NON-NATIVE SPECIES DISTRIBUTIONS IN SPACE AND TIME: INTEGRATING ECOLOGICAL THEORY AND PREDICTIVE MODELING

by

Tyler Jacob Brummer

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Land Resources and Environmental Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2012

APPROVAL

of a thesis submitted by

Tyler Jacob Brummer

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Lisa J Rew

Approved for the Department of Land Resources and Environmental Sciences

Tracy Sterling

Approved for The Graduate School

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

Tyler Jacob Brummer

April 2012

DEDICATION

I would like to dedicate this thesis to my outstanding advisor, family and friends for their continual support and encouragement throughout my graduate studies. Without their consistently positive energy the process would have seemed insurmountable.

ACKNOWLEDGEMENTS

I would like to acknowledge first and foremost my advisor Lisa Rew. Without her mentoring, support, encouragement, friendship and tireless work, this thesis would not have been completed and my development as a scientist would have been much diminished. I would like to thank Bruce Maxwell for providing a collaborative environment that forced me to grow as a thinker and raise my personal bar of achievement. Matt Lavin provided countless hours of discussions about ecology, challenged my world view and was great company at our home away from home in Atomic City, Idaho. And thanks to Megan Higgs who fostered my knowledge of statistics and was always available to improve my understanding of statistical modeling and inference. I could not have asked for a more involved and high quality committee.

I owe a debt to my unbeatable field crews that put up with long days, inclement weather, and isolation in the desert of Idaho. Alex Wing, Kim Taylor, Rose Marks, Elizabeth Usher, Kevin Shields, Tyler Schober and Dan Campbell helped collect one of the best datasets a graduate student could ask for and provided weeks of entertainment in a town of about 24 people. Finally, thanks to my fellow graduate students Fred Pollnac, Melissa Bridges, Tanya Skurski and Patrick Lawrence who provided moral support and great camaraderie over the last three years.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES - CONTINUED

Figure Page

ABSTRACT

Invasive plant species are perceived as a problem globally, but management occurs locally. Theoretical developments concerning the distribution of plant species and invasions have generally focused on coarse scales, with relatively little work performed at finer scales relevant to local landscape management. The majority of such model predictions are static reflections of current conditions. As species invasions are a temporally variable process, the need for tools to predict invasions through both space and time are vital. Thus, this thesis explored distribution models to predict non-native species occurrence, investigated sources of uncertainty in these models, and quantified the key drivers of non-native species metapopulation dynamics.

Sampling methodology and sample size requirements to inform logistic regression models used to predict invasive species realized distributions were evaluated with simulation and empirical data from two multi-species surveys. Transect sampling was the most efficient way to generate species occurrence data and consequently landscape scale species distribution models. Empirical and simulation modeling results indicated minimum sample sizes of between 0.06 and 0.23% of the study area to maximize model predictive ability, independent of site characteristics. However, landscape scale species distribution models were more predictive at sites with steeper environmental gradients and for species at their ecophysiological range limits. Detection error in the plant species surveys was also quantified, as well as its effects on predictions and uncertainty of species distributions. Detection error did not practically alter predictions, nor model based uncertainty estimates if the detection rate was greater than 87%. However, at lower detection rates care needs to be taken when interpreting response variables and prediction certainty. Finally, multi-season repeat survey data were used to investigate the key spatial and temporal drivers of non-native species colonization and extinction. The drivers of non-native species colonization and extinction were sometimes simple but other times resulted from complex interactions between dispersal, disturbance, habitat and temporal climatic variation. Overall, these results demonstrate the need for reliable species occurrence records and monitoring data to fully characterize species distributions at present and their dynamics resulting in future distributions.

CHAPTER ONE

INTRODUCTION TO THESIS

Darwin (1859) recognized and noted exotic (non-native, non-indigenous) species invasions, and nearly a century later the field of invasion ecology was popularized by Elton (1958) and has since become a sub-discipline of the field of ecology (Davis et al., 2001; Pysek & Hulme, 2009; Richardson & Pysek, 2008). The impetus to study exotic species has varied from understanding ecological processes, mitigating the effects of human impacts on ecosystem structure and function, to reducing direct economic losses due to undesirable plant species. A large economic investment is made annually to reduce the abundance and distribution of exotic species, but there remains a lack of consensus about the ecological processes that predict invasive species success, the impacts of these species and how society should best manage these species in dynamic systems. Presented here is a brief introduction to the theory of why exotic species become invaders, the impacts of these species in their introduced ranges, invasion as a process, and invasive plant management.

Why do Species Become Invasive?

Generally, there are two dominant schools of thought about the importance of plant species' native versus non-native status in determining the success of exotic plants in new ranges. The first school places much importance on a species biogeographical origin (*sensu* Hierro et al., 2005), citing theories such as Darwin's naturalization hypothesis (Darwin, 1859), enemy release (Liu & Stiling, 2006), anthropogenic induced adaptations to invade (AIAI) (Hufbauer et al., 2012), evolution of increased competitive ability (EICA) (Blossey & Notzold, 1995), occupation of empty niches (MacDougall et al., 2009; Mack et al., 2000), and novel weapons (Callaway & Ridenour, 2004). Evidence for EICA and novel weapons remains weak (Hierro et al., 2005), AIAI is supported mostly by theory and a few suggestive examples, while examples of empty niches have been shown (e.g. *Bromus tectorum* in the Great Basin of the United States) and linked to community assembly theory (Shea & Chesson, 2002).

The second school of thought maintains that the dichotomy between native and non-native species is more anthropological in origin than biological (Davis et al., 2001). By removing the non-native/native designations to plant species and evaluating them as simply a part of the greater species pool, research and theory suggest that invasive species are simply passengers, responding to the increase in disturbance frequency and intensity, pre- and post- industrialization. Davis (2009) suggested these species are bound by community assembly theory, in that both neutral (Hubbell, 2001; MacArthur & Wilson, 1967) and niche (Grime, 1977; MacDougall et al., 2009) related processes determine their distribution and abundance regardless of biogeographical origin. The implicit assumption being that processes such as enemy release have little overall effect and that adaptation is unlikely to occur at such short temporal scales as suggested by EICA and AIAI.

The reality is likely a combination of these theories. Stochastic neutral processes of dispersal and establishment, resource-based processes (i.e. the niche), and differential fitness due to a release from coevolutionary pressures interact, depending on the species and the circumstances of the specific introduction. Regardless of the cause of a species invading a new range, ca. 2100 vascular plant species have been classified as exotic in the conterminous US (Vitousek et al., 1997) and vary in their distribution and abundance and stage of invasion. Thus, in the hope of understanding the future distribution and abundance of exotic species, a framework for studying invasion as a spatio-temporal process is necessary.

Invasion as a Process

Theoretical work on biological invasions has attempted to develop a conceptual working model to unify how scientists view invasions. Important stages have been compartmentalized to facilitate reductionist research and target management. Blackburn et al. (2011) proposed a framework for biological invasions that unifies those of Richardson et al. (2011) and Williamson (1996). They proposed a stage-based barrier model with four stages: transport, introduction, establishment, and spread. The barriers described were geography, captivity or cultivation, survival, reproduction, dispersal and environment. A species is considered invasive if it makes it to the spread stage and can be subject to invasion failure at any one of the other three stages. They suggest management can be preventative in the first two stages, while containment can occur during the establishment stage. If the species becomes invasive, mitigation is the suggested management strategy. Eradication is stated to be a theoretical option throughout all four stages.

While this general framework serves as an organizing mechanism, most management of invasive species occurs during the spread phase. Thus, we must understand the processes that drive the spread phase in order to predict the outcome of introductions. Some of the original work in developing mechanistic spread models for introduced species focused on dispersal, disturbance, and environmental characteristics and their interactions as the key drivers of invasion (Higgins et al., 1999). This fits into a general framework of the factors influencing invasion success suggested by Catford et al. (2009), who proposed that invasions can be explained under three broad processes of propagule pressure, abiotic, and biotic factors. Most research, especially empirical studies or simulation studies parameterized with empirical data, have concluded that invasions are context specific (Daehler, 2003; Higgins et al., 1999; Williamson & Fitter, 1996) but that individual invasions can be predictable (Higgins et al., 2001). This implies the need for site and species specific data to understand the drivers of local invasions and the dynamics of these invasions to better inform management. Generalizations or general models have generally failed to predict the spatial and temporal dynamics of invaders thus far (Hastings et al., 2005).

Developments in general population ecology on source-sink (Pulliam, 1988) and metapopulation dynamics (Hanski, 1998, 1999) provided both theoretical and model-based frameworks to understand processes important to the extinction, persistence or spread of species on the landscape. The need for efficient ways to gather data to parameterize interpretable models about species' occupancy patterns spurred the development of occupancy based models of occurrence and metapopulation dynamics

(MacKenzie, 2006). Such models allow for a rigorous, statistical evaluation of factors influencing a species metapopulation dynamics, allowing hypotheses about key drivers to be tested and predictions to be made in both space and time. Although these models have been applied to numerous studies on animals, few examples exist of using this approach to explain plant invasions and predict future distributions.

The Impact of Invasions: Perceptions, Evidence, and Legislation

Given that species invade and that anthropogenic drivers are increasing migration rates, why is there the perception that these introductions need to be managed? A number of impacts of invasive species are suggested in the primary literature including impacts on fire regimes (Brooks et al., 2004), altering various biogeochemical properties of ecosystems and biotic properties (Vilà et al., 2011), and increasing extinctions of native species (Gurevitch & Padilla, 2004) but for the most part quantitative evidence is tenuous (Gurevitch & Padilla, 2004; Smith et al., 2006) and meta-analyses are subject to the shortfalls of the original experimental design of the aggregated studies. An alternative to the species centric view of negative impacts is that these species are simply passengers (Bauer, 2012; Crosby, 1986) or indirect drivers (Bauer, 2012; Vitousek et al., 1997) of ecosystem change as opposed to the direct cause. This view is strongly promoted by Davis (2009) and others that subscribe to invasion theory grounded in resource and disturbance theory (see Catford et al., 2009 for review).

Although contentious, the idea that invasive species negatively impact ecosystems and have direct economic consequences is widely accepted by society at large. Direct economic impacts due to yield loss associated with undesirable plant species has been estimated at $27 billion per annum in the United States agricultural sector with direct herbicide costs estimated at $4 billion per year. The economic impact of invasive species in rangelands was estimated to be $1 billion annually, coupled by a cost of $5 billion per year in treatment (Pimentel et al., 2005). These lines of evidence coupled with aggressive public outreach campaigns culminated in a presidential executive order (#13112) issued in 1999 by President Bill Clinton that mandated the management of federally listed exotic species. Thus, federal agencies that operate on public land were legislated to manage these species subject to appropriations from the federal government. And, any further development of a site has to consider the consequences of construction on invasive plant populations under the National Environmental Policy Act.

Given the Mandate to Manage, How do We Solve Such a Vast Problem?

The previous sections have demonstrated that plant invasions are a function of different ecological mechanisms and have the potential to impact ecosystems in their introduced range. But, such invasions are driven by complex ecological processes, and have a social and legislative mandate to be managed. However, the question remains of how to best physically manage invasive species. Although federal agencies are mandated to manage federally listed noxious plant species, annual budgets are finite and the number of infested acres is large. A body of research has developed in recent decades developing management prioritization frameworks (Rew et al., 2007), quantitative decision support tools (Hauser & McCarthy, 2009; Regan et al., 2011), and cost-benefit analysis for different aspects of the management process (Maxwell et al., 2009).

Rew et al. (2007) proposed a population based prioritization framework for invasive species where management goals are defined, inventory/survey of target species conducted, populations are selected for monitoring, and the most invasive populations are targeted for management. Maxwell et al. (2009) evaluated alternative management strategies, suggesting monitoring to improve management efficiency. Although finding recent invaders and executing early detection rapid response (EDRR) management may be perceived as the most effective strategy, most species on the noxious weed lists are well established and may be passed the EDRR stage. Indeed they are in the spread stage of the invasion as defined by Blackburn et al. (2011). Consequently, federal management agencies such as the United States National Park Service, Forest Service, Fish and Wildlife Service, Bureau of Land Management as well as many state agencies are tasked with developing plans to target available resources to decrease federal or state listed species distribution and abundance. But, the crux of this task resides in knowing where the target species are on the landscape, and which species and populations within a species to prioritize.

Habitat suitability, species distribution, probability of occurrence and occupancy modeling are all slight variations of tools that relate species occurrence (and non-occurrence) to environmental covariates, to make predictions about species locations. Popularized in the 1990’s with increased computation power coupled with user-friendly geospatial analysis software, species distribution modeling has exploded in popularity

(Araújo & Guisan, 2006; Elith & Leathwick, 2009; Franklin, 2009; Guisan & Thuiller, 2005; Guisan & Zimmermann, 2000). The major application of these tools has been to model native species distributions from herbaria records and floristic studies and predict country to continent wide distributions at course spatial resolutions as a function of bioclimatic variables (Graham et al., 2004). These studies rely on presence only data, as unbiased studies that would include absence data are generally not available. Further, their sampling frames are generally so large (e.g. 1 km^2), it is difficult for modelers to assume absence even when the species is not observed. These types of models have been used to inform conservation and environmental policy by predicting the potential effects of climate on species distribution (Engler et al., 2009; Pearson & Dawson, 2003).

Species distribution models have more recently been promoted as tools to be used in the context of invasive species management frameworks (Crossman & Bass, 2008; Giljohann et al., 2011; Rew et al., 2007; Stohlgren & Schnase, 2006). Modeling non-native species realized distributions from presence and absence data using logistic regression techniques can provide estimates of mean occurrence of the target species and can be used to prioritize species and their populations within a site for management (e.g. target the least frequent species as the likelihood of eradication is higher or target populations in the most suitable habitat under the assumption that they are likely to be sources of further invasion). These models can also spatially differentiate where the species is currently using to decrease search area when management is enacted. However, these models have been criticized for violating certain assumptions of the species distribution modeling framework, namely the assumption that the species is in equilibrium with its environment (Araújo & Pearson, 2005). Bioclimatic models implicitly assume that species distributions are a function of abiotic ecophysiological controls over a species individual fitness. While perhaps a valid approach at broad geographic extents (continent to global), to predict invasive species distributions at landscape scales, the focus must be redirected and the interpretation of the model refined. As outlined in the section above, there are a few ecological processes that are known to drive invasions and through intelligent covariate construction many of these processes can be accounted for. Resource gradients will differentially favor species based on their inherent ecological niche (e.g. wet to dry; warm to cold; disturbed to stable). Although measuring precipitation and temperature and other direct controls on species distribution at fine enough spatial resolutions to make useful predictions for managers is logistically infeasible, landscape derived covariates that are collinear with those mechanistic controls can be derived from digital elevation models and remote sensing data. Such variables can be broadly defined as habitat.

Dispersal is another key driver of species distribution and is highly dependent on the stage of the invasion (Václavík & Meentemeyer, 2012) as the degree to which the species has been able to explore the landscape will vary. Although difficult to account for natural dispersal processes (but see Václavík & Meentemeyer, 2009), it has been shown that roads are a major vector for species invasions (Gelbard & Belnap, 2003; Parendes & Jones, 2000; Seipel et al., 2012) as well as other anthropogenic disturbances. The effects of these can be accounted for with proximity variables from available geospatial data.

Although applying species distribution modeling at landscape scales is a promising management tool, applications at fine resolutions are limited (but see Crossman & Bass, 2008; Rew et al., 2005; Shafii et al., 2003; Zimmermann et al., 2007) and thorough evaluations of the utility of the tools in a multi-species framework in different landscapes does not yet exist. Further, efficient ways to sample for species occurrence and minimum sample size requirements have not been suggested in terms of modeling species distributions at landscape scales and fine resolutions. Thus, the goal of chapter two of this thesis was to evaluate sampling strategies and model predictive performance of landscape scale species distribution models in two environmentally disparate areas.

Predictions in the Face of Uncertainty

Predictions from models are plagued by sources of uncertainty (see reviews Barry & Elith, 2006; Rocchini et al., 2011). Prediction errors can be a function of the data or the model and can have serious implications when the models are being used to inform management. Mis-specified models (e.g. missing covariates) can reduce model predictive performance and result in residual autocorrelation. The degree to which missing covariates will decrease a models predictive power will be dependent on how important the missing processes are. Tools exist (e.g. auto-logistic regression) to account for this residual correlation (Betts et al., 2009; Dormann, 2009; Dormann et al., 2007; Hoeting et al., 2000) but easy implementation on large lattice data remains elusive due to the heavy computational requirements. Although these issues are important, model specification is a function of the modeler’s knowledge of the species and system, and available geospatial data, all of which can be hard to control or fix.

Arguably the largest source of uncertainty in predicting species distributions arises from the data used to parameterize the model, and as the colloquial saying goes: “garbage in, garbage out”. Sources of data uncertainty can vary from bias in site selection, observer, sampling time, etc. and can lead to biased representations of species distributions (Rocchini et al., 2011). Imperfect detection is a source of data uncertainty that has gained much recent attention in site occupancy studies (MacKenzie et al., 2002) but has little been considered in the context of species distribution modeling.

Although the concept of imperfect detection has been considered more by animal researchers due to the mobile nature of those organisms, imperfect detection in plant studies has been shown to bias vital rates in demography studies (Alexander et al., 2009; Kéry & Gregg, 2003), population size (Alexander et al., 1997), extinction rates (Kéry, 2004) and to alter management strategies (Regan et al., 2011; Regan et al., 2006). Detection in plant studies has been shown to be a function of time of day, weather, observer experience, species abundance, herbivory, and phenology (Alexander et al., 2009; Chen et al., 2009; Garrard et al., 2008; Moore et al., 2011). Although these relationships have been shown, few have considered the potential for detection error to bias predictions of species distribution (but see Rota et al., 2011). Thus, the third chapter of this thesis evaluated the influence of detection error, in a multi-visit within season survey, on predictions of non-native species distributions and associated uncertainty.

Moving from Static to Dynamic – Predicting Spread in Space and Time

Although the species distribution modeling approach is invaluable to understanding the realized distributions of non-native species at the landscape scale, management of such species also needs the ability to predict where they will be in the future. Due to the size and scope of non-native species invasions, it is rare that a practitioner or researcher will know the stage of the invasion of all the target species at a site. Further, just because an invasive species is present, does not imply that it is currently spreading. Thus, a way to quantitatively assess a species invasion status, make spatially and temporally explicit predictions about metapopulation invasions, and potentially identify sources of new populations could be invaluable to aide prioritization of species and populations to manage.

Based on the ecological theory associated with invasions reviewed in the first section, we can hypothesize that the key drivers for landscape scale invasions will have to do with dispersal, habitat and disturbance (Catford et al., 2009; Coutts et al., 2010; Higgins et al., 1999). It is possible to spatially quantify these drivers using geospatial software, existing datasets from management agencies, digital elevation models, and remote sensing data. The key to being able to predict spread is to gather reliable species occurrence data across these variables and at different times (probably years) to understand how the species is responding to each potential driver of invasion, and the interactions between the different drivers. Metapopulation theory (Hanski, 1998) and statistical models and tools (Fiske & Chandler, 2011; MacKenzie, 2006) have been developed over the last decade or so and provide a framework to test hypotheses about what the key drivers are of landscape scale invasions. Instead of a strictly mechanistic approach, stochastic metapopulation dynamics theory uses patch size and area, and connectivity between suitable patches to explain population colonization and extinction patterns.

While the metapopulation concept may theoretically apply to plant populations (Eriksson, 1996; Ehrlen & Eriksson 2003), little empirical work has been done to determine the relative importance of interpatch migration versus local dispersal, recruitment and mortality in structuring regional groups of populations (Freckleton and Watkinson, 2002). Thus, the fourth chapter of this thesis aimed at elucidating the drivers of non-native species colonization and extinction at landscape scales. Not only can this information improve our fundamental understanding of plant regional dynamics but can help target management and make predictions about potential future distributions of these species.

Study Sites

The primary study site for the following three chapters was the Idaho National Laboratories in east-central Idaho, USA. This was a federal research site closed to general public access and was approximately 490 km^2 and data were collected between 2009 and 2011. The elevation at the site ranged from 1461 – 1638 m. Mean annual precipitation was approximately 220 mm and mean growing season precipitation (May-June) was approximately 61 mm (55 year averages). The topographic variation can be described as rolling hills that are generally a result of ancient lava flow and lava tubes and ancient lake beds that are generally flat and clay rich. This area was fairly homogenous with only a few dominate soil substrates (clay rich playas, volcanic parent material loam soils and sandy soils). The site was dominated by Wyoming big sagebrush (*Artemisia tridentata* Nutt. ssp. *wyomingensis* Beetle & Young), (Shive et al., 2011) and *Poa secunda* J. Presl and a diversity of other shrubs, grasses and forbs. Anthropogenic disturbance was restricted to adjacent to the facilities and paved roads (including one US highway and one restricted access paved road) and primitive two-track roads that are used primarily for research and firefighting activities. The site was historically grazed by sheep in the late 1800's and early 1900's and there were a few attempts at settlement of the area. Residual evidence of settlement include water diversion features and landscape leveling for cultivation. Restricted seeding of *Agropyron cristatum* (L.) Gaertn. around paved roads has been employed over the last few decades to create "green strips" to act as fire breaks. There was an 18 year chronosequence of fires at the site, with the most recent burning 43,700 hectares in 2010. These fires are generally restricted to the southeast portion of the site and are primarily thought to be anthropogenic in origin. Presence and absence of all non-native plant species was recorded from 2009 – 2011 in a designed survey. Ten species were chosen for the purposes of this thesis. Those were *Agropyron cristatum* a perennial bunch forming Poaceae, *Alyssum desertorum, Bromus tectorum*, *Carduus nutans* L. a biennial to perennial Asteraceae, *Halogeton glomeratus* (M. Bieb.) C. A. Mey. an annual Chenopodiaceae, *Salsola kali* L. an annual Chenopodiaceae, *Sisymbrium altissimum* L. an annual to biennial Brassicaceae, *Tragopogon dubius* Scop.

an annual to biennial Asteraceae, *Lepidium perfoliatum* L. an annual to biennial Brassicaceae, and *Descurainia sophia* (L.) Webb ex Prantl an annual to biennial Brassicaceae. Species names and naming authorities from USDA Plants Database (USDA & NRCS, 2012).

The second dataset used in chapter two was collected over three years in the Northern Range of Yellowstone National Park, USA. The study area was approximately 1,200 km^2 and data were collected from 2001 to 2004. Elevation at the site ranged from 1587 – 2832 m and dominant vegetation varies from sagebrush steppe (*Artemisia tridentata* Nutt.) to lodgepole pine forest (*Pinus contorta* Douglas ex Loudon), riparian, meadow, and alpine vegetation. Precipitation was highly variable (380 – 2,000 mm annually) and anthropogenic disturbance is concentrated around a few main paved roads, the town of Mammoth Hot Spring, WY and trail systems. Further site data can be found in Rew et al. (2005). Data on the occurrence of all non-native species was collected at this site over the three years. Eight species were chosen for analysis in chapter two and included *Alyssum desertorum* Stapf an annual Brassicaceae, *Bromus inermis* Leyess. a perennial, rhizomatous Poaceae, *Bromus tectorum* L. an annual and winter annual Poaceae, *Cirsium arvense* (L.) Scop. a perennial, clonal Asteraceae, *Cynoglossum officinale* L. a binennial Boraginaceae, *Linaria dalmatica* (L.) Mill. a perennial, clonal Schropulariaceae, *Melilotus officinalis* (L.) Lam. an annual to perennial Fabaceae, and *Phleum pretense* (L.) a perennial, rhizomatous Poaceae. These species were chosen as they had a range of growth habits, habitat preferences, and relative abundances on the landscape.

References

Alexander, H.M., Slade, N.A., & Kettle, W.D. (1997) Application of mark-recapture models to estimation of the population size of plants. Ecology, 78, 1230-1237.

Alexander, H.M., Slade, N.A., Kettle, W.D., Pittman, G.L., & Reed, A.W. (2009) Detection, survival rates and dynamics of a cryptic plant, *Asclepias meadii*: applications of mark-recapture models to long-term monitoring studies. Journal of Ecology, 97, 267-276.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. Journal of Biogeography, 33, 1677-1688.

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. Ecography, 28, 693-695.

Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. Journal of Applied Ecology, 43, 413-423.

Bauer, J.T. (2012) Invasive species: “back-seat drivers” of ecosystem change? Biological Invasions.

Betts, M.G., Ganio, L.M., Huso, M.M.P., Som, N.A., Huettmann, F., Bowman, J., & Wintle, B.A. (2009) Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review”. Ecography, 32, 374-378.

Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U., & Richardson, D.M. (2011) A proposed unified framework for biological invasions. Trends in Ecology & Evolution, 26, 333-339.

Blossey, B. & Notzold, R. (1995) Evolution of increased competative ability in invasive nonindigenous plants: a hypothesis. The Journal of Ecology, 83, 887-889.

Brooks, M.L., D'Antonio, C.M., Richardson, D.M., Grace, J.B., Keeley, J.E., DiTomaso, J.M., Hobbs, R.J., Pellant, M., & Pyke, D. (2004) Effects of invasive alien plants on fire regimes. BioScience, 54, 677-688.

Callaway, R.M. & Ridenour, W.M. (2004) Novel weapons: invasive success and the evolution of increased competative ability. Frontiers in Ecology and the Environment, 2, 436-443.

Catford, J.A., Jansson, R., & Nilsson, C. (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Diversity and Distributions, 15, 22-40.

Chen, G., Kéry, M., Zhang, J., & Ma, K. (2009) Factors affecting detection probability in plant distribution studies. Journal of Ecology, 97, 1383-1389.

Coutts, S.R., Klinken, R.D., Yokomizo, H., & Buckley, Y.M. (2010) What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management? Biological Invasions, 13, 1649-1661.

Crosby, A.W. (1986) Ecological imperialism: the biological expansion of Europe, 900-1900 Cambridge University Press.

Crossman, N.D. & Bass, D.A. (2008) Application of common predictive habitat techniques for post-border weed risk management. Diversity and Distributions, 14, 213-224.

Daehler, C.C. (2003) Performane comparisons of co-occurring native and alien invasive plants: implications for conservation and restoration. Annual Review of Ecology, Evolution, and Systematics, 34, 183-211.

Darwin, C. (1859) On the Origin of Species. John Murray.

Davis, M.A. (2009) Invasion Biology Oxford University Press, Oxford, New York.

Davis, M.A., Thompson, K., & Grime, J.P. (2001) Charles S. Elton and the dissociate of invasion ecology from the rest of ecology. Diversity and Distributions, 7, 97-102.

Dormann, C.F. (2009) Response to Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review”. Ecography, 32, 379-381.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30, 609-628.

Ehrlen, J. & Eriksson, O. (2003) Large-scale spatial dynamics of plants: a response to Freckleton & Watkinson. Journal of Ecology, 91, 316-320.

Eriksson, O. (1996) Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. OIKOS, 77, 248-258.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 40, 677-697.

Elton, C.S. (1958) The ecology of invasions by plants and animals. The University of Chicago Press, Chicago and London.

Engler, R., Randin, C.F., Vittoz, P., Czáka, T., Beniston, M., Zimmermann, N.E., & Guisan, A. (2009) Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? Ecography, 32, 34-45.

Fiske, I. & Chandler, R. (2011) unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43, 1-23.

Franklin, J. (2009) Mapping species distributions: spatial inference and prediction Cambridge University Press, Cambridge, UK.

Freckleton, R.P. & Watkinson, A.R. (2002) Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. Journal of Ecology, 90, 419-434.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A., & Wintle, B.A. (2008) When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. Austral Ecology, 33, 986-998.

Gelbard, J.L. & Belnap, J. (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology, 17, 420-432.

Giljohann, K.M., Hauser, C.E., Williams, N.S.G., & Moore, J.L. (2011) Optimizing invasive species control across space: willow invasion management in the Australian Alps. Journal of Applied Ecology, 48, 1286-1294.

Graham, C., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution, 19, 497-503.

Grime, J.P. (1977) Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. The American Naturalist, 111, 1169-1194.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993-1009.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. Ecological Modelling, 135, 147-186.

Gurevitch, J. & Padilla, D. (2004) Are invasive species a major cause of extinctions? Trends in Ecology & Evolution, 19, 470-474.

Hanski, I. (1998) Metapopulation dynamics. Nature, 396, 41-49.

Hanski, I. (1999) Metapopulation ecology. Oxford University Press, Oxford.

Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B.A., Moore, K., Taylor, C., & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. Ecology Letters, 8, 91-101.

Hauser, C.E. & McCarthy, M.A. (2009) Streamlining 'search and destroy': cost-effective surveillance for invasive species management. Ecology Letters, 12, 683-692.

Hierro, J.L., Maron, J.L., & Callaway, R.M. (2005) A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. Journal of Ecology, 93, 5-15.

Higgins, S.I., Richardson, D.M., & Cowling, R.M. (2001) Validation of a spatial simulation model of a spreading alien plant population. Journal of Applied Ecology, 38, 571-584.

Higgins, S.I., Richardson, D.M., Cowling, R.M., & Trinder-Smith, T.H. (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conservation Biology, 13, 303-313.

Hoeting, J.A., Leecaster, M., & Bowden, D. (2000) An improved model for spatially correlated binary responses. Journal of Agriculture, Biological, and Environmental Statistics, 5, 102-114.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography Princeton University Press, Princeton, New Jersey.

Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O., & Estoup, A. (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. Evolutionary Applications, 5, 89-101.

Kéry, M. (2004) Extinction rate estimates for plant populations in revisitation studies: importance of detectability. Conservation Biology, 18, 570-574.

Kéry, M. & Gregg, K.B. (2003) Effects of life-state on detectability in a demographic study of the terrestrial orchid *Cleistes bifaria*. Journal of Ecology, 91, 265-273.

Liu, H. & Stiling, P. (2006) Testing the enemy release hypothesis: a review and meta-analysis. Biological Invasions, 8, 1535-1545.

MacArthur, R.H. & Wilson, E.O. (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey.

MacDougall, A.S., Gilbert, B., & Levine, J.M. (2009) Plant invasions and the niche. Journal of Ecology, 97, 609-615.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., & Fakhri, B. (2000) Biotic invasions: causes, epidemiology, global consequences, and control. Ecological Applications, 10, 689-710.

MacKenzie, D.I. (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence Elsevier, Amsterdam; Boston.

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

Maxwell, B.D., Lehnhoff, E., & Rew, L.J. (2009) The rationale for monitoring invasive plant populations as a crucial step for management. Invasive Plant Science and Management, 2, 1-9.

Moore, J.L., Hauser, C.E., Bear, J.L., Williams, N.S.G., & McCarthy, M.A. (2011) Estimating detection-effort curves for plants using search experiments. Ecological Applications, 21, 601-607.

Parendes, L.A. & Jones, J.A. (2000) Role of light availability and dispersal in exotic plant invasion along roads and streams in the H. J. Andrews Experimental Forest, Oregon. Conservation Biology, 14, 64-75.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology & Biogeography, 12, 361-371.

Pimentel, D., Zuniga, R., & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics, 52, 273-288.

Pulliam, R.H. (1988) Sources, sinks, and population regulation. The American Naturalist, 132, 652-661.

Pysek, P. & Hulme, P.E. (2009) Invasion biology is a discipline that's too young to die. Nature, 460, 324.

Regan, T.J., Chadès, I., & Possingham, H.P. (2011) Optimally managing under imperfect detection: a method for plant invasions. Journal of Applied Ecology, 48, 76-85.

Regan, T.J., McCarthy, M.A., Baxter, P.W.J., Dane Panetta, F., & Possingham, H.P. (2006) Optimal eradication: when to stop looking for an invasive plant. Ecology Letters, 9, 759-766.

Rew, L.J., Lehnhoff, E.A., & Maxwell, B.D. (2007) Non-indigenous species management using a population prioritization framework. Canadian Journal of Plant Science, 87, 1029-1036.

Rew, L.J., Maxwell, B.D., & Aspinall, R. (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Science, 53, 236-241.

Richardson, D.M. & Pysek, P. (2008) Fifty years of invasion ecology - the legacy of Charles Elton. Diversity and Distributions, 161-168.

Richardson, D.M. et al. (2011). A compendium of essential concepts and terminology in biological invasions. In Fifty Years of Invasion Ecology: The Legacy of Charles Elton (ed D.M. Richardson), pp. 409-420. Wiley-Blackwell.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. Progress in Physical Geography, 35, 211-226.

Rota, C.T., Fletcher, R.J., Evans, J.M., & Hutto, R.L. (2011) Does accounting for imperfect detection improve species distribution models? Ecography, 34, 659-670.

Seipel, T., Kueffer, C., Rew, L.J., Daehler, C.C., Pauchard, A., Naylor, B.J., Alexander, J.M., Edwards, P.J., Parks, C.G., Arevalo, J.R., Cavieres, L.A., Dietz, H., Jakobs, G., McDougall, K., Otto, R., & Walsh, N. (2012) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. Global Ecology and Biogeography, 21, 236-246.

Shafii, B., Price, W.J., Prather, T.S., Lass, L.W., & Thill, D.C. (2003) Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Science, 51, 748-751.

Shea, K. & Chesson, P. (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution*, 17, 170-176.

Shive, J.P., Forman, A.D., Aho, K., Hafla, J.R., Blew, R.D., & Edwards, K.T. (2011). Vegetation Community Classification and Mapping of the INL Site. Gonzales-Stoller Surveillance, LLC.

Smith, R.G., Maxwell, B.D., Menallad, F.D., & Rew, L.J. (2006) Lessons from agriculture may improve the management of invasive plants in wildland systems. Frontiers in Ecology and the Environment, 4, 428-434.

Stohlgren, T.J. & Schnase, J.L. (2006) Risk analysis for biological hazards: what we need to know about invasive species. Risk Analysis, 26, 163-173.

USDA & NRCS (2012) The PLANTS Database (ed N.P.D. Team), Vol. 2012, pp. USDA, NRCS. 2012. The PLANTS Database (http://plants.usda.gov, 28 February 2012). National Plant Data Team, Greensboro, NC 27401-4901 USA., Greensboro, NC.

Václavík, T. & Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling, 220, 3248-3258.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Diversity and Distributions, 18, 73-83.

Vilà, M., Espinar, J.L., Hejda, M., Hulme, P.E., Jarošík, V., Maron, J.L., Pergl, J., Schaffner, U., Sun, Y., & Pyšek, P. (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters, 14, 702-708.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmanek, M., & Westbrooks, R. (1997) Introduced species: a significant component of human-caused global change. New Zealand Journal of Ecology, 21, 1-16.

Williamson, M. & Fitter, A. (1996) The varying success of invaders. Ecology, 77, 1661-1666.

Williamson, M.H. (1996) Biological Invasions. Chapman & Hall, London.

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., & Blackard, J.A. (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. Journal of Applied Ecology, 44, 1057-1067.

CHAPTER TWO

MAXIMIZING THE UTILITY OF LANDSCAPE SCALE SPECIES DISTRIBUTION MODELS

Contributions of Authors and Co-Authors

Manuscript in Chapter 2

Author: Tyler J. Brummer

Contributions: Conceived the study, developed simulation model, performed analysis and wrote the manuscript.

Co-author: Bruce D. Maxwell

Contributions: Obtained funding in conjunction with Lisa Rew, provided assistance in conceptual development and revisions to the manuscript.

Co-author: Megan D. Higgs

Contributions: Provided consultation on statistical methods and ideas and extensive revisions especially in methodological clarification and inference.

Co-author: Lisa J. Rew

Contributions: Obtained funding for the project, developed ideas, discussed results and revised and edited the manuscript at all stages.

