect not only absolute extinction risk but the relative risk associated with different types of spatial patterns.

Several limitations of this statistical framework and the equations presented here are important to note. First, this approach does not account explicitly for the geometry of habitat loss [56–59]. Within the sampling framework used here, the species-level spatial abundance distribution gives the number of individuals of a species that are expected to remain in an area A following habitat loss. To the extent that any given spatial abundance distribution is believed to apply only to regularly shaped habitats, such as rectangles or circles, the resulting predictions of both extinction metrics will also apply only when the area A remaining after habitat loss takes these shapes. In the event that the spatial abundance distribution depends on shape and habitat is instead lost in a regular shape, metrics similar to the "endemics-area relationship" [27, 48, 57, 60] may be more applicable and can be derived using the same general methods described here.

Second, and relatedly, this approach applies to the generic case in which suitable habitat area shrinks from A_0 to A. The case in which several patches of habitat remain after loss is not explicitly treated here. In this scenario, two adjustments may be needed to improve the predictions of the two extinction metrics presented here. First, information about the overlap in species occurrence within remnant patches (i.e., beta diversity) could be used in concert with the species-level spatial abundance distribution, applied to each individual patch, to determine the landscape-wide forms of both extinction metrics. Second, a species-level spatial abundance distribution that itself accounts for the number and configuration of remnant habitat patches could also be used. In general, the prediction of spatial pattern in multi-patch landscapes remains an ongoing area of research [e.g., 30, 61–64], and future developments in this area will improve the accuracy of these extinction metrics when used in complex, fragmented landscapes.

Third, as previously noted, factors such as habitat heterogeneity and ecological mechanisms, such as competition, mutualism, predation, and others, are incorporated in this approach only insofar as they affect the shapes of the underlying species abundance and species-level spatial abundance distributions. Understanding how these factors combine to produce known forms of these underlying distributions is thus critical to understanding how the proposed extinction metrics themselves are influenced by these factors. For example, while events such as the extinction of a predator or competitor may tend to reduce a target species' extinction risk, extinction of a mutualist may have the opposite effect. The potential for dynamic community-wide phenomena, such as coextinction, trophic cascades, and extinction cascades, adds further uncertainty to the prediction of extinction rates using the static statistical approach presented here [65–67].

Fourth, the sampling framework proposed here does not explicitly account for the temporal dimension of extinction processes. In particular, the possibility of extinction debt [68–71], time delayed extinctions that occur following habitat loss, is not directly addressed, and thus the extinction metrics proposed here may ultimately understate the total number of extinctions that follow habitat loss and climate change. This concern, however, may be partially addressed by the ability to select a minimum critical abundance for protection, which could be set to a population size that is presumed to have a negligible probability of subsequent extinction within a target time horizon.

The two new metrics presented here, the extinction-area relationship and the probabilistic species-area relationship, provide, in explicitly probabilistic terms, a new approach for predicting extinction risks and rates in landscapes undergoing habitat loss and climate change. The sampling framework presented here demonstrates that both metrics can be derived from two well-studied ecological distributions, the species abundance distribution and species-level spatial abundance distribution, and that both metrics can be parameterized using limited empirical data. While further research may improve on the specific results in this chapter, I suggest that in the absence of data to the contrary, the equations presented here, based on the predictions of an extensively tested ecological theory, may be useful for generating rich and useful first-pass predictions of extinction risks and rates in landscapes that are experiencing habitat loss and climate change and where little is known about resident species.

Acknowledgements

I thank John Harte, William Lidicker, Jr., and Adina Merenlender for helpful comments on drafts of this chapter.

References

[1] Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis* (Island Press, Washington, D.C.), p 137.

[2] Elith J, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.

[3] Franklin J (2010) *Mapping Species Distributions: Spatial Inference and Prediction* (Cambridge University Press, Cambridge, United Kingdom), p 338.

[4] Beissinger SR, McCullough DR (2002) *Population Viability Analysis* (University of Chicago Press, Chicago, IL), p 577.

[5] Carroll C, Noss RF, Paquet PC, Schumaker NH (2003) Use of population viability analysis and reserve selection algorithms in regional conservation plans. *Ecological Applications* 13:1773–1789.

[6] Akçakaya HR (2004) *Species Conservation and Management: Case Studies, Volume 1* (Oxford University Press, Oxford, United Kingdom), p 533.

[7] Pressey RL, Johnson IR, Wilson PD (1994) Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. *Biodiversity and Conservation* 3:242–262.

[8] McDonnell MD, Possingham HP, Ball IR, Cousins EA (2002) Mathematical methods for spatially cohesive reserve design. *Environmental Monitoring and Assessment* 7:107–114.

[9] Williams JC, ReVelle CS, Levin SA (2004) Using mathematical optimization models to design nature reserves. *Frontiers in Ecology and the Environment* 2:98–105.

[10] Sarkar S, et al. (2006) Biodiversity conservation planning rools: present status and challenges for the future. *Annual Review of Environment and Resources* 31:123–159.

[11] Moilanen A (2007) Landscape Zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation* 134:571–579.

[12] Ball I, Possingham H, Watts M (2009) in *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*, eds Moilanen A, Wilson K, Possingham H (Oxford University Press, Oxford, United Kingdom), pp 185–195.

[13] Rosenzweig ML (1995) *Species Diversity in Space and Time* (Cambridge University Press, Cambridge, United Kingdom), p 436.

[14] Drakare S, Lennon JJ, Hillebrand H (2006) The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters* 9:215–227.

[15] McGuinness KA (1984) Equations and explanations in the study of species-area curves. *Biological Reviews* 59:423–440.

[16] Tjø rve E (2003) Shapes and functions of species-area curves: a review of possible models. *Journal of Biogeography* 30:827–835.

[17] Dengler J (2009) Which function describes the species-area relationship best? A review and empirical evaluation. *Journal of Biogeography* 36:728–744.

[18] Pimm S, Russell G, Gittleman J, Brooks T (1995) The future of biodiversity. *Science* 269:347–350.

[19] Thomas CD, et al. (2004) Extinction risk from climate change. *Nature* 427:145–148.

[20] Vuuren DPV, Sala OE, Pereira HM (2006) The future of vascular plant diversity under four global scenarios. *Ecology and Society* 11:25.

[21] Wilson KA, McBride MF, Bode M, Possingham HP (2006) Prioritizing global conservation efforts. *Nature* 440:337–40.

[22] Murdoch W, et al. (2007) Maximizing return on investment in conservation. *Biological Conservation* 139:375–388.

[23] Wilson KA, et al. (2007) Conserving biodiversity efficiently: what to do, where, and when. *PLoS Biology* 5:e223.

[24] Desmet P, Cowling R (2004) Using the species-area relationship to set baseline targets for conservation. *Ecology and Society* 9:11.

[25] Koh LP, Ghazoul J (2010) A matrix-calibrated species-area model for predicting biodiversity losses due to land-use change. *Conservation Biology* 24:994–1001.

[26] May RM, Lawton JH, Stork NE (1995) in *Extinction Rates*, eds Lawton JH, May RM (Oxford University Press, Oxford, United Kingdom), pp 1–24.

[27] Smith AB (2010) Caution with curves: caveats for using the species-area relationship in conservation. *Biological Conservation* 143:555–564.

[28] Harte J, Kitzes J (2012) in *Saving a Million Species: Extinction Risk from Climate Change*, ed Hannah L (Island Press, Washington, D.C.), pp 73–86.

[29] Harte J, Smith AB, Storch D (2009) Biodiversity scales from plots to biomes with a universal species-area curve. *Ecology Letters* 12:789–797.

[30] Harte J (2011) *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* (Oxford University Press, Oxford, United Kingdom), p 264.

[31] Harte J, Zillio T, Conlisk E, Smith A (2008) Maximum entropy and the state-variable approach to macroecology. *Ecology* 89:2700–2711.

[32] Morris WF, Doak DF (2002) *Quantitative Conservation Biology: Theory and Practice of Population Viability Analysis* (Sinauer Associates, Sunderland, MA), p 480.

[33] Traill L, Bradshaw C, Brook B (2007) Minimum viable population size: a meta-analysis of 30 years of published estimates. *Biological Conservation* 139:159–166.

[34] Flather CH, Hayward GD, Beissinger SR, Stephens Pa (2011) Minimum viable populations: is there a 'magic number' for conservation practitioners? *Trends in Ecology and Evolution* 26:307–316.

[35] He F, Legendre P (2002) Species diversity patterns derived from species-area models. *Ecology* 83:1185–1198.

[36] Tjø rve E, Kunin WE, Polce C, Calf Tjørve KM (2008) Species-area relationship: separating the effects of species abundance and spatial distribution. *Journal of Ecology* 96:1141–1151.

[37] Pielou EC (1976) *Ecological Diversity* (John Wiley & Sons, New York, NY), p 176.

[38] Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, Princeton, NJ), p 375.

[39] Fisher RA, Corbet AS, Williams CB (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42–58.

[40] Preston FW (1962) The canonical distribution of commonness and rarity: part I. *Ecology* 43:185–215.

[41] May RM (1975) in *Ecology and Evolution of Communities*, eds Cody ML, Diamond JM (Belknap Press, Cambridge, MA), pp 81–120.

[42] Green JL, Plotkin JB (2007) A statistical theory for sampling species abundances. *Ecology Letters* 10:1037–1045.

[43] Haegeman B, Etienne RS (2010) Entropy maximization and the spatial distribution of species. *The American Naturalist* 175:E74–E90.

[44] Conlisk J, Conlisk E, Kassim AR, Billick I, Harte J (2012) The shape of a species' spatial abundance distribution. *Global Ecology and Biogeography*.

[45] Coleman BD (1981) On random placement and species-area relations. *Mathematical Biosciences* 54:191–215.

[46] Pielou EC (1977) *Mathematical ecology* (John Wiley & Sons, New York, NY), p 385.

[47] Zillio T, He F (2010) Modeling spatial aggregation of finite populations. *Ecology* 91:3698–3706.

[48] He F, Hubbell SP (2011) Species-area relationships always overestimate extinction rates from habitat loss. *Nature* 473:368–371.

[49] McGill BJ, Nekola JC (2010) Mechanisms in macroecology: AWOL or purloined letter? Towards a pragmatic view of mechanism. *Oikos* 119:591–603.

[50] Bolker BM (2008) *Ecological Models and Data in R* (Princeton University Press, Princeton, NJ), p 408.

[51] Nedelman J, Wallenius TED (1986) Bernoulli trials, poisson trials, surprising variances, and Jensen's Inequality. *The American Statistician* 40:286–289.

[52] Fernández M, Williams S (2010) Closed-Form expression for the poisson-binomial probability density function. *IEEE Transactions on Aerospace and Electronic Systems* 46:803–817.

[53] Jaynes ET (2003) *Probability Theory: The Logic of Sciece* (Cambridge University Press, Cambridge, United Kingdom), p 727.

[54] White EP, Thibault KM, Xiao X (2012) Characterizing species-abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology*.

[55] Volkov I, Banavar JR, Hubbell SP, Maritan A (2003) Neutral theory and relative species abundance in ecology. *Nature* 424:1035–1037.

[56] Ney-Nifle M, Mangel M (1999) Species-area curves based on geographic range and occupancy. *Journal of Theoretical Biology* 196:327–342.

[57] Kinzig AP, Harte J (2000) Implications of endemics-area relationships for estimates of species extinctions. *Ecology* 81:3305–3311.

[58] Ney-Nifle M, Mangel M (2000) Habitat loss and changes in the species-area relationship. *Conservation Biology* 14:893–898.

[59] Seabloom EW, Dobson AP, Stoms DM (2002) Extinction rates under nonrandom patterns of habitat loss. *Proceedings of the National Academy of Sciences of the United States of America* 99:11229–34.

[60] Harte J, Kinzig AP (1997) On the implications of species-area relationships for endemism, spatial rurnover, and food web patterns. *Oikos* 80:417–427.

[61] Chave J, Leigh EG (2002) A spatially explicit neutral model of β-diversity in tropical forests. *Theoretical Population Biology* 62:153–168.

[62] Soininen J, McDonald R, Hillebrand H (2007) The distance decay of similarity in ecological communities. *Ecography* 30:3–12.

[63] Morlon H, et al. (2008) A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters* 11:904–917.

[64] Kraft NJB, et al. (2011) Disentangling the drivers of diversity along latitudinal and elevational gradients. *Science* 333:1755–1758.

[65] Pace ML, Cole JJ, Carpenter SR, Kitchell JF (1999) Trophic cascades revealed in diverse ecosystems. *Trends in Ecology and Evolution* 14:483–488.

[66] Terborgh J, et al. (2001) Ecological meltdown in predator-free forest fragments. *Science* 294:1923–1926.

[67] Dunn RR, Harris NC, Colwell RK, Koh LP, Sodhi NS (2009) The sixth mass coextinction: are most endangered species parasites and mutualists? *Proceedings of the Royal Society B* 276:3037–3045.

[68] Tilman D, May R, Lehman C, Nowak M (1994) Habitat destruction and the extinction debt. *Nature* 371:65–66.

[69] Kuussaari M, et al. (2009) Extinction debt: a challenge for biodiversity conservation. *Trends in ecology & evolution* 24:564–571.

[70] Halley JM, Iwasa Y (2010) Neutral theory as a predictor of avifaunal extinctions after habitat loss. *Proceedings of the National Academy of Sciences of the United States of America* 108:2316–2321.

[71] Tanentzap AJ, Walker S, Theo Stephens RT, Lee WG (2012) A framework for predicting species extinction by linking population dynamics with habitat loss. *Conservation Letters*.

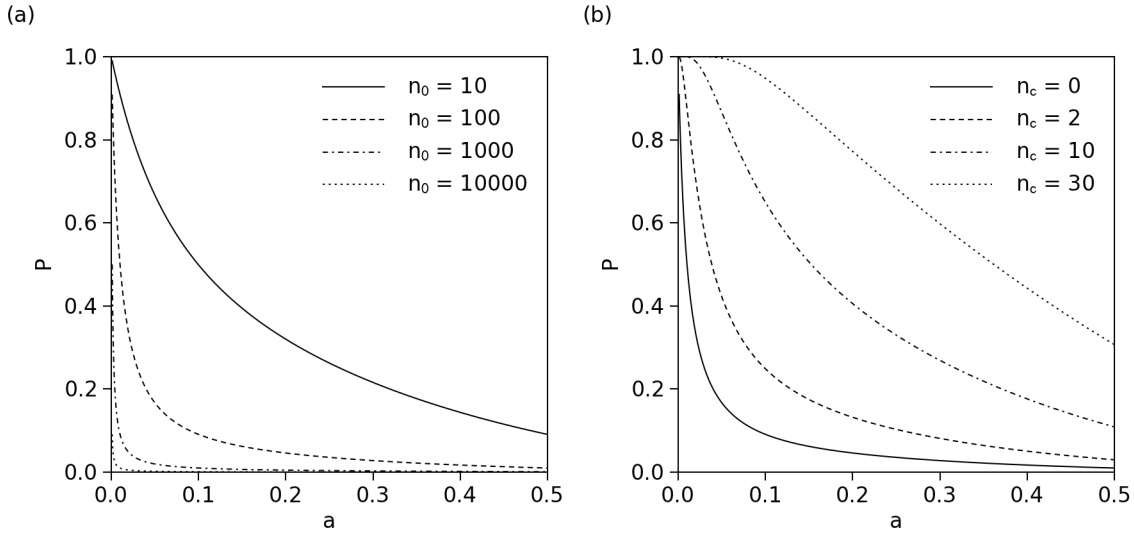

Figure 3.1: Extinction-area relationships, giving the probability, P, of species extinction as a function of the fraction of area remaining after habitat loss, a, for different values of (a) initial abundance, n_0, with critical abundance $n_c = 0$, and (b) critical abundance, n_c, for a species with initial abundance $n_0 = 100$.