Manuscript Information Page

Tyler J. Brummer, Bruce D. Maxwell, Megan D. Higgs, Lisa J. Rew
Journal Name: Biological Invasions
Status of Manuscript:
___Prepared for submission to a peer-reviewed journal
_x_Officially submitted to a peer-reviewed journal
___Accepted by a peer-reviewed journal
___Published in a peer-reviewed journal
Published by Springer
Submitted March 6, 2012

Abstract

Species distribution models (SDMs) and associated maps can be useful tools for managing rare or invasive plant species. The utility of SDMs and subsequent distribution maps requires reliable presence and absence data along with use of ecologically appropriate variables, all collected at scales commensurate with management. Thus, we simulated different presence and absence sampling methods at the 10 meter resolution for their effectiveness to capture plant metapopulation distributions on a landscape (10,000 ha) when the distributions were driven by known variables. The simulations indicated that transects were most efficient at generating probability of species occurrence maps when time efficiency was accounted for across four theoretic species distribution patterns. We also resampled empirical data for sixteen non-native plant species at two sites (one with steep, and the other with shallow environmental gradients) to generate predictive power accumulation curves as a function of sample size. There was greater predictive performance for species at their ecophysiological range limits than on a species-level niche generalist/specialist dichotomy. Model performance was best where there were steeper environmental gradients indicating increased importance of environmental filtering. Both simulation and empirical analysis suggested sampling between 0.06% and 0.23% of a study area to maximize predictive performance independent of site characteristics.

Keywords: generalized linear model; invasive species distribution model; logistic regression; management; non-indigenous plant species; plant; resolution; sampling; scale; species distribution model; transect

Introduction

Species distribution modeling has traditionally focused on relating species occurrence data to ecophysiological variables at coarse spatial resolutions to make country to continent scale predictions (see reviews Elith and Leathwick 2009; Richards et al. 2007). Herbaria records, floristic studies and incidental sightings have been used to model species distributions (Graham et al. 2004). Predicting invasive species distributions has focused on using data from the native range to predict potential distributions in the introduced range, or using survey data from the introduced range with some accountability for non-equilibrium or dispersal constraints at course spatial scales (Lozier and Mills 2011; Peterson 2003; Václavík and Meentemeyer 2009, 2012). In contrast, most management of invasive species occurs within units of smaller spatial scale (e.g. refuges, wildlife conservation units, national parks, etc.). Coarse scale predictions (e.g. 1 km and greater) may give only a handful of pixel values for a particular management unit, which does not aide in decreasing search areas or targeting populations within a site for management (Rew et al. 2007). Species distribution modeling at continent and country scales can be applied to landscape scales (*sensu* Pearson and Dawson 2003) to estimate the current realized distribution and inform management. Smaller extent (< 200 km) and finer resolution (10 - 30 m) were explored in a few studies (Fleishman et al. 2002; Rew et al. 2005; Shafii et al. 2003; Václavík et al. 2010). However, an evaluation of the effects of sample size and sampling method on landscape scale SDMs is needed to quantify the potential predictive ability of these models across species and sites so search and population prioritization recommendations can be made to managers.

There are major advantages for managers to collect data on invasive species distribution on-site in their management area. Managers are more likely to believe the data and resulting predictions. Locally parameterized SDMs are more likely to reveal the relative importance of invasion driving processes (e.g. ecological filtering and dispersal) that can then be manipulated for more efficient management. And, a sampling plan can facilitate effective site-specific monitoring of management outcomes as well as metapopulation dynamics (invasion progress). However, currently plant data do not readily exist at scales necessary to inform models to make predictions that practitioners can incorporate into management planning and implementation (e.g. 10 m^2 resolution). Ecophysiological variables that drive species distributions that would require a massive network of instrumentation do not exist at fine resolution where they are likely to act. However, the result of ecological filtering and dispersal limitations can be determined by using surrogate landscape predictor variables that are collinear with mechanistic variables, or represent best approximations of processes (reviewed in Franklin 2009). Environmental variation (resource gradients) can be effectively described using variables derived from digital elevation models (DEMs), and remotely sensed data (e.g. land cover) (Zimmermann et al. 2007). Anthropogenic disturbance, which has often been found to be a key driver of invasion (Hierro et al. 2005; Hufbauer et al. 2012), can be captured using proximity variables to human activity. Similarly, natural disturbances such as fire have been recorded in spatial databases and can be utilized in combination with local manager observations to model non-indigenous plant species (NIS) distributions on the landscape. Finally we can characterize dominant dispersal vectors such as roads (Gelbard and Belnap 2003; Parendes, Jones 2000; Seipel et al. 2012; Wilson et al. 2009) and test the expectation that species spread directionally from these vectors to account for potential dispersal limitation for more recently introduced species.

Although there are sufficient ways to characterize resource gradients and anthropogenic dispersal mechanisms that structure species occurrence at landscape scales, SDMs tend to lack inclusion of biotic interactions (Pearson and Dawson 2003), source sink dynamics (Pulliam 1988) and stochastic metapopulation dynamics (Hanski 1999). The degree to which ecological filtering and large scale dispersal patterns are structuring local species distributions may vary by species and site. Complementarily, biotic interactions, source sink dynamics, and random small scale primary and secondary dispersal processes may be more important in structuring distributions at landscape scales, which would naturally limit the usefulness of models built on the principles of ecological filtering and landscape scale dispersal patterns. But, no empirical evaluation of the performance of these models has been undertaken to suggest advantages or disadvantages of modeling landscape scale species distributions.

Model building and prediction requires quality species presence and absence data. Such data collection is costly, especially in the eyes of invasive plant managers whose incentives are generally to manage populations as found rather than make observations, especially of species absence. Thus, efficient and cost effective data collection methods are imperative, and determining the minimum amount of sampling necessary to generate predictive maps from these models is needed. An evaluation of these questions has not been undertaken with the expressed goal of creating maps of species occurrence at the scale that management decisions are made.

Our objectives were to (1) determine an efficient sampling strategy to gather species occurrence data and establish sample size requirements to reach maximum achievable predictive power for landscape scale distribution models and, (2) to quantify maximum model performance of 16 introduced species from two sites and offer ecological hypotheses for variable model performance across sites and species.

Methods

Simulation Space

To evaluate sample size and methods, a simulation study was constructed using the statistical package R (R Development Core Team, 2011). To generate a realistic topographic space to conduct simulations, a 10 km x 10 km area of the northern range of Yellowstone National Park was selected. Five spatial covariates at a 10 m x 10 m resolution were considered for the simulation modeling exercise. These were distance to road, slope, cosine of aspect (the amount that a cell faced north), sine of aspect (the amount that a cell faced east), and elevation. All covariates with the exception of distance to road were derived from a 10 m resolution digital elevation model. Distance to road was calculated at a 10 m scale in ArcMap 9.3 from a road shapefile. The resulting simulation space was 1000 by 1000 grid cells, with each cell representing a 10 m x 10 m area.

Plant Distributions

Four theoretical plant distribution patterns were used in the simulation model (Figure 1). The distributions were designed to emulate different species that were tightly bound by environmental conditions (specialists) versus not tightly bound (generalists), varied in their relationship to distance to road and were either high or low frequency. Pattern 1 exemplified a specialist non-native species with a strong negative association with distance to road (Strong Negative, Low frequency, Specialist- SNLS); pattern 2 also showed a strong negative association with distance to road but with a "generalist" pattern (Strong Negative, Low frequency generalist - SNL); pattern 3 showed increased presence with distance from road, and had a low frequency (Weak Positive, Low frequency generalist - WPL); and pattern 4 had a weak positive association to the road but a high frequency on the landscape (Weak Positive, High frequency generalist - WPH). The proportions of presences for the distribution patterns were 0.036, 0.036, 0.037 and 0.225, for the SNLS, SNL, WPL and WPH, respectively. The SNLS and SNL distribution patterns are most analogous to non-native species (Rew et al. 2005), whereas the WPL and WPH would be most representative of a non-native species that is no longer dispersal limited or native plant species in equilibrium with their environment.

These theoretical plant distribution patterns were generated for the sampling simulations using generalized linear models (GLMs) with the logit link function (logistic regression) and were used to create probabilities of occupancy for each grid cell in the simulation space. Four probabilities of occupancy maps were created to represent the four different species, and these were considered as "truth" for the simulation study to compare to probability of occupancy maps predicted from sampling procedures. Regression coefficient estimates (Table 1) for each simulated distribution pattern were derived from survey data collected in the Northern Range of Yellowstone National Park (Rew et al. 2005).

The model form used to generate the probability of occupancy at each grid cell was: $logit(p_i) = \beta_0 + \beta_1 cosine(aspect)_i + \beta_2 sine(aspect)_i + \beta_3 slope_i + \beta_4 Distancetoroad_i + \beta_5 elevation_i$ Equation 1

where p_i is the probability of occupancy at location i. Values for the coefficients used to generate the four probability of occupancy surfaces are provided in Table 1. This binary logistic regression model (used for the simulation and empirical study) assumed independence and thus did not account for residual autocorrelation. This is not a problem for the simulation study, because the data were generated under independence. However, for the empirical study, it is possible there was leftover spatial dependence not accounted for by the covariates. We do not expect that our conclusions from the investigation would change even if spatial autocorrelation was explicitly modeled, as we focused on point predictions and are confident we have included the most important spatial covariates. We also focus on the non-spatial model because it is a method feasible for managers to employ.

Each grid cell on the map was assigned a presence or absence using a Bernoulli trial based on the probability of occupancy calculated from Equ. 1. For a single run of the simulation a new realization of presence/absence was generated for each grid cell across the map using Equ. 1. The simulated map was then sampled using four different survey methods. These binary data were then used to obtain estimates of the regression coefficients in equation 1 using binary logistic regression. The fitted logistic regression model was used to predict the probability of occupancy for every grid cell on the map for each survey method, sample size and species type combination. The predicted probabilities were then compared to the "true" probabilities that generated the binary data from which the sample was taken. The ability of the model and sampling design to classify presence and absence was also assessed using area under the receiver operator characteristic curve (AUC) as well as omission and commission error rates.

Survey Methods

Four survey methods, two random and two clustered, were evaluated. Here, clustered is defined by a group of adjacent plots (10 m x 10 m cells) being surveyed at a randomly chosen location. Rew et al. (2006) proposed a "targeted" transect approach for surveying introduced species where 2 km long x 10 m wide belt transects that originated from random points on the road and traveled perpendicular to the road were used to assess species presence and absence using 10 m x 10 m plots. This method was designed to emulate a stratified simple random sample approach, stratified on distance to road, but to increase sampling efficiency by having 200 contiguous plots and will be referred to as transects, but may more appropriately be considered linear clusters. The stratification variable was distance to road because of the empirically demonstrated decline of non-native species occurrence as distance to roads increases (Parendes, Jones 2000; Seipel et al. 2012). As the 200 plots are arranged in a linear cluster, there is the potential for underrepresentation of certain covariate patterns that could generate an unrepresentative sample and misestimate mean occupancy rates and predict presence/absence poorly. Thus, this method was compared to a simple random sample and stratified simple random sample (e.g. equal size simple random sample of points within each 10 m wide strata defined using distance from the road). The fourth survey strategy was randomly placed rectangular clusters that were 20 plots by 10 plots in size. This strategy was employed to represent a technique commonly used by many land managers, where polygons are haphazardly assessed for species occurrence. Managers are potentially biased with site selection but our application of the method provides a first approximation of the common behavior of managers. All strategies were restricted to 2 km from roads to ensure the results were comparable.

Each of the four plant patterns (SNLS, SNL, WPL, WPH) were assessed using the four surveying methods. This was performed using 20 different sampling intensities, from 400 to 8000 plots in 400 plot intervals (2 to 40 clusters of 200 plots in 2 cluster intervals for clustered designs). Fifty replicates were run per distribution/survey method/sample size combination. Transects were laid out perpendicular to the road and could terminate outside the simulated management area. If transects ended outside the simulated management area they were truncated at the edge of the simulation space and only the number of plots within the management area were used in the comparative analysis.

Sample Size and Method Performance Evaluation

The influence of sample size and method on the ability to predict presence/absence and estimate mean occupancy rate was evaluated using five metrics. Area under the receiver operator characteristic curve (AUC) was used as a measure of classification accuracy (Pearce and Ferrier 2000). The AUC was calculated from a validation data set of 50,000 random plots from the simulation space. The other metrics used were omission and commission error as calculated from the optimal binary classification threshold (threshold that maximizes sensitivity and specificity), mean residual error based on sample and true probability of occupancy (to assess sampling bias), and sum of mean square prediction error between the data generating probabilities of occupancy (truth) and the predicted probabilities obtained via the models fitted from simulated sample data.

To understand the distribution of the five performance metrics over 50 realizations, the range of the middle 95% of the metrics were constructed using the 0.025 and 0.975 quantiles (i.e. excluding the most extreme 5% of the data). These plots were used to visually assess the sampling intensity at which the clustered designs performed comparably to simple random sampling and/or stratified random sampling. As a conservative measure of the sample size at which maximum predictive performance was achieved, the sample size at which 95% of mean maximum AUC was consistently achieved was calculated for each sampling method and plant distribution.

Efficiency

We defined efficiency as the ratio of sample size to optimize a model performance metric, penalized for sampling time between two sample methods:

$$\frac{(\#\ plots\ to\ optimize\ criteria\ for\ simple\ random\ sampling)\ x\ (time\ to\ sample)}{(\#\ plots\ to\ optimize\ criteria\ for\ other\ survey\ method)\ x\ (time\ to\ sample)}$$

This ratio compares the number of plots needed to optimize a model performance metric relative to simple random sampling and weights the number of plots by the time needed to collect the data. The relative times to sample based on cluster versus random plot sampling were taken from a simulation accounting for logistical cost (Rew et al. 2006).

Empirical Data

Field data were collected at two sites, The Northern Range of Yellowstone National Park, Wyoming, USA (YNP) and the Idaho National Laboratories, Idaho, USA (INL). The INL study area was 490 km^2 and YNP was 1200 km^2. Data were collected in 2001-2004 at YNP and 2009-2010 at the INL, using the targeted transect approach as outlined above, with Trimble GeoExplorer and GeoXT GPS units. Covariates were as exhaustive as possible at both sites, using 10 m resolution digital elevation model derived variables, as well as, vegetation cover, proximity variables (e.g. distance to roads and streams) (see Online Resource 3 for full list). Variables were meant to account for as much environmental variation as possible as well as potential dispersal processes as the plants modeled were introduced species at various stages of invasion. The INL had shallow environmental gradients with one dominant vegetation type and an elevation range of 1461 – 1638 m, while YNP had steep environmental gradients with multiple vegetation types and an elevation range of 1587 – 2832 m. The INL had one public and one restricted paved road as well as research facilities and lightly traveled two-track fire access roads throughout the site, while YNP had only one main public paved road but walking trails distributed through the site.

Resampling Methods

To generate model performance curves as a function of sample size, a resampling cross validation procedure was implemented. At each run of the resampling simulation, the data were initially split into training (75% of the transects) and test data (25% of the transects). Random selection of the training and test data was carried out via whole transects even though the observation unit was 10 m x 10 m pixels. This was to avoid over-fitting the data (inflating predictive performance and decreasing variation in predictions) which could arise if validating points with adjacent points in the training data. A random sample (2 to 78 transects in increments of 2 for INL and 4 to 288 transects in increments of 4 for YNP) of the training data was then selected and used to fit a binary logistic regression model and predict to the test data. Area under the receiver operator characteristic curve was then calculated and stored. Twenty replicates at each sample size were completed to quantify the variation in predictions. Eight species at each location were tested for a total of 16 species across two sites (Table 4). Two species, *Bromus tectorum* L. and *Alyssum desertorum* Stapf, occurred at both sites.

Results

Simulation Study

Simple random sampling (SRS) and stratified simple random sampling methods (StRS) performed best at small sample sizes (< 2000 plots – 0.2% of study area) compared to transects and rectangular clusters when looking at both the mean and spread of the five model performance metrics considered for the four simulated plant distributions (Figure 2, Online Resources 1-2). Visual assessment of the 95% interval boundaries for transects showed they almost completely overlapped with the 95% interval boundaries for SRS and StRS between 1200 and 2400 plots (6 and 12 transects) depending on the metric considered and the distribution type (Figure 2, Online Resources 1-2). Rectangular clusters required more samples for the 95% intervals to reach near equivalency with SRS and StRS, requiring sample sizes between 4000 and 8000 plots (20 to 40 clusters).

The mean AUC became asymptotic for all the simulated distributions. SRS met the assumptions of the modeling method most closely and has been shown to need the fewest plots relative to other sampling methods when generating models of plant species distribution (Reese et al. 2005). Thus, the mean AUC for simple random sampling at a sampling intensity of 8000 plots was considered the maximum achievable AUC. Maximum achievable AUC varied by plant distribution pattern. The SNLS, SND, WPL, and WPH distributions had maximum mean AUC values of 0.90, 0.67, 0.65, and 0.58 respectively. Both simple random and stratified random sampling reached 95% of maximum AUC at the lowest sample size considered (400 plots) for the SNLS distribution (Table 2). Transects reached 95% of maximum AUC with 400 more plots (2 transects) compared to SRS for the WPL, SND and SNLS distributions and 800 more plots (4 transects) for the WPH pattern (Table 2). The rectangular cluster method always required a greater sample size to reach 95% of maximum AUC compared to transect sampling – needing between 400 and 1200 more sample plots (2 to 6 clusters) depending on the distribution (Table 2).

Commission error (model predicts occurrence where species was absent) was always higher than omission error (model predicts absence where species was present) at the optimal threshold with minimum commission error being 0.23, 0.42, 0.38, and 0.36 for the SNLS, SNL, WPL and WPH distributions, respectively. Minimum omission error was, 0.004, 0.13, 0.014 and 0.10 for the SNLS SNL, WPL, and WPH patterns respectively. Simple and stratified random sampling had the narrowest spread in commission and omission error at small sample sizes (< 2000 – 4000 plots depending on distribution). Both cluster methods had the largest spread in both error types at low sample sizes but the 95% intervals did become near equivalent with simple and stratified random sampling at sufficiently large sample sizes (Online Resource 1). In general, rectangular clusters performed worse than transects for all scenarios based on omission and commission, taking more samples for the 95% intervals to reach equivalency with that of random point methods.

Comparison of mean prediction error between predicted probabilities of occupancy from the simulated sample and true probability of occupancy (Online Resource 2) was used to assess if one sampling method systematically under- or over-predicted probability of species occupancy. It showed that SRS and StRS had no prediction bias on average as the 95% intervals were centered around 0 whereas, both cluster methods consistently under-predicted probability of occurrence at smaller sample sizes (< 1200 – 2000 plots depending on distribution and sample method). The 95% intervals for mean prediction error associated with transects became centered around 0 (indicating no bias) at sample sizes between 800 and 1600 plots depending on distribution, and mean prediction error for rectangular clusters became centered on 0 between 1200 and 2200 (0.12% and 0.22% of the study area) plots. The high frequency species distribution (WPH) differed from the low frequency distributions as there were instances where mean occupancy rate was over-predicted at low sample sizes, whereas the low frequency distributions rarely over-predicted occupancy rates.

The efficiency of the stratified random sampling, transect sampling, and rectangular cluster sampling based on maximizing AUC were calculated relative to simple random sampling (Table 3). The efficiency of stratified random sampling ranged between 0.75 and 1.10 times as efficient as simple random sampling. The cluster methods were more efficient than the random sampling methods with transects between 2.5 and 4.6 times as efficient as SRS depending on distribution and rectangular cluster sampling between 1.67 and 3.75 times as efficient. Transect sampling was always more efficient than rectangular cluster sampling for all four distributions.

Empirical Study

Gross sampling effort (Table 4) required to maximize the predictive performance of local species distribution models was greater for the species in YNP (72 to 128 transects) versus INL (14 to 56 transects). However, when normalized to the percent of the area sampled, results were surprisingly consistent ($\bar{x}$ = 0.15, SD = 0.05), and there was no detectable difference between sites (Fig. 5). Maximum mean AUC was variable across sites ($\bar{x}$ = 0.81, SD = 0.09) and there was evidence of a difference in maximum AUC between sites (Fig. 6) with an estimated mean difference of 0.09 with an associated 95% confidence interval of 0.0011 to 0.179. YNP had higher mean AUC ($\bar{x}$ = 0.857, SD = 0.08) than INL ($\bar{x}$ = 0.766, SD = 0.084).

Maximum AUC decreased as a function of a species' landscape frequency (Fig. 7) with the most frequent species, *Alyssum desertorum* and *Bromus tectorum* at the INL, having the lowest maximum AUC of 0.660 and 0.685 respectively. The trend held even when the two high frequency species were removed from the analysis. Although, this trend was evident, there was a large amount of variation in maximum predictive performance, especially for the low frequency species.

Discussion

Maps of local to landscape scale distributions of non-native plant species have the potential to greatly increase management efficiency when integrated into invasive plant species management planning and execution (Evangelista et al. 2008; Higgins et al. 1999; Rew et al. 2005; Shafii et al. 2003). Widespread adoption of landscape scale distribution models and maps is slow, as data do not exist to make predictions at resolutions commensurate with on-the-ground management. The perception is that data are too expensive to acquire, plus there has been little quantification of the predictive capacity of species distribution models at the scale that management is implemented. In this study, we determined an efficient way to sample, implemented that methodology in the field, and evaluated the predictive performance of different non-native plant species models at two different sites and sample sizes.

How to Sample

Adoption of local to landscape scale species distribution models is contingent upon having an efficient way to collect data to build models and their associated maps. The simulation model demonstrated that the two random sampling methods maximized model performance with the least amount of sampling for all four plant species distribution patterns when strictly looking at the number of plots required. There was also no clear difference between simple random sampling and stratified random sampling even for the pattern that strongly associated a species distribution with the road.

Rew et al. (2006) demonstrated using targeted transects was five times faster than gathering data using random sampling when walking cost was included in sampling efficiency. However, they did not consider the effects of using clustered data for fitting SDMs and predicting probability of occurrence to the entire management area. Our extension of that initial work found that when model predictive potential was factored into the number of samples needed, transects were 2.5 to 4.6 times more efficient than simple random sampling. Randomly placed rectangular clusters of plots were more efficient than random sampling of the same number of plots; however, they were less efficient than transects for all species distribution patterns. Transects are also better at sampling the environmental variation than rectangular clusters, because of the inherent spatial autocorrelation of geographic data (Elith and Leathwick 2009).

The sample size at which the model performance metrics associated with transects became equivalent to simple random sampling was lower for the two patterns strongly associated with the road compared to the two distributions that were weakly associated with the road. This indicated that if there was an *a priori* reason to believe a species was associated with the road, transects of plots perpendicular to the road could very effectively capture that relationship for modeling the species distribution. More samples were needed if there was a weak/positive association of plant occupancy to distance to road, though the transect method still proved to be an efficient way to collect the data as transects were still 4.6 to 4.1 times more efficient than simple random sampling for the WPL and WPH distribution patterns respectively (Table 3).

The rectangular cluster method proved to be less efficient than transects. However, our models showed that with sufficient data, the predictive performance of models using rectangular clusters did approach random sampling methods for all plant distribution patterns. Much data exists in management agencies that emulates characteristics of the rectangular cluster sampling (e.g. many full census polygons across the landscape) that could be aggregated to generate species distribution maps. However, our results demonstrate the importance of being cautious about using models/maps generated from clustered data at low sample sizes as the proportion of instances under which the model performed poorly were much greater than simple and stratified random sampling and transect sampling. The mean prediction error showed that for the three low frequency species distributions, using clustered data at small sample sizes led to a high probability of severely underestimating occupancy rates (Online Resource 2). For the higher frequency species, there was a chance of both over and under prediction at low sample sizes but the mean response was to under-predict occupancy rates. Furthermore, there was no systematic bias in the site selection for our rectangular clusters, but there is likely bias in manager databases. However, Wisz et al. (2008) found that biased sampling along roads produced good models, indicating using biased data at large enough sample sizes may be robust. Regardless, it is our conclusion that NIS distribution models and subsequent maps will be improved if managers incorporate the targeted transect method identified here.

How Much to Sample

Practitioners are generally limited by money and time. Therefore, having some general guidelines as to necessary sample size could inform the planning and implementation of data collection to generate species distribution models (SDMs) for local management units. Most studies of sample size effects on SDMs were at coarse scales and with presence only methods and bioclimatic predictors, suggesting 20 to 100 points were adequate (Hernandez et al. 2006; Stockwell and Peterson 2002; Wisz et al. 2008). One study recommends at least 50 occurrences to estimate species response curves when using presence and absence data with logistic regression (Coudun and Gégout 2006) but little information exists about sample size requirements and SDM performance at landscape scales and fine (10 m) resolutions.

Our simulation study showed that under transect sampling, the range at which predictive performance of the models reached 95% of maximum was from 0.08% to 0.48% of the landscape sampled. The two plant patterns strongly associated with the road (most NIS-like) needed the lowest amount of area sampled at 0.08% and 0.28% for the specialist (SNLS) and generalist (SNL) patterns respectively. Whereas the two patterns less strongly associated with the road needed ~0.46% (0.44% for WPH and 0.48% for WPL) of the area to be sampled for maximum predictive performance. Regardless, this demonstrates that the amount of area necessary to produce reliable maps may be attainable by practitioners within their logistic and budget constraints.

Empirical data corroborated our simulation study for transects, actually suggesting smaller sample sizes should effectively maximize model predictive performance. Although gross sampling effort was greater for YNP, after normalizing to the size of the study area, there was no difference in the amount of sampling required to maximize model predictive performance. The results suggested sampling on average 0.15% of the landscape with some species requiring up to 0.23% to maximize model performance. These values fit within the realm of the simulation study results and again, represent attainable sampling densities for practical implementation.

Effect of Site

The two sites considered in the empirical portion of this study were markedly different in their environmental heterogeneity. The INL had relatively shallow environmental gradients with only 177 m of elevation change while YNP had 1,245 m of elevation difference. YNP had multiple vegetative cover types while the INL was characterized primarily by one vegetation cover. Although there was no significant difference in the amount of data required to maximize the performance of models, maximum predictive performance was significantly higher in YNP. It follows that steeper environmental gradients would lead to more intense and detectable ecological filtering, thus greater discriminatory power of the models based on these principles. That is indeed what we found, but the similarity in sampling requirements facilitates making sample size recommendations, independent of site characteristics.

Effect of Species Prevalence

Predictable effects of species characteristics on model performance can guide the applied use of species distribution models (Franklin 2009). Studies have shown better model performance of species with narrower ecological tolerances and that are rarer on the landscape (e.g. Loiselle et al. 2008; Luoto et al. 2005; McPherson and Jetz 2007); however, these have been primarily carried out at the regional, continental and global scales. The effectiveness of species distribution models for specialists and/or locally rare species has not been considered at smaller scales. We found model performance decreased as species prevalence increased, suggesting a trend similar to studies of species at greater grain size and extents. However, we did not find a consistent trend for niche specialist versus generalist species across our two study sites suggesting that the generalist/specialist dichotomy works for coarse grain size and large geographic extents. Thus, one may generalize that knowledge of where a species is within its ecophysiological range is more important for local (finer grain) prediction of occurrence relevant to management than knowledge about its ecophysiological limits.

B.tectorum is generally considered a generalist and to exhibit broad environmental tolerances to ecophysiological variables. Consequently, when *B. tectorum* was compared with *Tamarix chinensis*, a specialist, the *B. tectorum* model gave poorer predictions than the *T. chinensis* model at coarse spatial resolutions for the whole USA (Evangelista et al. 2008). Conversely, in our empirical study, we had two species (one in common with the previous study) that exemplified generalist characteristics, *B. tectorum* and *A. desertorum*, and are associated with Great Basin vegetation. These species were the most frequent and had the least predictive models at the INL, but had low frequencies and much more predictive models at YNP. Thus, the generalist/specialist categorization did not work across sites. However, YNP has very limited habitats characteristic of the Great Basin, while INL is fully characterized by Great Basin vegetation. This indicated that YNP was on the limits of these two species' ecophysiological range, while the INL was located in the core of their ecophysiological range, thus their behaviors switched between sites from being specialists at YNP to generalists at INL subsequently being more predictable at YNP versus INL. Thus, it is more important to know where a species is along its ecophysiological niche space (limits vs. optima) to predict the potential discrimination capacity of landscape scale SDMs. Luoto et al. (2005) had similar findings with butterfly species in regional scale models.

We suggest that continental scale bioclimatic SDMs could be used to understand whether a species is at its ecophysiological core or limits at a landscape scale relevant to local management. Consequently, if a bioclimatic model predicts low probability of occurrence, one may expect that species to be a low frequency specialist versus if the bioclimatic model predicts high probability one would expect it to be a high frequency generalist within smaller management units (i.e. the landscape scale).

Conclusions

We presented a logistically feasible and efficient way for managers to sample for non-native plant species occurrence data and predict species distribution at scales relevant to on-the-ground management. Randomly placed, targeted transects that are 2 km long (or extend to the edge of a management area) and 10 m wide and perpendicular to roads can be used to efficiently generate data to inform models to predict species distribution. Models will be most predictive in areas with steeper environmental gradients where environmental filtering is more important, but can still predict better than random in areas that are less environmentally heterogeneous. Species at their range limits and that are rare on the landscape are easier to accurately predict than species at the core of their ecophysiological range, and practitioners should aim to sample between 0.06% and 0.23% of their study area to ensure model performance is maximized for predicting species distribution.

Acknowledgements

Funding for this investigation was provided mainly by Stoller Corporation, project number GSS-0300-1100008, but with additional support from CERL-ERDC W9132T-09 and NRI 2009-55320-05033 Special thanks to Matt Lavin for ecological discussions, suggestions and moral support. Thanks to land management agency people at YNP: Mary

Heckner, Roy Rankin, Christie Hendrix and at INL: Roger Blew, as well as all the field crews and technicians that facilitated data collection and collation:, Frank Dougher, Judit Barroso, Jerad Corbin, Matthew Hulbert, , Mara Johnson, Rebecca Kennedy, Erik Lehnhoff, Ben Levy, Amanda Morrison, Charles Repath, John Rose, Kit Sawyer, Alex Wing, Rose Marks, Elizabeth Usher, Kim Taylor, Tyler Schober, Kevin Shields, Dan Campbell.

Tables and Figures

Table 2.1: Parameters and associated coefficients used to generate simulated species patterns, and the landscape frequency of each simulated species. The four plant distribution patterns were SNLS – Strong Negative Low frequency Specialist, SNL – Strong Negative Low frequency generalist, WPL – Weak Positive Low frequency generalist, and WPH – Weak Positive High frequency generalist.

	Distribution			
Parameter	SNLS	SNL	WPL	WPH
Intercept	19.96617	-1.11600	-5.09034	1.05506
Cosine(aspect)	-0.69414	0.25040	-0.14827	-0.05384
Sine(aspect)	-0.59102	0.19490	0.69982	0.29343
Slope	0.00323	0.00616	0.00750	-0.00898
Distance to road	-0.00100	-0.00051	0.00013	0.00007
Elevation	-0.01129	-0.00040	0.00065	-0.00109
Landscape Frequency	0.036	0.037	0.037	0.225

Table 2.2: Number of plots at which 95% of the maximal mean AUC was achieved for the four sampling strategies (SimpleRS = simple random sample, StratRS = Stratified Random sample, RectCluster = Rectangular clusters) and four plant distribution patterns (Strong Negative, Low frequency, Specialist- SNLS; Strong Negative, Low frequency generalist – SNL; Weak Positive, Low frequency generalist – WPL; Weak Positive, High frequency generalist – WPH). Values in parenthesis are the number of transects to achieve the stated number of plots. Note that 100 plots represent 0.01% of the area.

	Sampling Strategy			
Distribution	SimpleRS	StratRS	Transects	RectCluster
SNLS	400	400	800 (4)	1200 (6)
SNL	2400	2800	2800 (14)	3200 (16)
WPL	4400	4000	4800 (24)	6000 (30)
WPH	3600	4800	4400 (22)	5200 (26)

Table 2.3: Efficiency of sampling strategies relative to simple random sampling for four species distribution patterns. The three sampling strategies compared to simple random sample were StratRS = Stratified Random sample, and RectCluster = Rectangular clusters. The four plant distribution patterns were Strong Negative, Low frequency, Specialist- SNLS; Strong Negative, Low frequency generalist – SNL; Weak Positive, Low frequency generalist – WPL; Weak Positive, High frequency generalist – WPH.

	Efficiency (Relative to SRS)		
Distribution	StratRS	Transects	RectClusters
SNLS	1.00	2.50	1.67
SNL	0.86	4.29	3.75
WPL	1.10	4.58	3.67
WPH	0.75	4.09	3.46

Table 2.4: Amount of data required to reach 95% of the maximum predictive capacity of landscape scale species distribution models for 16 introduced species at two sites. The sites are Yellowstone National Park, Wyoming, USA (YNP) and the Idaho National Laboratories, Idaho, USA (INL). Number of transects were normalized to the percent of the study area sampled (% sampled) to facilitate comparisons between sites. Area under the receiver operator characteristic curve (AUC) was used as the model performance metric. The proportion of presences in the dataset is also reported (Freq.).

Site	Species	Transects	Ha Sampled	% Sampled	Max AUC	Freq.
YNP	*Alyssum desertorum*	72	144	0.12	0.964	0.039
YNP	*Bromus inermis*	100	200	0.17	0.844	0.044
YNP	*Bromus tectorum*	112	224	0.19	0.888	0.063
YNP	*Cirsium arvense*	128	256	0.21	0.730	0.050
YNP	*Cynoglossum officinale*	84	168	0.14	0.852	0.005
YNP	*Linaria dalmatica*	108	216	0.18	0.872	0.034
YNP	*Melilotus officinalis*	100	200	0.17	0.944	0.002
YNP	*Phleum pratense*	76	152	0.13	0.760	0.229
INL	*Agropyron cristatum*	26	52	0.11	0.893	0.165
INL	*Alyssum desertorum*	14	28	0.06	0.660	0.704
INL	*Bromus tectorum*	26	52	0.11	0.685	0.819
INL	*Carduus nutans*	42	84	0.17	0.843	0.027
INL	*Halogeton glomeratus*	56	112	0.23	0.833	0.046
INL	*Salsola kali*	54	108	0.22	0.789	0.050
INL	*Sisymbrium altissimum*	30	60	0.12	0.728	0.212
INL	*Tragopogon dubius*	36	72	0.15	0.702	0.156

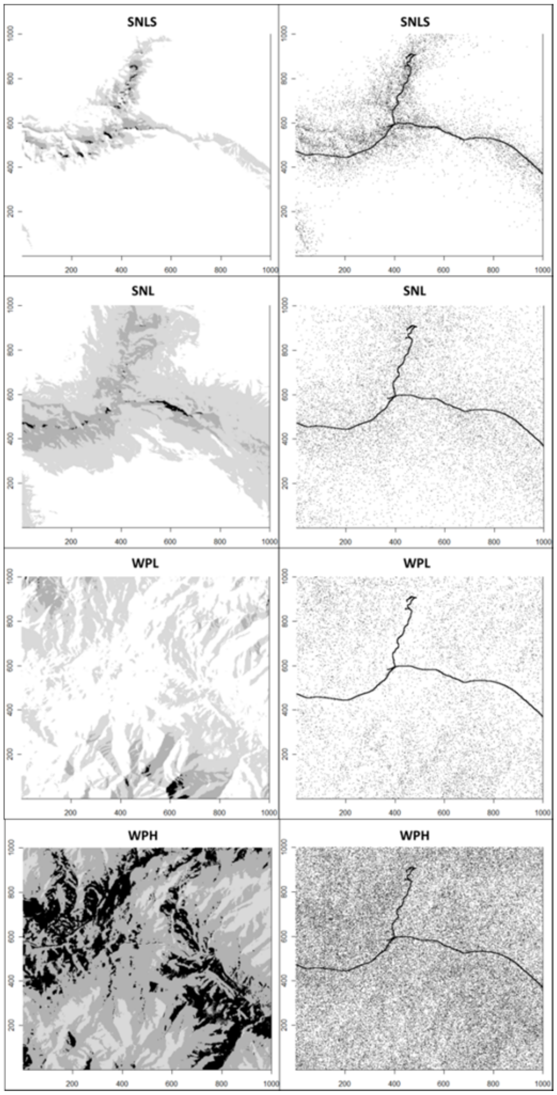


Figure 2.1: Management area maps of the four plant probability of occupancy surfaces (left panel) and examples of corresponding stochastic realizations (right panel). The left panel shows the distribution of the true probabilities of occurrence (PO) and is classified into four gray scale classes with lighter colors being lower PO and darker being high PO. Values for PO start at 0 and go to a maximum of 0.45 for the top and bottom panels and 0.15 for the middle two panels. The right panel is one stochastic realization of the species distribution (white indicates absence, black indicates presence). The road is shown as a black line through each stochastic realization. The four plant patterns were SNLS – Strong Negative Low frequency Specialist, SNL – Strong Negative Low frequency generalist, WPL – Weak Positive Low frequency generalist, and WPH – Weak Positive High frequency generalist.