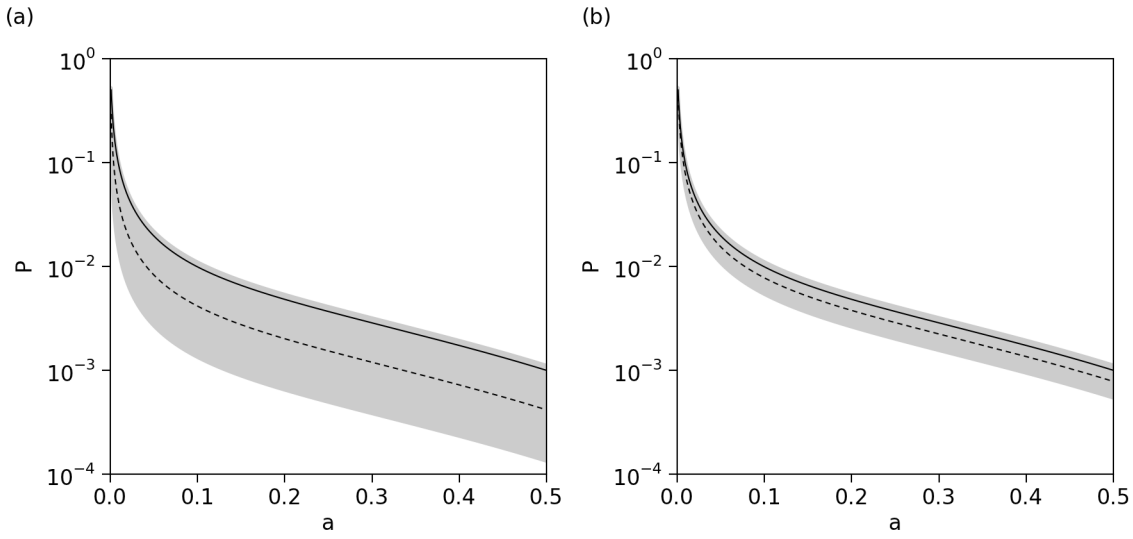

Figure 3.2: Example extinction-area relationships predicted from (a) 5 and (b) 50 simulated samples of a species' abundance in small plots of area $a_{\mathrm{m}} = 0.001$. Plots show the probability, P, of species extinction as a function of the fraction of area remaining after habitat loss, a. Estimated maximum likelihood extinction-area relationships are shown with a dashed line, with gray indicating 95% confidence intervals around this estimate. The true extinction-area relationship, based on $n_{\mathbf{0}} = 1000$, is shown in solid line. $n_{\mathrm{c}} = 0$ for all calculations.

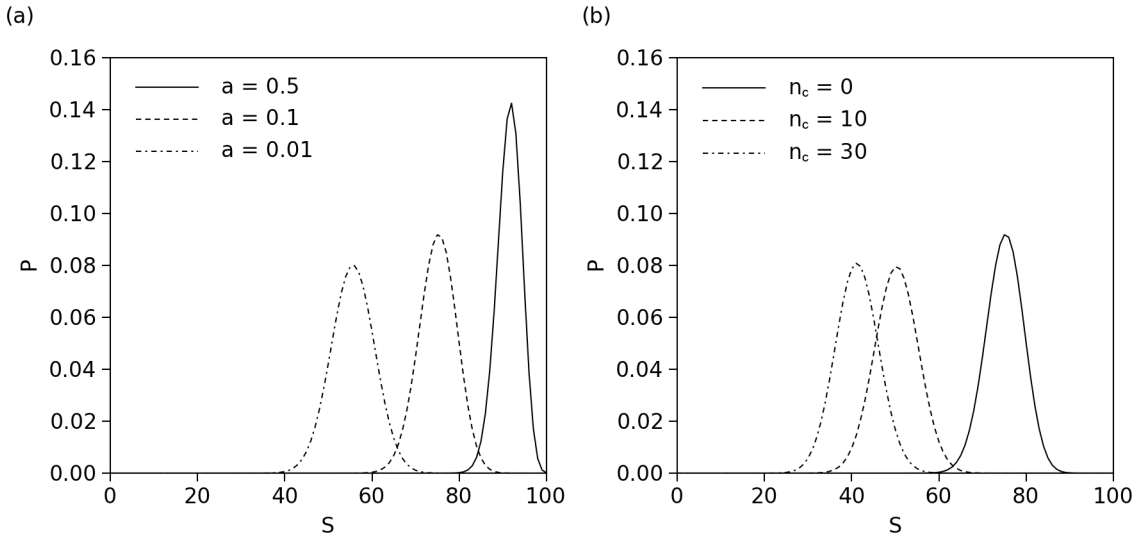

Figure 3.3: Probabilistic species-area relationships for a community with $S_0 = 100$ and $\mu_\phi = 10{,}000$. (a) The probability, P, of protecting S species at three different levels of a, with $n_c = 0$. (b) The probability, P, of protecting S species at three different levels of n_c, with $a = 0.1$

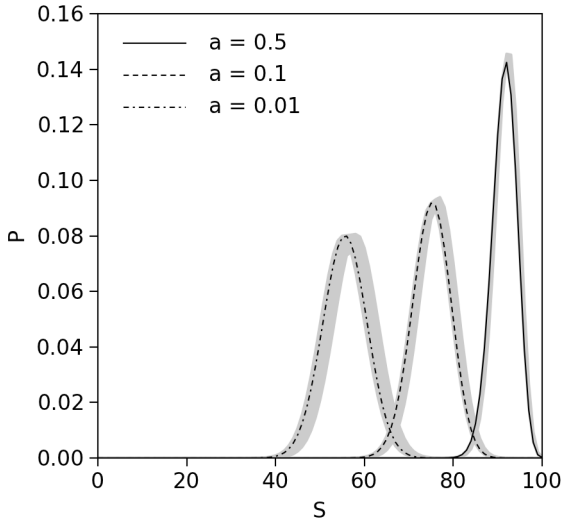

Figure 3.4: Example probabilistic species-area relationships predicted from a simulated sample of 10 presence-absence surveys of S in measurement plots of area $a_{\mathrm{m}} = 0.00001$. Plot shows the probability, P, of protecting S species at three different levels of a. Gray areas show 95% confidence interval of estimated $\rho(S)$ curves. True curves with $S_0 = 100$ and $\mu_\phi = 10,000$ are shown in black lines. $n_{\mathrm{c}} = 0$ for all calculations.

Chapter 4

California bats avoid roads

Abstract

Although the negative impacts of roads on terrestrial vertebrate and bird populations are well documented, there have been few studies of the road ecology of bats. Here, I present the results of acoustic surveys of bat activity along ten 300 m transects bordering three large highways in northern California. Nightly counts of bat passes were analyzed with generalized linear mixed models to determine the relationship between bat activity and distance from a road. Total bat activity recorded at points adjacent to roads is found to be approximately one-half of the level observed at 300 m. Statistically significant decreases in activity near roads are also found for the Brazilian free-tailed bat (*Tadarida brasiliensis*), big brown bat (*Eptesicus fuscus*), hoary bat (*Lasiurus cinereus*), and silver-haired bat (*Lasionycteris noctivagans*). This road effect is temperature dependent, with hot days both increasing total foraging activity at night and reducing the difference between activity levels near and far from roads. Decreases in bat activity in the vicinity of roads may have implications for both bat conservation and the provision of insect control services.

Introduction

Roads represent one of the most extensive human impacts on the biosphere [1, 2]. With respect to wildlife, roads are known to cause direct mortality due to vehicle collisions, serve as barriers to physical movement and gene flow across landscapes, and decrease the use of otherwise suitable habitat [3, 4]. Two recent comprehensive reviews of the road effects literature have found that most studied species showed decreased abundances or activity levels in the vicinity of large roads, with effects extending to several hundred meters from the roadside and beyond [5, 6].

The vast majority of road ecology studies, however, have examined the effects of roads on terrestrial vertebrates and birds, and there have been few studies of the effects of roads on bat populations. Several species of bats, including the endangered Indiana bat (*Myotis sodalis*), have been identified in roadkill surveys [7, 8], suggesting that vehicle collisions

may be a source of bat mortality. The presence of vehicles on a road has also been found to serve as a barrier during foraging for some species of bats [9, 10], and traffic noise has been found to reduce the foraging effectiveness of the greater mouse-eared bat (*Myotis myotis*) [11, 12]. Lower levels of total bat activity and activity of the common Pipistrelle (*Pipistrellus pipistrellus*) have also been observed within 1 km of a large road in Cumbria, United Kingdom [13].

Here, I present the results of acoustic surveys documenting the effects of three large roads on bat activity in northern California salt marshes. Consistent with previous research, decreased bat activity levels are observed near roads, with bat activity at points adjacent to roads approximately one-half of the level at control points located 300 m from a road. The four most common individual species in these landscapes also show decreased activity levels near roads. This road effect is found to be temperature dependent, with higher daily maximum temperatures both increasing total bat activity and decreasing the difference in activity between points adjacent to a road and far from a road.

Methods

Survey Locations and Methods

Acoustic surveys were conducted during August and September of 2010 and 2011 at three sites bordering the San Francisco Bay and adjacent to a major highway: Don Edwards San Francisco Bay National Wildlife Refuge (California State Road 84), Hayward Shoreline Interpretive Center (California State Road 92), and San Pablo Bay National Wildlife Refuge (California State Road 37). These sites were specifically selected due to their relatively homogeneous salt marsh habitat, which allowed transects to be established within a single, relatively uniform vegetation cover. All three roads have moderate to high traffic volumes, with annual average daily traffic of 55,000 (SR84), 86,000 (SR92), and 33,500 (SR37) vehicles in 2010 [14].

Sampling was conducted along 300 m transects running perpendicular to each highway, with acoustic bat detectors placed at points located approximately 0-20 m, 100 m, and 300 m from the edge of the road. Four transects were established at Don Edwards, four at San Pablo Bay, and two at Hayward Shoreline. The dominant vegetation cover along transects consisted of pickleweed (*Salicornia spp.*) (6 transects), invasive grasses and forbs (2 transects), and sand (2 transects). Due to equipment constraints and failures, the final sampling design was unbalanced, with each combination of transect and distance sampled for a total of 4-10 nights. Transects were located a minimum of 300 m apart with the exception of two transects at Hayward Shoreline, where the two transect points at 100 m from the road were separated by approximately 120 m.

In 2010, surveys were conducted using AnaBat SD2 zero-crossing bat detectors (Titley Scientific Pty Ltd, Lawton, Queensland, Australia). In 2011, surveys were conducted

using Songmeter SM2BAT 192kHz full-spectrum bat detectors (Wildlife Acoustics Inc., Concord, MA). Microphones were mounted on 1 m poles to avoid ground noise and reduce recordings of reflected calls. Directional microphones were used with AnaBat detectors and were pointed approximately perpendicular to the transect line and at a 45 degree upward angle. Omnidirectional microphones were used with SM2BAT detectors, and a 6" x 6" foam block was attached to each pole below these microphones to reduce recordings of ground-level insects and noise. The AnaBat detectors were set to autotrigger, saving separate data files for each bat pass, while the SM2BAT detectors recorded data continuously. Data were collected from sunset to sunrise each night. Although the length of nights varied by as much as one half hour during the course of the study, no clear trends in the number of recorded bat passes with calendar date were observed for either year.

Data Processing and Species Identification

In order to use comparable methods for data analysis across both years, full-spectrum recordings made in 2011 were converted to 8-division zero-crossing format using wac2wav 3.3.0 [15]. A minimum spacing of five seconds was used to separate passes, consistent with the default five second setting used by the AnaBat detectors to separate passes. A single pass was thus defined as a temporally contiguous set of calls with no more than a five second gap between successive calls.

Call data were processed using AnaLookW 3.8s [16]. A filter was used to identify high quality calls, defined as data sequences with a characteristic frequency between 5 and 60 kHz, a duration between 2 and 50 milliseconds, a body over 1 millisecond, and a *Qual* index less than 0.3 [17, 18]. All parameters were calculated within AnalookW. Passes containing no high quality calls were excluded from subsequent analysis. Counts of the number of passes recorded per night were used as a measure of bat activity.

To allow statistical analysis of individual species activity, a random forest classifier was constructed to classify each pass according to the species that produced it, following the methods described in [18]. All classification analysis was carried out in Python 2.7.2 using scikit-learn 0.11 [19]. To construct the classifier, a reference library was assembled for 15 bat species known to occur in the San Fransisco Bay area [20] (Table 4.1). The reference library consisted of AnaBat recordings of bat passes made by known, expert-identified individuals of each species (W. Rainey, personal communication). For each bat pass in the reference library, high quality calls were identified by applying the filter described above, and eleven features describing the shape of each high quality call were extracted using AnaLookW [see 18].

The resulting set of reference calls and associated features was used to train a random forest classifier, a machine learning algorithm that fits a set of classification trees each built from a bootstrapped sample from a training data set [21, 22]. Among machine learning algorithms, random forests are considered to be highly robust to correlation between predictor variables and do not require specific distributional assumptions about the variation

in call features among or between species [21]. The random forest approach was also previously recommended for identifying bat calls to the species level in a comparative analysis of several machine learning algorithms [18].

A classifier consisting of 100 trees was trained on 80% of reference calls, selected using stratified random sampling to achieve proportional representation of each species, with 20% of calls held out for testing. All trees were fully built, with three randomly drawn features used to split each node. The use of larger ensembles of trees did not substantially improve classification accuracy. Ten replicate fits of the classifier had a mean accuracy (proportion of calls in the testing set correctly identified) of 0.8411 (SD 0.0013), similar to the accuracy achieved by random forest classifiers constructed for southwestern and Florida bats [18] (Table 4.1).

The fitted random forest classifier was used to predict the probability that each high quality call recorded during field surveys was made by each species. These probabilities were summed across all calls within each pass, and the species with the highest summed probability for a pass was identified as the species associated with that pass. Summed counts of the total number of passes per night were used for all subsequent species-level analysis.

Statistical Analysis

Statistical analyses were carried out using generalized linear mixed models (GLMMs) [23, 24] using R 2.14.2 [25] with the packages lme4 0.999375-42 [26] for Poisson models and glmmADMB 0.7.2.12 [27, 28] for negative binomial models.

The statistical models included three main variables of interest that could potentially influence total bat activity: distance from road (*dist_road*, three levels: 0-20 m, 100 m, and 300 m), presence of a light within 100 m of a sampling point (*light_100*, two levels: False, True), and daily maximum temperature (*temp_max*, continuous). The latter two variables were included due to substantial prior research that shows that light sources [29–32] and ambient temperature [33–36] can affect bat activity levels. Lights were present within 100 m of four out of thirty sampled locations. Daily maximum temperature data were drawn from records at the KOAK weather station at the Oakland International Airport [37]. As ambient temperature at locations around the margins of San Francisco Bay are very similar, and all three sampled sites were immediately adjacent to the bay, temperature data recorded at this station were used as an estimate of daily maximum temperature at all sampled locations.

In addition to these three main variables, four additional variables representing site (*site*, three levels: DOED, HAYW, SAPA), year (*year*, two levels: 2010, 2011), night (*night*, random effect), and transect (*transect*, random effect) were also included in statistical models. The site and year variables were coded as fixed effects, as only three and two levels of these variables were present, respectively. Because sampling was carried out

using different detector models in 2010 and 2011, the estimated effect of year also includes the effect of detector model. Additional interaction terms between *dist_road* and *site* and between *dist_road* and *temp_max* were added to account for the biologically-reasonable possibility that road effects may vary by site and with temperature. With all variables and interactions, the final models had a total of 14 estimated fixed effect parameters (Table 4.2) plus normally distributed crossed random effects for night and transect.

As appropriate for the analysis of count data, exploratory analyses were conducted using Poisson generalized linear mixed models [24, 38]. Initial models containing all predictor variables described above indicated that nightly count data were overdispersed, with the ratio of summed squared Pearson residuals to residual degrees of freedom (accounting for fixed effects only) equal to 13.67 for a model of the total number of passes and ranging from 2.57 to 6.06 for five individual species models (see *Results*). To account for this overdispersion, negative binomial generalized linear mixed models, specified with type 1 variance structure within the glmmADMB package and using a log-link function, were used for subsequent modeling.

In specification of the model matrix, sum contrasts were used for the variables *road*, *year*, and *light_100*, such that the sum of estimated coefficients for all levels of each variable are equal to zero, as there is, philosophically, no control level of these variables. Treatment contrasts were used for *dist_road*, with counts at 300 m from the road serving as the control level. Temperature data were centered on the mean maximum daily temperature (75° F).

Given this model matrix, the estimated intercept gives the number of bat passes expected at a distance of 300 m at the mean daily maximum temperature. Estimates of the number of passes at 0 m and 100 m are obtained by adding the *dist_road_0* and *dist_road_100* coefficients, respectively. The addition of the coefficient *site1* gives estimates for San Pablo Bay, the addition of the coefficient *site2* gives estimates for Don Edwards, and the subtraction of both coefficients gives estimates for Hayward Shoreline. Estimates of passes in 2010 are obtained by adding the coefficient *year*, and estimates of passes in 2011 are obtained by subtracting this coefficient. Estimates of passes at locations within 100 m of a light are obtained by subtracting the coefficient *light*, and estimates of passes at locations more than 100 m from a light are obtained by adding this coefficient. Similar logic applies to the interpretation of the coefficients of the interaction terms.

The significance of each predictor variable was tested using likelihood ratio tests of the full model containing all predictors and a reduced model omitting only that predictor, carried out using the *anova* command of the package glmmADMB. Reduced models used to test the direct effects of road, distance from road, and temperature necessarily excluded the interaction terms that included those variables. Reduced models used to test interaction terms retained all direct effects.

Fitted statistical models were used to generate predictions of the expected number of recorded bat passes as a function of specified values of predictor variables. Expected

numbers of bat passes were calculated as the matrix product of the vector of fixed effect coefficients and $\mathbf{X_p}$, a vector of specified levels of each fixed effect. Standard errors for predictions were calculated as $\mathbf{X_p V}(\hat{\beta})\mathbf{X_p^T}$, where $\mathbf{V}(\hat{\beta})$ is the covariance matrix of the estimated fixed effect coefficients. This method of estimating standard errors accounts only for uncertainty in the estimates of fixed effects.

Results

A total of 8,308 bat passes with at least one high quality call were recorded across 174 detector-nights of sampling. On 8 of 10 sampled transects, the highest mean nightly bat activity occurred either at 100 m or 300 m from the road (Fig. 4.1), while the highest activity occurred at 0-20 m at two transects at Don Edwards (see *Discussion*).

Of all recorded passes, more than 90% were identified by the random forest classifier as either *Tadarida brasiliensis*, the Brazilian free-tailed bat (40%), *Eptesicus fuscus*, the big brown bat (25%), *Lasionycteris noctivagans*, the silver-haired bat (15%), *Lasiurus cinereus*, the hoary bat (7%), or *Myotis yumanensis*, the Yuma myotis (5%) (Table 4.1). Statistical models of bat activity, measured as a nightly count of bat passes, were fit for the total number of bat passes made by all species and for each of these five species separately.

The results of negative binomial generalized linear mixed models of nightly bat activity are shown in Table (4.2). Total bat activity is positively correlated with distance from a road, with average recorded bat activity of 16.5 passes per night at 0-20 m, 22.9 passes per night at 100 m, and 32.7 passes per night at 300 m at a mean daily maximum temperature of 75°F. A likelihood ratio test shows that this effect is statistically significant. Four of the five individual species examined show higher activity away from roads, while *M. yumanensis* shows no significant trend in activity with distance from a road (Fig. 4.2).

Maximum daily temperature prior to the recording night significantly influences bat activity, with an increase in maximum daily temperature from 75 to 85 degrees F leading to an increase in total nightly recorded bat passes from 32.7 to 43.9 at 300 m from a road. A significant interaction between distance from a road and temperature is also observed for the model of total bat activity, with the difference between activity levels at 0-20 m and at 300 m decreasing with an increase in temperature, as indicated by the positive coefficient on the interaction of *dist_road_0* and *temp_max*. This same qualitative pattern is observed for the five individual species, although the distance and temperature interaction terms are not statistically significant in the species-level models.

The number of recorded bat passes varies significantly by site for all species except *M. yumanensis*. Overall, transects at Hayward Shoreline have substantially higher total bat activity than those at Don Edwards or San Pablo Bay. The combination of main effects and interaction terms confirms that the strongest road effect is observed at the Hayward Shoreline site (see also Fig. 4.1), while at Don Edwards, unlike the other two roads, higher

total bat activity and higher activity for *E. fuscus*, *L. noctivagans*, and *M. yumanensis* is observed near roads than at 300 m.

The effect of year is significant for all species except *M. yumanensis*, with more *T. brasiliensis* and *L. cinereus* passes recorded in 2010 with the AnaBat detectors and more *E. fuscus* and *L. noctivagans* passes recorded in 2011 with the SM2BAT detectors. These differences may reflect either year-to-year variability in activity levels or different sensitivities of these two detector models to the calls of different species. The total number of passes recorded is not significantly different between years. The presence of a light within 100 m of a sampling point appeared to decrease total bat activity and activity for all five species, but this effect was not statistically significant in any model.

Discussion

These results demonstrate that total bat activity and the activity of four common bat species is consistently depressed near three large California highways as compared to control points 300 m from these roads. Possible explanations for this road effect include differences in traffic noise, illumination, wind speeds, or prey availability between points near and far from roads. Prey availability could plausibly be influenced by local temperature gradients (road surfaces are likely warmer than natural vegetation cover at night) or by other differences in microhabitat within transects. While this study was not designed to examine the causes of the observed road effect, distinguishing between these causes will be an important future area of research, as this knowledge will be needed to inform efforts to mitigate the negative impacts of roads on bat populations.

Field observations along the sampled transects suggest that small-scale differences in vegetation structure may be an important factor in determining bat activity levels. In particular, the two transects at Don Edwards at which bat activity was significantly higher near the road had notably heterogeneous vegetation cover, with relatively prominent stands of tall, invasive grasses found immediately adjacent to the road and pickleweed-dominated marsh habitat found away from the road. This observation suggests that differences in microhabitat, perhaps through an effect on prey availability, may influence bat activity strongly enough to override the otherwise repulsive effect of roads.

Although previous studies have documented increases in bat activity with increasing ambient temperature, a significant interaction between temperature and road avoidance behavior in bats has apparently not been previously described. On the hottest days of the study (daily maximum 95°F), models for all species predict approximately constant activity across all distances from a road, with road effects increasing in magnitude as temperature decreases. This pattern could potentially result from crowding if higher overall bat densities lead bats to spread out across the landscape, increasing activity in otherwise less desirable habitat, or if road surfaces themselves retained more heat following hot days, perhaps increasing insect activity near the roads during the subsequent night. The

statistically and biologically significant effect of temperature on total bat activity, and the significant interactions between temperature and the road effect, suggests that future road ecology studies for bats should explicitly incorporate temperature effects. Future studies might be specifically advised to improve on the daily maximum temperature variable used here by instead recording nightly temperature data at each sampling point.

Bats provide several valuable ecosystem services [39], and the species *T. brasiliensis* is known to provide an economically valuable pest control service in the southern United States [40, 41]. The finding that *T. brasiliensis* shows decreaseed activity near roads thus has specific implications for agricultural landscapes. In particular, if the observed road effect is driven by abiotic factors and not by prey density, this finding would suggest that insect control services may be reduced along the margins of large roads, potentially leading to higher pest densities in these areas compared to areas away from roadsides.

While this study has documented the negative effects of roads on bat activity, the net effect of large roads on bat populations may be more mixed. The availability of high quality roosts is often considered to be the most important limiting factor for bat populations, and highway overpasses and bridges often serve as important roosting sites for California bats [42]. In areas with high levels of habitat transformation, however, foraging habitat may also play a role in limiting populations [43], and an integrated strategy involving roosting, foraging, and other required habitat types may be needed for successful long-term bat conservation [44]. Given the ubiquitous nature of roads in many landscapes [2], conservation efforts should be directed towards identifying methods for preserving the beneficial aspects of road infrastructure while minimizing the negative consequences.

Overall, this study demonstrates that northern California bats display decreased activity levels in the vicinity of major roads and that this road effect is temperature dependent, with the strongest road effect occurring on nights following cold days. Roads have the potential to influence bat populations and the provision of ecosystem services in human-altered landscapes, but very little is known about the road ecology of bats. Particularly important directions for future study include conducting road effects research in habitats other than salt marsh, investigating the road ecology of species other than those present in northern California, and determining the specific factors that cause the observed road effects.

Acknowledgements

I thank John Harte, William Lidicker, Jr., and Adina Merenlender for helpful comments on drafts of this chapter. I also thank Adina Merelender and William Rainey for providing the equipment used for field surveys, and William Rainey for providing the reference call library used for species classification.

References

[1] Sanderson E, Jaiteh M, Levy M, Redford K (2002) The human footprint and the last of the wild. *BioScience* 52:891–904.

[2] Watts RD, et al. (2007) Roadless space of the conterminous United States. *Science* 316:736–738.

[3] Forman R (1998) Roads and their major ecological effects. *Annual Review of Ecology and Systematics* 29:207–231.

[4] Ree RVD, Jaeger JAG, van der Grift EA, Clevenger AP (2011) Effects of roads and traffic on wildlife populations and landscape function: road ecology is moving toward larger scales. *Ecology and Society* 16:48.

[5] Fahrig L, Rytwinski T (2009) Effects of roads on animal abundance: an empirical review and synthesis. *Ecology and Society* 14.

[6] Benítez-López A, Alkemade R, Verweij PA (2010) The impacts of roads and other infrastructure on mammal and bird populations: A meta-analysis. *Biological Conservation* 143:1307–1316.

[7] Russell A, Butchkoski C, Saidak L, McCracken G (2009) Road-killed bats, highway design, and the commuting ecology of bats. *Endangered Species Research* 8:49–60.

[8] Lesinski G, Sikora A, Olszewski A (2011) Bat casualties on a road crossing a mosaic landscape. *European Journal of Wildlife Research* 57:217–223.

[9] Kerth G, Melber M (2009) Species-specific barrier effects of a motorway on the habitat use of two threatened forest-living bat species. *Biological Conservation* 142:270–279.

[10] Zurcher AA, Sparks DW, Bennett VJ (2010) Why the bat did not cross the road? *Acta Chiropterologica* 12:337–340.

[11] Schaub A, Ostwald J, Siemers BM (2008) Foraging bats avoid noise. *The Journal of Experimental Biology* 211:3174–3180.

[12] Siemers BM, Schaub A (2011) Hunting at the highway: traffic noise reduces foraging efficiency in acoustic predators. *Proceedings of the Royal Society B* 278:1646–1652.

[13] Berthinussen A, Altringham J (2012) The effect of a major road on bat activity and diversity. *Journal of Applied Ecology* 49:82–89.

[14] CalTrans (2012) Traffic Data Branch. http://traffic-counts.dot.ca.gov/.

[15] Wildlife Acoustics (2012) Wildlife Acoustics. http://www.wildlifeacoustics.com/wa-php/downloads.php.

[16] Corben C (2012) ANABAT. http://users.lmi.net/corben/anabat.htm.

[17] Rainey WE, Ingersoll T, Corben CJ, Pierson ED (2009) Using acoustic sampling of bat assemblages to monitor ecosystem trends. (U.S. Geological Survey, Point Reyes, CA), Technical report.

[18] Armitage DW, Ober HK (2010) A comparison of supervised learning techniques in the classification of bat echolocation calls. *Ecological Informatics* 5:465–473.

[19] Pedregosa F, et al. (2011) Scikit-learn: machine learning in Python. *Journal of Machine Learning Research* 12:2825–2830.

[20] California Department of Fish And Game (2010) California Wildlife Habitat Relationships Home. http://www.dfg.ca.gov/biogeodata/cwhr/.

[21] Breiman L (2001) Random forests. *Machine Learning* 45:5–32.

[22] Cutler DR, et al. (2007) Random forests for classification in ecology. *Ecology* 88:2783–2792.

[23] Bolker BM, et al. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* 24:127–135.

[24] Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed Effects Models and Extensions in Ecology with R* (Springer, New York, NY), p 574.

[25] R Development Core Team (2012) R: A language and environment for statistical computing.

[26] Bates D, Maechler M, Bolker B (2011) lme4: Linear mixed-effects models using S4 classes.

[27] Fournier DA, et al. (2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 27:233–249.

[28] Skaug H, Fournier D, Nielsen A, Magnusson A, Bolker B (2012) glmmADMB: Generalized linear mixed models using AD Model Builder.

[29] Jong J, Ahlen I (1991) Factors affecting the distribution pattern of bats in Uppland, central Sweden. *Holarctic Ecology* 14:92–96.

[30] Blake D, Hutson AM, Racey PA, Rydell J, Speakman JR (1994) Use of lamplit roads by foraging bats in southern England. *Journal of Zoology, London* 234:453–462.

[31] Hickey MBC, Acharya L, Pennington S (1996) Resource partitioning by two species of Vespertilionid bats (Lasiurus cinereus and Lasiurus borealis) feeding around street lights. *Journal of Mammalogy* 77:325–334.

[32] Stone EL, Jones G, Harris S (2009) Street lighting disturbs commuting bats. *Current Biology* 19:1123–1127.

[33] Anthony ELP, Stack MH, Kunz TH (1981) Night roosting and the nocturnal time budget of the little brown bat Myotis lucifugus: effects of reproductive status, prey density, and environmental conditions. *Oecologia* 51:151–156.

[34] Avery MI (1985) Winter activity of Pipistrelle bats. *Journal of Animal Ecology* 54:721–738.

[35] Richards G (1989) Nocturnal activity of insectivorous bats relative to temperature and prey availability in tropical Queensland. *Wildlife Research* 16:151–158.

[36] Rydell J (1989) Feeding activity of the northern bat Eptesicus nilssoni during pregnancy and lactation. *Oecologia* 80:562–565.

[37] National Climatic Data Center (2012) Global Surface Summary of Day. http://www.ncdc.noaa.gov/cgi-bin/res40.pl.

[38] O'Hara RB, Kotze DJ (2010) Do not log-transform count data. *Methods in Ecology and Evolution* 1:118–122.

[39] Kunz TH, de Torrez EB, Bauer D, Lobova T, Fleming TH (2011) Ecosystem services provided by bats. *Annals of the New York Academy of Sciences* 1223:1–38.

[40] Cleveland CJ, et al. (2006) Economic value of the pest control service provided by Brazilian free-tailed bats in south-central Texas. *Frontiers in Ecology and the Environment* 4:238–243.