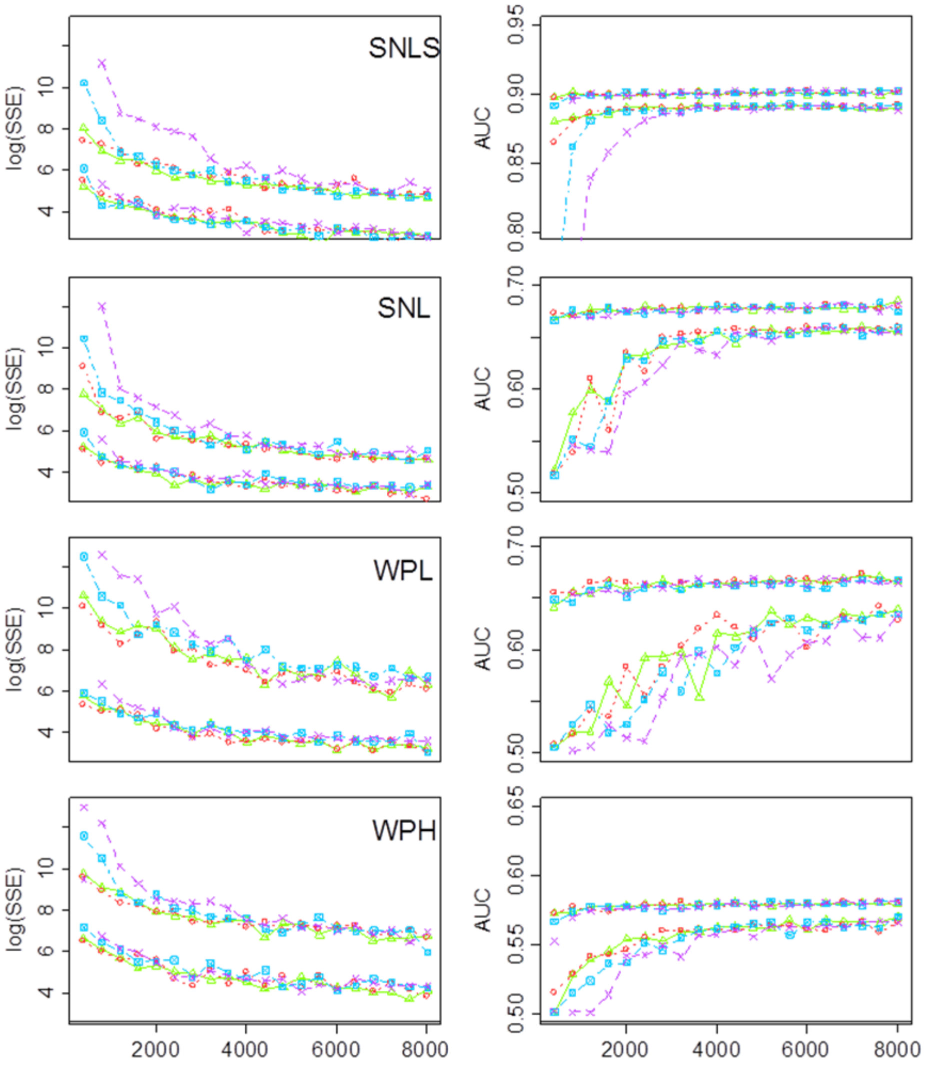


Figure 2.2: Left panel shows the 95% data intervals (0.025 and 0.975 percentiles) for the distribution of the logarithm of the sum of squared error between the true probabilities and the predicted probabilities obtained from the model fit to simulated data as a function of sample size for the four plant patterns. The right panel shows 95% data interval boundaries for the distribution of AUC values for each plant distribution pattern as a function of sample size. The four plant patterns were SNLS – Strong Negative Low frequency Specialist, SNL – Strong Negative Low frequency generalist, WPL – Weak Positive Low frequency generalist, and WPH – Weak Positive High frequency generalist. The four sampling methods were– Simple Random Sample (red circles), Stratified Random Sample (green triangles), Transects (blue squares) and Rectangular Clusters (purple X's)

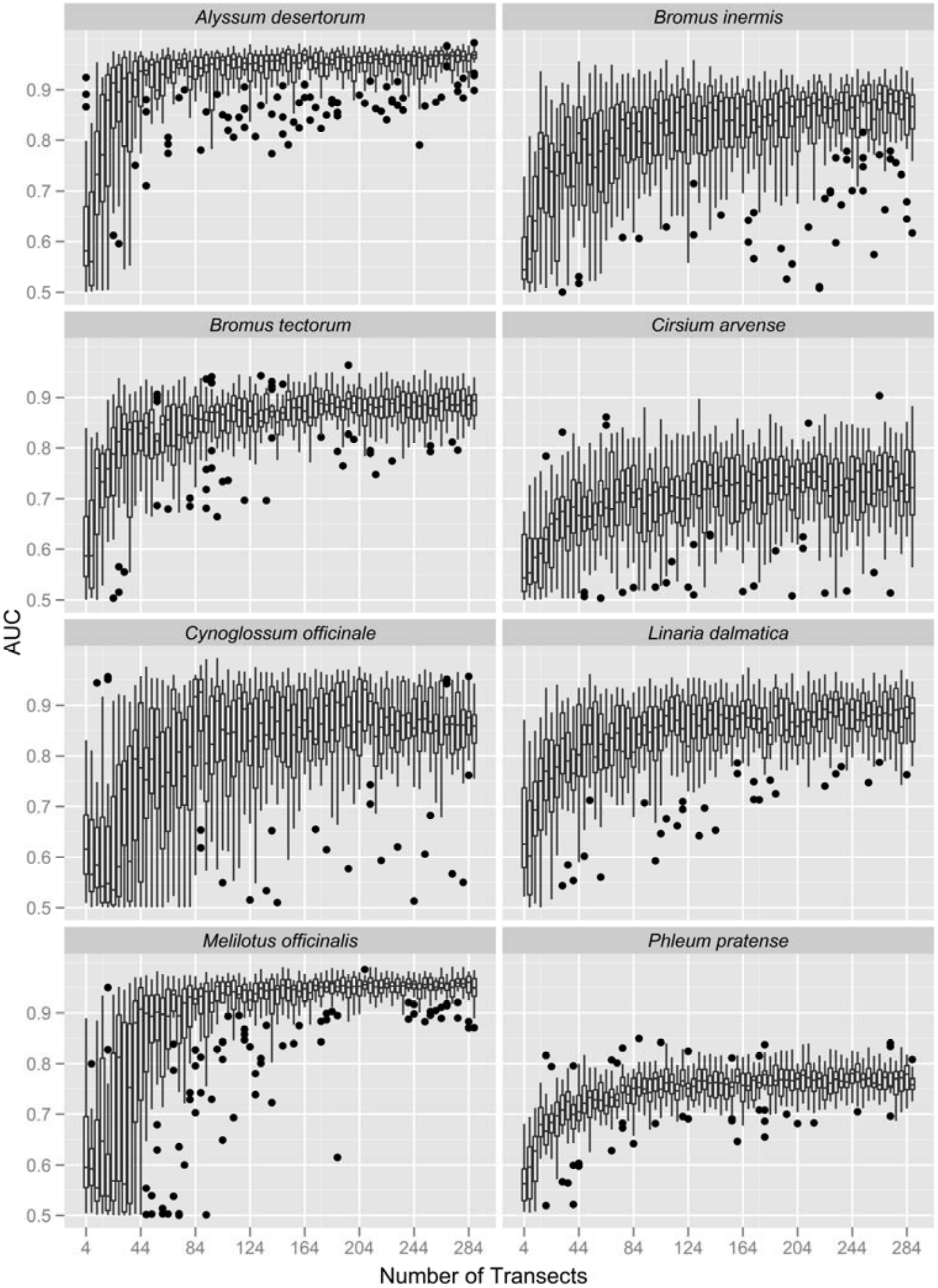


Figure 2.3: Boxplots of area under the receiver operator characteristic curve (AUC) as a function of sample size for eight species from Yellowstone National Park from a cross-validation resampling procedure. Sample size ranged from 4 to 288 transects and increased in increments of 4 transects. There were 20 resampling cross validation replicates at each sample size.

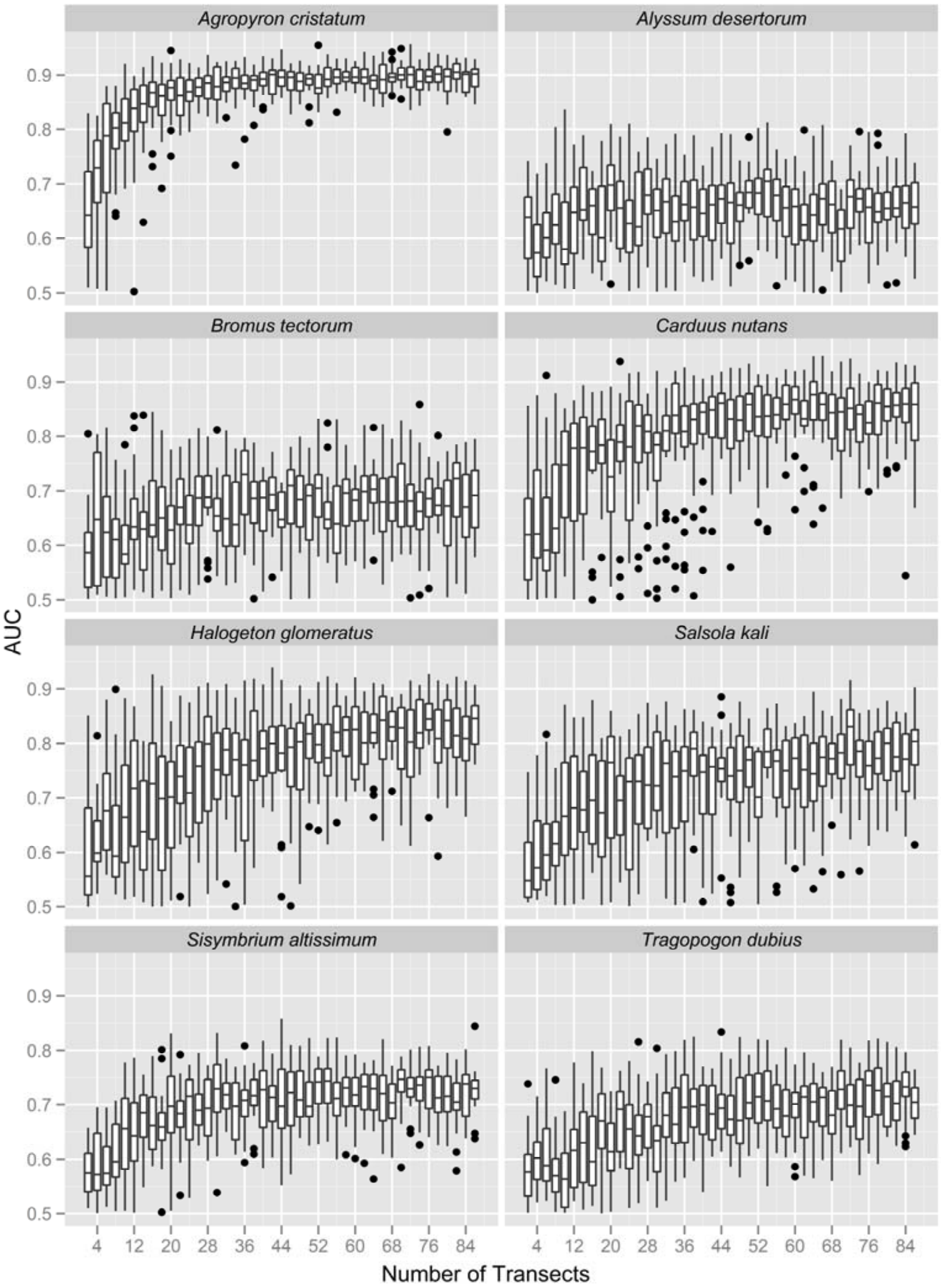


Figure 2.4: Boxplots of area under the receiver operator characteristic curve (AUC) as a function of sample size for eight species from the Idaho National Laboratories from a cross-validation resampling procedure. Sample size ranged from 2 to 86 transects and increased in increments of 2 transects. There were 20 resampling cross validation replicates at each sample size.

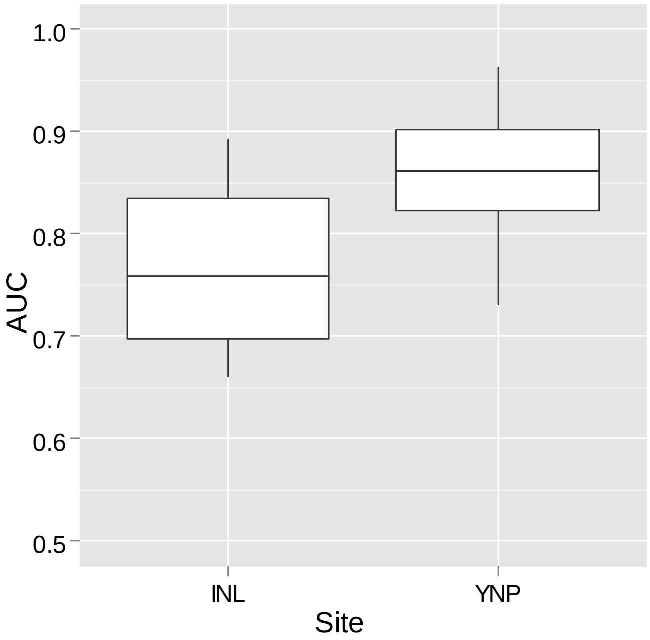


Figure 2.5: Boxplots of the maximum AUC for 16 NIS species from two sites. Models of species in Yellowstone National Park (YNP) had overall higher maximum predictive performance compared to models of species from the Idaho National Laboratories (INL) (p = 0.047 from a two-sample t-test).

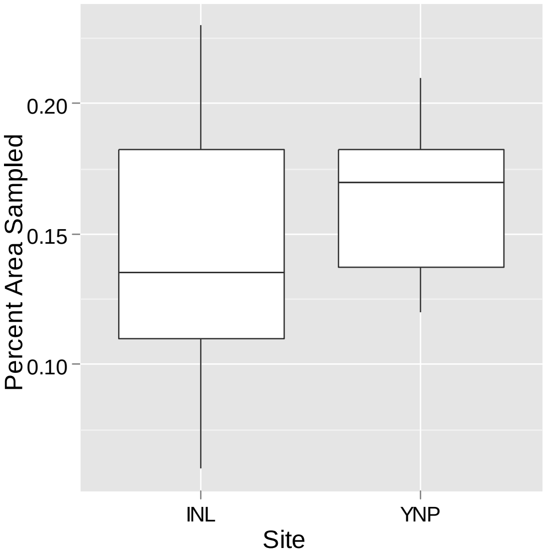


Figure 2.6: Boxplots of the percent area necessary to sample to maximize the predictive performance of models for 16 NIS species at two sites. There was no difference (p = 0.47 from Welch two-sample t-test) between species at Idaho National Labs (INL) compared to species at Yellowstone National Park (YNP).

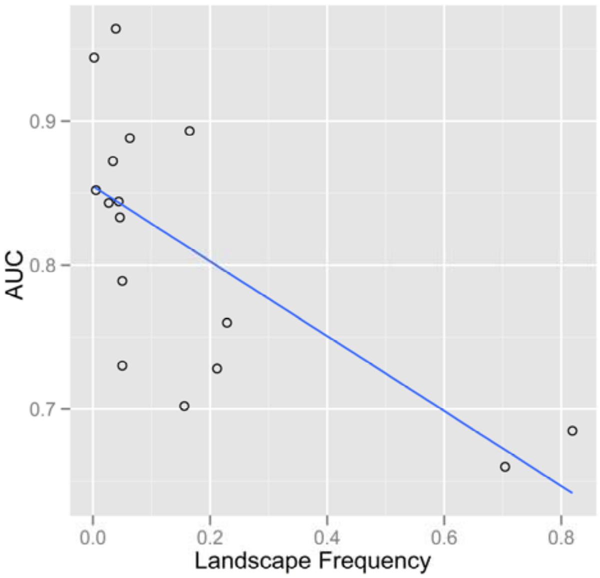


Figure 2.7: Maximum area under the receiver operator character curve as a function of landscape frequency for 16 NIS species across two sites (YNP and INL). The predicted regression line from a simple linear regression had a significant negative slope ($p = 0.003$) indicating decreased model maximum performance with for species of greater frequency independent of site.

References

Coudun C, Gégout J-C (2006) The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. Ecological Modelling 199:164-175

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677-697

Evangelista PH, Kumar S, Stohlgren TJ, et al. (2008) Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions 14:808-817

Fleishman E, Mac Nally R, Fay JP, et al. (2002) Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. Conservation Biology 15:1674-1685

Franklin J (2009) Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge, UK

Gelbard JL, Belnap J (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology 17:420-432

Graham C, Ferrier S, Huettman F, et al. (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution 19:497-503

Hanski I (1999) Metapopulation ecology. Oxford University Press, Oxford

Hernandez PA, Graham CH, Master LL, et al. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773-785

Hierro JL, Maron JL, Callaway RM (2005) A biogeographical appraoch to plant invasions: the importance of studying exotics in their introduced and native range. Journal of Ecology 93:5-15

Higgins SI, Richardson DM, Cowling RM, et al. (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conservation Biology 13:303-313

Hufbauer RA, Facon B, Ravigné V, et al. (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. Evolutionary Applications 5: 89-101

Loiselle BA, Jørgensen PM, Consiglio T, et al. (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? Journal of Biogeography 35:105-116

Lozier JD, Mills NJ (2011) Predicting the potential invasive range of light brown apple moth (Epiphyas postvittana) using biologically informed and correlative species distribution models. Biological Invasions 13:2409-2421

Luoto M, Pöyry J, Heikkinen RK, et al. (2005) Uncertainty of bioclimate envelope models based on the geographical distribution of species. Global Ecology and Biogeography 14:575-584

McPherson JM, Jetz W (2007) Effects of species ecology on the accuracy of distribution models. Ecography 30:135-151

Parendes LA, Jones JA (2000) Role of light availability and dispersal in exotic plant invasion along roads and streams in the H. J. Andrews Experimental Forest, Oregon. Conservation Biology 14:64-75

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling 133:225-245

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology & Biogeography 12:361-371

Peterson T (2003) Predicting the Geography of Species' Invasions via Ecological Niche Modeling. The Quarterly Review of Biology 78:419-433

Pulliam RH (1988) Sources, Sinks, and Population Regulation. The American Naturalist 132:652-661

Reese GC, Wilson KR, Hoeting JA, et al. (2005) Factors affecting species distribution predictions: a simulation modeling experiment. Ecological Applications 15:554-564

Rew LJ, Lehnhoff EA, Maxwell BD (2007) Non-indigenous species management using a population prioritization framework. Canadian Journal of Plant Science 87:1029-1036

Rew LJ, Maxwell BD, Aspinall R (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Science 53:236-241

Rew LJ, Maxwell BD, Dougher FL, et al. (2006) Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species. Biological Invasions 8:523-539

Richards CL, Carstens BC, Lacey Knowles L (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography 34:1833-1845

Seipel T, Kueffer C, Rew LJ, et al. (2012) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. Global Ecology and Biogeography 5:236-246

Shafii B, Price WJ, Prather TS, et al. (2003) Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Science 51:748-751

Stockwell DRB, Peterson T (2002) Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1-13

R Development Core Team (2011) R: A language and environment for statistical computing. 2.13.1 edn. R Foundation for Statistical Computing, Vienna, Austria,

Václavík T, Kanaskie A, Hansen EM, et al. (2010) Predicting potential and actual distribution of sudden oak death in Oregon: Prioritizing landscape contexts for early detection and eradication of disease outbreaks. Forest Ecology and Management 260:1026-1035

Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling 220:3248-3258

Václavík T, Meentemeyer RK (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Diversity and Distributions 18:73-83

Wilson JRU, Dormontt EE, Prentis PJ, et al. (2009) Something in the way you move: dispersal pathways affect invasion success. Trends in Ecology & Evolution 24:136-144

Wisz MS, Hijmans RJ, Li J, et al. (2008) Effects of sample size on the performance of species distribution models. Diversity and Distributions 14:763-773

Zimmermann NE, Edwards TC, Moisen GG, et al. (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. Journal of Applied Ecology 44:1057-1067

CHAPTER THREE

DETECTION ERROR IN PLANT SURVEYS: TO CORRECT OR NOT TO CORRECT

Contributions of Authors and Co-Authors

Manuscript in Chapter 3

Author: Tyler J. Brummer

Contributions: Conceived study, collected and analyzed data and wrote manuscript.

Co-author: Bruce D. Maxwell

Contributions: Obtained funding for the project in conjunction with Lisa Rew, assisted in study design, discussed results and provided revisions and edits.

Co-author: Subhash R. Lele

Contributions: Provided discussions on statistical methods, discussed results and provided comments on manuscript. (Revise if needed)

Co-author: Lisa J Rew

Contributions: Obtained funding for the project, assisted in study design, developed ideas, discussed results and revised and edited the manuscript at all stages.

Manuscript Information Page

Tyler J. Brummer, Bruce D. Maxwell, Subhash R. Lele, Lisa J. Rew
Journal Name: Journal of Applied Ecology
Status of Manuscript:
_x_Prepared for submission to a peer-reviewed journal
___Officially submitted to a peer-reviewed journal
___Accepted by a peer-reviewed journal
___Published in a peer-reviewed journal
Published by British Ecological Society

Abstract

1) Imperfect detection is a potential source of uncertainty in predictive models of non-native plant probability of occurrence employed for understanding and management. Detection error has been little investigated using empirical data. We investigated the influence of imperfect detection and modelling method on landscape scale predictive maps and the uncertainty associated with those maps.
2) Using multi-visit data from a multispecies non-native plant survey, we evaluated empirical detection error rates using a hierarchical site occupancy model. Detection rates ranged from 0.24 to 0.94, however the majority of detection rates were greater than 0.75. Accounting for detection error did not substantially change the rank of species occurrence.
3) Habitat maps generated using logistic regression models that ignored, or corrected for detection, were compared to hierarchical model results. Practical changes to predicted probability of occurrence values were not observed associated with changes in detectability of 10% due to season or observer. Probability of occupancy predictions were robust to choice of modelling method for four of the eight species studied, while predictions changed substantially for three of the species.
4) Uncertainty was greater when the hierarchical model was used for three of eight species but was unchanged for four of eight species. Predictions and uncertainties for species with detection rates higher than 87% were invariant to modelling method.

5) *Synthesis and applications:* Weighing the cost of sampling more sites versus resampling sites within season to quantify detection bias is a dilemma managers are faced with when trying to model non-native plant species' probability of occurrence. As long as the magnitude of change in detection over the sampling period is low (e.g. 10% in our case) predictive maps will likely be practically unbiased. However, if detection is below 87% the uncertainty in predictions may be underestimated if data are modelled via logistic regression assuming perfect detection instead of using multi-visit data to account for imperfect detection. This may lead to false confidence in the data and maps that drive decision making. Habitat relationships also may change depending on modelling method for species with <87% detection rate and is further exacerbated when species are low frequency. We suggest investing effort in initial training to ensure >87% detection rates and, to schedule sampling during the peak growing season to prevent costly investment in a multi-visit sampling design.

Introduction

Local scale predictions of plant species occupancy, especially non-native plant species, are becoming an accepted tool to aide in management planning and implementation (Rew et al., 2007). Easily interpretable logistic regression models using species presence and absence data and geo-referenced abiotic environmental predictor variables can, for many species, effectively differentiate a gradient of probability of occupancy from survey data (Franklin et al., 2009; Syphard & Franklin, 2009). Maps produced from these models can help prioritize species and populations of a species for management and dramatically decrease and direct the future search area for target species (Miller et al., 2010; Rew et al., 2007).

Imperfect detection can influence estimates of occupancy and has been extensively studied in wildlife research. MacKenzie et al. (2002) developed a very intuitive hierarchical binomial mixed model that models detection and occupancy hierarchically under a sampling design where sites are visited two or more times. This approach has been applied to numerous wildlife survey data (e.g. Gu & Swihart, 2004; Kroll et al., 2008; Smith et al., 2007), but has been less extensively used to correct for imperfect detection in plant studies. However, imperfect detection in plant studies has been shown to bias estimates of vital rates in demography studies (Alexander et al., 2009; Kéry & Gregg, 2003), population size (Alexander et al., 1997), extinction rates as well as covariates associated with extinction (Kéry, 2004), and can alter management strategies in the case of invasive plants (Regan et al., 2011; Regan et al., 2006). But, detection error has not been explicitly incorporated into empirical studies of plant probability of occurrence modeling that aim to generate maps of species distribution.

A few empirical studies of the drivers of plant detectability have shown an increase in detection rates with increased search effort, and a variety of other factors including time of day, weather, observer experience, abundance, herbivory, species phenology and patch size (Alexander et al., 2009; Chen et al., 2009; Garrard et al., 2008; Moore et al., 2011). Studies have demonstrated a broad range of detection error rates (0.09 to 0.91) for a few plant species using a few quadrat sizes (0.04 to 4.5 ha)

(Alexander et al., 1997; Chen et al., 2009; Regan et al., 2006; Shefferson et al., 2001). Even fewer studies have evaluated how ignoring detection error may bias predictions of species distribution (but see Rota et al., 2011). This begs the questions: (1) is detection error pervasive enough to sufficiently bias local scale models of species probability of occupancy, and consequently alter management decisions? and (2) is the cost of collecting data to estimate detection error worth the reduction in bias?

Failure to correct for uniform detection error if present, will negatively bias estimates of probability of occupancy (Gu & Swihart, 2004) (i.e. we will assume species are less abundant than they really are). In the case of non-indigenous plant species (NIS), this underestimate could change management priorities and practices. Further, systematic changes in the detection rate throughout a sampling period may bias coefficient estimates of variables related to occupancy (Gu & Swihart, 2004), altering predictions (maps) that management decisions are centered around. Thus, an empirical investigation of detection rates for NIS is necessary to understand the degree to which detection error may influence probability of occupancy predictions for NIS.

One final dimension rarely considered is the effect of ignoring detection error on the uncertainty of predictions (i.e. confidence interval widths). Maps that predict the mean probability of occupancy at all sites of the study/management area are the end product that will be integrated into management planning (Rew et al., 2007). Rocchini et al. (2011) stated that accompanying these predictions should also be “maps of ignorance” where the certainty of predictions is also communicated. The data requirements to effectively disentangle both detection and occupancy processes hierarchically while maintaining acceptable uncertainty in predictions (Mackenzie & Royle, 2005) may be prohibitive for NIS due to the characteristics of the species (regional and local rarity) as well as the added cost in data collection. However, failure to explicitly incorporate uncertainty due to detection error may provide false confidence in estimated probability of occupancy rates and in some cases bias habitat coefficient estimates. Thus, we investigated how naïve estimates of occupancy probability compared to those corrected for detection, and how the uncertainty in both estimates differed, using empirical plant occurrence data.

Our goals were to determine how well we could detect a range of non-native plant species and how ignoring, or incorporating, detection error affected the interpretation and management implications from probability of plant occurrence maps.

Methods

Study Site

Data were collected within a 20 km east-west and 30 km north-south area at the Idaho National Laboratories (INL) in south-central Idaho, USA. The study area was classified as a sagebrush steppe ecosystem, dominated by *Artemisia tridentata* Nutt. ssp. *wyomingensis* Beetle & Young and various *Chrysothamnus* species. Annual precipitation averaged 22 cm with a mean annual temperature of approximately 5.5°C and the primary growing season was between late May and early July. Elevation ranged between 1461 m and 1638 m.

Data Collection Methods

Data on the presence and absence of non-native plant species were collected during late May through late July over three growing seasons (2009 to 2011). Data collection to investigate detection error was nested within a larger non-native plant survey that aimed to predict the distribution of all target species at the study area. Non-native plant occurrence was recorded along transects, originating from randomly chosen start points along roads and traveled perpendicular to roads. The transects were 2 km long by 10 m wide and presence and absence was recorded in each 10 m x 10 m plot along the transect. Thus, a transect consisted of a linear cluster of 200 plots, perpendicular to roads (Rew et al., 2005; Rew et al., 2006). The perpendicular arrangement was used to emulate a stratified random sampling approach, stratified on distance to road, while maintaining logistical feasibility by collecting 200 adjacent plots. Distance to road was used as a stratification variable because other studies have empirically demonstrated that NIS occurrence decreases as distance to road increases (Parendes & Jones, 2000; Seipel et al., 2012).

To assess detection error associated with the survey, a subset of the plots surveyed in each of the years were repeated near the end of the sampling period within each season. In 2009 five entire transects, and in 2010, 10 entire transects were repeated for species occurrence. Because of the clustered nature of this sampling (200 plots along a transect, but few transects) a different approach was devised in 2011 that had more spatial and temporal independence in the repeated plots. Thirty-seven full 2 km transects were completed for eight target species as part of the 2011 survey (visit 1). To evaluate detection error, two - 200 m segments (20 contiguous plots) were sampled a second time (visit 2) for each transect. The location of the 200 m segments along the original transects was randomly chosen, one between 0 and 1 km and the other between 1 and 2 km from the road. Segments were relocated using sub-meter accuracy Trimble® GeoXT® global positioning units. This resulted in 74-200 m segments being visited twice, or 1477 plots with 2 visits each and is the only data used in the following analyses. This latter approach allowed covariates for detection rate to be assessed in a hierarchical framework (MacKenzie *et al.*, 2002).

Statistical Models and Data Used

Data were used in four different ways depending on modelling and comparison objectives. The two-visit data gave us four possible detection histories: 00 if not observed both visits, 10 if observed the first visit but not the second, 01 if not observed the first visit but observed the second and 11 if observed at both visits.

Three datasets were derived from these data for different comparisons; data from just Visit 1, data from just Visit 2, and the maximum of detection the history (combining Visit 1 and 2 with presences overriding absences). These three sets of data were used in logistic regression models that did not explicitly model the probability of detection. Finally, both the Visit 1 and 2 data were used explicitly in hierarchical detection/occupancy models that explicitly modeled detection error and used that information to correct estimates of occupancy.

The four different treatments of the data (Visit 1, Visit 2, the maximum the visits or “Combined” and the explicit use of both visits) along with two modelling methods (logistic regression and hierarchical detection occupancy) gave us five unique combinations of data and modelling method and will be referred to as Models 1 through 5 throughout the paper. Model 1 used Visit 1 in a logistic regression, Model 2 used Visit 2 in a logistic regression, and Model 3 used Combined data in a logistic regression. Model 4 used Visit 1 and 2 in a hierarchical occupancy model, with the detection process as intercept only, and Model 5 was the same as Model 4 but allowed detection to vary as a function of covariates. Depending on the objectives of each analysis, covariates were either excluded or included for the occupancy process in these five different models. Statistical analyses were carried out in the R statistical package (R Core Development R Development Core Team, 2012). Hierarchical occupancy models *sensu* MacKenzie et al. (2002) were used as facilitated by the "unmarked" package in R (Fiske & Chandler, 2011).

Mean Detection Rates

To evaluate the range of detection rates for the different species over the three years we analyzed the two-visit data using a hierarchical detection/occupancy model (Model 4) with intercept only terms for both processes. Data from 2009 and 2010 were analyzed together and data from 2011 was analyzed separate due to the different sampling methodologies.

Assessing Bias Due to Detection Error

To assess bias in estimates of mean occupancy due to detection error, we analyzed the two-visit plant occurrence data from 2011 using Models 1 through 4, with an intercept only term for the occupancy process. Back transformed estimates of the mean probability of occupancy and detection were determined and empirical 95% confidence intervals calculated from a non-parametric bootstrapping routine.

To assess for changes in detection rate during the sampling period and the effect of observer, covariates were added to the detection piece of the hierarchical model (Model 4), but only 2011 data were used. Detection covariates considered were: days since first sample day, observer and a quadratic term for sample day to test for a potential optimum detection time. Sites were always observed by one of two “expert” observers. Both observers had two seasons of experience in the system and previous plant identification experience. Sampling occurred over a 25 day period in June of 2011. Each observer was accompanied by a data recorder. Six combinations of the covariates (Table 3.3) were incorporated into the detection portion of Model 4 and support for each model was assessed using Akaike’s Information Criterion (AIC) (Burnham & Anderson, 2004).

Assessing Changes in Predictions Due to Detection Error

To assess the potential for detection error to bias probability of occupancy estimates at the site level (10 m x 10 m plot = site), a number of covariates were considered for the occupancy process. Occupancy covariates used included elevation, landscape curvature, distance to paved and unpaved roads and years since wildfire. These were determined *a priori* and derived from available data sources. Elevation was determined from a 10 m resolution Digital Elevation Model (DEM). Landscape curvature was calculated from the DEM to quantitatively describe the concavities and convexities at the site at a 250 m^2 and 640 m^2 scale. Distances to paved and unpaved roads were calculated at a 10 m resolution in ArcMap 9.3. Whether the site was burned or unburned and time since last fire were also used as covariates in the model. There was an 18 year fire chronosequence at the study area.

Covariates were incorporated into Models 1-5, and probability of occupancy maps generated. The covariates differed by species as different species showed different occupancy patterns. Generally, each species had two to four predictors associated with probability of occupancy. Maps were visually assessed to determine changes in predictions and covariate relationships.

Assessing Changes in Uncertainty between Models

Uncertainty associated with site level probability of occupancy for Models 3-5 with covariates for the occupancy process was assessed by calculating 95% confidence intervals for each 10 m x 10 m cell within the study area. Model 3 was used to compare with Models 4 and 5 as these data were most comparable because information from both visits were incorporated. Thus this tested for the effect of implicitly versus explicitly incorporating detection error into probability of occurrence models. Covariates were the same between the three models within a species, but differed between species. Confidence intervals were calculated using a non-parametric bootstrapping procedure. Predictions of mean probability of occurrence were made at each of 1000 bootstrap runs and empirical 95% confidence intervals were calculated. Predictions were restricted to the sampled sites (n = 1477) and the confidence interval widths across all sites were calculated. The distribution of confidence interval widths for Models 3-5 for the eight species sampled in 2011 were summarized using histograms.

To generate a visually rich example of how uncertainty and predictions compared between Models 3 and 4, predicted probability of occurrence and 95% confidence interval widths were calculated for the entire study area for two example species, *Sisymbrium altissimum* L. and *Descurainia sophia* (L.) Webb ex Prantl. These predictions were made to unsampled sites so the associated 95% confidence intervals are likely optimistically narrow (Efron, 1986), but they facilitated comparisons between models and allowed us to understand how uncertainty related to various covariate patterns within models and how predictions differed between an implicit versus explicit treatment of detection error.