[41] Boyles JG, Cryan PM, McCracken GF, Kunz TH (2011) Economic importance of bats in agriculture. *Science* 332:41–42.

[42] Johnston D (2004) California bat mitigation techniques, solutions, and effectiveness. (California Department of Transportation (Caltrans), Office of Biological Studies and Technical Assistance, Sacramento, CA), Technical Report 2394.

[43] Hutson AM, Mickleburgh SP, Racey PA (2001) *Microchiropteran Bats: Global Status Survey And Conservation Action Plan* (IUCN/SSC Chiroptera Specialist Group, Gland, Switzerland and Cambridge, United Kingdom), p 258.

[44] Pierson E (1998) in *Bat Biology and Conservation*, eds Kunz TH, Racey PA (Smithsonian Institution Scholarly Press, Washington, D.C.), pp 309–325.

Species	N_{ref}	P	R	N_{obs}
Antrozous pallidus	353	0.78	0.78	50
Corynorhinus townsendii	148	0.86	0.77	6
Eptesicus fuscus	884	0.83	0.89	2094
Eumops perotis	204	0.98	0.99	51
Lasionycteris noctivagans	365	0.82	0.78	1258
Lasiurus blossevillii	98	0.86	0.64	155
Lasiurus cinereus	174	0.91	0.79	584
Myotis californicus	220	0.75	0.54	23
Myotis evotis	169	0.83	0.83	2
Myotis lucifugus	219	0.75	0.84	196
Myotis thysanodes	184	0.89	0.91	0
Myotis volans	198	0.79	0.62	26
Myotis yumanensis	650	0.87	0.96	441
Parastrellus hesperus	267	0.96	0.98	46
Tadarida brasiliensis	304	0.86	0.72	3376
Total	4437			8308

Table 4.1: Species present near San Francisco Bay and included in species classifier. N_{ref} is the number of passes for each species included in the reference library. P is precision, the fraction of all calls in the testing data set identified as a species that were made by that species, and R is recall, the fraction of all calls in the testing data set made by a species that were identified as that species. Overall species classification accuracy, the fraction of all calls in the testing data set identified correctly, was 84%. N_{obs} is the number of passes recorded during the study that were identified as each species.

Coefficient	TOTAL		TABR		EPFU		LANO		LACI		MYYU	
Intercept	3.4870	**	2.6711	**	0.2755	**	-0.0029	**	1.0646	**	0.1514	**
dist_road		**		**		**		**		**		.
dist_road_100	-0.3560		-0.2624		-0.4410		-0.5513		-0.3051		0.0530	
dist_road_0	-0.6853		-0.6459		-0.5909		-1.0851		-0.5279		0.2250	
temp_max	0.0294	*	0.0238	.	0.0485	**	0.0283	*	0.0268	.	-0.0070	.
dist_road:temp_max		*		.		.		.		.		.
dist_road_100:temp_max	-0.0238		-0.0255		0.0408		0.0066		-0.0283		0.0122	
dist_road_0:temp_max	0.0277		0.0176		0.0654		0.0866		0.0161		0.0400	
site		**		**		**		**		*		.
site1	-0.3937		-0.4017		-0.0068		-1.1843		-0.1025		0.5127	
site2	-1.0882		-0.8379		-3.5913		-2.0287		0.1845		-0.8154	
dist_road:site		**		**		**		*		*		.
dist_road_100:site1	0.5899		0.8801		0.1765		0.5549		0.6708		-0.2036	
dist_road_0:site1	0.3101		0.4048		0.0225		0.2496		0.0976		-0.1368	
dist_road_100:site2	-0.3017		-0.4784		0.3632		-0.1876		-0.6392		0.1887	
dist_road_0:site2	1.0328		0.5164		2.1542		1.3285		0.4569		1.1560	
year	0.1384	.	0.4238	*	-0.8726	**	-0.6150	*	0.4750	**	-0.4024	.
light	0.2129	.	0.0849	.	0.2592	.	0.4402	.	0.2159	.	0.1132	.

Table 4.2: Results of negative binomial generalized linear mixed models of counts of nightly bat passes as a function of distance from the road, maximum daily temperature, site, year, and presence of a light within 100 m (fixed effects), and night and transect (random effects, not shown). * indicates p-values below 0.05 and ** indicates p-values below 0.01 as determined by a likelihood ratio test. The interpretation of levels of categorical variables is described in *Methods*. Distance coefficients are relative to a control distance of 300 m. Coefficients are in units of log(passes). TABR: *Tadarida brasiliensis* (Brazilian free-tailed bat), EPFU: *Eptesicus fuscus* (big brown bat), LANO: *Lasionycteris noctivagans* (silver-haired bat), LACI: *Lasiurus cinereus* (hoary bat), MYYU: *Myotis yumanensis* (Yuma myotis).

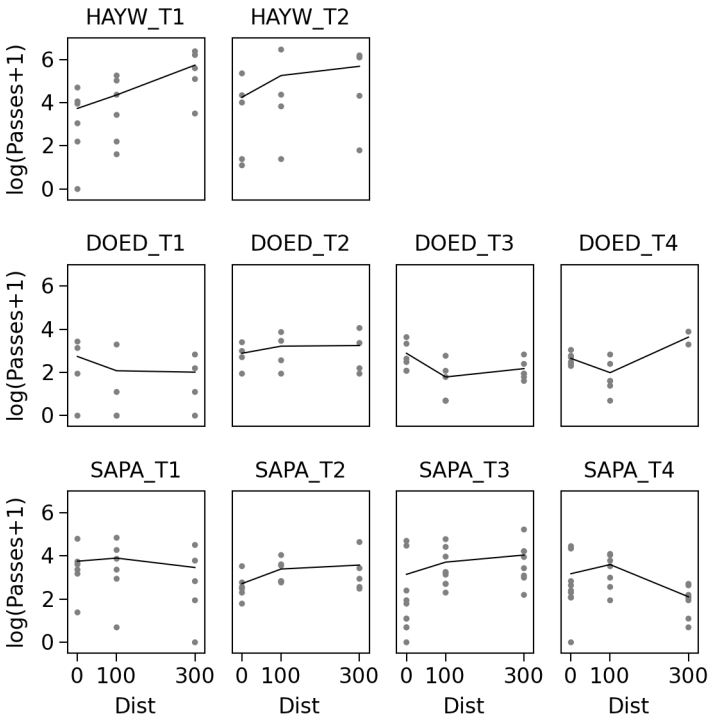

Figure 4.1: Total nightly counts of bat passes at each transect (HAYW: Hayward Shoreline, DOED: Don Edwards, SAPA: San Pablo Bay) as a function of distance from a road. Lines show mean nightly counts of total bat passes at each distance.

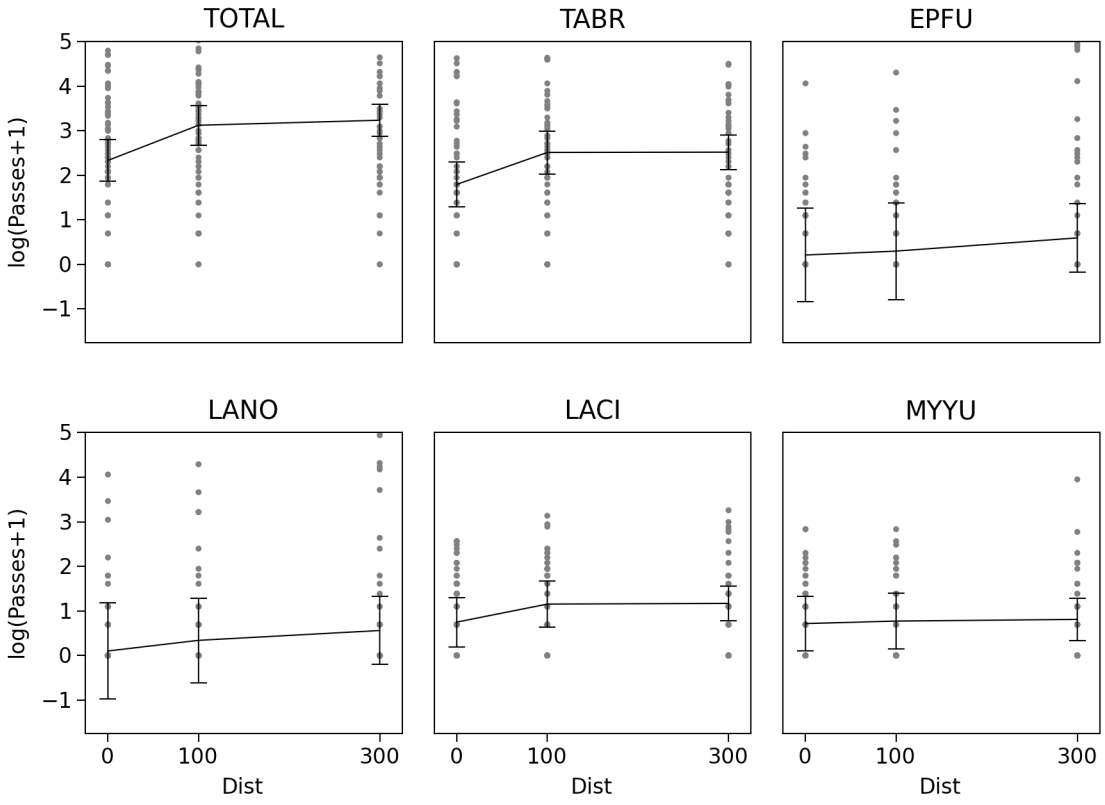

Figure 4.2: Predicted total and species-specific nightly counts of bat passes as a function of distance from a road. Lines connect mean predicted activity levels across distances, and error bars represent standard errors of predictions. Predictions are for a maximum daily temperature of 65°F. Gray points are observed data, with observed nightly counts greater than 150 not shown. TABR: *Tadarida brasiliensis* (Brazilian free-tailed bat), EPFU: *Eptesicus fuscus* (big brown bat), LANO: *Lasionycteris noctivagans* (silver-haired bat), LACI: *Lasiurus cinereus* (hoary bat), MYYU: *Myotis yumanensis* (Yuma myotis).

Chapter 5

The global wildlife footprint: linking bird and mammal population losses to economic consumption

Abstract

Although the proximate causes of biodiversity loss, such as land use change, overharvesting, and invasive species, are increasingly well understood, quantitative analyses of the relationship between biodiversity loss and the indirect, upstream economic activities that drive this loss remain uncommon. Here, I present the results of a global "wildlife footprint" analysis that links decreases in global wildlife populations due to land use activities to 57 categories of economic consumption in 113 regions of the world. Current land use practices are found to decrease global terrestrial bird and mammal populations by 20–25% below the baseline levels that would be present in the absence of human land use activities. This decrease represents ongoing losses of 26 billion individual birds and 1.1 trillion individual mammals, or approximately 4 wild birds and 160 wild mammals per person alive today. The largest wildlife population losses occur in regions with high baseline wildlife populations, while the global wildlife population losses driven by economic consumption within a region is closely correlated with that region's human population size. Consumer purchases from agricultural sectors lead to the majority of wildlife population losses, with one dollar spent by a final consumer in several agricultural sectors displacing on the order of 0.1–1 individual birds or mammals. These results provide a new perspective on the global drivers of biodiversity loss and highlight a potentially important approach for incorporating biodiversity considerations into economic decision making.

Introduction

Despite growing recognition of the importance of biodiversity to natural and human communities, global biodiversity loss continues at a rapid pace [1]. In response, substantial efforts have been directed towards understanding the extent and location of the direct, on-the-ground causes of biodiversity loss, such as land use change, climate change, invasive species, over-exploitation, and pollution [2]. These direct causes, however, are themselves

ultimately driven by the upstream economic production and consumption activities of sectors and individual consumers. Deforestation in the tropics, for example, is not practiced for its own sake, but rather ultimately is driven by demands, both local and global, for heating, wood-based products, agricultural products such as beef and palm oil, and other goods and services.

To date, however, relatively little effort has been directed towards studying these indirect economic drivers of biodiversity loss. Building on recent advances in consumption-based environmental accounting [3–5], I here calculate a first estimate of a global "wildlife footprint" that allocates losses of global terrestrial bird and mammal populations due to human land use activities to economic consumption in 113 regions and 57 economic sectors. In the spirit of the carbon, water, and ecological footprints [4–9], the wildlife footprint measures the total wildlife population losses caused indirectly by the everyday activities of human populations.

Throughout, the term "populations" should be understood to refer to an unweighted or weighted count of a number of individual organisms, not to a count of the number of reproductively isolated groups within a species. Metrics based on wildlife population losses arguably provide more sensitive measurements of human impacts on biodiversity than metrics based on species extinctions [10, 11] and highlight potentially important changes in the status of common as well as rare species [12, 13]. The population-based approach used here thus complements existing estimates of species extinction risks and rates [14–16]. This analysis focuses on decreases in global wild bird and mammal populations, as these taxa are a common focus of conservation concern and are relatively well characterized at the global scale.

This analysis addresses three key questions. First, in what regions are human land use activities causing the greatest losses of wildlife populations, and what types of land use activities are the proximate causes of these losses? Second, which regions are indirectly driving wildlife population losses through their consumption demands, and are these the same regions in which the losses are occurring (i.e., to what extent is the global wildlife footprint traded internationally)? Third, which economic sectors play the largest roles in promoting wildlife population losses, and what is the wildlife population loss associated with each dollar spent by a consumer in each sector (i.e., what is the wildlife population-intensity of each economic sector)?

Addressing the first question requires creating global maps of estimated wildlife populations in the absence of human land use activities. Baseline, pre-industrial age wildlife populations are measured using two metrics: a count of a total number of individual organisms at a location and a fractional-range metric equal to the sum across all species of the fraction of each species' range found at a location. These two measurements are closely related to the first two IUCN criteria for assessing a species' extinction risk [16] and offer complementary perspectives on the value of different species. Specifically, the individual approach assigns the highest wildlife footprints to activities that affect

widespread, abundant species, while the fractional-range approach assigns the highest wildlife footprints to activities that affect rare, range-restricted species. These baseline wildlife population maps are then combined with global maps of the human appropriation of net primary productivity (HANPP) related to four major human land use activities (crop production, grazing, forestry, and infrastructure areas) to estimate the wildlife population losses occurring at any location and their direct cause.

The second and third questions are answered using an environmentally-extended multi-regional input-output analysis [17]. For this analysis, wildlife population losses are initially allocated to the economic sector, within a specific geographic region, that is the proximate cause of the losses. A matrix of monetary flows between sectors and end consumers is then used to determine the upstream consumption activities, both from a regional and a sectoral perspective, that drive global wildlife population losses. The total, global losses of wildlife populations due to economic consumption in a region is defined as that region's wildlife footprint. This approach explicitly accounts for trade in embodied wildlife footprint between regions, as the wildlife footprint of any region is equal to the wildlife population losses within that region plus imports and minus exports of embodied wildlife population losses in traded goods and services.

Throughout, wildlife population losses are defined as an ongoing rather than a static phenomenon. All calculations reflects the premise that losses do not occur only once as a result of the initial conversion of natural habitat, but rather on an ongoing basis as continually occupied habitat fails to "produce" wildlife in each year that it is occupied. For example, the complete occupation of a habitat that would support 100 individual birds in the absence of humans is considered to represent a population loss of 100 individual birds in each time period that the habitat is occupied, regardless of the date that the habitat was first converted for human uses. This approach has the conceptual advantage of recognizing the possibility for the recovery of wildlife populations on currently occupied lands and the practical advantage of being feasible to calculate given existing global datasets.

Methods

Temporal aspects of population loss

Any effort to measure the causes of wildlife population losses must first determine a method for accounting for the temporal dimension of population loss. A "marginal approach" defines population loss as occurring only once, at the instant at which a natural land cover is converted to a human-dominated state. This approach implicitly assumes that continuing occupation of existing transformed land involves no associated continuing wildlife population loss, thus ignoring the potential for the recovery and restoration of existing human-used lands. An "ongoing approach" defines population loss as a process that occurs on a continuous basis, such that habitat occupation in any time period, regardless of the date of the original conversion of habitat, continues to represent a loss. This approach

implicitly assumes unlimited potential for recovery and restoration of existing human used lands. The marginal approach is consistent with considering wildlife populations to be a stock, while the ongoing approach is consistent with considering wildlife populations to be a flow.

Consider, for example, a plot of forest containing habitat for 100 birds that is cleared for cropland. The marginal approach allocates 100 birds of population loss to the initial actor who clears the plot, while any future actor who continues to harvest crops from the plot is presumed to cause zero population loss, even though such continual harvest likely prevents additional birds from recruiting to the plot. Conversely, the ongoing approach allocates a wildlife footprint of 100 lost birds to the actor who harvests the plot in any time period, regardless of when the original plot was harvested.

The marginal and ongoing approaches to measuring population loss represent two valid and internally-consistent approaches to assigning responsibility for losses across time, as would hybrid approaches that mix the two in a manner akin to depreciation of a capital investment in finance. This analysis uses the ongoing approach, for which losses in any given year can be estimated from maps of global baseline wildlife populations and existing land uses. The marginal approach would instead require maps of global baseline wildlife populations and land clearing, with each unit of cleared land assigned to the activity responsible for the clearing.

While conceptual and not physical in nature, the differences between the ongoing and marginal calculations of population loss are substantial, and comparison of the two approaches will be an important area for future investigation. In general, the ongoing approach will find the largest wildlife population losses to be occurring in regions with the largest baseline wildlife populations and highest intensity of human land use, while the marginal approach will find the largest wildlife population losses to be occurring in regions with the largest baseline wildlife populations and current land conversion activity.

Spatially-explicit baseline wildlife population maps

To begin the analysis, spatially-explicit maps of baseline global wildlife populations in the absence of human land use activities were estimated on a global 5 arc-minute grid.

Global species range data for 5,276 terrestrial mammal species, published by IUCN [16], and 2,080 threatened bird species, published by Birdlife International [18], were used for baseline wildlife population analysis. Baseline population estimates using the fractional-range approach were calculated for both taxa, for each grid cell, as the sum over all species of the fraction of each species' total range that occurs within that grid cell. As no global range maps for all bird species were available, only threatened birds were considered for this fractional-range approach. To the extent that population density is approximately constant across a species' range, the fractional-range metric can be interpreted as a rarity-weighted population metric.

Baseline wildlife population estimates using the individual approach were calculated differently for birds and mammals. For birds, the baseline number of individuals in a cell was calculated from an estimate of breeding bird population density within different land cover classes and a map of potential vegetation classes in the absence of human land use activities [19, 20].

For mammals, the baseline number of individuals in a cell was calculated by assigning each species a population density, taken to be constant across that species' range and zero otherwise, and, for each grid cell, summing across all species the product of cell area and population density for each species. Mammal population densities were drawn from the PanTHERIA database [21] when available ($n = 908$, 17% of species). When no measured population density for a species was unavailable, population density was estimated from adult body mass (*Appendix*, $n = 2749$, 52% of species). Mammal population densities were capped at a maximum of 10,000 individuals per km^2. This approach estimates a total global baseline mammalian biomass of approximately 10^{12} kg, comparable to previous estimates on the order of 10^{10}–10^{11} kg for all wild mammals [22] and 10^{11} kg for all mammals with body sizes greater than 44 kg [23].

The range maps and density database used to construct baseline population maps include only extant or recently extant species, and, as such, baseline population levels should be interpreted as approximately the populations that were present immediately prior to the industrial age (i.e., several hundred years before present). Estimates of wildlife population loss based on these baseline maps thus do not account for wildlife harvesting, landscape conversion, and other human activities that occurred during pre-industrial management of landscapes.

Spatially-explicit wildlife population loss maps

To create maps of global wildlife population losses, baseline population maps were combined with maps that estimate the proportional losses of wildlife populations in each grid cell. These proportional losses were estimated using maps of the human appropriation of net primary productivity (HANPP) [24–27]. Although several definitions of HANPP have been proposed, this study follows [27] in defining HANPP as the difference between the amount of net primary productivity (NPP) (t C / ha yr) that would be available in a grid cell in the absence of human activities, NPP_0, and the amount of NPP that remains in the grid cell following human harvest and land conversion.

The fractional wildlife population loss in each grid cell due to each of four different human land use activities, crop production, grazing, forestry, and infrastructure areas (buildings and roads), is calculated as the ratio of the HANPP associated with that activity to NPP_0, multiplied by the fraction of grid cell area associated with each land use [28]. Activities wth negative associated HANPP in a grid cell, indicating that total available NPP in the cell is increased due to that activity, are presumed to cause zero population loss. HANPP and NPP_0 for each land use activity were calculated as in [27] and are for the year 2000.

This use of HANPP data to estimate wildlife population losses presumes a linear relationship between NPP availability and wildlife population sizes. In the context of this analysis, this assumption can be viewed as reflecting an individuals-area relationship, the scaling of the number of individual organisms with area, which is linear when population density is constant across a species' range. If a species, rather than a population, perspective were to be used to calculate a wildlife footprint, this assumption of linearity would be poor, as species richness is well known to increase with area as approximately a power law [29–31]. As a consequence of this assumption of linearity, these calculations are agnostic to the intensity of land use, as the 100% occupation (i.e., zero NPP remaining in a grid cell) of 10 hectares of land for human uses, for example, is calculated to lead to the same population loss as a 50% occupation of 20 hectares of land. This assumption is important both philosophically and practically in this analysis, as data on the configuration and intensity of land uses within grid cells was not available.

This HANPP-based analysis provides a conservative estimate of wildlife population losses, as it measures only the quantity of NPP available and does not consider changes in land cover or the quality of the NPP remaining in a cell (e.g., if the NPP embodied in crop residue remaining after harvest is equivalent to the NPP of a former, natural ecosystem, wildlife population losses are taken to be zero). This approach also does not include other non-NPP related disturbances that could lead to wildlife population losses.

Regional and sectoral drivers of wildlife population loss

An environmentally-extended multi-regional input-output analysis [17] is used to link estimated wildlife population losses in each grid cell to upstream economic sectors and human consumption activities. A multi-regional input-output table covering 57 sectors in 113 regions was extracted from the Global Trade Analysis Project (GTAP) database v7 [32] following [3]. The GTAP v7 database has a reference year of 2004, and all monetary figures are in units of 2004 U.S. dollars.

For both the individual and fractional-range measures of wildlife populations, for both birds and mammals, the direct population loss caused by each land use activity (crop production, pasture, forestry, and infrastructure areas) within each region was calculated as the sum of wildlife population loss due to each activity occurring within all grid cells in that region.

A direct requirements vector, f^d, with 6441 elements giving the direct population loss caused by each sector within each region, was then constructed by allocating the population loss caused by each of the four major land use activities within each region to the relevant economic sectors within that region (*Appendix*). The sum of f^d across all sectors within a region gives the total wildlife population loss that occurs within the borders of that region, and the sum across all regions for a single sector gives an estimate of the total wildlife population loss caused directly by the activities of a given economic sector.

A direct intensity vector, reflecting the wildlife population loss directly caused to create

one dollar of output from a given sector within a given region, was calculated as the direct requirements vector divided element-wise by a vector of total sectoral economic output, x.

$$F^D = f^d / x \tag{5.1}$$

A total intensity vector, F^T, reflecting the total wildlife population loss caused throughout all sectors worldwide to produce one dollar of output from a given sector within a region, was calculated using Eq. (5.2).

$$F^T = F^D(\mathbf{I} - \mathbf{A})^{-1} \tag{5.2}$$

In Eq. (5.2), $\mathbf{I}$ is the identity matrix and $\mathbf{A}$ is a matrix of input-output coefficients [3, 17]. The total intensity vector F^T measures the wildlife footprint-intensity of final consumption in each sector, or, equivalently, the total downstream wildlife population loss caused by an end consumer spending one dollar to purchase goods or services from each sector.

The total wildlife population loss associated with final consumption in each sector in each region can be calculated as the product of F^T and the total final demand by households, government, and industry (c) for goods and services from each sector within each region.

$$f^c = F^T c \tag{5.3}$$

The final consumption vector f^c measures the total wildlife population loss that is caused directly and indirectly by the consumption of all goods and services sold by each sector in each region. The sum of f^c across all sectors within a region gives an estimate of the global wildlife population loss, occurring anywhere in the world, that is associated with economic consumption in each region, defined as that region's wildlife footprint. The sum across all regions provides a similar metric for each sector.

The difference between the sum of f^d and f^c for a given region represents the difference between wildlife population loss that occurs within a region and global population loss driven by consumption of residents of that region, giving net trade in embodied wildlife footprint for that region.

Results

For both bird and mammal communities, using both the individual and fractional-range measurement approaches, baseline populations in the absence of human land use activities are largest in the tropics, decreasing towards the poles, with the highest levels found in the Amazon basin and southeast Asia (Fig. 5.A1). As measured by counts of individuals,

large baseline bird and mammal populations are also found in continental Europe and the western United States. Locations with very high summed fractional-ranges occur throughout the coastal tropics, where large numbers of endemic, range-restricted species are found. Globally, baseline populations of approximately 120 billion individual birds and 5.4 trillion individual mammals are estimated [see also 20, 22, 23].

Based on a spatially-explicit analysis of the human appropriation of net primary productivity (HANPP), crop production leads to the vast majority of global wildlife population losses due to land use activities, with grazing, forestry, and infrastructure areas locally important in some regions (Fig. 5.A2). Central North America, continental Europe, south Asia, and east Asia contain large areas in which high intensity of land use leads to greater than 50% losses of baseline wildlife populations.

Combining baseline wildlife population and HANPP estimates, global ongoing wildlife population losses are estimated as 22% of bird individuals, 24% of bird fractional-ranges, 21% of mammal individuals, and 22% of mammal fractional-ranges. In total, 26 billion individual birds and 1.1 trillion individual mammals are currently displaced by human land use activities. These 1.1 trillion individual lost mammals represent an estimated total biomass of approximately 200 billion kg, which, given a global human population of 7 billion and an individual human mass on the order of 50–70 kg [23, 33], implies that each human currently alive displaces approximately one-half his or her own mass in wild mammals.

From a wildlife footprint perspective, seven nations, Brazil, China, India, Indonesia, the Philippines, Russia, and the United States, have the most significant global impact under several methods of measurement. For both birds and mammals and using both the individual and fractional-range measurements of wildlife populations, 38–50% of all global ongoing wildlife population losses occur within the borders of these nations, and 37–47% of all global ongoing wildlife population losses are driven by consumption demands of residents of these nations. For comparison, these seven nations contain 51% of the global human population and represent 38% of global GDP. It is notable that among these seven nations, wildlife population losses within Brazil and Russia are relatively low (9–23%), indicating that these regions have large remaining, intact wildlife populations that could be benefited by conservation efforts. Wildlife population losses within India, in contrast, are currently 59–71%, suggesting that India has relatively less potential for experiencing large additional losses in the future.

Using log-log linear regressions (Table 5.1), the wildlife population losses occurring within each region are found to be best predicted by baseline wildlife populations in that region ($R^2 = 0.64$–0.87), indicating that the regions with the largest wildlife populations to lose are those that generally experiences the largest losses. Regional GDP and human population size predict regional wildlife population losses less strongly. Together, baseline wildlife population size, human population size, and GDP explain 80–92% of variation in the wildlife population losses occurring within regions. Conversely, the regional wildlife

footprint (i.e., the global wildlife population losses due to a region's consumption) is most strongly correlated with a region's human population size ($R^2 = 0.65$–0.90). The slope of the relationship between human population size and total regional wildlife footprint is less than one, indicating a potential economy of scale in meeting human demands with lowered impact on wild populations.

The per-person wildlife footprint is relatively poorly predicted by per-person GDP and baseline biodiversity (Table 5.1). The largest per-person wildlife footprints occur in different regions depending on whether the wildlife footprint is measured in units of bird individuals (Australia, Botswana, Iceland and Liechtenstein), bird fractional-ranges (Brunei Darussalam and Timor-Leste, Panama, Ecuador), mammal individuals (Kazakhstan, Luxembourg, Belarus), or mammal fractional-ranges (Madagascar, Panama, Costa Rica) (Fig. 5.2).

The difference between wildlife population losses occurring within a region and the regional wildlife footprint provides a measure of net international trade in embodied wildlife footprint for that region. Depending on the method of baseline wildlife population measurement, an estimated 24–26% of the global wildlife footprint is traded internationally, amounting to total international trade flows of 6 billion embodied lost individual birds and 281 billion embodied lost individual mammals.

The importance of international trade for each region can be measured by calculating the fraction of a region's wildlife footprint that involves wildlife population losses occurring outside of the borders of that region, a measure of imports of embodied wildlife footprint (Fig. 5.3). Regions with large total wildlife footprints and large imported wildlife footprints include the United States, western Europe, and Japan. Conversely, the wildlife footprints of Russia, Brazil, India, China, and Australia are largely associated with wildlife population losses occurring within the borders of those regions themselves.

The economic sectors whose output is associated with the largest total wildlife footprint include those that are large in monetary terms and linked directly to activities that affect land use, including all food products sectors, services (including government), construction, and trade (retail purchases) (Fig. 5.A4). Consumer purchases of products from food-related sectors involve the highest wildlife footprint per dollar, on the order of 0.1 individual birds and 1 individual mammal per 2004 U.S. dollar spent in these sectors by final consumers. Patterns in the relative wildlife footprint associated with different sectors are broadly consistent across all four measures of wildlife population losses.

Discussion

This analysis of the global wildlife footprint due to human land use activities highlights the regions and economic sectors that are the most significant drivers of global wildlife population losses. Globally, the largest wildlife population losses occur in areas with high

estimated baseline wildlife populations, while the regions that drive the most wildlife population losses through their consumption are those with large human populations. Controlling for human population size, the list of regions with the largest per-person wildlife footprint varies greatly depending on taxonomic group (birds or mammals) and the method used to measure wildlife population losses (individuals or fractional-ranges).

On a sectoral basis, agricultural sectors are associated with the largest wildlife footprints, a finding that reflects the land use basis of this analysis. The wildlife footprint intensity of sectoral output can be interpreted as the average wildlife population losses associated with a consumer purchase of one dollar of goods or services from a given sector. These intensities indicate that, from the perspective of a developed nation consumer, seemingly minor daily consumption activities can have significant impacts on wildlife populations. For example, a person who purchases $10 of bovine meat products displaces approximately 0.1 wild birds and 6 wild mammals for one year (Table 5.2).

The results presented here should be understood as a first-pass estimate of the global wildlife footprint and interpreted in light of the many limitations of existing data sets and the uncertainties in the calculation methods used here, of which three are especially important. First, estimates of baseline wildlife populations rely on coarse approximations of bird and mammal ranges and population densities that could be improved with finer scale biodiversity data. Second, the assumption of a linear relationship between wildlife population sizes, area, and human appropriation of net primary productivity is an approximation that ignores the potential for feedbacks between habitat loss and wildlife population sizes. Third, multi-regional input-output analyses are linear models that are limited by the regional and sectoral resolution of data tables and disparities in the collection and standardization of raw data across regions [4], although newer input-output databases that partially address these shortcomings are currently being developed [34, 35].