Results

Detection rates, as estimated from the hierarchical occupancy/detection model, for the 2009 and 2010 data ranged from 0.24 to 0.94 across species (Table 1). The average detection rate for 11 species was 0.71. The data collected in 2011 had a mean estimated detection rate of 0.83 for eight species (Table 2). Detection rate increased as occupancy rate increased ($p = 0.01$ from linear regression model; Fig. 1) with no difference in detection between years ($p = 0.11$). Although this linear trend was evident, there was a large amount of variation in detection rates for species with low occupancy rates (< 0.15). *Kochia scoparia* (L.) Schrad., *Halogeton glomeratus* (M. Bieb.) C.A. Mey. (2009 and 2010) and *Lepidium perfoliatum* L. (2011), did not follow the expected trend and had high detection despite low occupancy as well as having low uncertainty associated with their detection estimates (CI widths of 0.05 and 0.09 respectively). *H. glomeratus* did demonstrate a year effect though, as in 2011 it had low detection (0.62) but similar occupancy to the previous years. Species such as *Lactuca serriola* L. and *Taraxacum officinale* F.H. Wigg. had relatively low detection rates (0.58 and 0.24) and low occupancy rates (0.12 and 0.05), following the more general trend in the dataset.

When species were ranked according to their mean occupancy rate for the five different model/dataset combinations, only one shift in rank occurred (Table 2). *Sisymbrium altissimum* and *D. sophia* changed rank occurrence when ranks from Model 1, Visit 1 data were compared to ranks from Models 2 (visit 2 data) and Models 3 and 4 (Combined). Species rank occurrence was the same for Models 2 - 4. Estimates of mean occupancy rate changed slightly depending on the dataset and modeling method (Table 2). For all eight species, estimates from Visit 1 data only were lower than the Combined data using logistic regression and the estimates from the hierarchical model. Occupancy rate estimates from Visit 2 compared to Visit 1 increased for three species, decreased for two species and remained static for three species. Estimates of mean occupancy rate remained the same for two species, increased by 0.01 for five species and increased by 0.02 for one species when the logistic with Combined data was compared to the hierarchical model (Table 2).

We evaluated six combinations of detection covariates to describe detection rates for eight species surveyed in 2011 (Table 3). Intercept only detection models described the detection rate best for four species: *Carduus nutans* L., *Bromus tectorum* L., *Alyssum desertorum* Stapf and *L. perfoliatum*. Models that included day since first sampling as a linear term were improvements over intercept only models for *S. altissimum* and *D.*

sophia. *Agropyron cristatum* (L.) Gaertn. detection rate was best described by a model with sample day as a linear term plus an observer effect. AIC indicated that a quadratic term for sample day fit best for *H. glomeratus*, however plotting the predicted values indicated a nonsensical relationship, possibly due to clustering in the sampling times.

For the three species where a sample day detection model was supported, detection rate increased from sample day 0 to sample day 25 by a magnitude of 0.10 to 0.12. Detection rate increased from 0.76 to 0.88 for *S. altissimum*, 0.80 to 0.90 for *A. cristatum and* 0.70 to 0.82 for *D. sophia* on sample days 0 to 25, respectively. The detection rate for *A. desertorum* was the only species that exhibited a negative trend in detection rate as a function of sample day, going from 0.96 to 0.93 on day 0 and 25, respectively, however the model with sample day was not supported (ΔAIC < 2) over an intercept only model (Table 3). Detection rate differed by 0.10 between the two observers for *A. cristatum*.

Empirical 95% bootstrap confidence intervals of the probability of occurrence were compared at the sampled sites for models 3 – 5 (Fig. 2). Four of the eight species had practically identical confidence intervals when Model 3 was compared to Models 4 and 5, but Model 3 had narrower confidence intervals for *C. nutans, H. glomeratus* and *D. sophia,* and *A. cristatum* had intermediate uncertainty (i.e. some estimates became more certain with the hierarchical models while some became more uncertain compared to the logistic regression model). No differences in the confidence interval widths were observed between Models 4 and 5 for five species, with the exceptions being *H. glomeratus* and *A. cristatum*, though the differences were slight. Model 5 did not converge for *B. tectorum* and the histograms for Models 4 and 5 completely overlapped for the other five species.

To determine how predictions mapped to the whole sample area would change if detection error was implicitly incorporated versus explicitly modeled, predictions from a logistic regression model from the maximum of the detection history was compared to predictions from the hierarchical occupancy detection model. The most illustrative predicted probability of occurrence maps are shown in Fig. 3, as well as their associated maps of uncertainty. For the three species where a detection rate model that included sample day was supported, there was no discernible difference between the predictions from Model 4 versus Model 5. When we compared the results of Models 3 and 4, *C. nutans, H. glomeratus* and *D. sophia* were the species that showed marked differences in distribution. This shift was also reflected in the difference in uncertainty in Fig. 2. Covariate relationships in the logistic regression on the maximum visits compared to the hierarchical model were very different, with an opposite relationship for distance to unpaved road. *A. cristatum* showed minor shifts between Model 3 and 4 with the same areas around the paved road highlighted as high probability of occupancy, but the effect of elevation became less apparent in the hierarchical model. The species that showed the greatest difference between Models 3 and 4 were low in occupancy rate and low in detection rate relative to other species. *L. perfoliatum* predictions and uncertainty were robust to modeling method but it had a high detection rate relative to *H. glomeratus* and *C. nutans*. *D. sophia* represented the worst case where predictions shifted between the two models even though occupation rate was fairly high. Detection rate was lower for *D. sophia* demonstrating the increase in prediction uncertainties as detection rate decreased.

D, sophia and *S.altissimum* probability of occupancy maps and associated 95% confidence interval width maps are shown in Fig. 3. *S. altissimum* demonstrates a case where predictions and uncertainty in predictions are robust to modeling method, while *D. sohpia* demonstrates the opposite, with marked changes in probability of occupancy patterns between model 3 and 4 as well as much greater uncertainty under Model 4 (Fig. 3).

Discussion

Detection error has the potential to inject bias into estimates of NIS occupancy but as previously stated, this problem has rarely been assessed for plant species. Plants arguably provide an easier taxon of organisms to model, as the sessile nature of plants allows assumptions of closure to be met within a growing season, making the survey task more straight forward (e.g. not inferring detection from animal sign or call). However, the potential for imperfect detection remains, which can bias occupancy estimates. Further, temporally non-constant detection can bias regression coefficients (Gu & Swihart, 2004) of habitat covariates, all of which could affect management decisions if not adequately dealt with.

Empirical Detection Rates

We showed that detection rates varied by species and ranged from very high (0.94) to relatively low (0.24) with an overall mean of 0.76. Detection was lowest for species rare on the landscape as one might expect, indicating the increasing uncertainty associated with predicting rare plant species occurrence. A general trend of increased detection rate as landscape frequency of a species increased, however detection rates were highly variable for species with estimated occupancies of less than 0.2. Qualitatively, species biological and ecological preferences had more bearing on detection at these low frequencies. For example, *K. scoparia* and *H. glomeratus* had high detection rates (> 0.87) while having relatively low occupancy rates (< 0.10) in 2009 and 2010 which was likely due to both species being restricted to certain ecologies (roadside for *K. scoparia* and clay-rich for *H. glomeratus*) and having contrasting phenologies to other plants in the community. Both increased species frequency and restricted ecologies favor a species' detection as the observer is either alert to its presence because of other presences or alert to the possibility of presence because of the expectation to find the species based on previous observations. Species more generalist in nature (no clear habitat preferences) and rarer on the landscape had lower rates of detection, with species like *L. serriola*, *T. officionale*, exemplifying this category of species.

Changes in Rank Abundance

One way some managers may prioritize species for management is targeting species of low abundance (i.e. mean occupancy rates) as their perceived likelihood of success in an eradication program is far higher. Thus, correct species rank occurrence is important. When we compared species rank occurrence only two of eight species changed rank when the results of the Visit 1 analysis were compared to the hierarchical model. This indicated that while there may be slight changes in the prioritization of two species, there would be little overall change in species to species prioritization with the multi-visit data. Further, rank species occurrence did not change between the second visit and the Combined data, thus, ensuring sampling in peak growing season would likely prevent the change in species rank occurrence that we observed.

Effect of Detection Error on Predictions/Uncertainty

Generally, if Visit 1 data were used to predict the probability of occurrence using a logistic regression model, the mean probability of occurrence was under predicted. For the three species with a seasonal effect, predictions from the second visit alone were greater than that of the first and using the maximum of the detection history in a logistic regression model always increased mean occurrence across all species. However, although mean occurrence differed between those three models, the pattern of occurrence was generally consistent (areas of high probability remained high while areas of low probability remained low). Thus, from a management perspective, areas that would be prioritized for population searches would be consistent, but the relative frequency managers would expect to encounter populations may differ.

When we explicitly incorporated detection error in a hierarchical framework and compared the predictions to those from the more implicit treatment of detection error in the logistic regression model that utilized the maximum of the detection history, we saw some consistencies and some inconsistencies. For *B. tectorum*, *A. desertorum, L. perfoliatum* and *S.altissimum,* there was effectively no difference between model predictions using the hierarchical model versus the logistic model with the maximum detection history (Fig. 3). Accordingly, the uncertainty in the estimates of probability of occupancy was nearly identical regardless of modeling method for these four species (Fig. 2). However, *H. glomeratus* and *C. nutans* demonstrated less consistency between models. Their occupancy patterns shifted dramatically between the logistic regression model and hierarchical model. These two species had detection rates of 0.62 and 0.75 respectively but very low occupancy rates (0.05 and 0.07 respectively). In contrast, *L. perfoliatum* had very low occupancy rate (0.04) but gave consistent predictions regardless of modeling method and had almost identical uncertainty in the prediction of the mean occupancy rate. The key difference for the latter species was high detectability. These results demonstrated the interplay between detection and occupancy rate in the certainty of model output. If a species was low frequency and only moderately detectable, uncertainty in the model predictions was high. However, if the species had similarly low frequency but the detection was high, predictions were robust to ignoring detection in the model and uncertainty was invariant to choice of modeling method.

Agropyron cristatum was an intermediate example of changes in predictions due to detection error. When detection error was accounted for hierarchically, the effects of elevation and landscape curvature became much less apparent and estimated occurrence rate became more uniform around the paved road. Although these changes were apparent, the dominant trend remained which was high occupancy around paved roads.

It should be noted that even if a species has a relatively high occurrence rate, detection error can play a large role in predicting its probability of occurrence in various habitats. For example, *D. sophia* had an estimated occupancy rate of 0.40 and a detection rate of 0.79. The relationship between one of the spatial covariates (distance to unpaved roads) reversed depending on the modeling method, with occupancy rate increasing as distance to unpaved road increased for the hierarchical model while the opposite was true in the logistic regression model. This dramatic shift was also reflected in the probability of occupancy uncertainty estimates with the confidence intervals being large in the hierarchical model while the logistic regression model that used the combined detection data showed much greater certainty. This shift demonstrated that without explicitly incorporating detection error, we may underestimate the uncertainty of our predictions in situations where detection rate is low regardless of overall occupancy rates.

Effect of Season/Observer on Predictions

A detection model that included sample day was supported for three of eight species considered, but when sample day (which acted as a surrogate for phonological changes within season) was included in the hierarchical detection/occupancy model, the differences in predicted probabilities of occurrence were negligible when compared to intercept only detection model. Further, when we investigated the uncertainty of the estimated probabilities of occurrence between a hierarchical model with just an intercept only term for detection versus one that incorporated sample day for these three species, the decreased uncertainty was negligible. Thus, our data suggest that as long as sampling is targeted within the peak growing season and systematic changes in detection during the period are low, predictions will be robust to changes in detectability up to ~10%. Further, this trend held for observer differences, as *A. cristatum* showed a difference in observer detection rates of 10% but no practical difference in predictions when an observer effect was incorporated into the hierarchical model.

Management Synthesis

As predictive models of species occupancy rates are increasingly used in land management planning, it is imperative that unbiased representations of species distributions are put forth to inform the process. Until now, the potential for imperfect detection to bias occupancy estimates for plant species had not been adequately evaluated with regard to predicting landscape scale species distributions. We found that the detection process is in fact imperfect and varied by species for the non-native plant species considered in this study. But, our data suggest that occupancy rates are on average high for plant species. Rank occurrence only changed once for two species when the early season data were modeled ignoring detection, indicating only a small risk of changing overall management priorities based on mean occurrence criteria.

Although mean estimates were shown to be negatively biased, we only found systematic changes in detection over the sampling period for three out of eight species and an observer effect for only one species. Even with these systematic differences in detection error for three species, regression coefficients were *not* sufficiently biased to alter the probability of occurrence maps generated from estimates of regression coefficients corrected and uncorrected for sample day and observer. Thus, as long as the systematic changes in detection are sufficiently small, predictions should be robust even if they are ignored. Our results also suggest that if species detection rate is greater than 0.85, model predictions and estimated uncertainties are robust to choice of modeling method (logistic regression assuming perfect detection versus hierarchical occupancy/detection model).

Thus, we suggest that targeting high detection rates, standardizing observers and ensuring sampling during the peak growing season will decrease uncertainty enough and expensive sampling schemes and complex modeling methods will not have to be employed. However, if the detection problem is large for a target species and/or a species is increasingly rare, the hierarchical framework and multi-visit sampling design can assist in correcting for detection error and more accurately estimate the uncertainty in mean occurrence rates. However, our experience suggests that it is likely that visiting more sites as opposed to sampling the same sites multiple times to determine detection error, may be more important for locating rare species and may further decrease the effects of detection error on estimating species-habitat relationships. As managers become more involved with gathering occupancy data, they may want to test their detection rates for different species at different times of the year to help inform their data collection, and ascribe levels of certainty to occupancy models and species distribution maps.

Acknowledgements

Funding for this project was provided by Gonzales-Stoller Surveillance, LLC under contract DE-NE0000300 with the U.S. Department of Energy. Thanks to Matt Lavin and Megan Higgs for discussions and ideas. Thanks to field data collection crew members: Alex Wing, Rose Marks, Elizabeth Usher, Kim Taylor, Tyler Schober, Kevin Shields and Dan Campbell.

Tables and Figures

Table 3.1. Occupancy and detection rates and associated 95% confidence intervals as estimated from intercept only hierarchical models for the combined 2009 and 2010 repeat transect data. A total of 15 - 200 plot transects were included in this evaluation.

	Occupancy			Detection		
Species	Mean	2.5%	97.5%	Mean	2.5%	97.5%
Carduus nutans	0.02	0.01	0.03	0.43	0.26	0.60
Kochia scoparia	0.05	0.04	0.05	0.87	0.83	0.92
Taraxacum officinale	0.05	0.01	0.10	0.24	0.04	0.44
Salsola kali	0.08	0.06	0.09	0.52	0.44	0.60
Halogeton glomeratus	0.10	0.09	0.11	0.88	0.85	0.91
Lactuca serriola	0.12	0.11	0.14	0.58	0.52	0.64
Agropyron cristatum	0.18	0.17	0.20	0.74	0.70	0.77
Descurainia sophia	0.28	0.27	0.30	0.79	0.77	0.82
Tragopogon dubius	0.29	0.27	0.31	0.77	0.74	0.80
Sisymbrium altissimum	0.30	0.29	0.32	0.78	0.75	0.80
Alyssum desertorum	0.70	0.68	0.72	0.94	0.93	0.95
Bromus tectorum	0.88	0.86	0.89	0.94	0.93	0.94

Table 3.2. Occupancy (Occ.) and detection rate estimates (Est.) and associated 95% confidence interval limits for eight species sampled in 2011. The first three full columns are estimates from an intercept only logistic regression model which implicitly assumes perfect detection from Visit 1, Visit 2 and Combined maximum of Visit 1 and Visit 2 data (Max Visits). The last two full columns are estimates off occupancy and detection rates from a hierarchical occupancy model.

	Occ.; Visit 1 Data			Occ.; Visit 2 Data			Occ; Max Visits			Occ. MacKenzie			Detection Rate		
Species	Mean	2.5%	97.5%	Mean	2.5%	97.5%	Mean	2.5%	97.5%	Mean	2.5%	97.5%	Mean	2.5%	97.5%
Lepidium perfoliatum	0.03	0.02	0.04	0.03	0.02	0.04	0.03	0.02	0.04	0.03	0.02	0.04	0.87	0.80	0.94
Carduus nutans	0.04	0.03	0.05	0.04	0.03	0.04	0.04	0.03	0.05	0.05	0.04	0.06	0.75	0.66	0.84
Halogeton glomeratus	0.04	0.03	0.05	0.04	0.03	0.06	0.06	0.05	0.07	0.07	0.05	0.08	0.62	0.53	0.72
Agropyron cristatum	0.18	0.16	0.20	0.19	0.17	0.21	0.21	0.19	0.23	0.22	0.19	0.24	0.84	0.81	0.88
Sisymbrium altissimum	0.32	0.29	0.34	0.32	0.30	0.35	0.36	0.34	0.38	0.37	0.34	0.39	0.87	0.85	0.89
Descurainia sophia	0.30	0.28	0.33	0.33	0.31	0.35	0.38	0.36	0.41	0.40	0.37	0.43	0.79	0.76	0.82
Alyssum desertorum	0.73	0.70	0.75	0.72	0.70	0.74	0.76	0.74	0.78	0.77	0.74	0.79	0.94	0.93	0.95
Bromus tectorum	0.88	0.86	0.90	0.87	0.85	0.89	0.92	0.91	0.94	0.92	0.91	0.94	0.94	0.94	0.95

Table 3.3. AIC evaluation of six different detection models for eight non-native plant species sampled in 2011. Occupancy was modeled as a constant for all models. Detection rate was modeled as intercept only, observer (obs) only, sample day (sd) as a linear term, sample day as a linear term and observer, sample day as quadratic (^2) with linear term, sample day as linear and quadratic and observer. Best models are bolded (ΔAIC < 2). The more parsimonious model was chosen of the bolded models.

	Lepidium perfoliatum		*Carduus nutans*		*Halogeton glomeratus*		*Agropyron cristatum*		*Sisymbrium altissimum*		*Descurainia sophia*		*Alyssum desertorum*		*Bromus tectorum*	
Detection Model	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC
intercept only	490.4	**0.0**	651.3	**0.0**	850.7	22.1	2000.3	11.5	2672.7	3.0	2967.8	4.7	2555.7	**0.7**	1909.1	**0.0**
observer (obs)	492.3	2.0	652.8	**1.6**	844.4	15.8	1991.9	3.0	2673.6	3.9	2969.4	6.3	2556.3	**1.3**	1911.0	**1.8**
sample day (sd)	491.4	**1.0**	653.1	**1.8**	851.2	22.5	1997.7	8.8	2669.8	**0.1**	2963.1	**0.0**	2554.9	**0.0**	2226.5	317.4
obs + sd	492.5	2.2	653.3	2.1	838.6	9.9	1988.8	**0.0**	2669.8	**0.0**	2964.3	**1.2**	2555.8	0.8	2223.9	314.8
sd + sd^2	493.8	3.4	654.6	3.3	847.8	19.2	2002.6	13.8	2674.1	4.3	2965.0	**1.9**	2563.9	9.0	1933.5	24.4
obs + sd + sd^2	495.2	4.9	654.8	3.6	828.6	**0.0***	1993.3	4.4	2675.6	5.8	2966.9	3.8	2561.7	6.7	1928.3	19.1

*investigation of model fitted values indicated nonsensical relationship

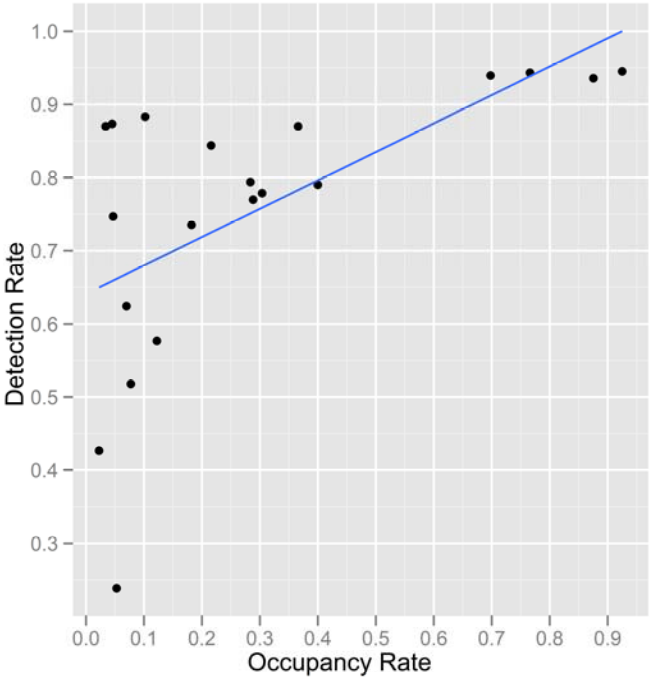


Figure 3.1. Estimated detection rate of the non-native species sampled in 2009, 2010 and 2011 as a function of occupancy rate. Detection rate increased as occupancy rate increased with no evident effect of year (p = 0.11). The fitted line is from a simple linear regression of detection rate as a function of occupancy rate.

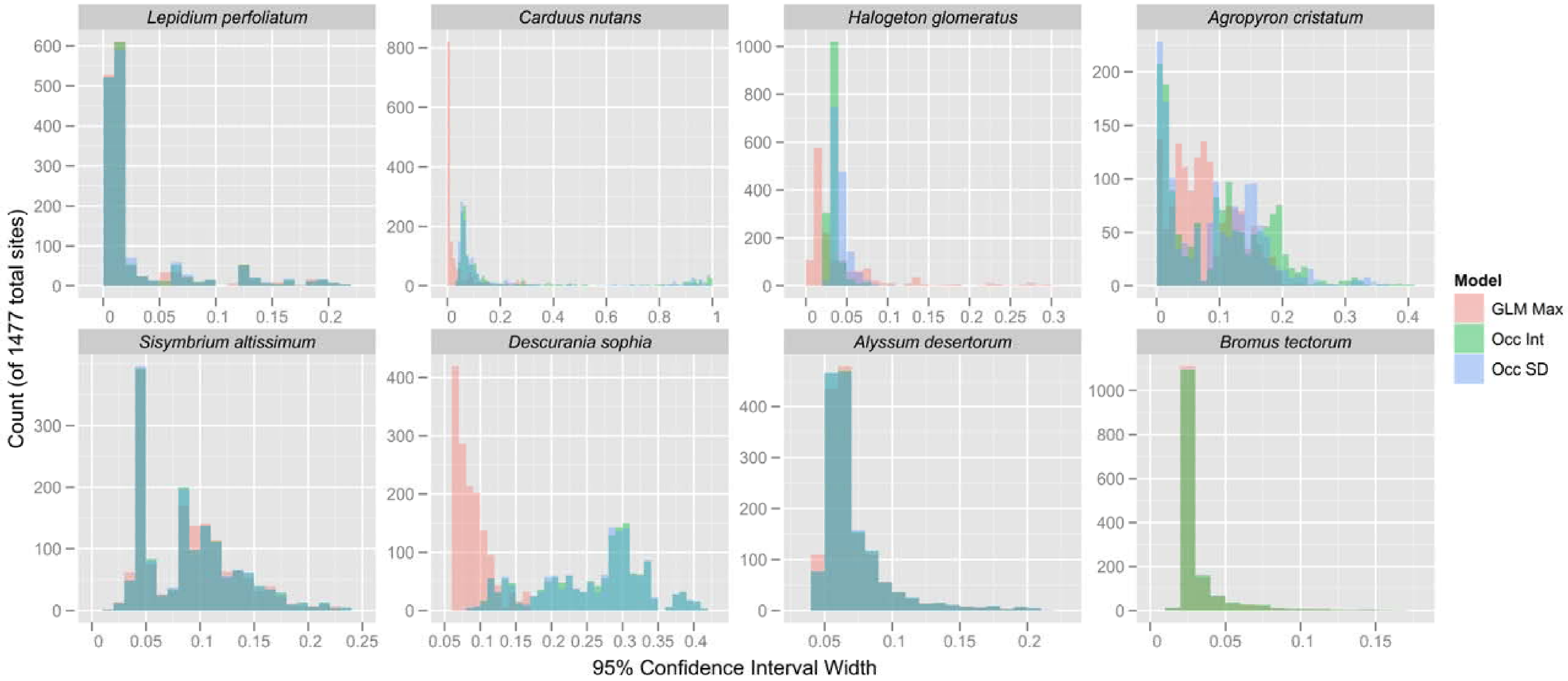


Figure 3.2. Histograms of the 95% bootstrap confidence interval widths of the probability of occurrence at the sampled locations from three different models for eight species. Red bars indicate a logistic regression model using the maximum of the detection history (GLM Max), green bars indicate a hierarchical occupancy model with an intercept only term for detection (Occ Int) and blue bars are a hierarchical occupancy model with sample day in the detection component (Occ SD). Occupancy was modeled as a function of covariates and consequently differed by species. Species are ordered by occupancy rate (low to high). The sample day model did not converge for *Bromus tectorum*. The bars are semi-transparent to facilitate comparison between histograms within a panel.

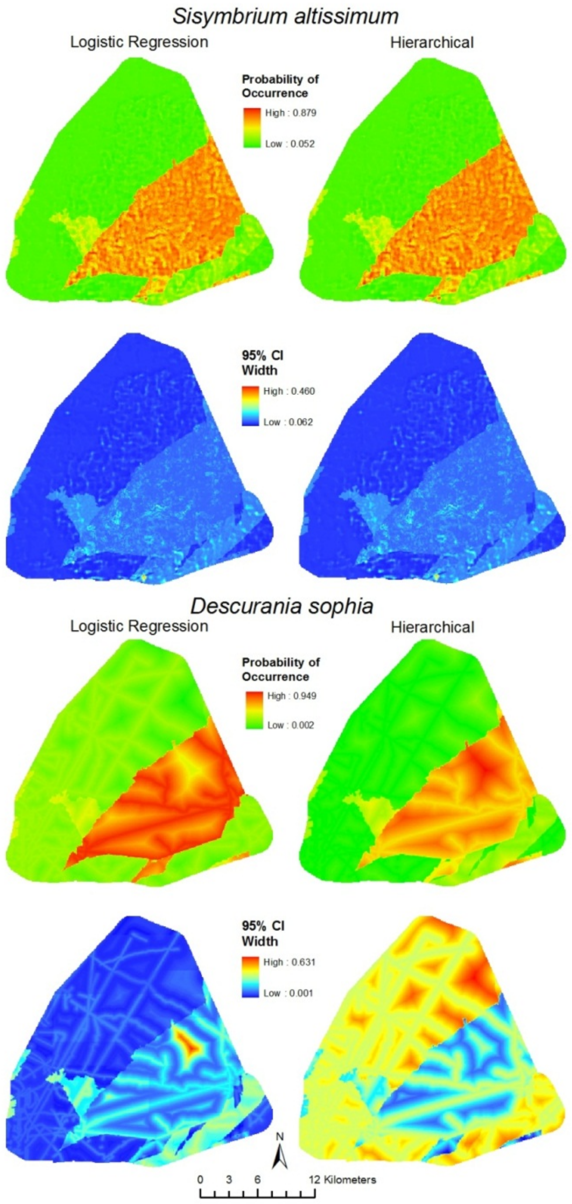


Figure 3.3. Probability of occupancy maps and associated 95% confidence interval (CI) widths for two species using two modeling methods. Maps in the left column are results from a logistic regression model while maps on the right are from a hierarchical occupancy model. Maps with the green to red gradient are probability of occupancy while the maps with the blue to red gradient are the width of the 95% confidence intervals about the mean probability of occupancy. The top four maps are results from *Sisymbrium altissimum* while the bottom four maps are results from *Descurainia sophia*. Probability of occupancy ranges from low (green), medium (yellow) and high (red) while confidence interval width goes from low (blue), medium (yellow), high (red). The range of values for each set of maps is displayed in the adjacent legend.

References

Alexander, H.M., Slade, N.A., & Kettle, W.D. (1997) Application of mark-recapture models to estimation of the population size of plants. Ecology, 78, 1230-1237.

Alexander, H.M., Slade, N.A., Kettle, W.D., Pittman, G.L., & Reed, A.W. (2009) Detection, survival rates and dynamics of a cryptic plant, *Asclepias meadii*: applications of mark-recapture models to long-term monitoring studies. Journal of Ecology, 97, 267-276.

Burnham, K.P. & Anderson, D.R. (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods Research, 33, 261-304.

Chen, G., Kéry, M., Zhang, J., & Ma, K. (2009) Factors affecting detection probability in plant distribution studies. Journal of Ecology, 97, 1383-1389.

Efron, B. (1986) How biased is the apparent error rate of a prediction rule? Journal of the American Statistical Association, 81, 461-470.

Fiske, I. & Chandler, R. (2011) unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43, 1-23.

Franklin, J., Wejnert, K.E., Hathaway, S.A., Rochester, C.J., & Fisher, R.N. (2009) Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. Diversity and Distributions, 15, 167-177.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A., & Wintle, B.A. (2008) When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. Austral Ecology, 33, 986-998.

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

Kéry, M. (2004) Extinction rate estimates for plant populations in revisitation studies: importance of detectability. Conservation Biology, 18, 570-574.

Kéry, M. & Gregg, K.B. (2003) Effects of life-state on detectability in a demographic study of the terrestrial orchid *Cleistes bifaria*. Journal of Ecology, 91, 265-273.

Kroll, A.J., Risenhoover, K., McBride, T., Beach, E., Kernohan, B.J., Light, J., & Bach, J. (2008) Factors influencing stream occupancy and detection probability parameters of stream-associated amphibians in commercial forests of Oregon and Washington, USA. Forest Ecology and Management, 255, 3726-3735.

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

Mackenzie, D.I. & Royle, J.A. (2005) Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology, 42, 1105-1114.

Miller, T.K., Allen, C.R., Landis, W.G., & Merchant, J.W. (2010) Risk assessment: Simultaneously prioritizing the control of invasive plant species and the conservation of rare plant species. Biological Conservation, 143, 2070-2079.

Moore, J.L., Hauser, C.E., Bear, J.L., Williams, N.S.G., & McCarthy, M.A. (2011) Estimating detection-effort curves for plants using search experiments. Ecological Applications, 21, 601-607.

Parendes, L.A. & Jones, J.A. (2000) Role of light availability and dispersal in exotic plant invasion along roads and streams in the H. J. Andrews Experimental Forest, Oregon. Conservation Biology, 14, 64-75.

R Development Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Regan, T.J., Chadès, I., & Possingham, H.P. (2011) Optimally managing under imperfect detection: a method for plant invasions. Journal of Applied Ecology, 48, 76-85.

Regan, T.J., McCarthy, M.A., Baxter, P.W.J., Dane Panetta, F., & Possingham, H.P. (2006) Optimal eradication: when to stop looking for an invasive plant. Ecology Letters, 9, 759-766.

Rew, L.J., Lehnhoff, E.A., & Maxwell, B.D. (2007) Non-indigenous species management using a population prioritization framework. Canadian Journal of Plant Science, 87, 1029-1036.

Rew, L.J., Maxwell, B.D., & Aspinall, R. (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Science, 53, 236-241.

Rew, L.J., Maxwell, B.D., Dougher, F.L., & Aspinall, R. (2006) Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species. Biological Invasions, 8, 523-539.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. Progress in Physical Geography, 35, 211-226.

Rota, C.T., Fletcher, R.J., Evans, J.M., & Hutto, R.L. (2011) Does accounting for imperfect detection improve species distribution models? Ecography, 34, 659-670.

Seipel, T., Kueffer, C., Rew, L.J., Daehler, C.C., Pauchard, A., Naylor, B.J., Alexander, J.M., Edwards, P.J., Parks, C.G., Arevalo, J.R., Cavieres, L.A., Dietz, H., Jakobs, G., McDougall, K., Otto, R., & Walsh, N. (2012) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. Global Ecology and Biogeography, 5, 236-246.

Shefferson, R.P., Sandercock, B.K., Proper, J., & Beissinger, S.R. (2001) Estimating dormancy and survival of a rare herbaceous perennial using mark-recapture models. Ecology, 82, 145-156.

Smith, J.B., Jenks, J.A., & Klaver, R.W. (2007) Evaluating detection probabilities for American Marten in the Black Hills, South Dakota. Journal of Wildlife Management, 71, 2412-2416.

Syphard, A.D. & Franklin, J. (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. Ecography, 32, 907-918.

CHAPTER FOUR

REGIONAL POPULATION DYNAMICS OF NON-NATIVE PLANT SPECIES

Abstract

Applying metapopulation dynamics theory to non-native plant species may improve the understanding of the relative importance of different processes that drive extinction, persistence or invasion. The relative importance of regional versus local processes on the regional dynamics of plant species remains unresolved and has yet to be applied to understanding species invasions. Species occurrence data for eight non-native plant species in a sagebrush steppe ecosystem were sampled in fixed plots over three seasons from 2009 to 2011 using a targeted transect approach. Habitat suitability with respect to environmental gradients was estimated using a logistic regression model. Disturbance variables such as paved and unpaved roads, and wildfire were collected from the local environmental surveillance contractor. The density of occupied neighboring cells was used as a covariate in colonization and extinction models for the eight species. Models for colonization and extinction ranged from simple main effects of a few predictors, to models requiring complex interactions between the habitat, disturbance, and neighborhood interpreted with reference to the temporal climatic dynamics to explain the observed colonization and extinction dynamics. Discussed in terms of more specific typologies of regional dynamics, our results suggest examples of classic metapopulations, regional ensembles and spatially extended populations, with most of the species falling within the regional ensemble or spatially extended populations categories.

Introduction

Regional scale dynamics of plant species are a relatively understudied topic in ecology with little consensus in the literature as to the dominant processes that influence population (patch) persistence or extinction. Classic metapopulation theory (Hanski, 1999) has been proposed to explain regional scale patterns of colonization and extinction in animals with a metapopulation being generally defined as a collection of local populations that have a non-zero chance of going extinct while new colonies can emerge in other suitable areas within breeding or migration distance (Levins, 1969). Although metapopulation theory has been applied to plant regional dynamics (Erikkson, 1996; Bullock et al., 2002; Ehrlen & Eriksson, 2003), it has been suggested that plants, although often observed to be patchy, do not strictly adhere to the traditional metapopulation concept (Freckleton & Watkinson, 2002; Ehrlen & Eriksson, 2003; Freckleton & Watkinson, 2003). Plant metapopulation dynamics have been primarily studied with native plant species (Eriksson, 1996).

Within the strict metapopulation framework, non-native invasive species may logically fall within the category of non-equilibrium metapopulations, where colonizations are greater than extinctions, allowing for metapopulation growth. However, there is little quantitative evidence to suggest whether regional or local processes are more important to colonization and extinction patterns for these species (Freckleton & Watkinson, 2002). Understanding the relative importance of regional versus local processes to colonization and extinction dynamics of non-native plant species is vital to predicting future potential occupancy patterns of non-native plants and also targeting management (Rew *et al.*, 2007; Maxwell *et al.*, 2009).