Several extensions of this analysis could further improve the utility of global wildlife footprint calculations. First, the wildlife population losses due to land use activities, as presented here, could be compared to estimated population losses due to other human impacts such as climate change, which may become a driver of biodiversity loss as important, if not more important, than land use change in the coming century [2]. The differences between a land use and climate change perspective could be substantial, as a climate-based analysis would find the largest biodiversity loss to be associated with the activities that lead to the largest emissions of greenhouse gases [4, 5]. Second, the aggregated bird and mammal population loss figures presented here could be further disaggregated to finer taxonomic levels, such as individual species, where data permit. The analysis presented here, for example, counts the loss of a single individual of all wildlife species to be equivalent (i.e., the loss of a single elephant is equivalent to the loss of a single shrew). While this approach allows the wildlife footprint to be measured in the single, straightforward unit of individual organisms, further extensions of this analysis may wish to focus on aggregate wildlife biomass losses or other weighted metrics.

Third, the monetary input-output tables used here embody an accounting philosophy in which financial flows and economic activity represent the links between direct wildlife population losses and the upstream activities that drive these losses. This approach could be compared to other methods, including those that use physical input-output tables, that attempt instead to directly measure physical resource flows [e.g., 36–38]. Physical flow accounting, however, is not an inherently preferable approach, as, for example, measurements of physical flows will not account for the wildlife footprint embodied in international trade in services.

Given the findings presented here, individuals and organizations may wish to examine ways to reduce the wildlife footprint of their consumption activities. This could be achieved, for example, by reducing purchases of products that lead to large wildlife population losses (such as meat products) or by sourcing products from regions where relatively lower wildlife population loss is associated with production, which will tend to be regions with low baseline wildlife populations and relatively efficient production practices. More generally, the quantification of the link between individual behavior and biodiversity consequences may help to bring biodiversity considerations more centrally into public discourse and decision making. Although more refined estimates of the wildlife footprint will be necessary for much fine-scale decision making, the overall framework presented here provides a useful path forward for future analyses.

Similar to the role that GDP plays in summarizing extensive and disparate economic data into a single headline indicator, the global wildlife footprint has the potential to serve as a useful aggregate measure of human impacts on global biodiversity. Despite the limitations and coarse nature of this analysis, results such as these that explicitly link human activities to global biodiversity outcomes have the potential to inform action to stem ongoing losses of global biodiversity. Continued efforts to quantify the relationship between economic activities and biodiversity loss will help conservation biologists to identify new opportunities for interventions, decision makers to evaluate the relationships between economic growth and wildlife populations, and the general public to better understand their personal roles in determining the fate of global biodiversity.

Acknowledgements

I thank my collaborators Eric Berlow, Erin Conlisk, Karl-Heinz Erb, John Harte, Katsunori Iha, Neo Martinez, Erica Newman, Cristoph Plutzar, and Adam Smith for their input on the analytical approaches used here. I especially thank Karl-Heinz Erb and Cristoph Plutzar for providing maps of global HANPP by land use activity and Erin Conlisk, Katsunori Iha, and Adam Smith for reviewing portions of the code used to generate the results presented here. I also thank all of my collaborators, Manisha Anantharaman, William Lidicker, Jr., Adina Merenlender, and Lydia Smith for helpful comments on drafts of this chapter.

References

[1] Butchart SHM, et al. (2010) Global biodiversity: indicators of recent declines. *Science* 328:1164–1168.

[2] Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis* (Island Press, Washington, D.C.), p 137.

[3] Peters G (2008) From production-based to consumption-based national emission inventories. *Ecological Economics* 65:13–23.

[4] Hertwich EG, Peters GP (2009) Carbon footprint of nations: a global, trade-linked analysis. *Environmental Science and Technology* 43:6414–6420.

[5] Davis SJ, Caldeira K (2010) Consumption-based accounting of CO2 emissions. *Proceedings of the National Academy of Sciences of the United States of America* 107:5687–5692.

[6] Wackernagel M, et al. (2002) Tracking the ecological overshoot of the human economy. *Proceedings of the National Academy of Sciences of the United States of America* 99:9266–9271.

[7] Hoekstra AY, Chapagain AK (2006) Water footprints of nations: water use by people as a function of their consumption pattern. *Water Resources Management* 21:35–48.

[8] Kitzes J, Wackernagel M (2009) Answers to common questions in Ecological Footprint accounting. *Ecological Indicators* 9:812–817.

[9] Hoekstra AY, Mekonnen MM (2012) The water footprint of humanity. *Proceedings of the National Academy of Sciences of the United States of America* 109:3232–3237.

[10] Ehrlich PR, Daily GC (1993) Population extinction and saving biodiversity. *Ambio* 22:64– 68.

[11] Ceballos G, Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296:904–907.

[12] Gaston KJ, Fuller RA (2008) Commonness, population depletion and conservation biology. *Trends in Ecology & Evolution* 23:14–9.

[13] Gaston KJ (2010) Valuing common species. *Science* 327:154–155.

[14] May RM, Lawton JH, Stork NE (1995) in *Extinction Rates*, eds Lawton JH, May RM (Oxford University Press, Oxford, United Kingdom), pp 1–24.

[15] Barnosky AD, et al. (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471:51–57.

[16] IUCN (2011) The IUCN Red List of Threatened Species. Version 2011.2.

[17] Miller R, Blair P (2009) *Input-output Analysis: Foundations and Extensions* (Cambridge University Press, Cambridge, United Kingdom) Vol. 328, Second edition, p 750.

[18] Birdlife International (2011) *Distribution maps of Birds of the World* (Cambridge, United Kingdom).

[19] Ramankutty N, Foley J (1999) Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13:997–1027.

[20] Gaston KJ, Blackburn TM, Klein Goldewijk K (2003) Habitat conversion and global avian biodiversity loss. *Proceedings of the Royal Society B* 270:1293–1300.

[21] Jones KE, et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90:2648.

[22] Smil V (2003) *The Earth's Biosphere: Evolution, Dynamics, and Change* (MIT Press), p 356.

[23] Barnosky AD (2008) Megafauna biomass tradeoff as a driver of Quaternary and future extinctions. *Proceedings of the National Academy of Sciences of the United States of America* 105:11543–11548.

[24] Vitousek PM, Ehrlich PR, Ehrlich AH, Matson PA (1986) Human appropriation of the products of photosynthesis. *BioScience* 36:368–373.

[25] Rojstaczer S, Sterling SM, Moore NJ (2001) Human appropriation of photosynthesis products. *Science* 294:2549–2552.

[26] Imhoff ML, et al. (2004) Global patterns in human consumption of net primary production. *Nature* 429:870–873.

[27] Haberl H, et al. (2007) Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 104:12942–12947.

[28] Erb KH, et al. (2007) A comprehensive global 5 min resolution land-use data set for the year 2000 consistent with national census data. *Journal of Land Use Science* 2:191–224.

[29] Rosenzweig ML (1995) *Species Diversity in Space and Time* (Cambridge University Press, Cambridge, United Kingdom), p 436.

[30] Drakare S, Lennon JJ, Hillebrand H (2006) The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters* 9:215–227.

[31] Harte J (2011) *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* (Oxford University Press, Oxford, United Kingdom), p 264.

[32] Narayanan GB, Walmsley TL (2008) *Global Trade, Assistance, and Production: The GTAP 7 Data Base* (Center for Global Trade Analysis, Purdue University).

[33] Hern WM (1998) How many times has the human population doubled? Comparisons with cancer. *Population and Environment* 21:59–80.

[34] Wiedmann T, Lenzen M, Turner K, Barrett J (2007) Examining the global environmental impact of regional consumption activities - Part 2: Review of input-output models for the assessment of environmental impacts embodied in trade. *Ecological Economics* 61:15–26.

[35] Wiedmann T, Wilting HC, Lenzen M, Lutter S, Palm V (2011) Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis. *Ecological Economics* 70:1937–1945.

[36] Hubacek K, Giljum S (2003) Applying physical input-output analysis to estimate land appropriation (ecological footprints) of international trade activities. *Ecological Economics* 44:137–151.

[37] Erb KH, Krausmann F, Lucht W, Haberl H (2009) Embodied HANPP: mapping the spatial disconnect between global biomass production and consumption. *Ecological Economics* 69:328–334.

[38] Haberl H, et al. (2012) Natural and socioeconomic determinants of the embodied human appropriation of net primary production and its relation to other resource use indicators. *Ecological Indicators* 23:222–231.

			POP		GDP		BWP		ALL
			R^2	Coeff	R^2	Coeff	R^2	Coeff	R^2
Total	Local	Bird Individs	0.37	1.52	0.07	0.56	0.87	0.94	0.88
		Mamm Individs	0.35	1.60	0.08	0.64	0.64	1.69	0.80
		Bird Range	0.35	0.38	0.05	0.12	0.86	0.62	0.92
		Mamm Range	0.55	0.57	0.08	0.18	0.86	0.69	0.92
	Cons	Bird Individs	0.90	0.88	0.44	0.52	0.40	0.24	0.94
		Mamm Individs	0.83	0.82	0.46	0.52	0.57	0.53	0.91
		Bird Range	0.65	0.94	0.33	0.56	0.61	0.97	0.91
		Mamm Range	0.79	0.88	0.31	0.47	0.61	0.75	0.91
Per-person	Cons	Bird Individs	-	-	0.31	0.17	0.01	0.02	0.52
		Mamm Individs	-	-	0.29	0.22	0.16	0.18	0.63
		Bird Range	-	-	0.06	0.17	0.40	0.47	0.55
		Mamm Range	-	-	0.05	0.10	0.38	0.34	0.62

Table 5.1: Results of log-log linear regressions of wildlife population losses and regional human population (POP), gross domestic product (GDP), baseline wildlife populations (BWP), and all three variables with interactions (ALL) ($n = 113$). Rows indicated by "Local" indicate losses occurring within the borders of a region, and rows indicated by "Cons" indicate the regional wildlife footprint (i.e., wildlife population losses occurring anywhere in the world due to the consumption of a region). Wildlife populations are measured in units of bird individuals, mammal individuals, bird fractional-ranges, and mammal fractional-ranges. Per-person regressions use per person GDP and BWP as predictor variables. $P < 0.02$ for all reported coefficients except for regression of per person consumption of bird individuals and BWP ($P = 0.2$). All logarithms are natural log. BWP estimates have a value of one added before log transformation to avoid log of zero errors.

Purchase	Bird Individuals	Mammal Individuals
US $10 on bovine meat products	0.13	6.57
US $10 on dairy products	0.05	2.57
US $10 on sugar	0.03	1.68
US $10 on vegetables, fruit, nuts	0.03	1.04
US $100 on general retail purchases	0.02	0.98
US $100 on petroleum products	0.01	0.70
US $100 on electronic equipment	0.01	0.66
US $100 on electricity (global average mix)	0.01	0.65
US $100 on air travel	0.01	0.64

Table 5.2: Estimates of the land use driven losses of bird and mammal individuals associated with various consumer purchases, calculated using global average wildlife population loss intensities (Eq. 5.2). Losses represent a one year decrease in global wild bird or mammal populations caused by each purchase. Results should be interpreted in light of the relatively coarse global resolution and various limitations of this study (see *Discussion*).

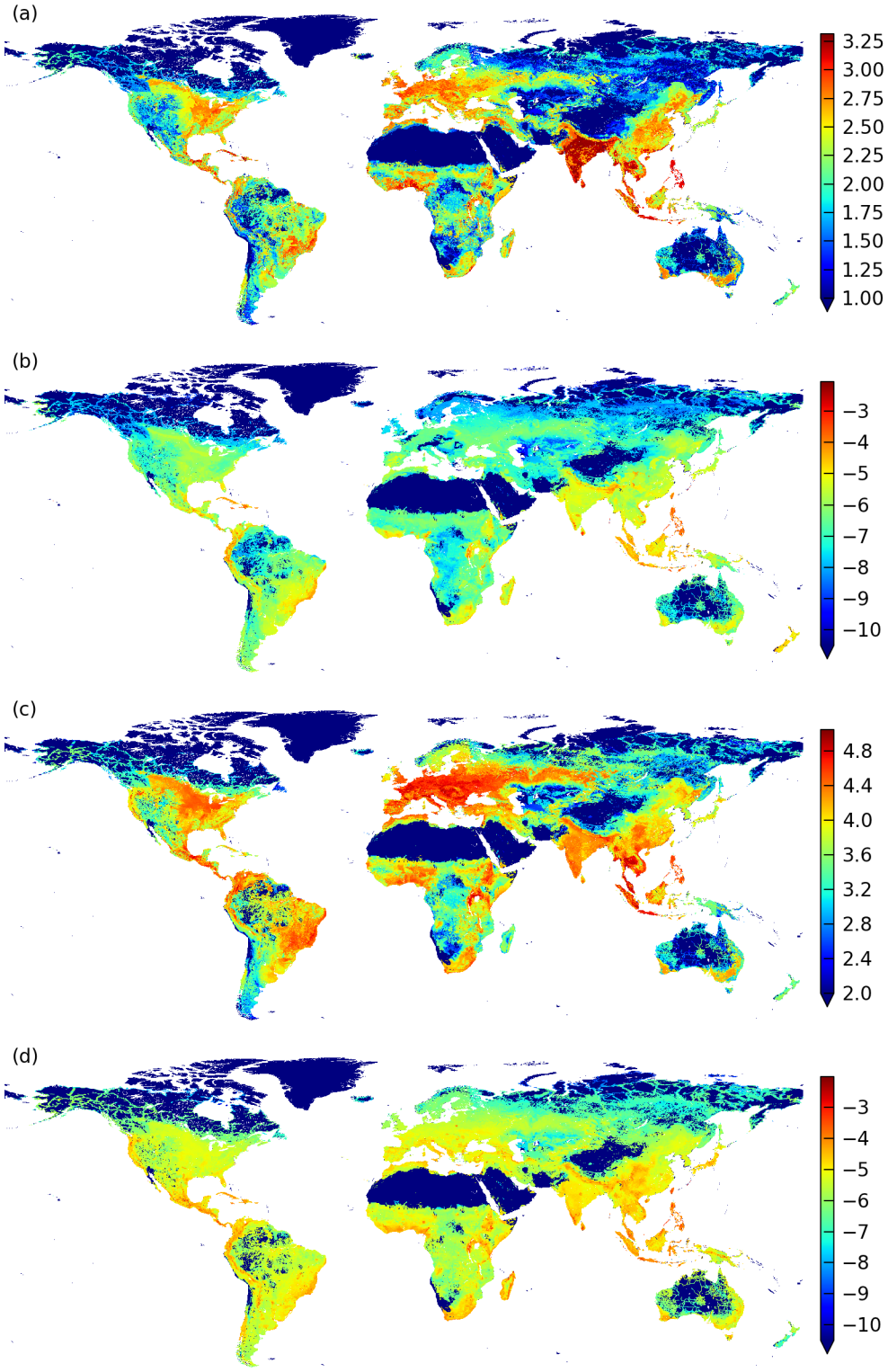

Figure 5.1: Wildlife population losses per km^2 caused by all human land use activities, measured in units of (a) bird individuals, (b) bird fractional-ranges, (c) mammal individuals, and (d) mammal fractional-ranges. All data are $\log_{10}$.