Freckleton and Watkinson (2002) reviewed the literature on plant regional and local dynamics and concluded that there were three primary types of regional dynamics that could arise based on the relative importance of regional versus local processes and the relative amount of suitable to unsuitable habitat. The primary regional processes that link discrete habitat patches were considered interpatch immigration and emigration while local processes were local dispersal, recruitment and mortality. The first classification they suggested was classic metapopulation dynamics (Erikkson, 1996; Hanski ,1998), and the special case of source-sink dynamics (Pulliam, 1988). These patterns were highly dependent on interpatch immigration and emigration to explain regional population extinction, persistence, or expansion. In addition to classic metapopulations, Freckleton and Watkinson (2002) added “regional ensembles” and “spatial extended populations” to fill out a continuum, where local processes become progressively more important relative to regional processes. They suggested regional ensembles and classic metapopulations should occur when unsuitable habitat is much greater than suitable habitat, but that interpatch migration is much less important in regional ensembles, with rare dispersal events followed by local dispersal, recruitment and mortality dominating regional dynamics. The third classification they suggested was spatially extended populations. These occur when suitable habitat is much greater than unsuitable habitat and regional population dynamics can be explained by the sum of local dispersal and recruitment-mortality dynamics. The authors also listed three types of local dynamics; “spatially structured” where a species occur in discrete but transient patches that are mainly limited by resources (e.g. gaps in a forest floor), “spatially unstructured” where the population is fairly continuous but local processes create some patchiness (e.g. management of weeds within agricultural fields), and finally “local populations” where the scale of the data collected was essentially local and so it is impossible to determine regional processes.

While these classifications seem logical, little quantitative evidence exists to suggest which of these classifications of plant regional dynamics exist. In the context of non-native plant species, predicting which type of spatial dynamics a species will ultimately exhibit in a system could both improve our understanding of regional scale invasion, but also aide in between species or population prioritization and suggest different management strategies depending on the type of dynamics predicted.

Studies of plant regional dynamics have primarily focused on native plant species (Eriksson, 1996) but studies of regional dynamics have great potential to provide insight into processes driving landscape scale invasions and predict future distribution patterns of species of concern. Non-native plant species are a subdivision of the plant species pool that reside somewhere along the non-equilibrium/equilibrium continuum depending on how long they have been in residence within a particular region. Their colonization and extinction rates also respond to regional disturbance and climate factors as native species do, causing variation in metapopulation trajectories through time. Where a species fits along this equilibrium/non-equilibrium gradient and how temporal climatic and disturbance drivers interact with suitable habitat, have large implications for our understanding of invasive species spread, metapopulation persistence or metapopulation extinction at landscape scales. Identification of the most important drivers of regional patterns of population dynamics has substantial implications for the active management of invasive species which generally occurs at landscape (regional) scales. Furthermore, management has traditionally focused on treating local populations, with the goal of extirpating the entire region's metapopulation, without much consideration to whether the management is actually affecting local or regional processes. Thus, an understanding of non-native species from a metapopulation perspective is vital for more effective and efficient management of undesirable species.

We conducted a multi-non-indigenous plant species metapopulation dynamics study over three years in a sagebrush steppe ecosystem. Our goal was to quantify the colonization and extinction dynamics of eight target species over two time transitions with respect to habitat, isolation (dispersal), disturbance, and annual climatic variation. Interspecific variability in the spatial and temporal drivers associated with colonization and extinction were used to predict the landscape level trajectories of each species. We classified our colonization and extinction results according to Freckleton and Watkinson's (2002) regional and local dynamics categories and determined the role of habitat configuration, and regional versus local processes.

Methods

Study Site

Data were collected within a 20 km east-west and 30 km north-south area at the Idaho National Laboratories (INL) in south-central Idaho, USA. The site is classified as a sagebrush steppe ecosystem, dominated by *Artemisia tridentata* Nutt. ssp. *wyomingensis* Beetle & Young, and various *Chrysothamnus* species. Annual precipitation was 22 cm with a mean annual temperature of approximately 5.5°C and the primary growing season was between late May and early July. Mean combined May and June precipitation (55 year average) was 61 mm. Elevation ranged between 1461 m and 1638 m.

This study was multi-year with stark differences between years. In 2009, the site experienced very high growing season precipitation (84 mm May and June combined) as compared to average, while 2010 was nearer average (57 mm May and June combined). In 2011, the site received only 44 mm of precipitation during May and June. Also, at the end of the sampling period in 2010, a 44,000 ha fire burned roughly one-third of the study area over the course of 24 hours. Coupled with the fire in 2010, there was an 18 year (as of 2011) fire chronosequence record at the site.

Sampling Scheme

Sampling occurred in 2009 through 2011. Targeted belt transects that were 2 km long by 10 m wide, started on random points along roads and traveled perpendicular to roads (see Rew *et al.*, 2006; Brummer *et al.* in review). Species presence and absence

(P/A) was recorded at every 10 m interval along each transect resulting in 200 - 10 m by 10 m plots along each transect where the presence or absence of target species was recorded. The clustered sampling scheme (200 contiguous plots) facilitated estimating the effects of neighborhood covariates, while simultaneously sampling the underlying environmental and disturbance covariate patterns. Sample size and species varied by season due to changing goals and varying logistical constraints on sampling. In 2009, the initial survey sampled 37 transects (7,358 - 10 m x 10 m plots) for all non-native plant species. In 2010, 14 (13 for *A. cristatum*) (2,611 plots) of these 37 transects were repeated and P/A for three target species, *Agropyron cristatum*, *Sisymbrium altissimum*, and *Carduus nutans,* was recorded.

In 2010, 75 new transects (14,999 plots) were sampled for all non-native plant occurrence. In 2011, 33 transects (6,550 plots) initiated in 2010 were resampled for eight target species occurrence. The 14 transects that had been sampled in 2009 and 2010 were also resampled for the occurrence of the eight target species (which included the three species that were resampled in 2010 on the 14 transects). This resulted in information on transitions for three species from 2009 to 2010 and 2010 to 2011, and transitions for eight species from 2010 to 2011.

Metapopulation Modeling

Metapopulation dynamics models were fit (sensu MacKenzie, 2006) using the R statistical Package (R Development Core Team, 2012) as facilitated by the contributed package ‘unmarked’ (Fiske & Chandler, 2011). This metapopulation model explicitly incorporated detection error assuming multiple site visits each year. Although it would have been ideal to revisit all the sites within season, this was logistically and monetarily infeasible as part of this study. To allow the colonization, extinction, and initial distribution components of the model to be estimated, the data were duplicated within each season to allow the model to estimate near 100% detection and then estimate the parameters of interest. While we recognize the shortcomings of this design and modeling method, we had to sacrifice visiting sites multiple times to get adequate coverage over the range of our predictor variables throughout the site.

Covariates for colonization, extinction and the initial occupancy pattern were carefully constructed based on hypothesized predictors that were determined *a priori*. Habitat suitability (HS) was one of the primary predictors and was calculated from the sum total of all data collected in 2009 and 2010 for all transects (113 transects or 22,357 - 10 m x 10 m plots). Habitat suitability was determined via logistic regression on the species P/A data using only environmental predictors. These were elevation, slope, landscape curvature (at 250 m^2 and 640 m^2 scales), sine of aspect, and cosine of aspect. Sample year was included to control for seasonal climate effects and the mean habitat suitability over both years was then predicted for each site sampled.

The effect of neighboring occupied cells on colonization and extinction rates was determined using a predictor called occupied neighborhood density (OND). This was calculated as the number of neighboring occupied cells divided by the total number of neighboring cells within a given radius. To determine the best neighborhood size, ONDs with neighborhood sizes from 10 m to 300 m in increments of 10 m were tested. Akaike's Information Criterion (AIC) was recorded from a model fit with each of the different neighborhood density sizes and extinction or colonization was tested independently for each neighborhood size. The neighborhood size was chosen that minimized the AIC for that species and process (colonization or extinction). The model where OND of differing sizes was varied contained other predictors thought to be important to the colonization and extinction processes of each species. This was to avoid picking up the entire spatial trend in this one variable as the other covariates were also spatial (e.g. habitat suitability). The optimal neighbor sizes chosen are reported in Table 4.1.

Finally, disturbance variables were considered. Low use, two-track roads existed on the site but their effect from a disturbance perspective was hypothesized to be locally concentrated immediately adjacent to the road. Thus, a binary variable was created that took the value of 1 if a cell was < 20 m from an unpaved road and 0 if not. There were two paved roads through the site, one US Highway and one public access-restricted paved road. Distance to paved road was considered as a predictor and calculated at a 10 m resolution. Wildfire was treated as a binary variable (1 if burned 0 if not) and the effect of time since fire was estimated for areas that had burned as there was an 18 year fire record. Finally, *Agropyron cristatum* had a variable unique to it. *A. cristatum* was planted along certain sections of paved road in “Green Strips” to act as a fire break. The green strip was treated as a binary variable (1 if green strip, 0 if not) and was derived from a vegetation layer created with site specific data (Shive *et al.*, 2011).

Inferential models were arrived at using manual model selection. An *a priori* model was determined for each species based on field observations of the potential drivers. Single variable removal and addition of variables stated above was conducted, to develop a suite of potential models. The model that minimized AIC was considered as the best inferential model. If the ΔAIC between a set of models for a species was less than two, generally the most parsimonious of the models was used for inference. However, the coefficient estimates for a model that had the absolute minimum AIC were evaluated to determine if the point estimates were biologically significant. Biologically sensible interactions were tested for certain variables, specifically, the interaction between OND and habitat suitability, and the effects of habitat suitability and OND in and out of burned areas.

Visualization

Colonization and extinction probabilities were predicted across a 2-dimensional surface over the range of habitat suitability and occupied neighborhood densities observed at the site for all but two species. Distance to paved road replaced habitat suitability for *A. cristatum* as habitat suitability was not retained after model selection. Occupied neighborhood density was the only variable retained for *C. nutans*, thus response surfaces were not required. The best models as determined by AIC model selection were used to predict the mean colonization and extinction probabilities and multiple panels were used to illustrate any effects of categorical or tertiary continuous covariates. Coupled with each one of these plots, surfaces of the prediction interval widths were created to quantitatively illustrate the uncertainty in the predicted values. These surface plots acted as a good proxy for where data actually existed to estimate the response surfaces, as not all combinations of the covariates considered actually existed in physical or biological space.

Results

The individual drivers of invasion and interactions between spatial and temporal processes were highly species specific. Thus, results were presented on a species basis. The three species with three years of data are presented first followed by the remaining five species.

Carduus nutans

C. nutans was rare on the landscape and the overall metapopulation size was predicted to be static between 2009 and 2010 (Fig 4.1) but decreasing from 2010 to 2011 (Fig 4.2). The best model for the 2009 to 2010 transition retained habitat suitability, occupied neighborhood density (OND), and the interaction between those variables for colonization and OND for extinction (Table 4.2). The best model for the 2010 to 2011 transition retained OND for both colonization and extinction and the effect of the Jefferson Burn for extinction (Table 4.3).

From 2009 to 2010, colonization was higher at low habitat suitability and high OND (Fig D.2). The habitat effect was highly uncertain as there were very few cases where OND and habitat suitability (HS) were both high (Fig C.1). Extinction decreased as occupied neighborhood density increased (Fig D.1). During the 2010 to 2011 transition, colonization probability increased as OND increased, and so did the uncertainty. There were very few colonization events during this transition. Extinctions decreased as OND increased (Fig C.3 and C.4), and areas that were burned by the Jefferson Fire had greater extinction probability (Fig D.3).

Agropyron cristatum

A. cristatum was predicted to have increased in both transitions despite progressively decreasing growing season precipitation (Fig 4.2; Fig 4.3). The best model for colonization from 2009 to 2010 retained just OND while the extinction model was a simple intercept only model (Table 4.2). The simplicity of the 2009 to 2010 transition was likely due to inadequate sampling as the sample size was substantially lower and the transects by chance did not cover drivers well such as paved roads and the green strip areas.

The best model for colonization and extinction between 2010 and 2011 retained the same variables of OND, “green strip” and distance to paved road. Habitat suitability was not retained as a predictor of colonization or extinction for *A. cristatum* (Table 4.3). Colonization probability decreased as distance to paved road increased, was higher in areas of greater occupied neighborhood density, and was slightly higher in green strip (Fig D.13). Extinction was influenced by the same variables but the effects were opposite. Extinction was greater as distance to paved roads increased and occupied neighborhood density decreased. Extinction was also slightly lower inside the green strip (Fig D.14).

Sisymbrium altissimum

S. altissimum colonization and extinction models for the 2010 to 2011 transition were the most complex of all the species studied. Habitat suitability, OND, and their interactions were retained in the best model and estimated to be different inside and outside the fire for both colonization and extinction. Further, time since last fire was important for colonization and extinction rates and extinction rates differed adjacent to the unpaved road (Table 4.3).

In unburned areas, *S. altissimum* colonization probability was low regardless of habitat suitability at low OND. At low to moderate habitat suitability (0 – 0.3) colonization probability increased as OND increased, but colonization was estimated to decrease as habitat suitability increased at high OND. However, when compared to the uncertainty and scatters of the data, these cases did not actually exist and thus this relationship was ignored. Colonization probability was greater in the burned areas and decreased as time since fire increased. Areas with high OND had the highest colonization probability while there was low colonization probability at low OND and high habitat suitability (Fig 4.5). Although it seemed counter intuitive to have higher colonization probability at lower habitat quality, this was likely due to the intense disturbance response that *S. altissimum* had to fire (mean occupation increased from 0.20 to 0.63 after the 2010 Jefferson Fire). Proportionally, more areas that would not have experienced colonization as predicted by habitat suitability experienced colonization due to the fire. Thus, low habitat suitability areas had high colonization probability relative to high habitat suitability areas within the Jefferson Fire.

Extinction probabilities for *S. altissimum* in unburned areas were greater at higher habitat suitability, and were lowest at the highest occupied neighborhood densities. Areas with lower OND and higher habitat suitability were estimated to have high extinction. Interactions also indicated that the highest extinction was estimated at low habitat suitability and low OND. In the fire, extinction probability was generally low, with the highest extinction rates being estimated at higher habitat suitability (Fig 4.6). Although this trend was evident, uncertainty also increased in these areas as there were fewer cases to estimate the relationship. As time since fire increased, extinction rate increased (Fig 4.6).

Contextualized in terms of total metapopulation size, *S. altissimum* experienced an extinction event from 2009 to 2010. 2009 was an extremely wet growing season, while 2010 was near normal and 2011 was below normal. Although, 2011 was below normal, an increase in the metapopulation was observed (mean colonization event). But, when parsed by burned and unburned, there was a net extinction event outside the fires from 2010 to 2011 (mean occupation of 0.129 to 0.101), and net extinction in burned areas excluding the Jefferson Fire, during the same transition (mean occupation of 0.319 to 0.301). Thus the colonization event present from 2010-2011 could be primarily attributed to the 2010 fire that occurred after sampling.

Descurainia sophia

The metapopulation size of *D. sophia* increased from 2010 to 2011 (Fig 4.2). The variables retained after model selection were the same for colonization and extinction and were habitat suitability, OND, fire, and time since fire (Table 4.3). Colonization probability was predicted to be higher at higher occupied neighborhood densities, and be lower at high habitat suitability, although habitat only had a weak effect (relative to OND) and there was a fair bit of uncertainty at high habitat suitability. Colonization probability was predicted to increase one year after fire and then decrease as time since fire increased (Fig D.5). *D. sophia* extinction was predicted to be highest at high habitat suitability and lowest at high occupied neighborhood densities (Fig D.6). Although the association with HS was predicted, the relationship was highly variable as evidenced by the data scatters and uncertainty plots (Fig D.6). Extinction was also predicted to decrease one year after fire and increase as time since fire increased (Fig C.6). *D. sophia* was predicted to be increasing in metapopulation size both inside and outside the fire, but the increase was of a greater magnitude in the Jefferson Burn.

Lepidium perfoliatum

L. perfoliatum was rare on the landscape but the metapopulation size was predicted to increase from 2010 to 2011 (Figure 4.2). The variables in the colonization model retained after model selection were habitat suitability, OND, and the effect of just the 16 year old fire (Table 4.3). The extinction model was the same with the addition of distance to paved road (Table 4.3). Colonization probability was predicted to be greater at higher occupied neighborhood density but lower at higher habitat suitability. Colonization was much greater inside the 16 year old fire, but no other fire response was retained (Fig D.7). *L. perfoliatum* extinction probability was higher at low OND and high habitat suitability, and increased as time since fire increased. Although it was estimated that extinctions were greater at high habitat suitability, there was a lot of variation in the data and the uncertainty was large. The trend of decreasing extinction probability with increasing OND had low uncertainty and was evident in the data scatter (Fig C.4; Fig D.8).

Halogeton glomeratus

H. glomeratus was rare on the landscape and the metapopulation size was estimated as decreasing between 2010 and 2011 (Fig 4.2). The colonization model retained habitat suitability, OND, and the interaction between those variables while the extinction model retained OND, the interaction between habitat suitability and OND and fire after model selection (Table 4.3). From the little colonization that occurred, it was estimated to be higher at high occupied neighborhood densities. Although it was predicted that colonization decreased at high habitat suitability the data scatter and uncertainty demonstrate that these cases did not exist and colonization was low overall (Fig C.5; Fig D.9). Extinction was high at high habitat suitability regardless of OND, and lowest at high OND. Although the model predicted high extinction at high OND and HS, only few cases of extinction were actually observed at high OND thus, most of the extinction dynamic was a function of OND (Fig C.5; Fig D.10).

Bromus tectorum

B. tectorum was abundant on the landscape and the metapopulation size was estimated as increasing between 2010 and 2011 (Fig 4.3). The dynamics of *B. tectorum* were very straight forward. The best model for colonization retained habitat suitability, OND, fire and time since fire while the extinction model retained just habitat suitability and OND (Table 4.3). Colonization was higher at higher habitat suitability and occupied neighborhood density, and low at low OND and habitat suitability regardless of fire. Colonization was higher one year after fire, and then decreased as time since fire increased (Fig 4.3). Generally there were few cases of high habitat suitability, low OND and low habitat suitability and high OND as evidence by the uncertainty plots and scatter but overall uncertainty in the predictions was low relative to other species (Fig 4.3; Fig C.6). Extinction probability was higher at lower OND and habitat suitability and effects of fire were not retained in the model, although overall extinction probability was low (Fig 4.4). These patterns in colonization and extinction were reflected in the overall metapopulation growth as *B. tectorum* experienced a large colonization event from 2010 to 2011, mostly due to the 2010 fire.

Alyssum desertorum

Mean population trends for the 2010 to 2011 transition suggested a slight decrease in mean metapopulation size but the variation was such that a difference was not statistically detectable. Colonization seemed to be higher in the fire and lower outside while the opposite was true for extinction. This suggested that had the fire not occurred, we may have seen a mean decrease in *A. desertorum* metapopulation size associated with decreases in growing season precipitation, but the decrease outside the burn were compensated for with increases in the fire.

The best model for *A. desertorum* colonization retained variables that estimated effects of habitat suitability (HS) differentially inside and outside the burned areas and also different effects of OND inside and outside the burned areas. The interaction between habitat suitability and OND was also retained and was different inside and outside the burned areas (Table 4.3). The extinction model retained similar variables, but without the effect of HS outside the burn nor the associated interaction between HS and OND outside the burn. Extinction was also affected by the burn (Table 4.3). Colonization probability was greater at higher habitat suitability, and at higher occupied neighborhood density in burned areas. Further, the interaction between OND and habitat suitability in the burned areas that indicated some synergism between the two variables. In unburned sites, colonization probability was higher at high habitat suitability and OND but the absolute colonization probabilities were lower on average and the interaction between habitat suitability and OND was opposite (Fig D.11). The model predicted high colonization in the fire at high OND regardless of habitat suitability but reviewing the uncertainty plots and the data scatter it was clear that these cases did not exist in the sample (Fig D.11; Fig C.7).

Extinction was generally affected by the same predictors as colonization except with generally opposite effects. Extinction was high at low OND and low habitat suitability in the unburned area, however there was generally little data in these combinations (Fig C.7; Fig D.12). In the burned area, extinction was higher at combinations of low OND and low habitat suitability as well but, the model predicted greatest extinction at high OND and low habitat suitability and at high habitat suitability and low OND. However the uncertainty and data scatter made it clear that again there were very few cases of these combinations of variables and thus was not a real observed effect. An interaction between habitat suitability and OND suggested low habitat suitability or low OND had high extinction probabilities (Fig C.7; Fig D.12).

Discussion

Predictors of Plant Regional Dynamics

Models to explain spatial variation in colonization and extinction ranged from very simple to very complex depending on the species. For the 2010 to 2011 transition, colonization and extinction for *D. sophia, B. tectorum, L. perfoliatum, C. nutans, H. glomeratus* and *A. cristatum* were best described with relatively simple models that only retained the covariates of occupied neighborhood density, habitat suitability, and/or various disturbance variables. However, the other two species (*A. desertorum* and *S. altissimum*) required complex interactions between environmental (i.e. habitat suitability), neighborhood and disturbance variables, for that same year. Many of the effects of the individual predictors of the spatial dynamics were species specific but a few general trends emerged. Seven of the eight species responded to fire with the four annual Brassicaceae having higher colonization and lower extinction in the fire, and colonization decreased and extinction increased as time since fire increased.

Occupied neighborhood density was the only consistent predictor, with greater colonization associated with greater OND, coupled with lower extinction at greater OND. Occupied neighborhood density was derived as a surrogate for isolation and dispersal processes. The fact that it was a consistently important predictor of colonization and extinction supports island biogeographic principles and neutral ecology (Hubbell, 2001; MacArthur & Wilson, 1967). Sites that were more isolated had a lower probability of colonization, and the extinction rate was lower when closer to sources of propagules. A site surrounded by more populations of the same species would likely build up large seedbanks, decreasing the chance of extinction the next season, while areas that were more isolated would have much fewer potential propagules, thus increasing the probability of extinction. Interestingly, the optimal size of the neighborhood for OND was highly variable and differed between colonization and extinction. There were few consistencies but neighborhood sizes ranged from 20 m to 300 m indicating some species may only have needed information on local isolation/dispersal while others were better predicted with information from larger scales, which were more regional in context.

Temporal information (2009-2010, 2010-2011) was only available for three species but suggested strong temporal dynamics for *S. altissimum* and *C. nutans,* while *A. cristatum* followed the same general (increasing metapopulation size) trajectory for both years. The interaction between temporal and spatial predictors was most epitomized by *S. altissimum* as the spatial predictors alone did not make ecological sense until they were interpreted in the context of their interactions between each other and the temporal climate patterns. The mean estimates for *S. altissimum* metapopulation size indicated a large extinction event from 2009 to 2010 and a colonization event from 2010 to 2011. The extinction event was associated with climate conditions, as there was an extremely wet growing season in 2009 followed by a near average year in 2010. However, growing season precipitation was even lower in 2011, thus we would have expected a metapopulation decrease if the trend continued, but saw a large colonization event due to a wildfire. Further, as the species was increasing our null expectation was for higher colonization in higher habitat suitability, but this trend was also not observed. Although these results seemed initially counterintuitive, when put in the context of the annual spatial predictors, primarily disturbance, it was clear that interactions between disturbance (in this case a large wildfire) and other spatial covariates were more important than seasonal climate.

This highlighted the importance of being wary of interpreting the main effects, as suggested by Higgins and Richardson's (1998) mechanistic approach to modeling the spread of invasive pines. They found that interactions were many times more important than main effects. Further, responses to disturbance and environment could not be generalized across species, as there was a range of metapopulation trajectories that were dependent on the introduced species traits. Thus, to truly predict responses to disturbance and environmental changes, a time series of species specific spatial occurrence data is necessary.

Metapopulations, Regional Ensembles, or Spatially Extended Populations?

Adopting Freckleton and Watkinson (2002) modified typology for plant regional dynamics, we tried to use the evidence from our observations, and colonization and extinction data to place each of the eight species into one of their three regional typologies: spatially extended populations, regional ensembles, and the classic metapopulation (Table 4.4). Freckleton and Watkinson (2002) suggested that spatially extended populations are likely to arise when suitable habitat is much greater than unsuitable habitat. Two species, *B. tectorum* and *A. desertorum*, immediately lent themselves to this classification as they had very high mean occupancy in the study region, but still had non-zero conditional probabilities of colonization and extinction.

Thus, most of the patterning we saw was likely only due to local variation in births and deaths and local immigration and emigration, instead of regional dispersal between habitat patches.

One species, *A. cristatum,* was planted at the site and appears to be spreading in spatial extent away from these areas. It is one of the few species we would consider to be truly invasive at the site. Habitat suitability did not affect colonization and extinction probabilities, which suggested the species was not restricted by specific habitat requirements. This species also does not possess adaptations to encourage long distance dispersal, suggesting regional scale dispersal to be limited. Thus, it would logically follow that this species would fall under the spatially extended population classification, with local processes dominating regional dynamics. Species with these types of dynamics would seemingly be most problematic to managers, as targeting suitable habitats is not a viable management strategy, and the potential for the species to occupy the entire site exists. However, this may also suggest strategies such as treating/controlling patch edges because they are likely areas with higher colonization rates (*sensu* Moody & Mack, 1988) and would be most influential on local dispersal. Management practices to control processes at the regional scale for spatially extended populations may be restricted to interventions to reduce the chance of rare long distance dispersal by vectors such as vehicles or mammals.

Three of the species in the study dispersed via tumbling, allowing for medium to long distance dispersal depending on the community canopy structure (*D. sophia, S. altissimum, H. glomeratus*). The spatial configuration of occurrence records suggested these species were dispersed throughout the site and thus regionally restricted dispersal, due to invasion stage, was not presumed to be influencing the regional dynamics. Interestingly, classification of these species into one of the three regional dynamic categories was complicated by wildfire (for *D. sophia* and *S. altissimum*) which has increased in frequency and spatial extent in recent decades and has been largely attributed to anthropogenic activities. Mature sagebrush (and other shrubs) provided a dispersal barrier for the tumbling dispersal mechanism; however, when those structures were removed after fire, that dispersal barrier was removed. Within intact sagebrush, suitable habitat was predicted to be much less than unsuitable habitat, suggesting *S. altissimum* and *D. sophia* did not fit within spatially extended populations within mature sagebrush stands. However, colonization was extremely high and suitable sites appeared greater than unsuitable sites in young fires. As time since fire increased, colonization decreased and extinction increased, suggesting a transition from a spatially extended population immediately after fire, to perhaps regional ensembles or classic metapopulations in mature sagebrush stands. Thus, successional stage or native plant diversity (Anderson & Inouye, 2001) seems to play a large role in determining the relative importance of regional versus local dynamics in structuring regional populations especially with more ruderal species as these two tumblers seem to exemplify. With the ability for the tumbling dispersal mechanism to bridge the gap between suitable sites in mature stands, it is not inconceivable that all three species fell within the traditional metapopulation classification and the gradient of suitable habitat certainly suggests regional dynamics should be considered when trying to explain the long term regional trajectories in occupancy.

One species that had the highest probability of demonstrating classic metapopulation dynamics was *C. nutans* (family: Asteraceae). The seeds of this species possess a relatively large pappus and can disperse long distances via wind suggesting regional connectivity between patches was realistic. Suitable habitat was also predicted to be much less abundant than unsuitable habitat, and was likely a function of water availability. Although the potential exists, longer term data would be necessary to support classic metapopulation characteristics as opposed to a regional ensemble of island populations. Local extinctions and colonizations were observed but regional dynamics remain to be supported with more temporal data. Whether the population exhibits characteristics of island populations or metapopulations may further be complicated by temporal climate variation. This species seems to be limited by water, thus prolonged drought periods may push the species into island population dynamics where dispersal between patches is low and existing patches move slowly toward extinction. However, in average to above average water years, previously unsuitable habitat may become suitable, and source-sink dynamics may emerge in the classic metapopulation framework.

These results demonstrate the complexities of categorizing species into a class of regional dynamics as they vary in space and time, and depend on the configuration and amount of suitable habitat. Suitable habitat can vary due to disturbance or climatic events as evidenced by the fire responses and temporal dynamics of different species in this study. Dispersal strategies and abilities of individual species also have a large influence on the potential regional dynamics, with different levels of regional connectivity expected depending on physical structure of the environment and the adaptations of particular species. Although variable, we believe certain species fall within certain categories with greater likelihoods and whether regional or local processes are more important can drastically change management strategies and expectations of the results of management if the goal is to eradicate undesirable species.

Tables and Figures

Table 4.1: Optimal neighborhood sizes (radii in meters) used to calculate occupied neighborhood density (OND). Neighborhoods from 10 to 300 meters were tested and the OND as calculated at different radii that minimized AIC of each model's colonization (Col.) and extinction (Ext.) component is reported. For two species (*A. desertorum* and *S. altissimum* 2010 to 2011), the optimal neighbor differed inside (i) and outside (o) the fire. A (-) signifies the variable was not retained after AIC model selection.

Optimal Neighborhood Sizes 10 to 11

Species	Col.	Ext.
Agropyron cristatum	50	120
Sisymbrium altissimum	i300 o30	i230 o30
Carduus nutans	70	40
Halogeton glomeratus	170	300
Lepidium perfoliatum	60	70
Alyssum desertorum	i300 o30	i150 o70
Bromus tectorum	50	300
Descurainia sophia	50	300

Optimal Neighborhood Sizes 09 to 10

Species	Col.	Ext.
Sisymbrium altissimum	20	90
Carduus nutans	40	60
Agropyron cristatum	120	-

Table 4.2: Coefficient estimates from spatial colonization and extinction models for the 2009 to 2010 transition for three non-native species. Models were arrived at using AIC and (-) indicates the variable was not retained in the model. Estimates reported are for the colonization probability (Col.) and extinction probability (Ext.). Variable abbreviations include habitat suitability (HS), occupied neighborhood density (OND), and distance to paved road (Dist. Rd.).

	Agropyron cristatum		*Sisymbrium altissimum*		*Carduus nutans*	
	Col.	Ext.	Col.	Ext.	Col.	Ext.
Intercept	-3.41	-0.908	-5.83	0.544	-4.96	0.434
HS	-	-	5.01	6.231	19.91	-
Dist. Rd.	-	-	-	-	-	-
OND	4.37	-	4.73	0.611	14.53	-2.38
Fire	-	-	-	1.191	-	-
HS*OND	-	-	-	-12.03	-160.8	-

Table 4.3: Coefficient estimates from spatial colonization and extinction models for the 2010 to 2011 transition for eight non-native species. Models were arrived at using AIC and (-) indicates the variable was not retained in the model. Estimates reported are for the colonization probability (Col.) and extinction probability (Ext.). Variable abbreviations include habitat suitability (HS), occupied neighborhood density (OND), and distance to paved road (Dist. Rd.). Note: binary fire variable differs for *L. perfoliatum*.

	Agropyron cristatum		*Sisymbrium altissimum*		*Carduus nutans*		*Halogeton glomeratus*		*Lepidium perfoliatum*		*Alyssum desertorum*		*Bromus tectorum*		*Descurainia sophia*	
	Col.	Ext.	Col.	Ext.	Col.	Ext.	Col.	Ext.	Col.	Ext.	Col.	Ext.	Col.	Ext.	Col.	Ext.
Intercept	-2.152	0.0805	-4.71	0.92	-6.28	2.31	-5.61	2.19	-6.34	1.077	-4.52	2.41	-3.828	4.11	-2.8	-0.011
HS	-	-	-	-	-	-	13.01	-	-16.3	11.004	-	-	3.977	-5.66	-2.91	2.18
HS*Fire	-	-	-4.75	5.037	-	-	-	-	-	-	3.27	2.73	-	-	-	-
HS*NoFire	-	-	2.59	-1.536	-	-	-	-	-	-	4.52	-	-	-	-	-
OND	4.955	-3.153	-	-	7.14	-5.00	9.56	-10.85	9.27	-3.68	-	-	3.478	-4.07	4.36	-1.31
Fire*OND	-	-	9.232	-7.537	-	-	-	-	-	-	1.32	1.6	-	-	-	-
NoFire*OND	-	-	11.069	-3.051	-	-	-	-	-	-	7.66	-5.82	-	-	-	-
HS*OND	-	-	-	-	-	-	-53.9	86.32	-	-	-	-	-	-	-	-
HS*Fire*OND	-	-	-11.72	12.7	-	-	-	-	-	-	8.63	-13.3	-	-	-	-
HS*NoFire*OND	-	-	-13.61	5.77	-	-	-	-	-	-	-6.97	-	-	-	-	-
Fire	-	-	4.976	-2.99	-	1.49†	-	-1.2	3.68*	-1.04*	-	-2.49	1.343	-	1.52	-1.089
Time Since Fire	-	-	-0.122	0.053	-	-	-	-	-	-	-	-	-0.088	-	-0.073	0.0596
Green Strip	0.2268	-1.786	-	-	-	-	-	-	-	-	-		-	-	-	-
Dist. Rd.	-0.001	0.00028	-	-	-	-	-	-	-	-0.002	-	-	-	-	-	-
Unpaved Rd.	-	-	-	1.2188	-	-	-	-	-	-	-	-	-	-	-	-

*Actual variable is 1 if fire is 16 years old and 0 if not

†Actual variable 1 if Jefferson Burn 0 if not

Table 4.4: Theorized regional dynamic classifications for the eight non-native species in the colonization and extinction study. Species are ordered based on their primary dispersal mechanism and two of eight species have fire dependent regional dynamics.

		Theorized Regional Dynamics			
Species	Dispersal adaptation	Spatially Extended Population	Regional Ensemble	Meta-population	Depend on Fire?
Bromus tectorum	Gravity	x			
Alyssum desertorum	Gravity	x			
Lepidium perfoliatum	Gravity	x			
Agropyron cristatum	Gravity	x			
Descurainia sophia	Tumbler		x	x	x
Sisymbrium altissimum	Tumbler		x	x	x
Halogeton glomeratus	Tumbler		x		
Carduus nutans	Wind - Pappus		x	x	

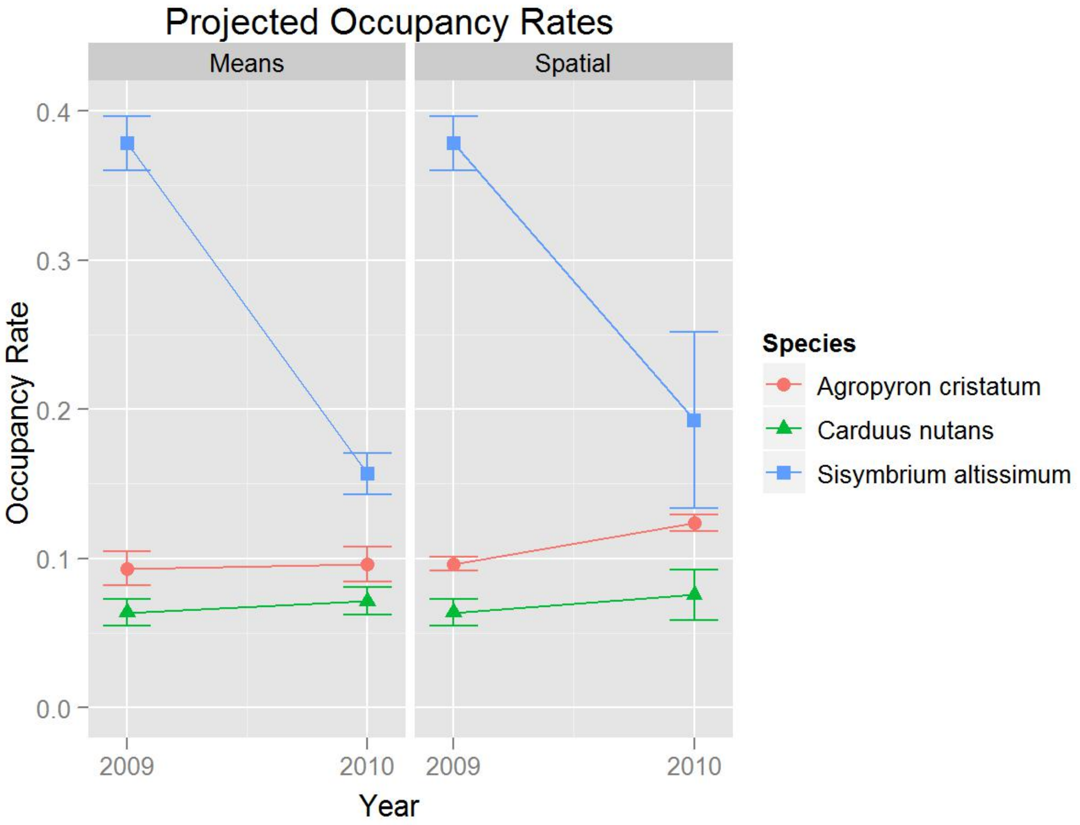


Figure 4.1: Projected mean occupancy rates from 2009 to 2010 for three species. The estimates in the left panel are naïve (intercept only terms for initial distribution, colonization and extinction). The estimates in the right panel are from spatially informed models. The 95% empirical bootstrap confidence intervals are shown.