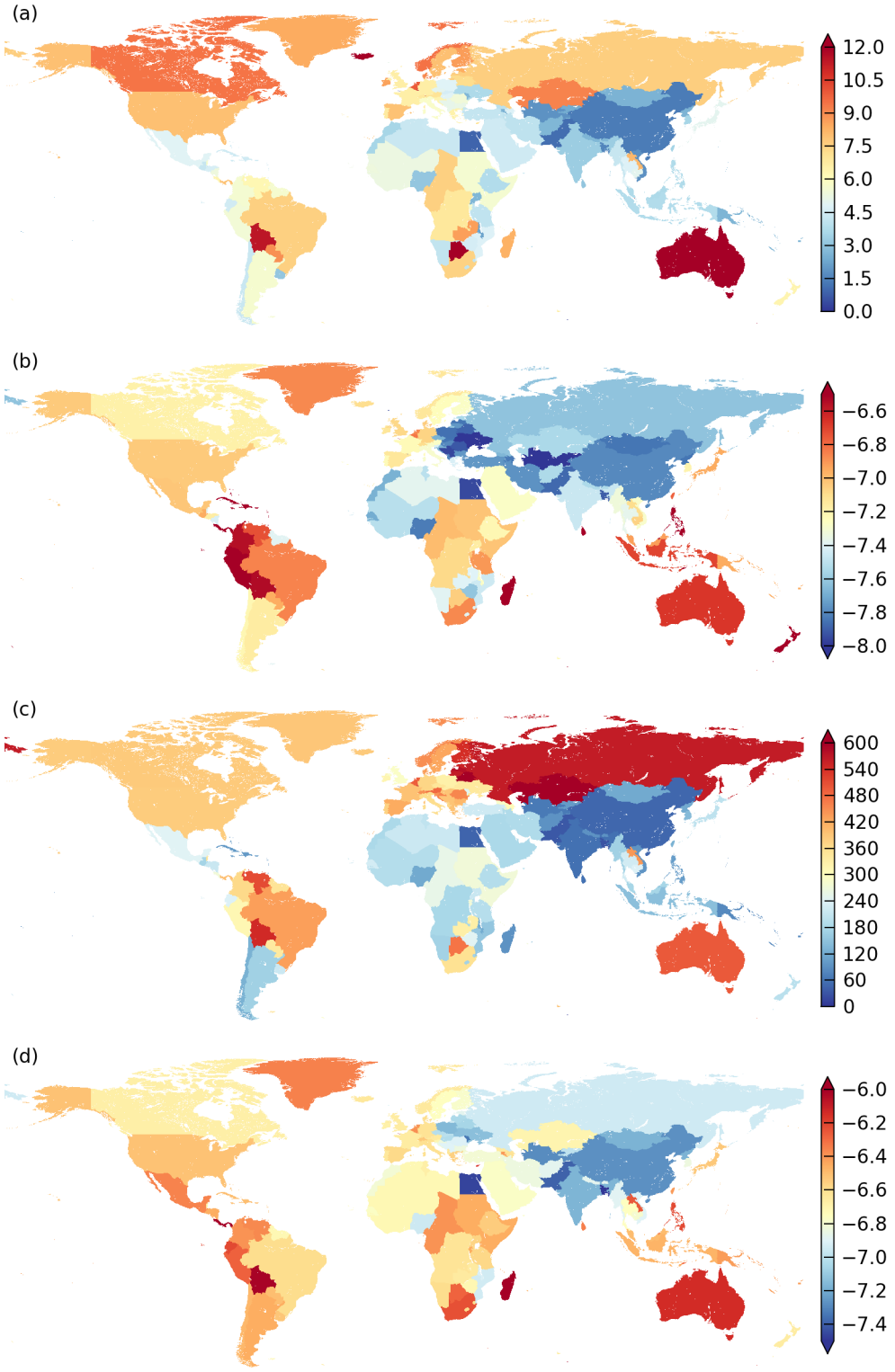

Figure 5.2: Per-person wildlife footprints by region, measured in units of (a) bird individuals, (b) bird fractional-ranges ($\log_{10}$ transformed), (c) mammal individuals, and (d) mammal fractional-ranges ($\log_{10}$ transformed).

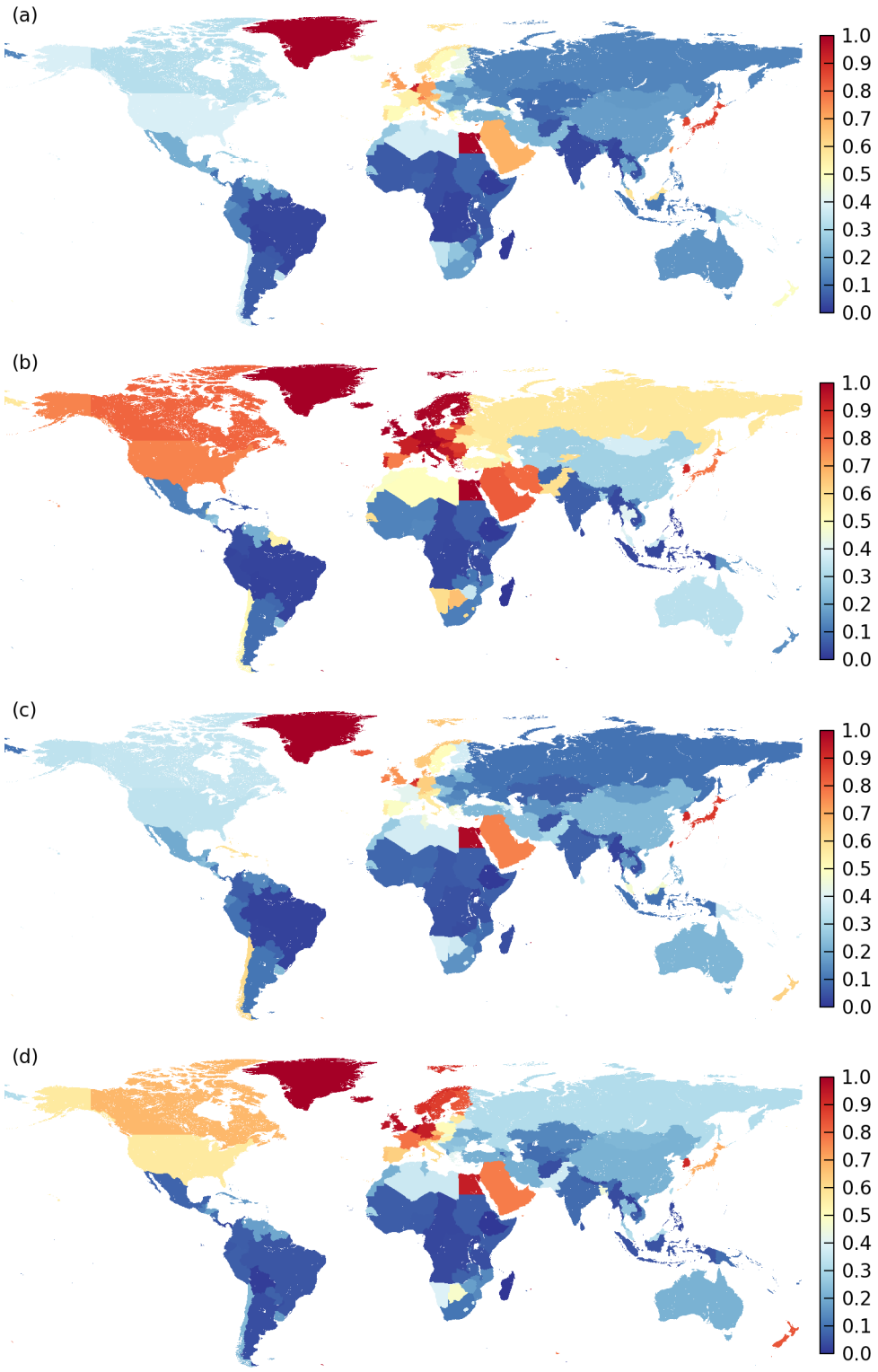

Figure 5.3: Imports of embodied wildlife footprint, measured as the fraction of a region's wildlife footprint that represents wildlife population losses occurring outside of that region's borders, measured in units of (a) bird individuals, (b) bird fractional-ranges, (c) mammal individuals, and (d) mammal fractional-ranges.

Appendix

Supplemental maps

Figure (5.A1) shows the global distribution of baseline wildlife populations, with populations measured as bird individuals, bird fractional-ranges, mammal individuals, and mammal fractional-ranges. Figure (5.A2) shows the global intensity of human land use activities leading to population losses. Figure (5.A3) shows the total wildlife footprint of each region.

Wildlife population losses by economic sector

Figure (5.A4) shows the global wildlife population losses associated with consumer purchases from 16 aggregated global economic sectors.

Initial allocation of wildlife population losses to economic sectors

Wildlife population losses are initially allocated to the sectors that are the direct causes of those losses (i.e., the sectors that are the first "point of entry" where embodied wildlife population losses enter the global economy). To allocate direct wildlife population losses to economic sectors, total wildlife population losses within each region are first calculated and divided among four major land use activities (crop production, grazing, forestry, and infrastructure areas) as described in *Methods*. The wildlife population losses associated with each of these four land use activities within each region are then allocated among the 57 economic sectors in each region as follows:

- Crop production. Wildlife population losses due to crop production were divided among eight crop producing sectors (1 Paddy rice, 2 Wheat, 3 Cereal grains nec, 4 Vegetables, fruit, nuts, 5 Oil seeds, 6 Sugar cane, Sugar beet, 7 Plant-based fibers, 8 Crops nec) based on the fraction of harvested area in each cell devoted to each crop group, as calculated from data in [1].
- Grazing. Wildlife population losses due to grazing were divided between two unprocessed livestock product sectors (9 Bovine cattle, sheep and goats, horses, 11 Raw milk) in proportion to the economic output of these sectors within a region.
- Forestry. All wildlife population losses due to forestry activities (harvest of wood) were allocated to the forestry sector (3 Forestry).
- Infrastructure areas. Wildlife population losses due to areas occupied by infrastructure (primarily buildings and roads) were divided among all 57 sectors in each region in proportion to the total economic output of each sector within that region.

This process was used to create the f^d vector described in *Methods*.

Calculation of mammal population density

Data from the PanTHERIA database [2] were used to determine population density for mammal species. Population density data (individs / km^2) were extracted directly from PanTHERIA (WR93 and WR95 datasets) for 908 species. For species without population density data but with adult body mass data (2,749 species), densities were estimated from an allometric relationship between body mass (g) and population density (individs / km^2) [see also 3–5].

$$\log_{10}\text{Density} = 3.88\ (\text{SE}\ 0.07) - 0.74\ (\text{SE}\ 0.02)\log_{10}\text{Mass} \tag{5.A1}$$

Eq. (5.A1) was constructed from the subset of all species in the PanTHERIA WR05 dataset with both population density and body mass data (927 species).

Species without population density or body mass data (1,619 species, representing 31% of mammals with IUCN range maps) were excluded from subsequent analyses using the mammal individuals measurement metric (all mammal range maps were used for mammal fractional-range analysis).

References

[1] Monfreda C, Ramankutty N, Foley, JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles* 22:1-19.

[2] Jones KE, et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90:2648-2648.

[3] Damuth, J (1981) Population density and body size in mammals. *Nature* 290:699-700.

[4] Peters RH, Raelson JV (1984) Relations between individual size and mammalian population density. *The American Naturalist* 124:498-517.

[5] Brown, JH (2004) Toward a metabolic theory of ecology. *Ecology* 85:1771-1789.

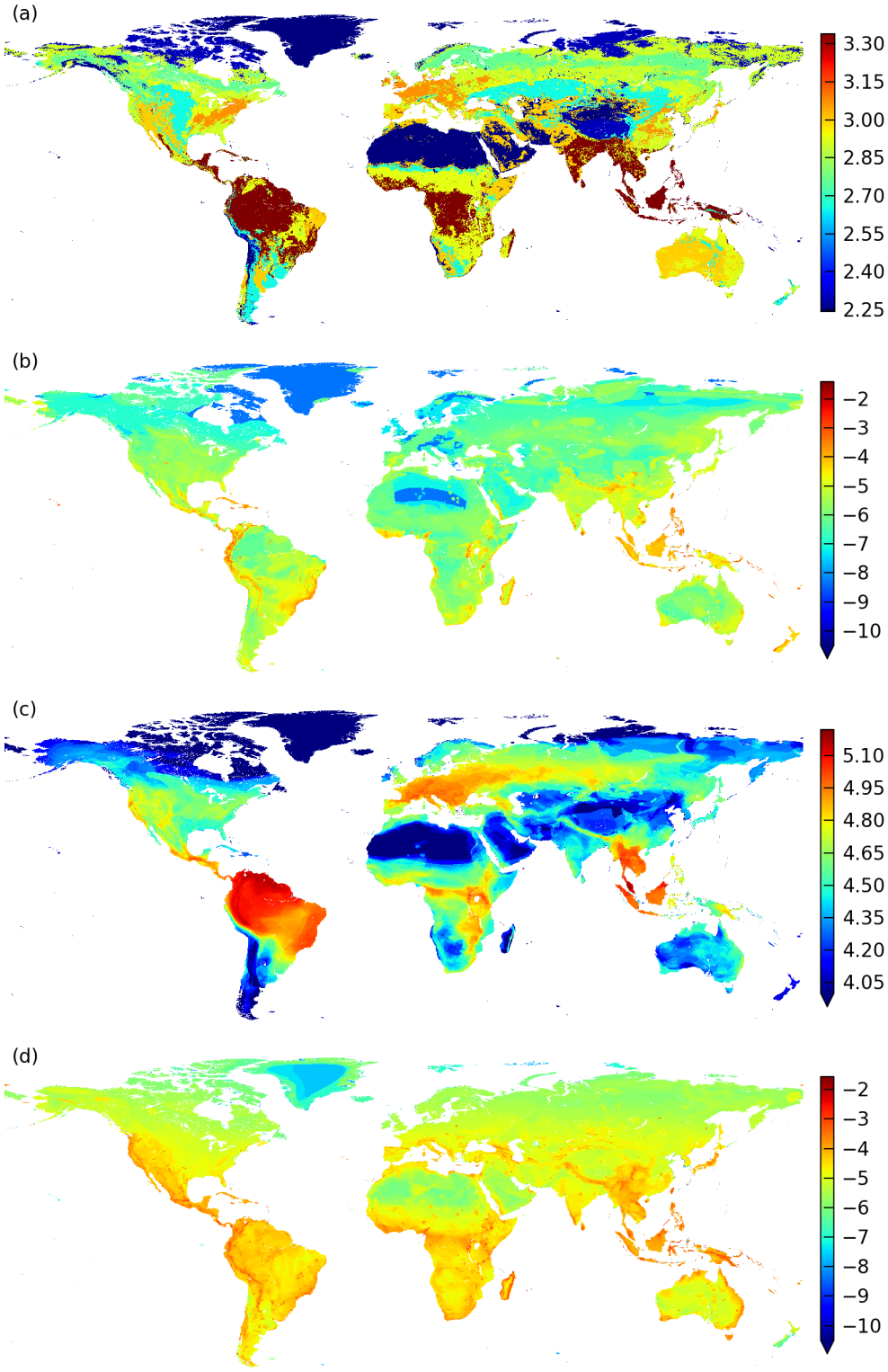

Figure 5.A1: Global estimated baseline wildlife populations per km^2 in the absence of human land use activities, measured in units of (a) bird individuals, (b) bird fractional-ranges, (c) mammal individuals, and (d) mammal fractional-ranges. All values are $\log_{10}$.

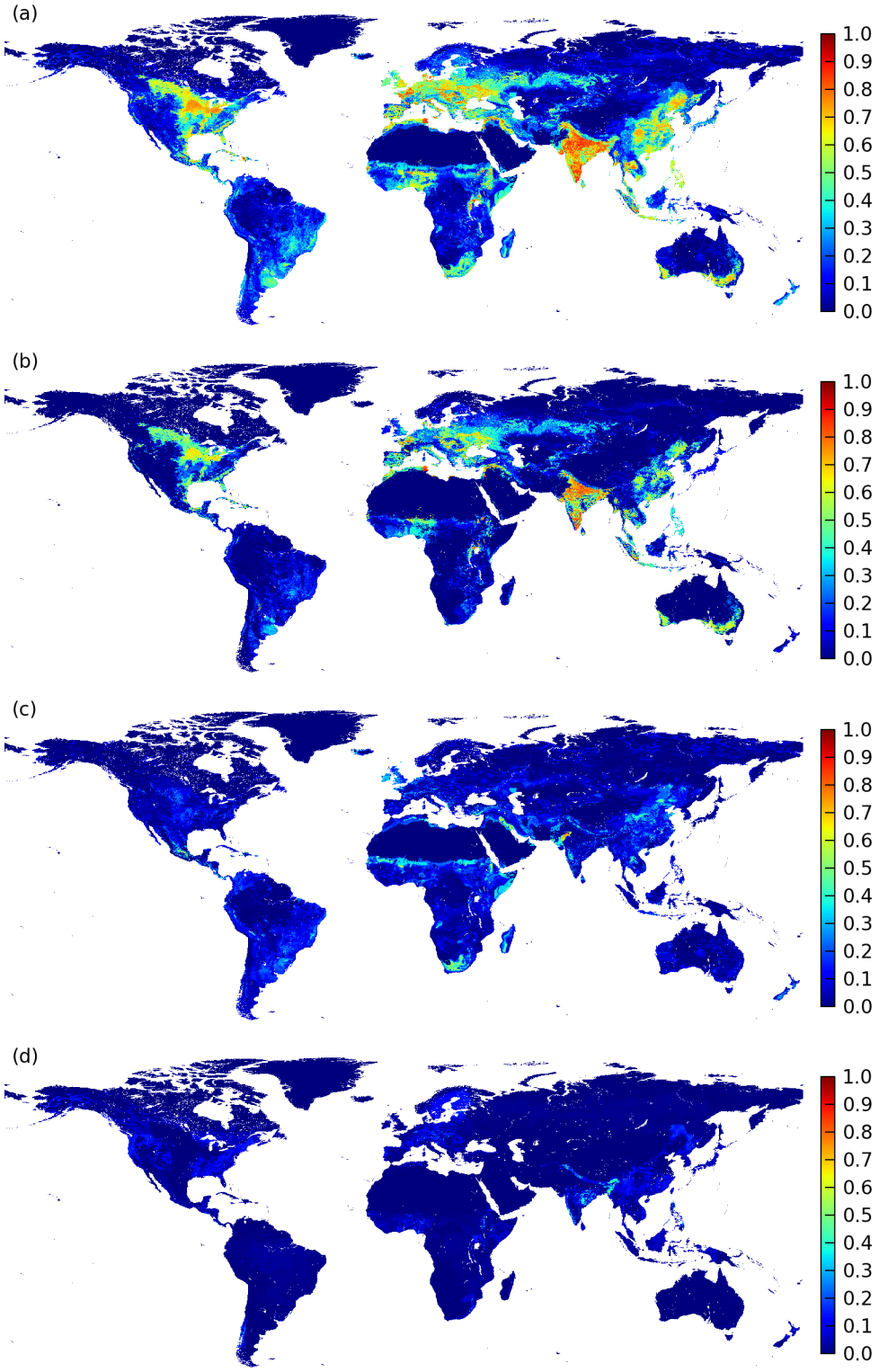

Figure 5.A2: Fractional wildlife population losses caused by different human land use activities. Maps show fractional losses for (a) all activities, (b) crop production, (c) grazing, and (d) forestry. The map of losses due to the presence of infrastructure areas (buildings and roads) is not shown.

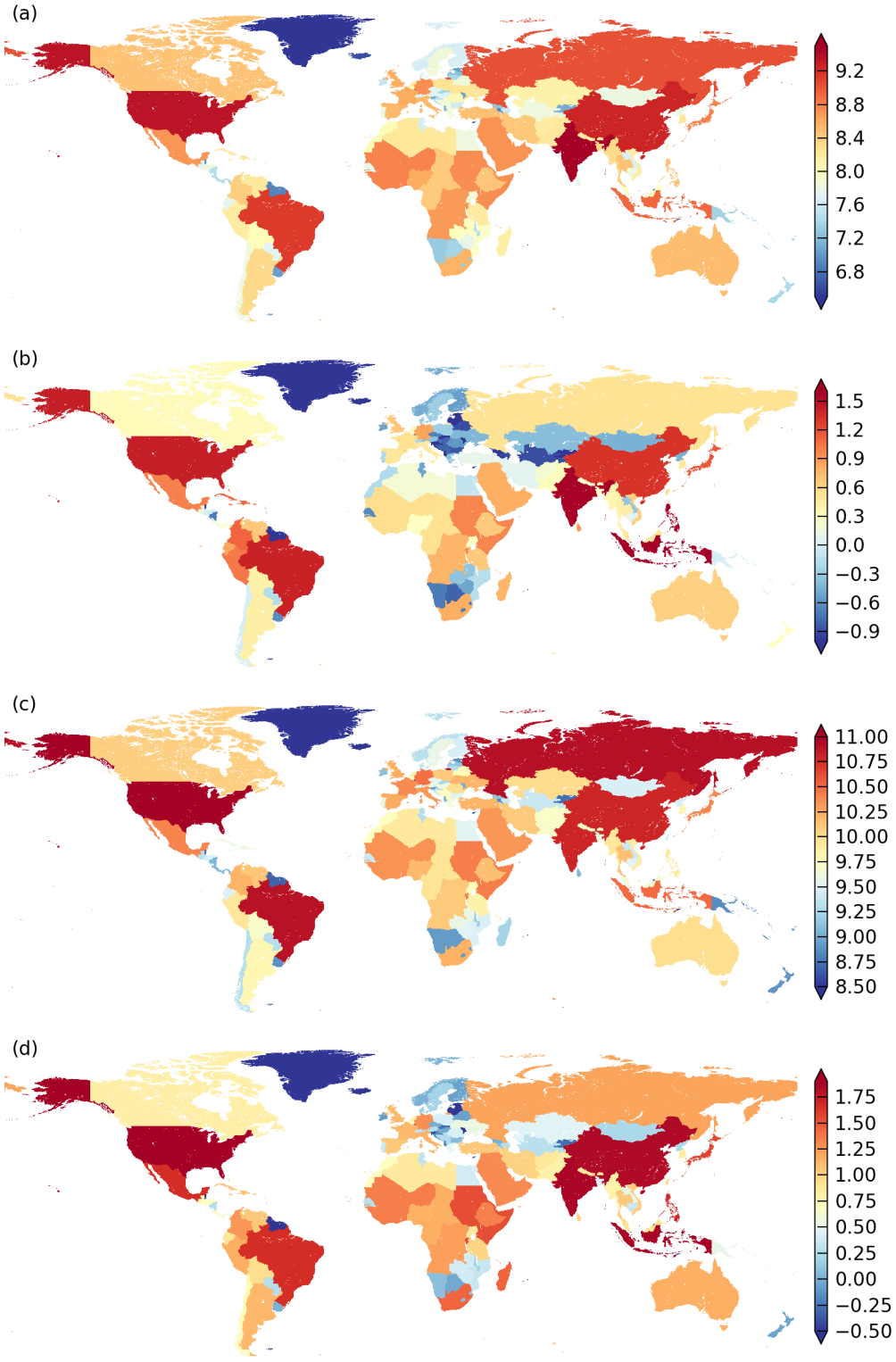

Figure 5.A3: Total regional wildlife footprint, measured in units of (a) bird individuals, (b) bird fractional-ranges, (c) mammal individuals, and (d) mammal fractional-ranges. All values are $\log_{10}$.

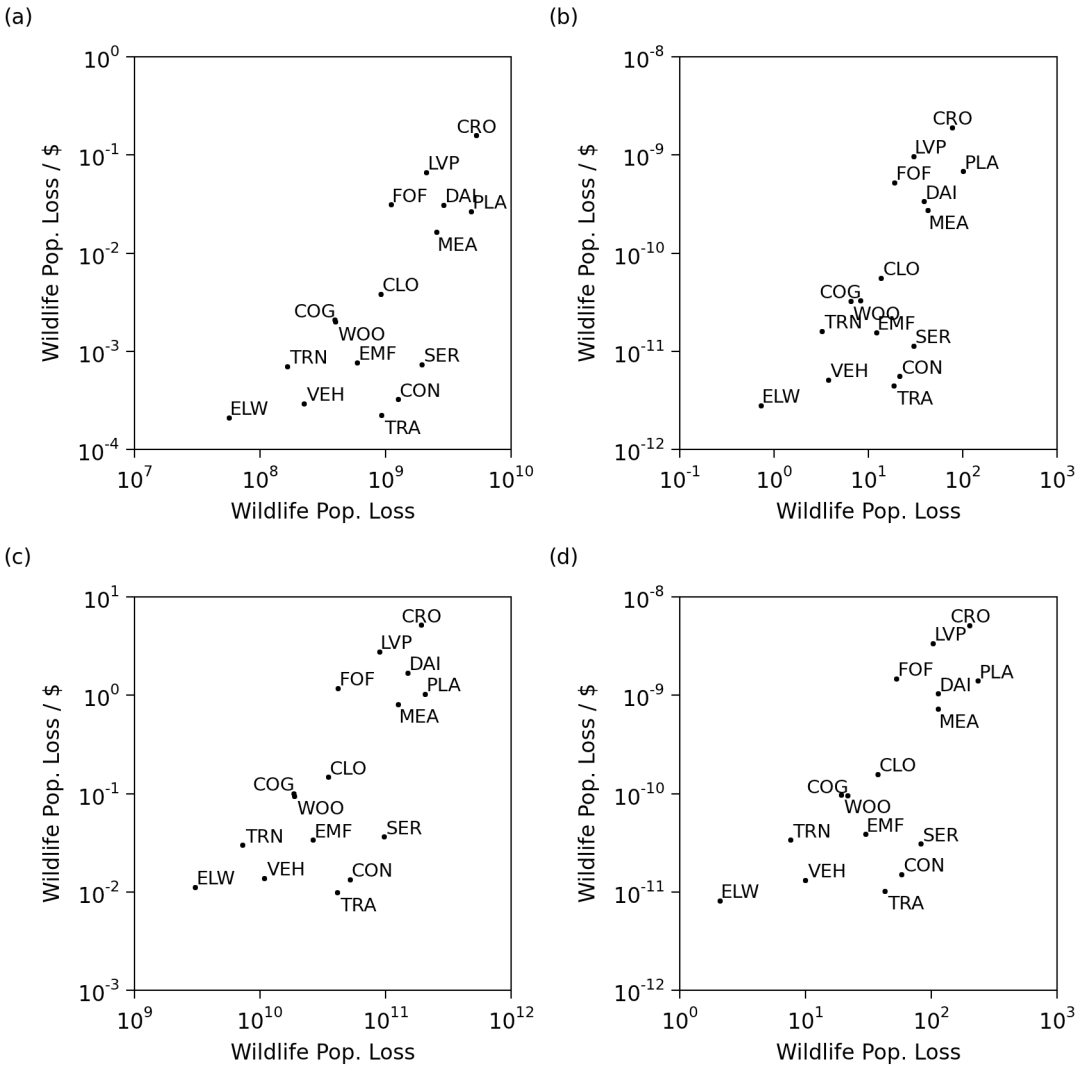

Figure 5.A4: Total wildlife population losses and wildlife population losses per dollar of final consumption in each of 16 aggregated global economic sectors, measured in units of (a) bird individuals, (b) bird fractional-ranges, (c) mammal individuals, and (d) mammal fractional populations. Sector abbreviations: CRO = Raw crop products, LVP = Raw livestock products, DAI = Dairy, PLA = Processed plant products, MEA = Processed meat products, CLO = Clothing and textiles, COG = Coal, oil, gas, minerals, chemicals, WOO = Wood and paper products, TRN = Transport, ELW = Electricity and water, VEH = Vehicles, EMF = Electronics and manufactured products, CON = Construction, TRA = Trade, SER = Services, government, and other.

Chapter 6

Conclusion

This dissertation has developed several new approaches in quantitative ecology that can be used to measure and predict species richness, abundance, and extinction in human-altered landscapes. A number of particularly important future extensions to these approaches are highlighted below.

The reserve design models presented in Chapter 2 could be improved by adding spatial autocorrelation to the proposed metapopulation models, which would generally decrease the value of reserve site clustering for all species. A second important improvement would be to adjust the models to allow individual organisms to use multiple habitat patches separated by short distances, a change that could perhaps substantially decrease predicted extinction risk for large bodied, low density species. Beyond these improvements, maps of landscape permeability could be used to modify species' dispersal probabilities between habitat patches when such maps are available. In addition to patch selection, the approach described here could also be used for linkage selection by using optimization methods to choose the set of linkages between patches whose presence most increases the probability of species' persistence. A final natural and practical extension of this analysis would be to integrate the proposed extinction-based reserve design approach with widely used representation-based tools, such as MARXAN, to create an integrated method that could be used to maximize the number of species expected to be present in a reserve network at a future time horizon. To accomplish this integration, however, the simulation methods presented here will likely need to be replaced with analytical approximations to increase the speed at which extinction risk can be estimated for many species.

The new extinction metrics presented in Chapter 3 are most limited by the lack of an explicit treatment of how extinction risks and rates are affected by the number and configuration of the habitat patches remaining after habitat loss. Incorporating habitat fragmentation into these metrics will likely require explicit macroecological predictions of patterns of beta diversity, predictions that have not yet been derived within the maximum entropy theory of ecology. Future research might be directed towards calculating these new extinction metrics using different forms of the underlying species abundance distribution and species-level spatial abundance distribution, with the specific goal of identifying the underlying

ecological and statistical assumptions to which extinction predictions are most sensitive. Further evaluation of these extinction metrics using empirical data should also be carried out where possible.

The study of the road ecology of bats presented in Chapter 4 could have been improved substantially if temperature measurements had been recorded at each individual sampling location. These data could replace the daily maximum regional temperature variable in statistical analyses, potentially leading to more robust inferences about the relationship between temperature and bat activity. Additional research should be directed towards attempting to identify the specific causes of the observed road effect, which might include differences in small scale habitat structure, abiotic conditions, or noise levels between points near and far from roads. Similar research should also be carried out in habitats other than salt marsh, and with species other than those present in northern California, to examine whether the road effect observed here is also present for other species and in other landscapes.

The accuracy of the global wildlife footprint calculations presented in Chapter 5 is limited by the availability, accuracy, and resolution of global-scale data on species' ranges, human appropriation of net primary productivity, and monetary flows between global economic sectors. In particular, the 57 sector level of aggregation in the monetary input-output tables used in this analysis limit the resolution with which different types of consumption activities can be distinguished, and future analyses may wish to incorporate higher resolution multi-regional input-output tables that are currently being developed. Given the communication value of footprint metrics and their ability to raise public awareness, a useful follow-up to this analysis would be to develop non-academic materials, such as an interactive website or literature directed towards general audiences, to popularize the concept of the wildlife footprint and the results presented here.