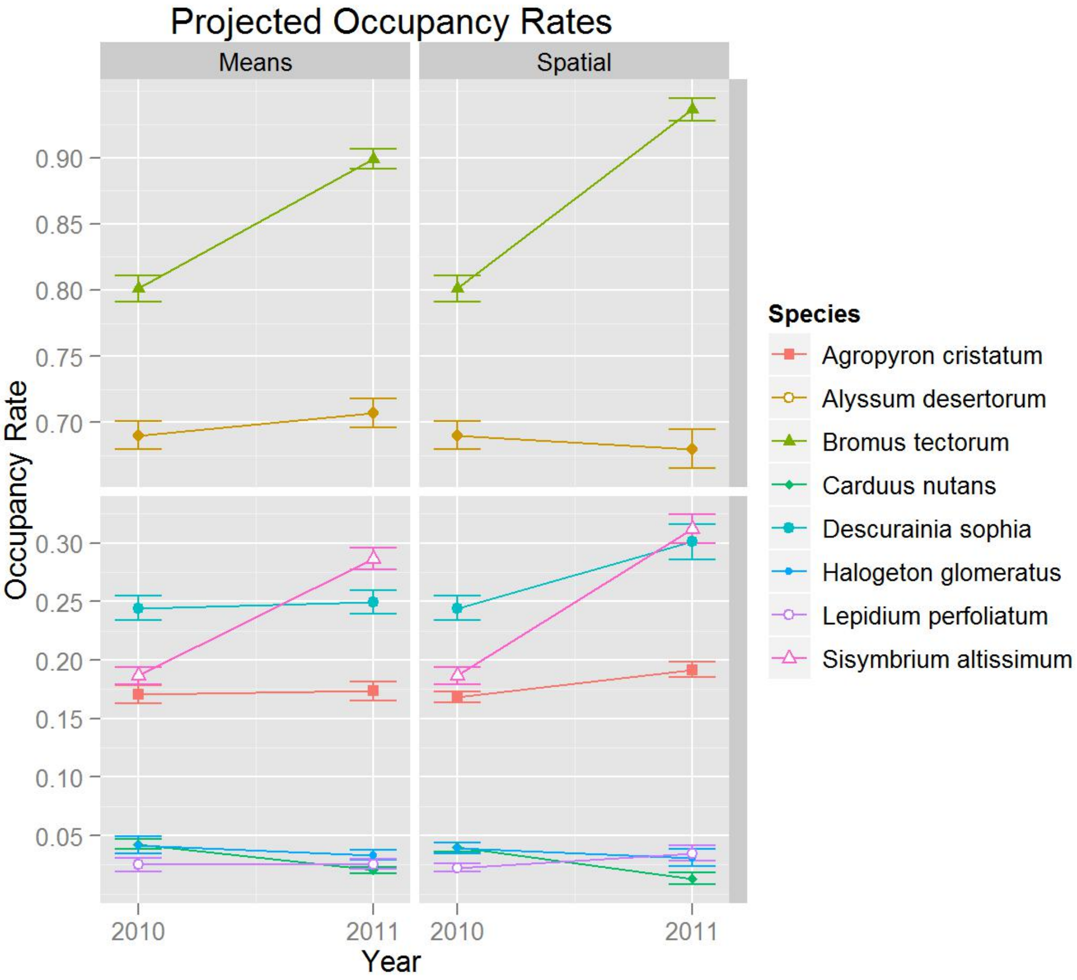


Figure 4.2: Projected mean occupancy rates from 2010 to 2011 for eight species. The estimates in the left panel are naïve (intercept only terms for initial distribution, colonization and extinction). The estimates in the right panel are from spatially informed models. The 95% empirical bootstrap confidence intervals are shown. Note the break in the y-axis to show more separation between species.

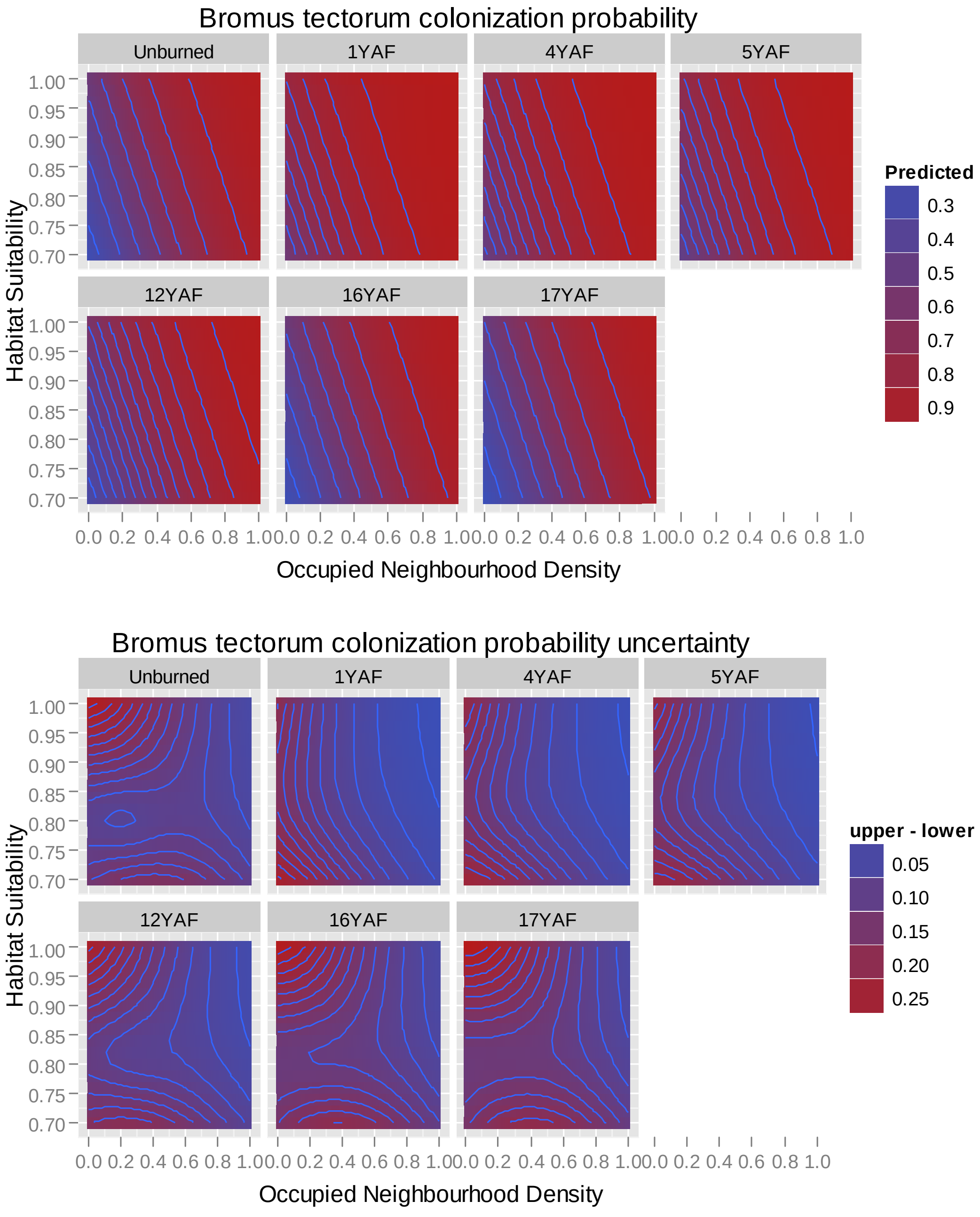


Figure 4.3: Predicted conditional probabilities of colonziation (upper7 panels) and associated uncertainty (lower 7 panels) for *Bromus tectorum*. Y-axis is habitat suitability and the x-axis is the occupied neighborhood density. The 7 panels indicate different burn status. The first panel is unburned, and the following panels are increasing burn age from one year after fire (1YAF) to 17 years after fire (17YAF). The surface is the predicted probability of colonization or extintion from low (blue) to high (red) and the lines indicate contours (isolines) in the response surface.

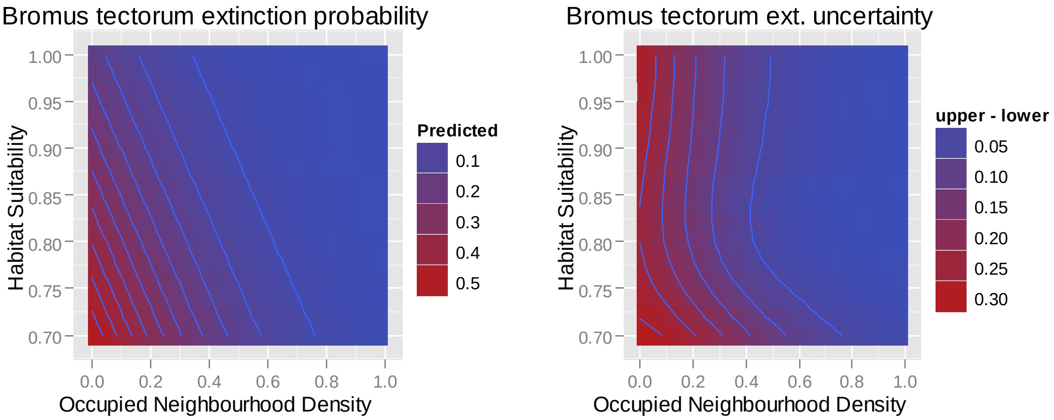


Figure 4.4: Predicted conditional probabilities of extinction and associated uncertainty for *Bromus tectorum*. Y-axis is habitat suitability and the x-axis is the occupied neighborhood density. The 7 panels indicate different burn status. The first panel is unburned, and the following panels are increasing burn age from one year after fire (1YAF) to 17 years after fire (17YAF). The surface is the predicted probability of colonization or extintion from low (blue) to high (red) and the lines indicate contours (isolines) in the response surface.

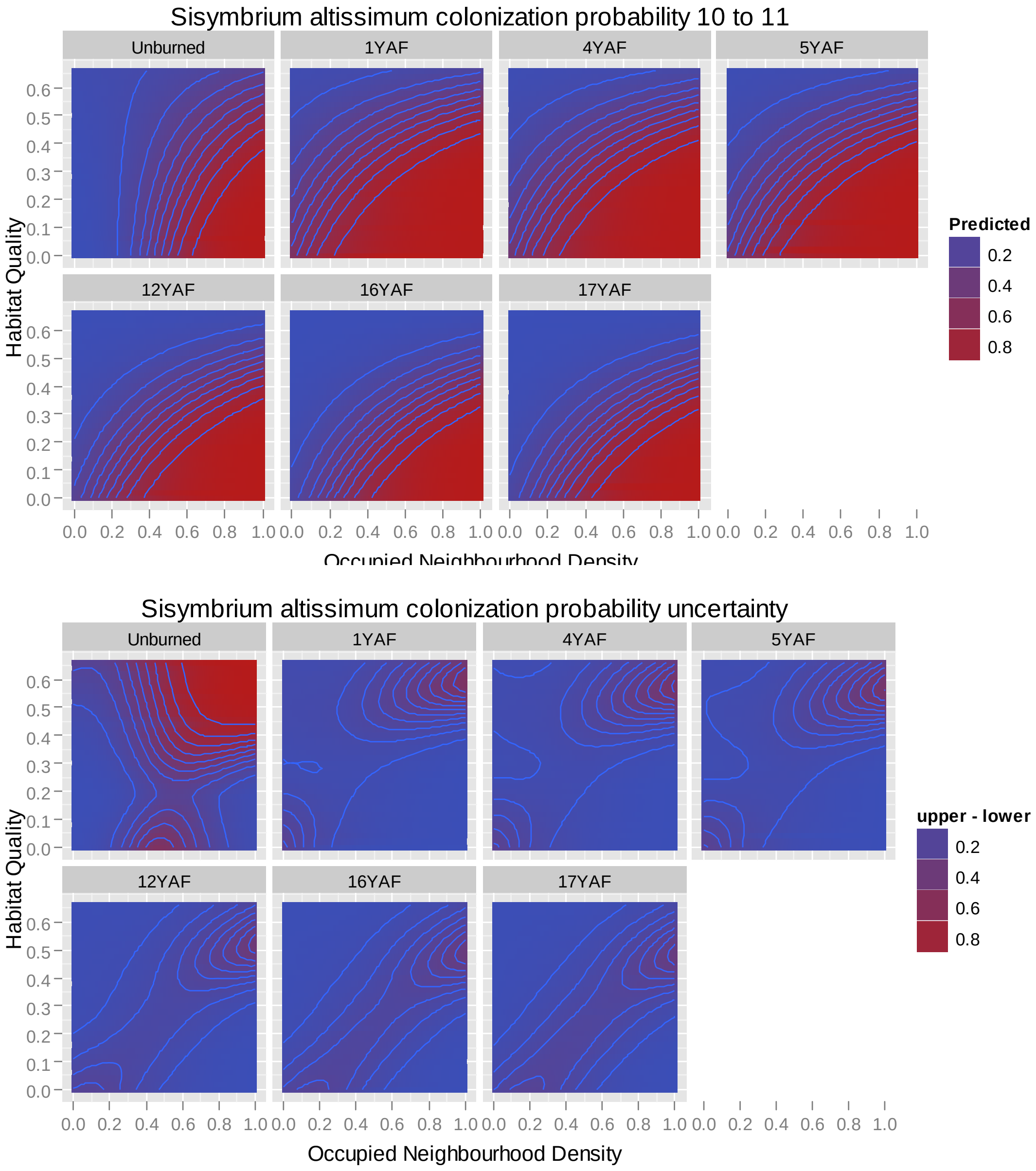


Figure 4.5: Predicted conditional probabilities of colonziation (upper 7 panels) and associated uncertainty (lower 7 panels) for *Sisymbrium altissimum*. Y-axis is habitat suitability and the x-axis is the occupied neighborhood density. The 7 panels indicate different burn status. The first panel is unburned, and the following panels are increasing burn age from one year after fire (1YAF) to 17 years after fire (17YAF). The surface is the predicted probability of colonization or extintion from low (blue) to high (red) and the lines indicate contours (isolines) in the response surface.

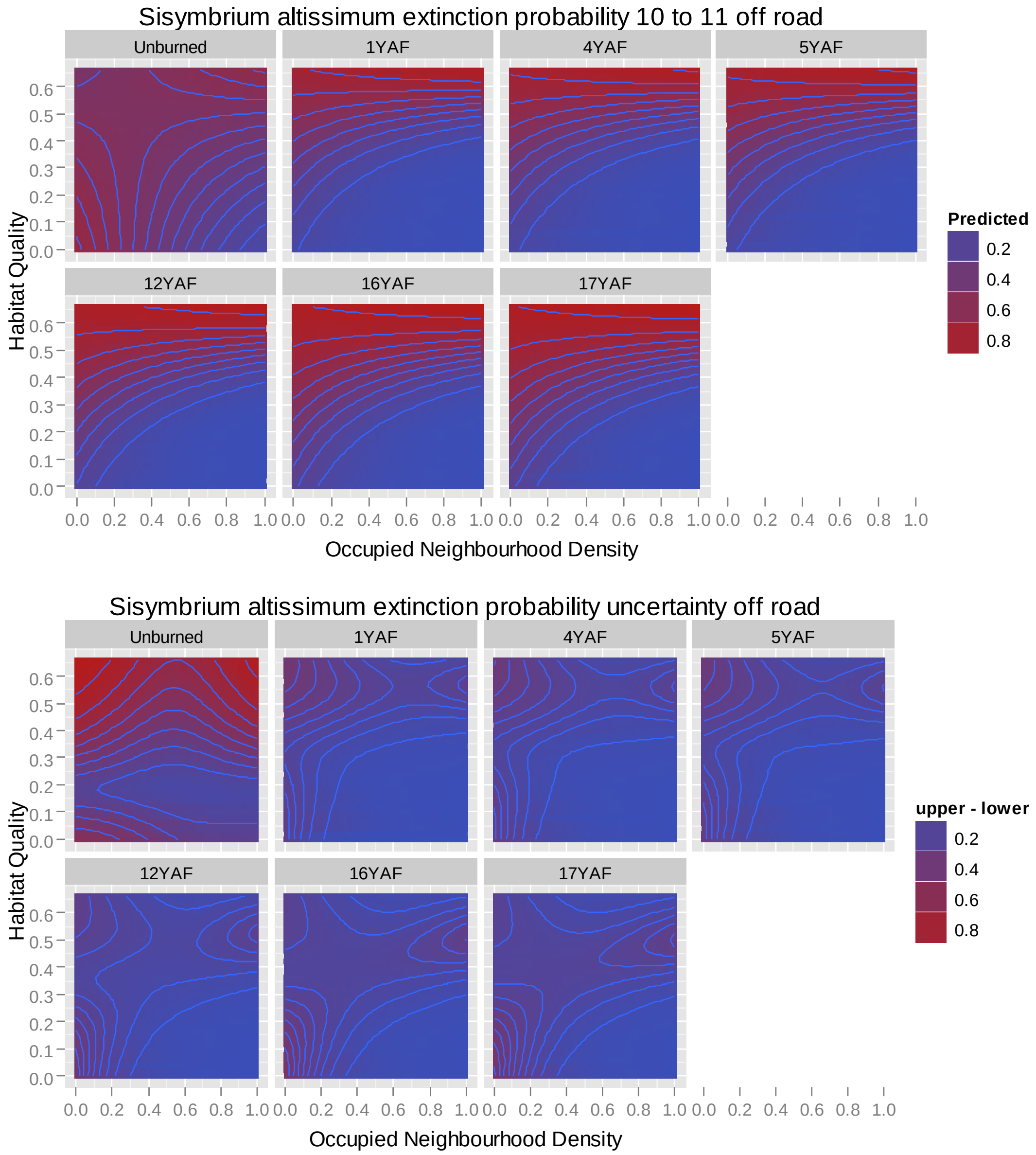


Figure 4.6: Predicted conditional probabilities of extinction (upper 7 panels) and associated uncertainty (lower 7 panels) for *Sisymbrium altissimum*. Y-axis is habitat suitability and the x-axis is the occupied neighborhood density. The 7 panels indicate different burn status. The first panel is unburned, and the following panels are increasing burn age from one year after fire (1YAF) to 17 years after fire (17YAF). The surface is the predicted probability of colonization or extintion from low (blue) to high (red) and the lines indicate contours (isolines) in the response surface.

References

Alpert, P., Bone, E. & Holzapfel, C. (2000) Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. *Perspectives in Plant Ecology, Evolution and Systematics*, **3**, 52-66.

Anderson, J.E. & Inouye, R.S. (2001) Landscape-scale changes in plant species abundance and biodiversity of a sagebrush steppe over 45 years. *Ecological Monographs*, **71**, 531-536

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677-1688.

Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U. & Richardson, D.M. (2011) A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, **26**, 333-339.

Bullock, J.M., Moy, I.L., Pywell, R., Coulson, S.J., Nolan, A.M. & Caswell, H. (2002) Plant dispersal and colonisation processes at local and landscape scales. *Dispersal Ecology* (ed. by J.M. Bullock, R.E. Kenward and R. Hails), pp. 279–302. Blackwell Science, Oxford.

Davis, M.A., Thompson, K. & Grime, J.P. (2001) Charles S. Elton and the dissociate of invasion ecology from the rest of ecology. *Diversity and Distributions*, **7**, 97-102.

Ehrlen, J. & Eriksson, O. (2003) Large-scale spatial dynamics of plants: a response to Freckleton & Watkinson. *Journal of Ecology*, **91**, 316-320.

Eriksson, O. (1996) Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. *OIKOS*, **77**, 248-258.

Fiske, I. & Chandler, R. (2011) unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, **43**, 1-23.

Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321-330.

Freckleton, R.P. & Watkinson, A.R. (2002) Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology*, **90**, 419-434.

Freckleton, R.P. & Watkinson, A.R. (2003) Are all plant populaiton metapopulations? *Journal of Ecology*, **91**, 321-324.

Hanski, I. (1998) Metapopulation dynamics. *Nature*, **396**, 41-49.

Hanski, I. (1999) *Metapopulation ecology*. Oxford University Press, Oxford.

Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B.A., Moore, K., Taylor, C. & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, **8**, 91-101.

Heger, T. & Trepl, L. (2003) Predicting biological invasions. *Biological Invasions*, **5**, 313-321.

Higgins, S.I. & Richardson, D.M. (1998) Pine invasions in the Southern Hemisphere: modelling interactions between organism, environment and disturbance. *Plant ecology*, **135**, 79-93.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography Princeton University Press, Princeton, New Jersey.

Kareiva, P., Mullen, A. & Southwood, R. (1990) Population dynamics in spatially complex environments: theory and data. *Philosophical Transactions: Biological Sciences*, **330**, 175-190.

Kolar, C.S. & Lodge, D.M. (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*, **16**, 199-204.

Koop, A.L., Fowler, L., Newton, L.P. & Caton, B.P. (2011) Development and validation of a weed screening tool for the United States. *Biological Invasions*, **14**, 273-294.

Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, **15**, 237–240.

MacKenzie, D.I. (2006) *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier, Amsterdam; Boston.

Maxwell, B.D., Lehnhoff, E. & Rew, L.J. (2009) The rationale for monitoring invasive plant populations as a crucial step for management. *Invasive Plant Science and Management*, **2**, 1-9.

Moody, M.E. & Mack, R.N. (1988) Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology*, **25**, 1009-1021.

R Development Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rew, L.J., Lehnhoff, E.A. & Maxwell, B.D. (2007) Non-indigenous species management using a population prioritization framework. *Canadian Journal of Plant Science*, **87**, 1029-1036.

Rew, L.J., Maxwell, B.D., Dougher, F.L. & Aspinall, R. (2006) Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species. *Biological Invasions*, **8**, 523-539.

Schreiber, S.J. (2010) Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 1907-1914.

Shive, J.P., Forman, A.D., Aho, K., Hafla, J.R., Blew, R.D. & Edwards, K.T. (2011) Vegetation Community Classification and Mapping of the INL Site. In. Gonzales-Stoller Surveillance, LLC

Simberloff, D. & Gibbons, L. (2004) Now you see them now you don't! - population crashes of established introduced species. *Biological Invasions*, **6**, 161-172.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmanek, M. & Westbrooks, R. (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, **21**, 1-16.

CHAPTER 5

CONCLUSION TO THESIS

The management of non-indigenous plant species can only be as effective as the information that informs it. Through much research over the last fifty years, description and quantification of biological invasions has improved our understanding of the underlying mechanisms and processes associated with invasions. Also, the vast increase in computation power over the last decade has given rise to incredible new tools for modeling and predicting species invasions. Knowledge of invasive species distributions at present and into the future can help inform management by prioritizing species for management and also populations within species for management based on a variety of criteria (Rew et al., 2007). However, models and predictions from models are only as good as the information used to inform them, and arguably all data contain some amount of uncertainty (Barry & Elith, 2006; Rocchini et al., 2011). Further, models and predictions from models are only as good as the ecological theory used to construct them and the degree to which assumptions are met by the modeler. Thus, this thesis had three goals: determine how to generate enough accurate and robust data to inform models of non-indigenous species (NIS) distribution (Chapter 2), evaluate the impact of sources of uncertainty on model predictions (Chapter 3), and to understand how different ecological process (local and regional) influenced predictions from models of species distribution to characterize the spatio-temporal dynamics of NIS (Chapters 2 and 4).

Most of the previous work predicting NIS spatial distributions has focused on country to global spatial scales to predict the potential range of a NIS in its new territory (often continent) (Elith & Leathwick, 2009; Richards et al., 2007; Thuiller et al., 2005). While this is an important endeavor for country and regional planning and prevention, actual management of NIS generally occurs at landscape scales within finite management units such as national parks and conservation units or in sections of private land. Thus, country to continent scale models fail to provide much information to managers concerned with smaller spatial extents. There have been a few studies that aim to predict landscape scale species distribution at fine resolution (Fleishman et al., 2002; Rew et al., 2005; Shafii et al., 2003; Václavík et al., 2010), but the adoption of these models has been slow, and recommendations for sampling strategies and sample sizes did not exist, nor did a comprehensive evaluation of the maximum predictive capacity of the models prior to this study.

Chapter 2 aimed to resolve the gaps in the existing research concerning landscape scale distribution models. Via simulation modeling we determined that transect sampling was 2.5 to 4.5 times more efficient than random sampling and both simulation and empirical data pointed to surprisingly consistent sample size recommendations (0.06% to 0.23%) to maximize the predictive capacity of landscape scale species distribution models. Surprisingly, sample size recommendations did not significantly differ between the two empirical sites, one with steep environmental gradients as compared to the other that had a relatively homogenous environment. The results did show quite a bit of variation in the maximum predictive power one could hope to achieve when using a species distribution model, with greater predictive power in areas that had steeper environmental gradients. Even though, predictive power was reduced in the more homogeneous site, models improved prediction over an assumed random distribution and provided satisfactory discrimination for a number of species.

Although these models were predictive and sample sizes were small enough to be achievable by managers, there were concerns about sources of uncertainty in the data that may bias model coefficients. Most species distribution modeling studies, especially those at large extents (continental and global) use herbarium records, or haphazardly collected occurrence records of species (Graham et al., 2004). These data have been collected over many years with little quality assurance or quality control and, are rife with potential sources of uncertainty ranging from observer bias to site selection bias to temporal climate variation. Our survey data in Chapter 2 represented our best effort to generate reliable data. The sampling frame was small enough such that species occurrence could be surveyed for the entire sampling frame area. There were always two observers at a given frame and sampling was targeted during the peak growing season. Given our best efforts though, there was still the potential for imperfect detection and we wanted to investigate how detection error could influence model predictions as understanding this would give us more confidence in the tools intended for managers to adopt and use in their management programs.

Chapter 3 investigated the effects of detection error on model predictions using a double survey sampling design within one season in 2011. We found that even given our best efforts, detection was imperfect and ranged from 0.24 to 0.94 with an average near

0.75. We fit models where detection error was explicitly incorporated versus those that assumed perfect detection and found that if detection was greater than 0.87 model predictions and uncertainty in predictions was not affected from a practical perspective. However, for species with detection rates lower than 0.87 habitat relationships sometimes switched and uncertainty in the naïve model could be greatly underestimated. Sampling sites twice in a season is costly from a management perspective, so our analysis will allow us to advise managers to spend more time training early in the season to achieve high detection rates. High detection rates would prevent them from having to sample frames twice to assess detection rates. However, if a given species has a low detection rate double sampling of frames would be necessary in order to deal with this source of uncertainty, and can be dealt with using hierarchical site occupancy models. Although detection error was found to be a problem in the study in Chapter 3 for certain species, the sample size used for the analysis was small relative to the recommendations for managers derived from the analysis in Chapter 2. It would follow that as sample size increases, the relative influence of detection error should decrease. From our results we would generally recommend visiting more sites to characterize the habitat relationships more confidently (i.e. increase sample size for all variable combinations) over revisiting sites within season to assess detection uncertainty. Nevertheless, with smaller data sets these sources of uncertainty need to be at least considered when the models and distribution maps from the models are informing management decisions.

Chapters 2 and 3 were focused on predicting realized distributions of NIS from data collected in one season. Such predictions are static. Some of the criticisms of species distribution modeling lies in the knowledge that natural systems are dynamic and constantly changing based on spatial and temporal processes. Consequently, the required assumption of a species being in equilibrium within the environment is generally incorrect regardless of a species origin (i.e. native or non-native). More definitely, NIS in the invasion stage (before maximum dispersion) are most certainly not in a state of equilibrium and thus static distribution maps are likely to become obsolete as the time from their creation increases. Thus, the importance of gathering monitoring data to help inform predictions of species distribution into the future is vital to increase the accurate application of these models.

Gathering temporal data on species occurrence can help us understand the regional dynamics of plant species and understand where on the equilibrium/non-equilibrium continuum they lie. Understanding the relative importance of local versus regional processes in the colonization and extinction dynamics are necessary to characterize regional trajectories of extinction, persistence or spread. Further, characterizing how these regional and local processes change depending on spatial and temporal drivers will allow us to target management and predict future species distributions. There has been much debate as to whether plants exhibit classical metapopulation dynamics (Hanski, 1997) or local processes coupled with rare stochastic dispersal events are more influential in regional dynamics (Ehrlen & Eriksson, 2003; Freckleton & Watkinson, 2002, 2003). Determining the dominant processes important to plant colonizations and extinctions can not only help our fundamental understanding of individual species regional dynamics but could dramatically change how management is performed. Thus, Chapter 4 aimed at measuring regional dynamics (colonization and extinction) of eight NIS and tried to parse out the most important drivers of local colonization and extinction to determine what types of regional dynamics were most likely.

Drivers of colonization and extinction were diverse and the effects quite different between the eight species studied. Wildfire was an important driver for a number of the species, but most were predicted to decrease in occurrence as time since fire increased, indicating active management post fire was not important and that successional dynamics would reduce the non-native ruderals over time. Occupied neighborhood density was the one ubiquitous predictor: greater colonization probability was associated with more neighbors and simultaneously lower extinction. Some species only needed a few predictors to explain colonization and extinction rates but two species required complex interactions between fire, habitat suitability, and occupied neighborhood density, all contextualized within the temporal climate patterns (i.e. wet year, average year, dry year) before the drivers became logical. This demonstrated the importance of both spatial and climate effects when trying to predict the future distributions of species and that to really be able to understand the regional dynamics, species specific data are necessary.

Freckleton and Watinkson (2002) suggested three categories of plant regional dynamics: classic metapopulations, regional ensembles, and spatially extended populations. The results of the distribution and dynamics of the eight species also suggested that each of the three possible regional dynamic patterns were present. The type of regional dynamics category could have large implications as to management strategies. Recommending treating source populations assumes a classic metapopulation and that source-sink dynamics are occurring, whereas recommending treating population edges assumes, and would be most effective for, spatially extended populations. Thus, further research to support or refute these categories could not only shed light on an understudied part of plant population ecology but drastically change management strategies and recommendations.

By exploring different aspects of the ecology and modeling of non-native plant species, this thesis will provide solid recommendations to land managers who wish to employ distribution modeling as a tool and investigated at least one source of uncertainty that can be used to bolster managers' confidence in using these tools. We have also added to the ecological understanding of regional dynamics of non-native species and identified important drivers of these dynamics. This research coupled with existing and future work will further improve our ability to make predictions about landscape scale NIS distribution and dynamics that will improve management effectiveness and efficiency.

Future Work

Perhaps the most stimulating component of my research was discovering how little ecologists know about plant regional dynamics. The research in Chapter 4 just started to suggest a variety of regional dynamics were likely but to quantitatively prove whether each type of regional dynamics exists will take some creativity and a longer time period of repeat data.

In the plant dynamics study, a predictor called occupied neighborhood density was used as a predictor of colonization and extinction but I would contend that that is primarily capturing a local process and not a regional one (although the varied neighborhood sizes suggested perhaps regional and local processes). To truly test the influence of regional dynamics there are two hurdles to overcome. The first is how to define suitable habitat. Metapopulation theory relies on conceptualizing suitable habitat as discrete patches and thus trying to accomplish this for these species is necessary. Currently we are using probability of occurrence based on the current realized distribution to define habitat suitability (based solely on environmental predictor variables). But from this we cannot calculate regional connectivity of habitat patches. Thus, one might take a resource selection probability function approach to define the habitat (thus removing the false absences from the habitat calculation) and create a binary classification threshold (perhaps by maximizing the sensitivity). Then those habitat patches could be used to calculate regional connectivity and isolation to test whether they are important predictors of colonization and extinction. Further, my study just quantified local colonizations and extinctions instead of patch level colonization and extinction. Patch level extinction and colonization may be more indicative of regional dynamics whereas local (10 m x 10 m frame/quadrat size) extinction may speak more to the local dynamics, possibly within a patch, independent of regional processes. With information on the relative importance of regional versus local processes in regional population dynamics, we would be able to more definitively classify regional dynamics of these species and thus suggest management strategies or make predictions of future occupancy patterns.

References

Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. Journal of Applied Ecology, 43, 413-423.

Ehrlen, J. & Eriksson, O. (2003) Large-scale spatial dynamics of plants: a response to Freckleton & Watkinson. Journal of Ecology, 91, 316-320.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology, Evolution, and Systematics, 40, 677-697.

Fleishman, E., Mac Nally, R., Fay, J.P., & Murphy, D.D. (2002) Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. Conservation Biology, 15, 1674-1685.

Freckleton, R.P. & Watkinson, A.R. (2002) Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. Journal of Ecology, 90, 419-434.

Freckleton, R.P. & Watkinson, A.R. (2003) Are all plant populaiton metapopulations? Journal of Ecology, 91, 321-324.

Graham, C., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution, 19, 497-503.

Hanski, I. (1997) Uniting two general patterns in the distribution of species. Science, 275, 397-400.

MacArthur, R.H. & Wilson, E.O. (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey.

Rew, L.J., Lehnhoff, E.A., & Maxwell, B.D. (2007) Non-indigenous species management using a population prioritization framework. Canadian Journal of Plant Science, 87, 1029-1036.

Rew, L.J., Maxwell, B.D., & Aspinall, R. (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Science, 53, 236-241.

Richards, C.L., Carstens, B.C., & Lacey Knowles, L. (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography, 34, 1833-1845.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. Progress in Physical Geography, 35, 211-226.

Shafii, B., Price, W.J., Prather, T.S., Lass, L.W., & Thill, D.C. (2003) Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Science, 51, 748-751.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G., Hughes, G.O., & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology, 11, 2234-2250.

Václavík, T., Kanaskie, A., Hansen, E.M., Ohmann, J.L., & Meentemeyer, R.K. (2010) Predicting potential and actual distribution of sudden oak death in Oregon: Prioritizing landscape contexts for early detection and eradication of disease outbreaks. Forest Ecology and Management, 260, 1026-1035.

REFERENCES CITED

References

Alexander, H.M., Slade, N.A., & Kettle, W.D. (1997) Application of mark-recapture models to estimation of the population size of plants. Ecology, 78, 1230-1237.

Alexander, H.M., Slade, N.A., Kettle, W.D., Pittman, G.L., & Reed, A.W. (2009) Detection, survival rates and dynamics of a cryptic plant, *Asclepias meadii*: applications of mark-recapture models to long-term monitoring studies. Journal of Ecology, 97, 267-276.

Alpert, P., Bone, E. & Holzapfel, C. (2000) Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. Perspectives in Plant Ecology, Evolution and Systematics, 3, 52-66.

Anderson, J.E. & Inouye, R.S. (2001) Landscape-scale changes in plant species abundance and biodiversity of a sagebrush steppe over 45 years. Ecological Monographs, 71, 531-536

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. Journal of Biogeography, 33, 1677-1688.

Araújo, M.B. & Pearson, R.G. (2005) Equlibrium of species' distributions with climate. Ecography, 28, 693-695.

Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. Journal of Applied Ecology, 43, 413-423.

Bauer, J.T. (2012) Invasive species: “back-seat drivers” of ecosystem change? Biological Invasions.

Betts, M.G., Ganio, L.M., Huso, M.M.P., Som, N.A., Huettmann, F., Bowman, J., & Wintle, B.A. (2009) Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review”. Ecography, 32, 374-378.

Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U., & Richardson, D.M. (2011) A proposed unified framework for biological invasions. Trends in Ecology & Evolution, 26, 333-339.

Blossey, B. & Notzold, R. (1995) Evolution of increased competative ability in invasive nonindigenous plants: a hypothesis. The Journal of Ecology, 83, 887-889.

Brooks, M.L., D'Antonio, C.M., Richardson, D.M., Grace, J.B., Keeley, J.E., DiTomaso, J.M., Hobbs, R.J., Pellant, M., & Pyke, D. (2004) Effects of invasive alien plants on fire regimes. BioScience, 54, 677-688.

Bullock, J.M., Moy, I.L., Pywell, R., Coulson, S.J., Nolan, A.M. & Caswell, H. (2002) Plant dispersal and colonisation processes at local and landscape scales. *Dispersal Ecology* (ed. by J.M. Bullock, R.E. Kenward and R. Hails), pp. 279–302. Blackwell Science, Oxford.

Burnham, K.P. & Anderson, D.R. (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods Research, 33, 261-304.

Callaway, R.M. & Ridenour, W.M. (2004) Novel weapons: invasive success and the evolution of increased competative ability. Frontiers in Ecology and the Environment, 2, 436-443.

Catford, J.A., Jansson, R., & Nilsson, C. (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Diversity and Distributions, 15, 22-40.

Chen, G., Kéry, M., Zhang, J., & Ma, K. (2009) Factors affecting detection probability in plant distribution studies. Journal of Ecology, 97, 1383-1389.

Coudun C, Gégout J-C (2006) The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. Ecological Modelling 199:164-175

Coutts, S.R., Klinken, R.D., Yokomizo, H., & Buckley, Y.M. (2010) What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management? Biological Invasions, 13, 1649-1661.

Crosby, A.W. (1986) Ecological imperialism: the biological expansion of Europe, 900-1900 Cambridge University Press.

Crossman, N.D. & Bass, D.A. (2008) Application of common predictive habitat techniques for post-border weed risk management. Diversity and Distributions, 14, 213-224.

Daehler, C.C. (2003) Performane comparisons of co-occurring native and alien invasive plants: implications for conservation and restoration. Annual Review of Ecology, Evolution, and Systematics, 34, 183-211.

Darwin, C. (1859) On the Origin of Species. John Murray.

Davis, M.A. (2009) Invasion Biology Oxford University Press, Oxford, New York.

Davis, M.A., Thompson, K., & Grime, J.P. (2001) Charles S. Elton and the dissociate of invasion ecology from the rest of ecology. Diversity and Distributions, 7, 97-102.

Dormann, C.F. (2009) Response to Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review". Ecography, 32, 379-381.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30, 609-628.

Efron, B. (1986) How biased is the apparent error rate of a prediction rule? Journal of the American Statistical Association, 81, 461-470.

Ehrlen, J. & Eriksson, O. (2003) Large-scale spatial dynamics of plants: a response to Freckleton & Watkinson. Journal of Ecology, 91, 316-320.

Eriksson, O. (1996) Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. OIKOS, 77, 248-258.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 40, 677-697.

Elton, C.S. (1958) The ecology of invasions by plants and animals. The University of Chicago Press, Chicago and London.

Engler, R., Randin, C.F., Vittoz, P., Czáka, T., Beniston, M., Zimmermann, N.E., & Guisan, A. (2009) Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? Ecography, 32, 34-45.

Evangelista PH, Kumar S, Stohlgren TJ, et al. (2008) Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions 14:808-817

Fiske, I. & Chandler, R. (2011) unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43, 1-23.

Fleishman E, Mac Nally R, Fay JP, et al. (2002) Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. Conservation Biology 15:1674-1685

Franklin, J. (2009) Mapping species distributions: spatial inference and prediction Cambridge University Press, Cambridge, UK.

Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. Diversity and Distributions, 16, 321-330.

Franklin, J., Wejnert, K.E., Hathaway, S.A., Rochester, C.J., & Fisher, R.N. (2009) Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. Diversity and Distributions, 15, 167-177.

Freckleton, R.P. & Watkinson, A.R. (2002) Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. Journal of Ecology, 90, 419-434.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A., & Wintle, B.A. (2008) When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. Austral Ecology, 33, 986-998.

Gelbard, J.L. & Belnap, J. (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology, 17, 420-432.

Giljohann, K.M., Hauser, C.E., Williams, N.S.G., & Moore, J.L. (2011) Optimizing invasive species control across space: willow invasion management in the Australian Alps. Journal of Applied Ecology, 48, 1286-1294.

Graham, C., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution, 19, 497-503.

Grime, J.P. (1977) Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. The American Naturalist, 111, 1169-1194.

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

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993-1009.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. Ecological Modelling, 135, 147-186.

Gurevitch, J. & Padilla, D. (2004) Are invasive species a major cause of extinctions? Trends in Ecology & Evolution, 19, 470-474.

Hanski, I. (1998) Metapopulation dynamics. Nature, 396, 41-49.

Hanski, I. (1999) Metapopulation ecology. Oxford University Press, Oxford.

Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B.A., Moore, K., Taylor, C., & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. Ecology Letters, 8, 91-101.

Hauser, C.E. & McCarthy, M.A. (2009) Streamlining 'search and destroy': cost-effective surveillance for invasive species management. Ecology Letters, 12, 683-692.

Heger, T. & Trepl, L. (2003) Predicting biological invasions. Biological Invasions, 5, 313-321.

Hernandez PA, Graham CH, Master LL, et al. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773-785

Hierro, J.L., Maron, J.L., & Callaway, R.M. (2005) A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. Journal of Ecology, 93, 5-15.

Higgins, S.I. & Richardson, D.M. (1998) Pine invasions in the Southern Hemisphere: modelling interactions between organism, environment and disturbance. Plant Ecology, 135, 79-93.

Higgins, S.I., Richardson, D.M., & Cowling, R.M. (2001) Validation of a spatial simulation model of a spreading alien plant population. Journal of Applied Ecology, 38, 571-584.

Higgins, S.I., Richardson, D.M., Cowling, R.M., & Trinder-Smith, T.H. (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conservation Biology, 13, 303-313.

Hoeting, J.A., Leecaster, M., & Bowden, D. (2000) An improved model for spatially correlated binary responses. Journal of Agriculture, Biological, and Environmental Statistics, 5, 102-114.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey.

Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O., & Estoup, A. (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. Evolutionary Applications, 5, 89-101.

Kareiva, P., Mullen, A. & Southwood, R. (1990) Population dynamics in spatially complex environments: theory and data. Philosophical Transactions: Biological Sciences, 330, 175-190.

Kéry, M. (2004) Extinction rate estimates for plant populations in revisitation studies: importance of detectability. Conservation Biology, 18, 570-574.

Kéry, M. & Gregg, K.B. (2003) Effects of life-state on detectability in a demographic study of the terrestrial orchid *Cleistes bifaria*. Journal of Ecology, 91, 265-273.

Kolar, C.S. & Lodge, D.M. (2001) Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution, 16, 199-204.

Koop, A.L., Fowler, L., Newton, L.P. & Caton, B.P. (2011) Development and validation of a weed screening tool for the United States. Biological Invasions, 14, 273-294.

Kroll, A.J., Risenhoover, K., McBride, T., Beach, E., Kernohan, B.J., Light, J., & Bach, J. (2008) Factors influencing stream occupancy and detection probability parameters of stream-associated amphibians in commercial forests of Oregon and Washington, USA. Forest Ecology and Management, 255, 3726-3735.

Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America, 15, 237–240.

Liu, H. & Stiling, P. (2006) Testing the enemy release hypothesis: a review and meta-analysis. Biological Invasions, 8, 1535-1545.

Loiselle BA, Jørgensen PM, Consiglio T, et al. (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? Journal of Biogeography 35:105-116

Lozier JD, Mills NJ (2011) Predicting the potential invasive range of light brown apple moth (Epiphyas postvittana) using biologically informed and correlative species distribution models. Biological Invasions 13:2409-2421

Luoto M, Pöyry J, Heikkinen RK, et al. (2005) Uncertainty of bioclimate envelope models based on the geographical distribution of species. Global Ecology and Biogeography 14:575-584

MacArthur, R.H. & Wilson, E.O. (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey.

MacDougall, A.S., Gilbert, B., & Levine, J.M. (2009) Plant invasions and the niche. Journal of Ecology, 97, 609-615.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., & Fakhri, B. (2000) Biotic invasions: causes, epidemiology, global consequences, and control. Ecological Applications, 10, 689-710.

MacKenzie, D.I. (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence Elsevier, Amsterdam; Boston.

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

Mackenzie, D.I. & Royle, J.A. (2005) Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology, 42, 1105-1114.

Maxwell, B.D., Lehnhoff, E., & Rew, L.J. (2009) The rationale for monitoring invasive plant populations as a crucial step for management. Invasive Plant Science and Management, 2, 1-9.

McPherson JM, Jetz W (2007) Effects of species ecology on the accuracy of distribution models. Ecography 30:135-151

Moody, M.E. & Mack, R.N. (1988) Controlling the spread of plant invasions: the importance of nascent foci. Journal of Applied Ecology, 25, 1009-1021.

Moore, J.L., Hauser, C.E., Bear, J.L., Williams, N.S.G., & McCarthy, M.A. (2011) Estimating detection-effort curves for plants using search experiments. Ecological Applications, 21, 601-607.

Parendes, L.A. & Jones, J.A. (2000) Role of light availability and dispersal in exotic plant invasion along roads and streams in the H. J. Andrews Experimental Forest, Oregon. Conservation Biology, 14, 64-75.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology & Biogeography, 12, 361-371.

Pimentel, D., Zuniga, R., & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics, 52, 273-288.

Peterson T (2003) Predicting the Geography of Species' Invasions via Ecological Niche Modeling. The Quarterly Review of Biology 78:419-433

Pulliam, R.H. (1988) Sources, sinks, and population regulation. The American Naturalist, 132, 652-661.

Pysek, P. & Hulme, P.E. (2009) Invasion biology is a discipline that's too young to die. Nature, 460, 324.

R Development Core Team (2011) R: A language and environment for statistical computing. 2.13.1 edn. R Foundation for Statistical Computing, Vienna, Austria,

Reese GC, Wilson KR, Hoeting JA, et al. (2005) Factors affecting species distribution predictions: a simulation modeling experiment. Ecological Applications 15:554-564

Regan, T.J., Chadès, I., & Possingham, H.P. (2011) Optimally managing under imperfect detection: a method for plant invasions. Journal of Applied Ecology, 48, 76-85.

Regan, T.J., McCarthy, M.A., Baxter, P.W.J., Dane Panetta, F., & Possingham, H.P. (2006) Optimal eradication: when to stop looking for an invasive plant. Ecology Letters, 9, 759-766.

Rew, L.J., Lehnhoff, E.A., & Maxwell, B.D. (2007) Non-indigenous species management using a population prioritization framework. Canadian Journal of Plant Science, 87, 1029-1036.

Rew, L.J., Maxwell, B.D., & Aspinall, R. (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. Weed Science, 53, 236-241.

Rew, L.J., Maxwell, B.D., Dougher, F.L. & Aspinall, R. (2006) Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species. Biological Invasions, 8, 523-539.

Richards CL, Carstens BC, Lacey Knowles L (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography 34:1833-1845

Richardson, D.M. & Pysek, P. (2008) Fifty years of invasion ecology - the legacy of Charles Elton. Diversity and Distributions, 161-168.

Richardson, D.M. et al. (2011). A compendium of essential concepts and terminology in biological invasions. In Fifty Years of Invasion Ecology: The Legacy of Charles Elton (ed D.M. Richardson), pp. 409-420. Wiley-Blackwell.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. Progress in Physical Geography, 35, 211-226.

Rota, C.T., Fletcher, R.J., Evans, J.M., & Hutto, R.L. (2011) Does accounting for imperfect detection improve species distribution models? Ecography, 34, 659-670.

Schreiber, S.J. (2010) Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. Proceedings of the Royal Society B: Biological Sciences, 277, 1907-1914.

Seipel, T., Kueffer, C., Rew, L.J., Daehler, C.C., Pauchard, A., Naylor, B.J., Alexander, J.M., Edwards, P.J., Parks, C.G., Arevalo, J.R., Cavieres, L.A., Dietz, H., Jakobs, G., McDougall, K., Otto, R., & Walsh, N. (2012) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. Global Ecology and Biogeography, 21, 236-246.

Shafii, B., Price, W.J., Prather, T.S., Lass, L.W., & Thill, D.C. (2003) Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Science, 51, 748-751.

Shea, K. & Chesson, P. (2002) Community ecology theory as a framework for biological invasions. Trends in Ecology and Evolution, 17, 170-176.

Shefferson, R.P., Sandercock, B.K., Proper, J., & Beissinger, S.R. (2001) Estimating dormancy and survival of a rare herbaceous perennial using mark-recapture models. Ecology, 82, 145-156.

Shive, J.P., Forman, A.D., Aho, K., Hafla, J.R., Blew, R.D., & Edwards, K.T. (2011). Vegetation community classification and mapping of the INL Site. Gonzales-Stoller Surveillance, LLC.

Simberloff, D. & Gibbons, L. (2004) Now you see them now you don't! - population crashes of established introduced species. Biological Invasions, 6, 161-172.

Smith, J.B., Jenks, J.A., & Klaver, R.W. (2007) Evaluating detection probabilities for American Marten in the Black Hills, South Dakota. Journal of Wildlife Management, 71, 2412-2416.

Smith, R.G., Maxwell, B.D., Menallad, F.D., & Rew, L.J. (2006) Lessons from agriculture may improve the management of invasive plants in wildland systems. Frontiers in Ecology and the Environment, 4, 428-434.

Stockwell DRB, Peterson T (2002) Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1-13

Stohlgren, T.J. & Schnase, J.L. (2006) Risk analysis for biological hazards: what we need to know about invasive species. Risk Analysis, 26, 163-173.

Syphard, A.D. & Franklin, J. (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. Ecography, 32, 907-918.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G., Hughes, G.O., & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology, 11, 2234-2250.

USDA & NRCS (2012) The PLANTS Database (ed N.P.D. Team), Vol. 2012, pp. USDA, NRCS. 2012. The PLANTS Database (http://plants.usda.gov, 28 February 2012). National Plant Data Team, Greensboro, NC 27401-4901 USA., Greensboro, NC.

Václavík T, Kanaskie A, Hansen EM, et al. (2010) Predicting potential and actual distribution of sudden oak death in Oregon: Prioritizing landscape contexts for early detection and eradication of disease outbreaks. Forest Ecology and Management 260:1026-1035

Václavík, T. & Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling, 220, 3248-3258.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Diversity and Distributions, 18, 73-83.

Vilà, M., Espinar, J.L., Hejda, M., Hulme, P.E., Jarošík, V., Maron, J.L., Pergl, J., Schaffner, U., Sun, Y., & Pyšek, P. (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters, 14, 702-708.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmanek, M., & Westbrooks, R. (1997) Introduced species: a significant component of human-caused global change. New Zealand Journal of Ecology, 21, 1-16.

Williamson, M. & Fitter, A. (1996) The varying success of invaders. Ecology, 77, 1661-1666.

Williamson, M.H. (1996) Biological Invasions. Chapman & Hall, London.

Wilson JRU, Dormontt EE, Prentis PJ, et al. (2009) Something in the way you move: dispersal pathways affect invasion success. Trends in Ecology & Evolution 24:136-144

Wisz MS, Hijmans RJ, Li J, et al. (2008) Effects of sample size on the performance of species distribution models. Diversity and Distributions 14:763-773

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., & Blackard, J.A. (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. Journal of Applied Ecology, 44, 1057-1067.

APPENDICES

APPENDIX A:

SAMPLING MAPS: IDAHO NATIONAL LABORATORIES

Figure A.1. Transects sampled in 2009 and 2010 at the Idaho National Laboratories. The yellow line is a 2 km buffer around the conservation management plan area, the black lines are roads, the orange polygons are facilities and the blue lines are 2 km long transects.

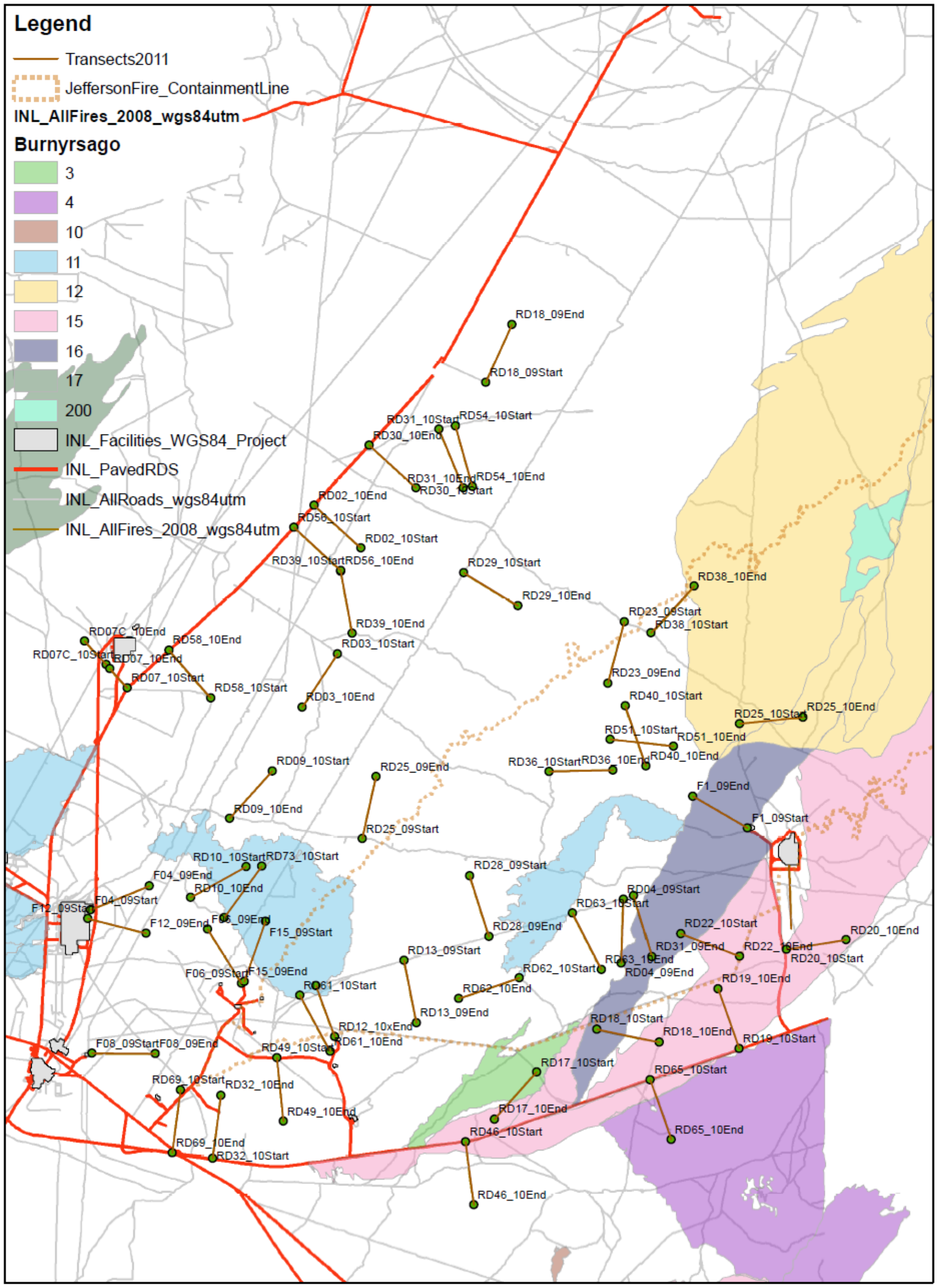


Figure A.2. Transects resampled in 2011 (brown lines with green start and end points). The red lines are paved roads, gray lines unpaved roads, and blue polygons facilities. The colored polygons are range fires that burned over the previous 18 years. The dotted brown line is the boundary of the Jefferson Fire that burned in 2010.

APPENDIX B:

SUPPLEMENTAL TABLES AND FIGURES

(CHAPTER 2)

Table B.1: Table of predictor variables used in the empirical resampling study. Sites are Yellowstone National Park (YNP) and the Idaho National Laboratories (INL).

Yellowstone National Park	Idaho National Laboratories
Distance to road	Burned (1 if burned, 0 if not)
Distance to trail	Time since last fire
Elevation	Slope
Slope	Cosine(Aspect)
Distance to streams	Sine(Aspect)
Trees (1 if trees, 0 if not)	Elevation
Annual Radiation	Curvature (9 Cell Window)
ISO 128	Curvature (21 Cell Window)
Burned (1 if burned, 0 if not)	Curvature (63 Cell Window)
Landsat Bands 1-5 and 7	Distance to Facilities
Normalized Difference Vegetation Index	Distance to Fire Boundaries
Ratio Vegetation Index	Binary two-track road (1 if adjacent to road 0 if not)
Green Normalized Difference Vegetation Index	Distance to paved roads
Specific Leaf Area Vegetation Index	Year Sampled
Normalized Difference Moisture Index	
Cosine(Aspect)	
Sine(Aspect)	
Year Sampled	

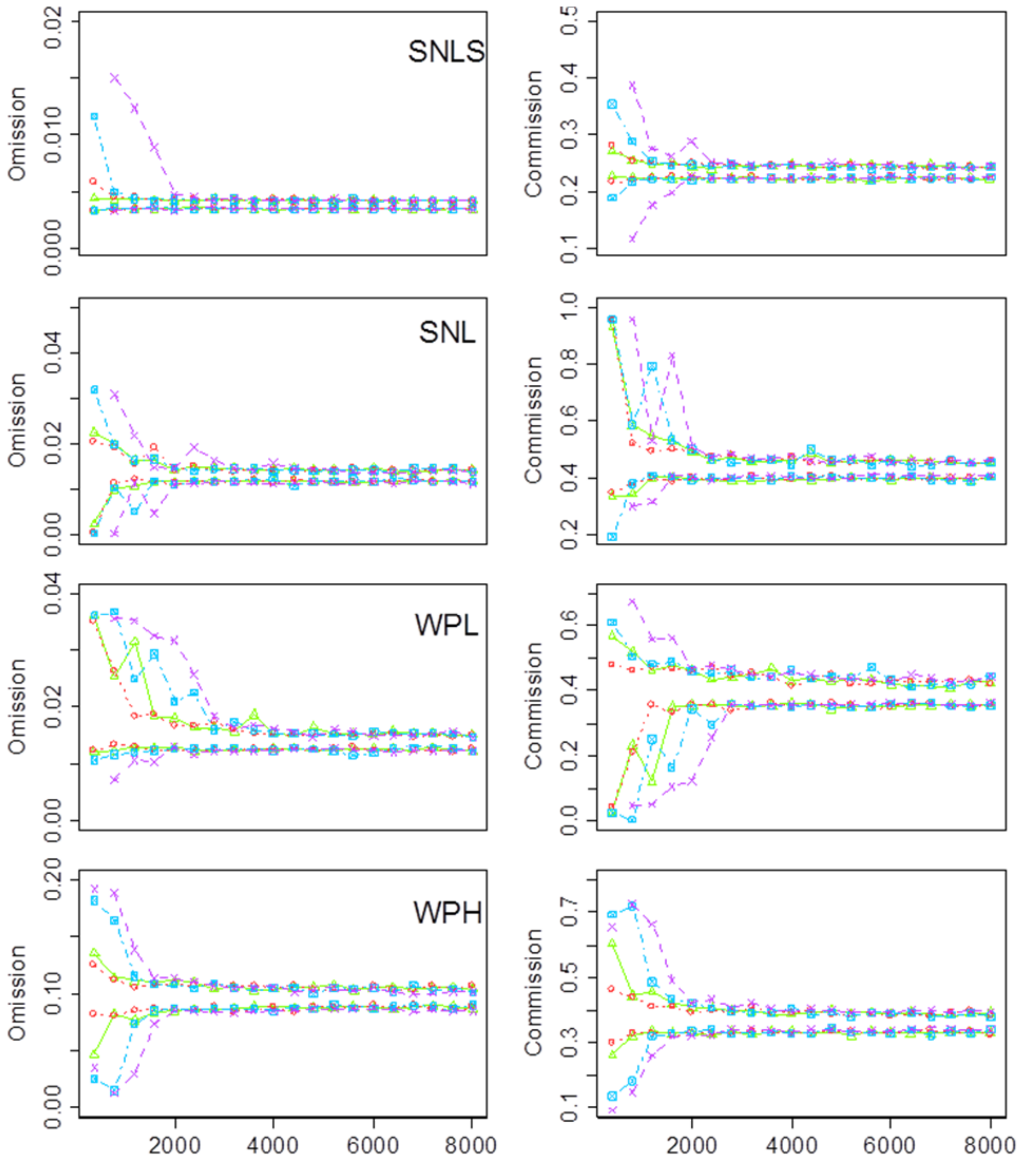


Figure B.1: Left panel shows the 95% interval boundaries for the distribution of the omission error as a function of sample size for the four plant patterns. The right panel shows 95% interval boundaries for distribution of commission error values for each plant pattern as a function of sample size. The four plant distribution patterns were SNLS – Strong Negative Low frequency Specialist, SNL – Strong Negative Low frequency generalist, WPL – Weak Positive Low frequency generalist, and WPH – Weak Positive High frequency generalist. The four sampling methods were– Simple Random Sample (red circles), Stratified Random Sample (green triangles), Transects (blue squares) and Rectangular Clusters (purple X's).

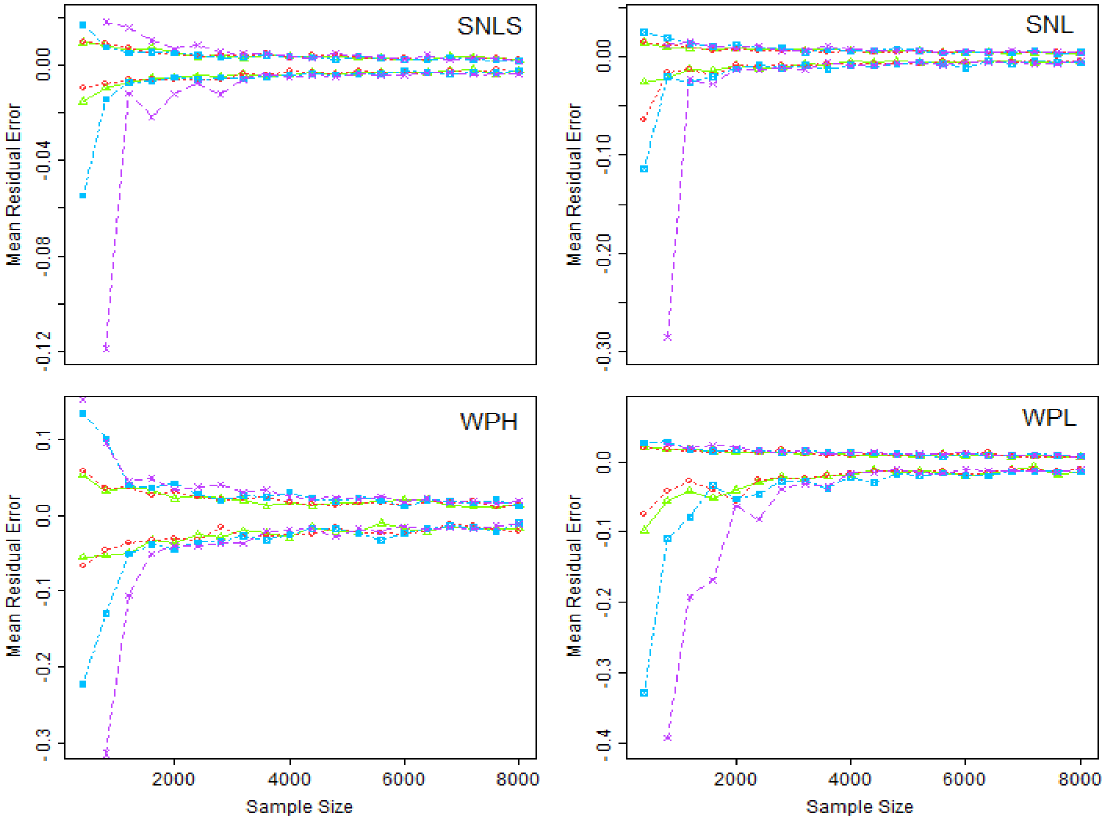


Figure B.2. 95% interval boundaries for the distribution of the mean residual error as a function of sample size for the four plant patterns. The four plant distribution patterns were SNLS – Strong Negative Low frequency Specialist, SNL – Strong Negative Low frequency generalist, WPL – Weak Positive Low frequency generalist, and WPH – Weak Positive High frequency generalist. The four sampling methods were– Simple Random Sample (red circles), Stratified Random Sample (green triangles), Transects (blue squares) and Rectangular Clusters (purple x’s).

APPENDIX C:

DATA SCATTERS OF COLONIZATION AND EXTINCTION

(CHAPTER 4)

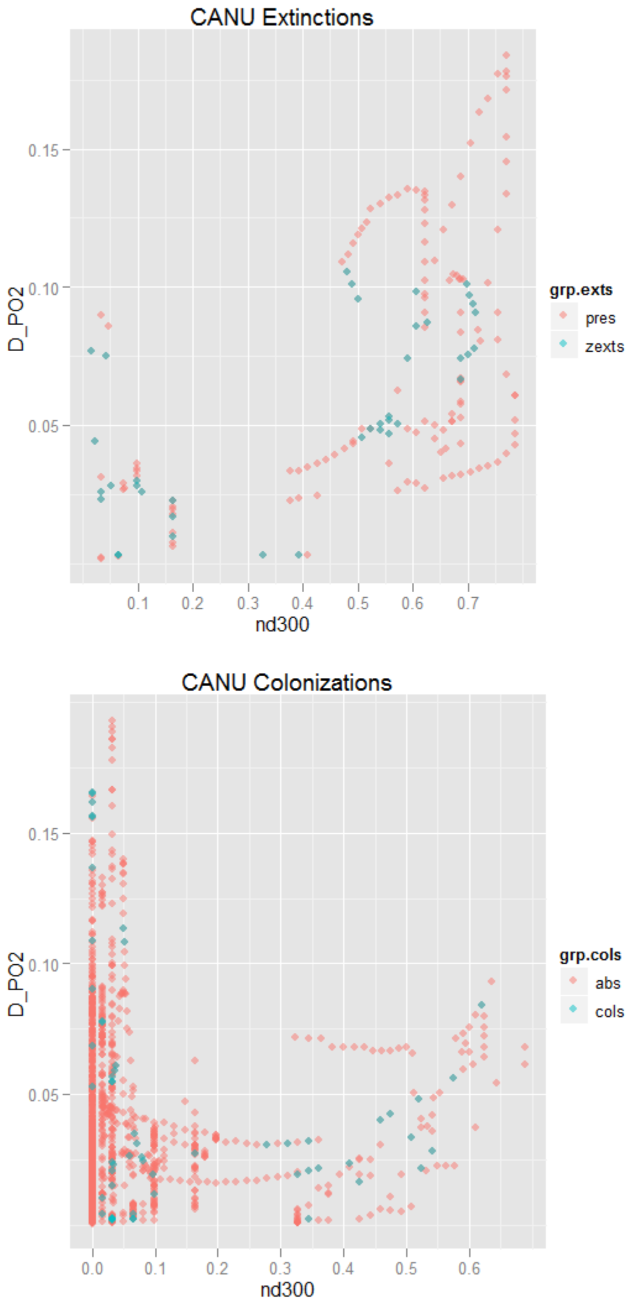


Figure C.1. Raw data scatters of colonization and extinction for *Carduus nutans*, 2009 to 2010 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2009 that went extinct in 2010 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis ('D_PO2') and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis ('nd300'). The right panel shows the absences in 2009 in red and the blue points are the absences in 2009 that became present in 2010 (colonizations). These points are arrayed on the same two axes as the left panel.

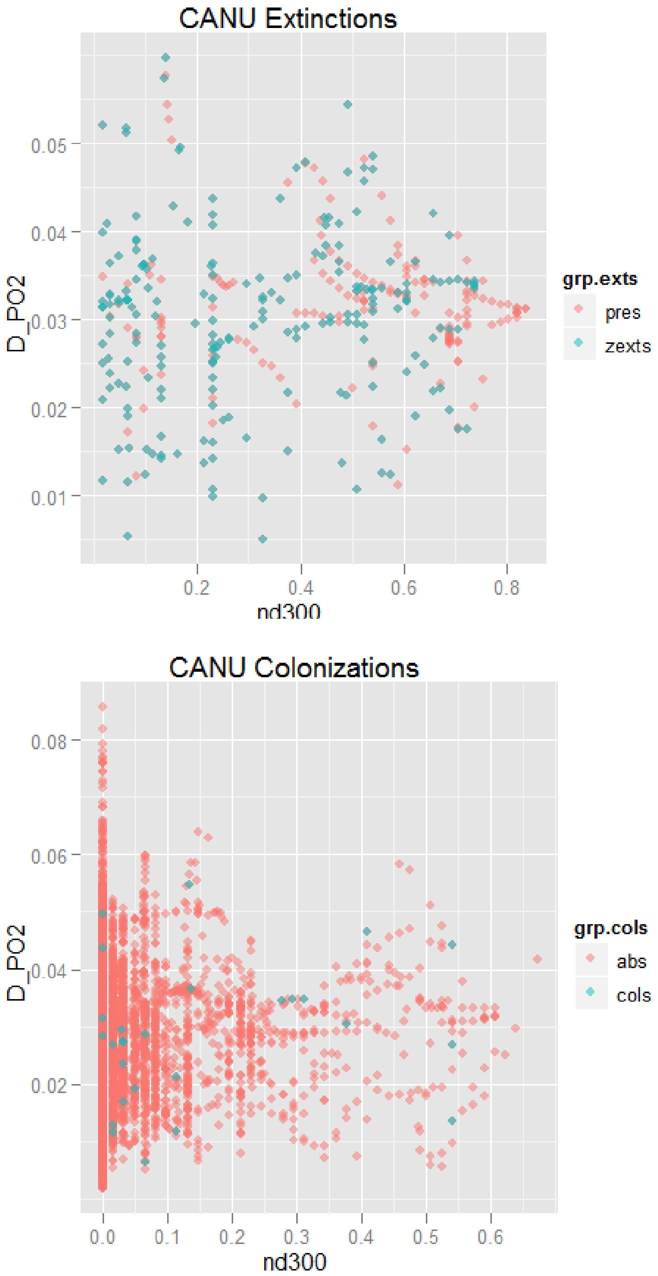


Figure C.2. Raw data scatters of colonization and extinction for *Carduus nutans*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2010 (colonizations). These points are arrayed on the same two axes as the left panel.

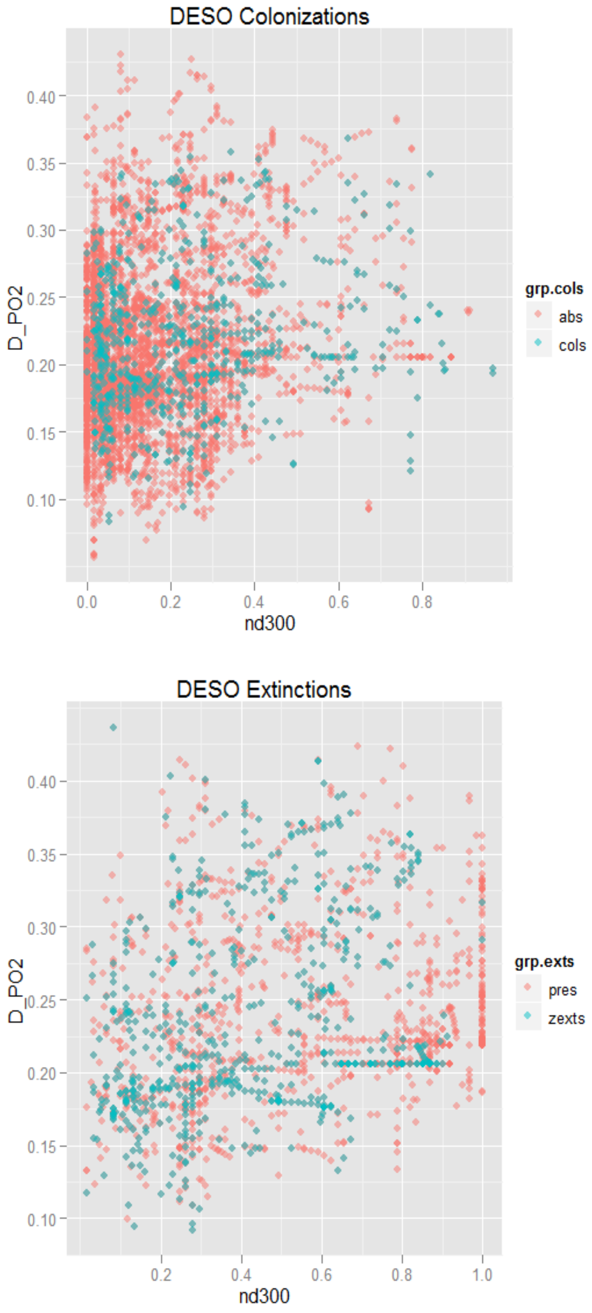


Figure C.3. Raw data scatters of colonization and extinction for *Descurainia sophia*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

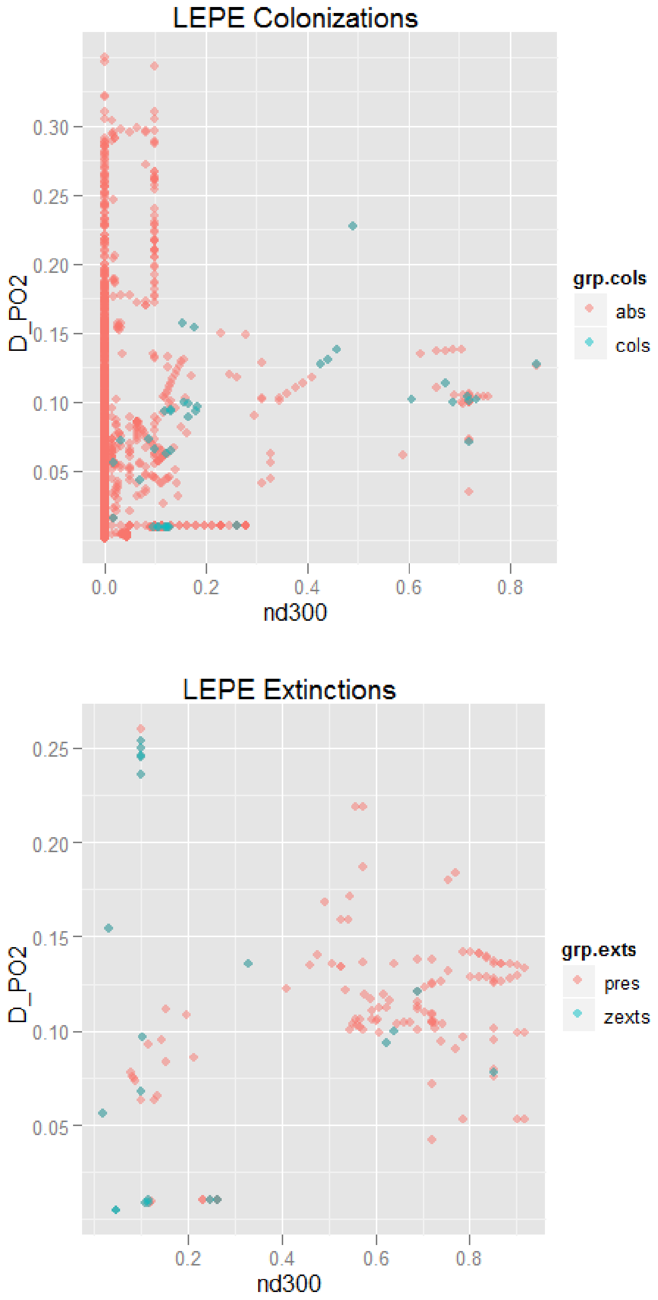


Figure C.4. Raw data scatters of colonization and extinction for *Lepidium perfoliatum*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

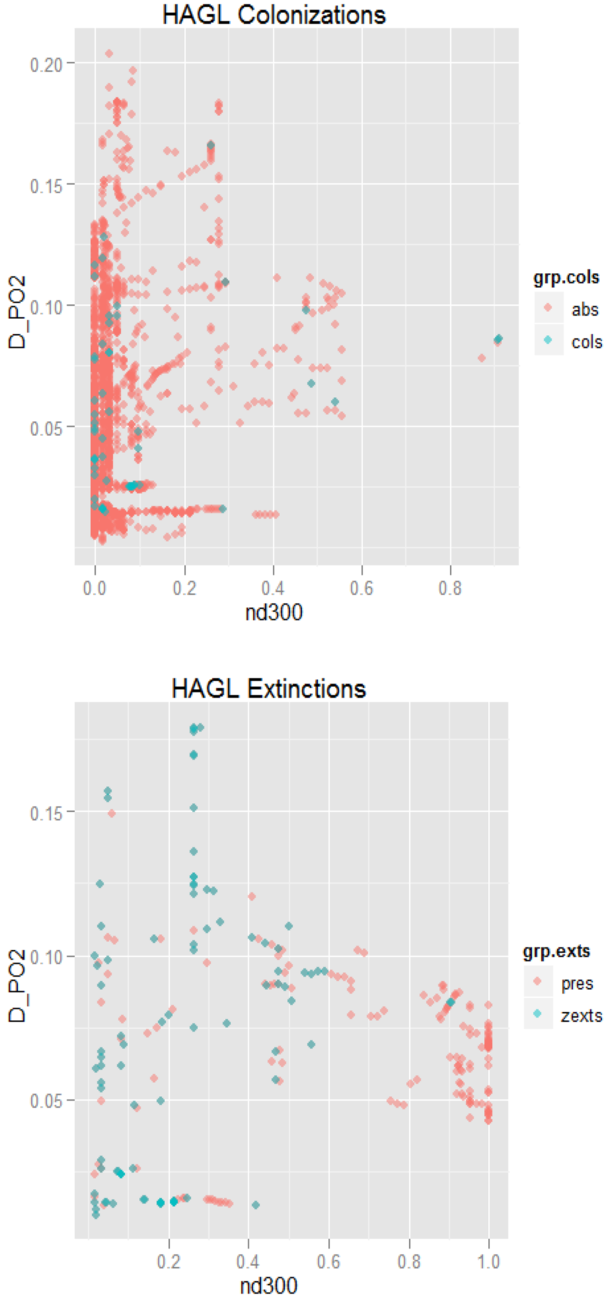


Figure C.5. Raw data scatters of colonization and extinction for *Halogeton glomeratus*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

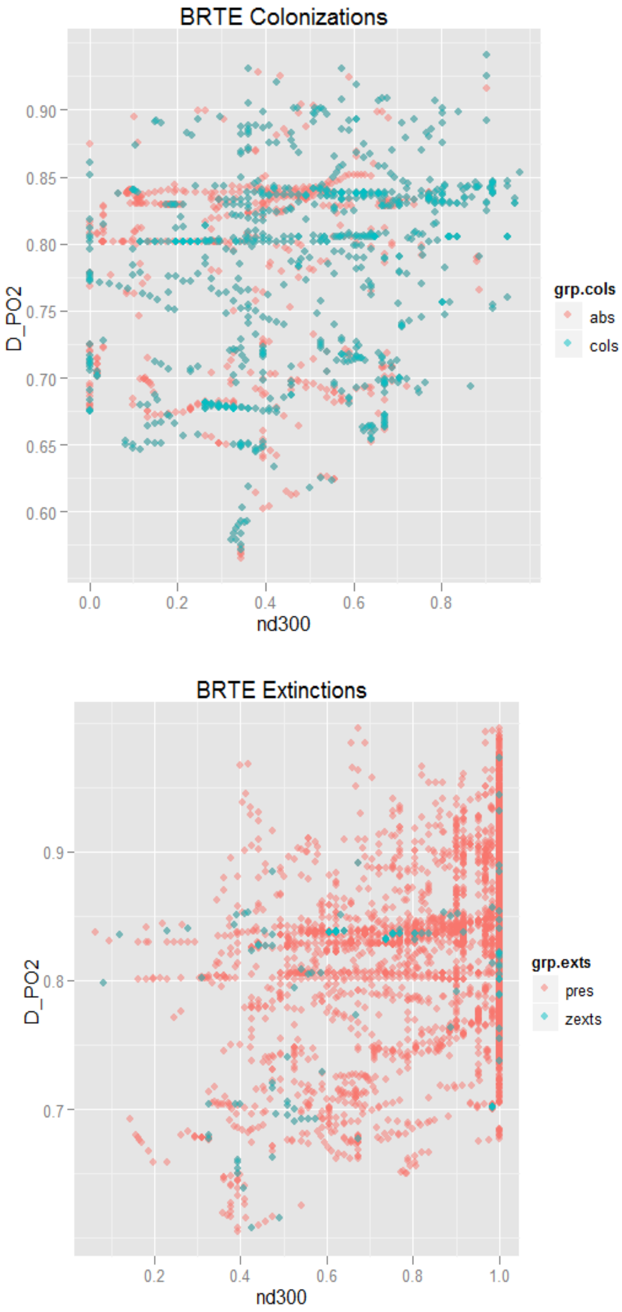


Figure C.6. Raw data scatters of colonization and extinction for *Bromus tectorum*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

Figure C.7. Raw data scatters of colonization and extinction for *Alyssum desertorum*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO2’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

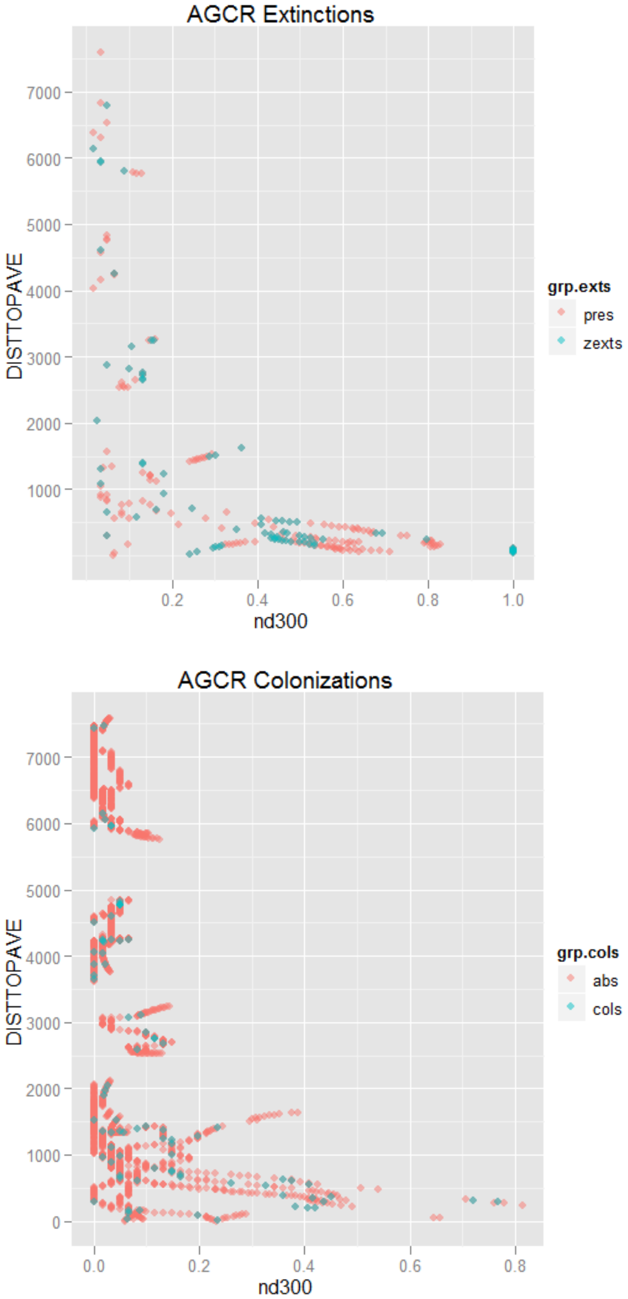


Figure C.8. Raw data scatters of colonization and extinction for *Agropyron cristatum*, 2009 to 2010 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2009 that went extinct in 2010 (extinctions). The points are arrayed based on two independent variables, distance to paved road on the y-axis (‘DISTTOPAVE’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2009 in red and the blue points are the absences in 2009 that became present in 2010 (colonizations). These points are arrayed on the same two axes as the left panel.

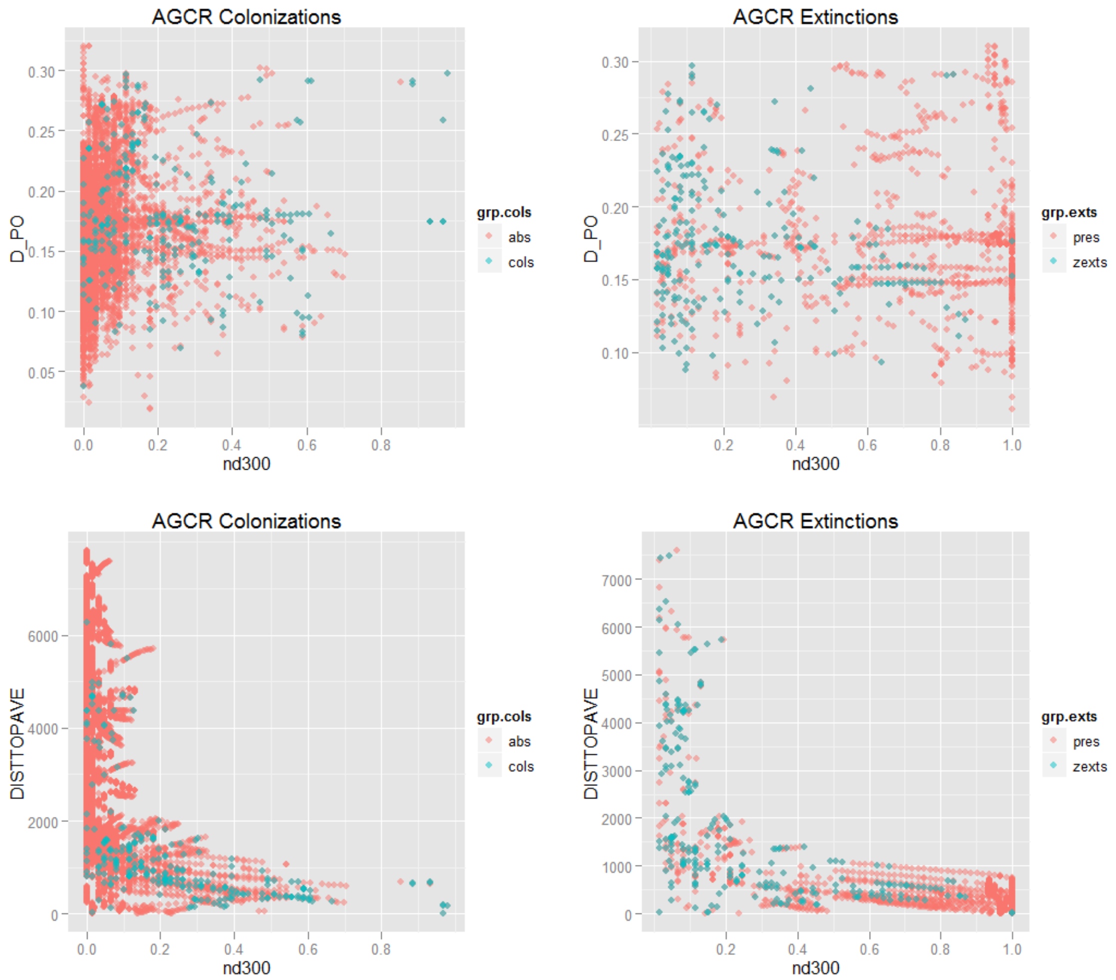
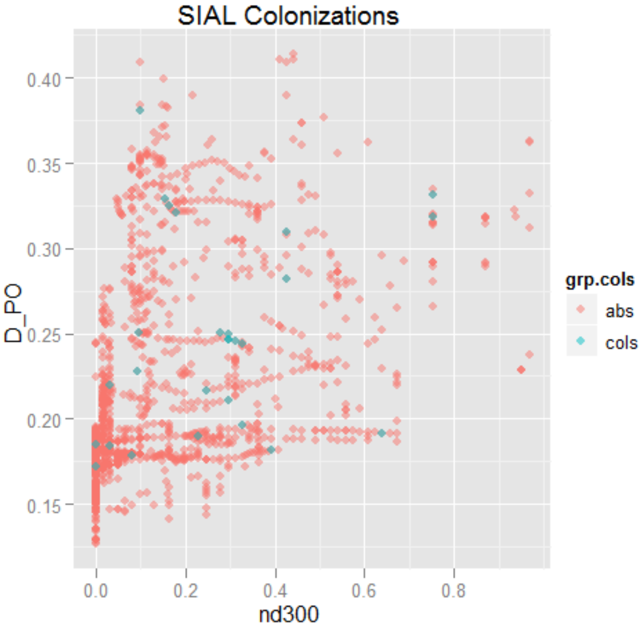


Figure C.9. Raw data scatters of colonization and extinction for *Agropyron cristatum*, 2010 to 2011 transition. The left column shows the presences in 2010 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability (‘D_PO’) on the y-axis and occupied neighborhood density on the x-axis in the first row, and distance to paved road on the y-axis (‘DISTTOPAVE’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’) in the second row. The right column shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.



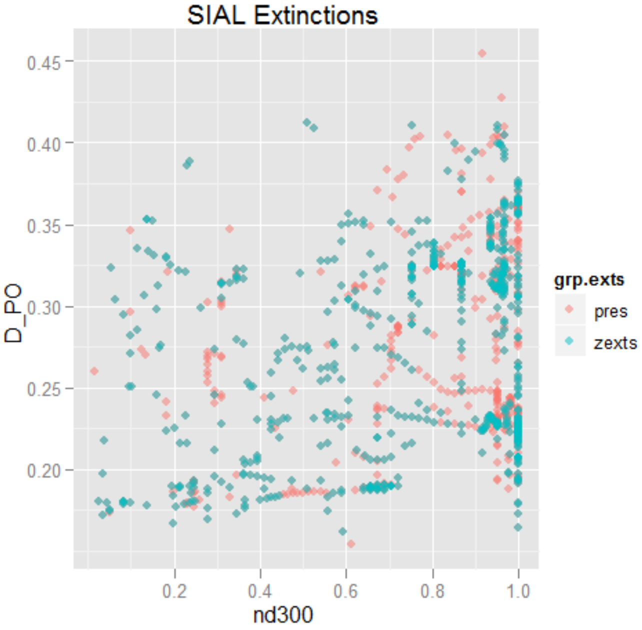


Figure C.10. Raw data scatters of colonization and extinction for *Sisymbrium altissimum*, 2009 to 2010 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2009 that went extinct in 2010 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis (‘D_PO’) and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis (‘nd300’). The right panel shows the absences in 2009 in red and the blue points are the absences in 2009 that became present in 2010 (colonizations). These points are arrayed on the same two axes as the left panel.

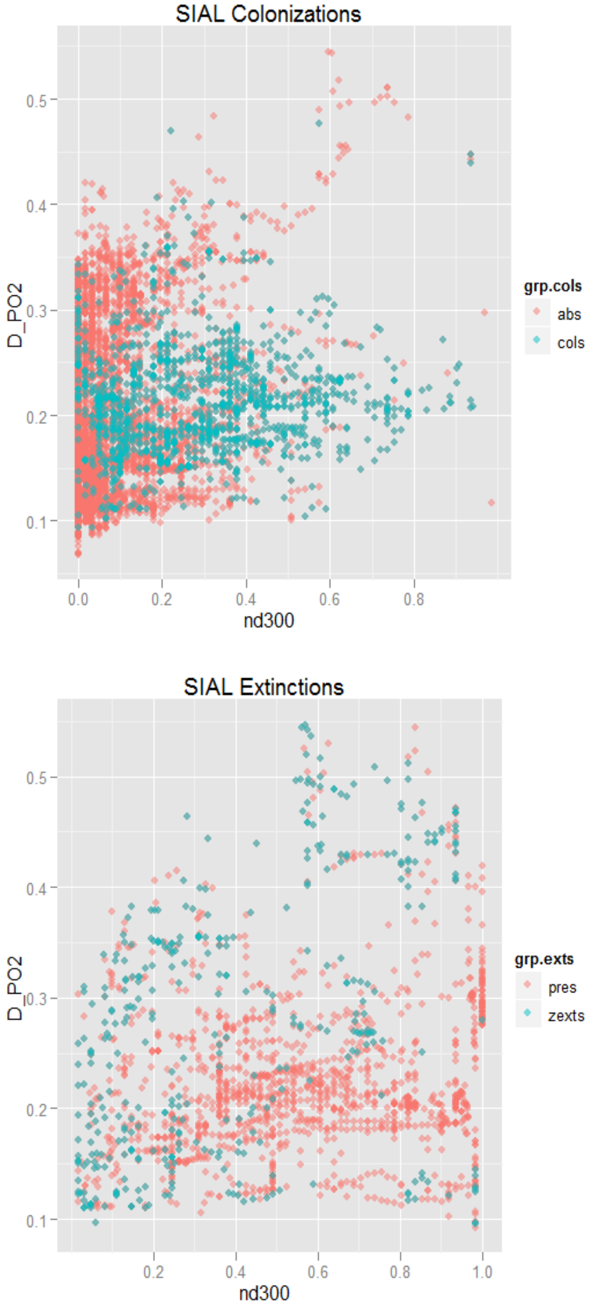


Figure C.11. Raw data scatters of colonization and extinction for *Sisymbrium altissimum*, 2010 to 2011 transition. The left panel shows the presences in 2009 in red and the blue points are the presences in 2010 that went extinct in 2011 (extinctions). The points are arrayed based on two independent variables, habitat suitability on the y-axis ('D_PO2') and occupied neighborhood density as calculated using a 300 m neighborhood size on the x-axis ('nd300'). The right panel shows the absences in 2010 in red and the blue points are the absences in 2010 that became present in 2011 (colonizations). These points are arrayed on the same two axes as the left panel.

APPENDIX D:

SUPPLEMENTAL COLONIZATION AND EXTINCTION RESPONSE SURFACES (CHAPTER 4)

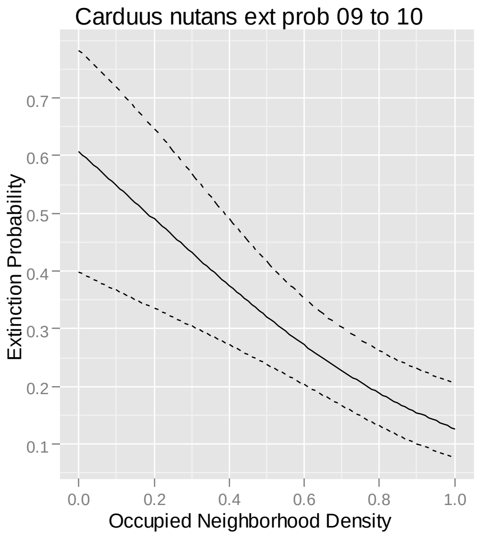


Figure D.1. Extinction probability from 2009 to 2010 for *Carduus nutans* as a function of occupied neighborhood density. Solid line is the estimated mean extinction probability and the dashed lines are the 95% prediction interval about the mean.

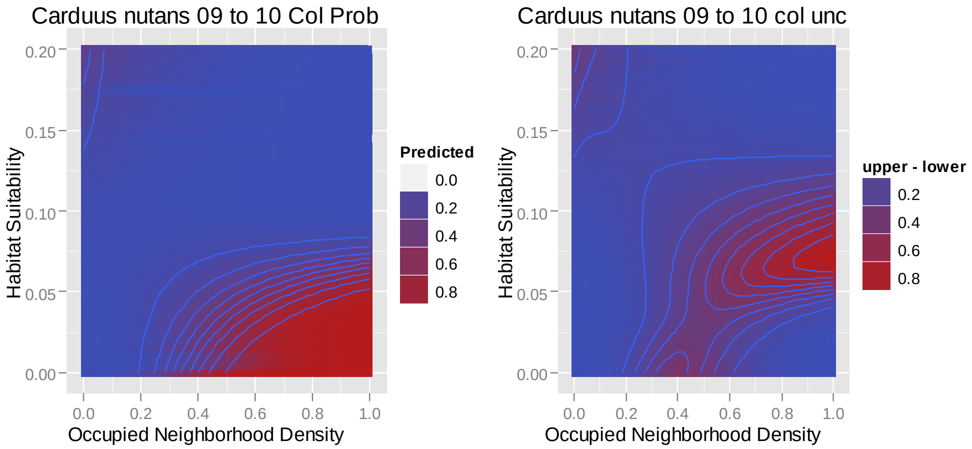


Figure D.2. Predicted colonization probability for *Carduus nutans* from 2009 to 2010 and associated uncertainty. Predictions are arrayed on two independent variables, occupied neighborhood density on the horizontal axis and habitat suitability on the vertical axis. The colors indicate low (blue) to high (red) probability in the left panel while blue colors in the right panel indicate narrow prediction intervals (greater certainty) and red indicates wider confidence intervals (greater uncertainty). Contours are isolines of either the probability or uncertainty.

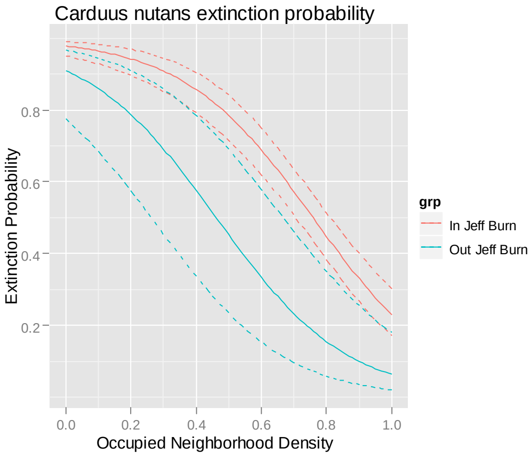


Figure D.3. Predicted extinction probability for *Carduus nutans* from 2010 to 2011. Solid lines indicate the predicted extinction probability as a function of occupied neighborhood density while the dashed lines are the 95% prediction intervals. The colors indicate the effect inside (red) and outside (blue) the Jefferson fire.

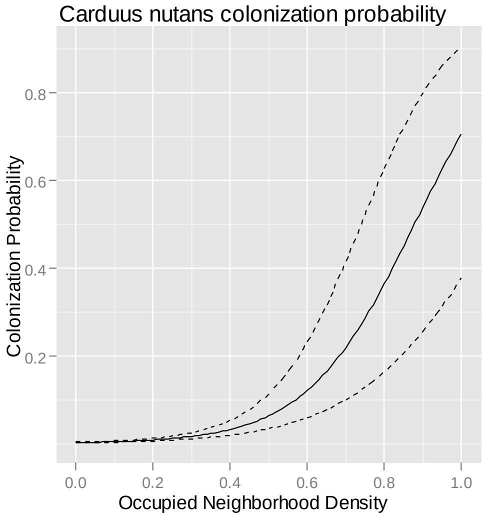


Figure D.4. Predicted colonization probability for *Carduus nutans* from 2010 to 2011 as a function of occupied neighborhood density. The 95% prediction intervals are the dashed lines.

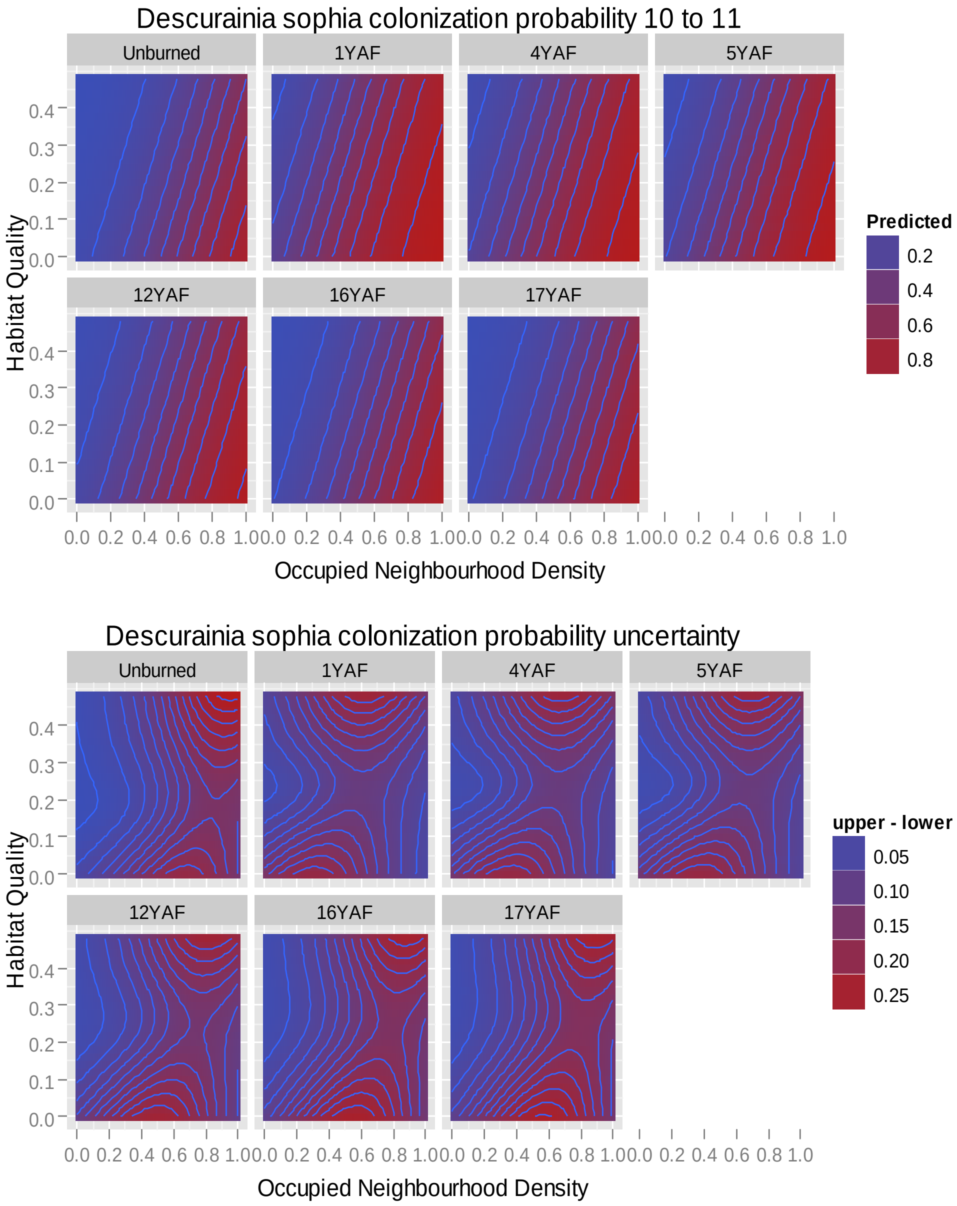
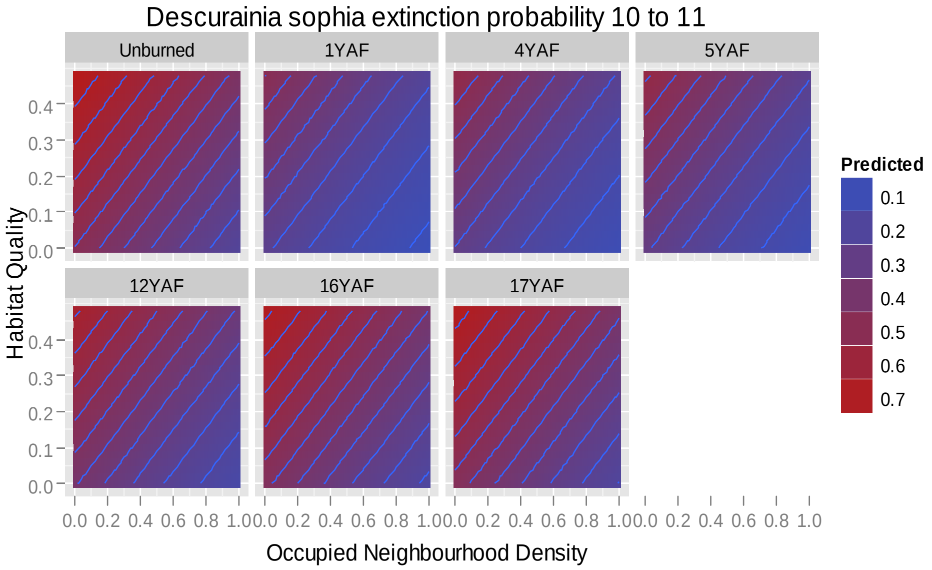


Figure D.5. Predicted colonization probability for *Descurainia sophia* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, habitat suitability on the vertical axis and occupied neighborhood density on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The seven panels indicate burn status. The first panel indicates unburned while the others range from one year after fire (1YAF) to 17 years after fire (17YAF).



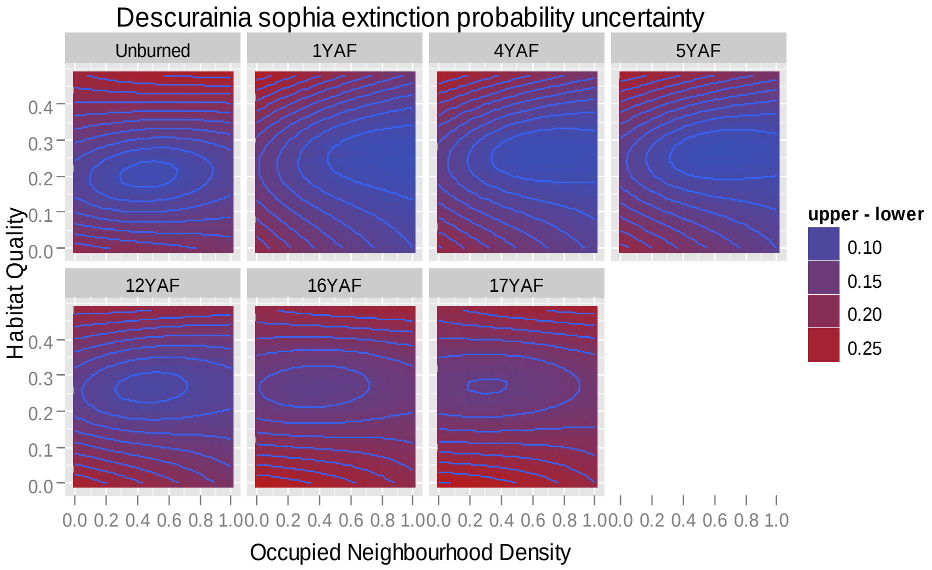


Figure D.6. Predicted extinction probability for *Descurainia sophia* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, habitat suitability on the vertical axis and occupied neighborhood density on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The seven panels indicate burn status. The first panel indicates unburned while the others range from one year after fire (1YAF) to 17 years after fire (17YAF).

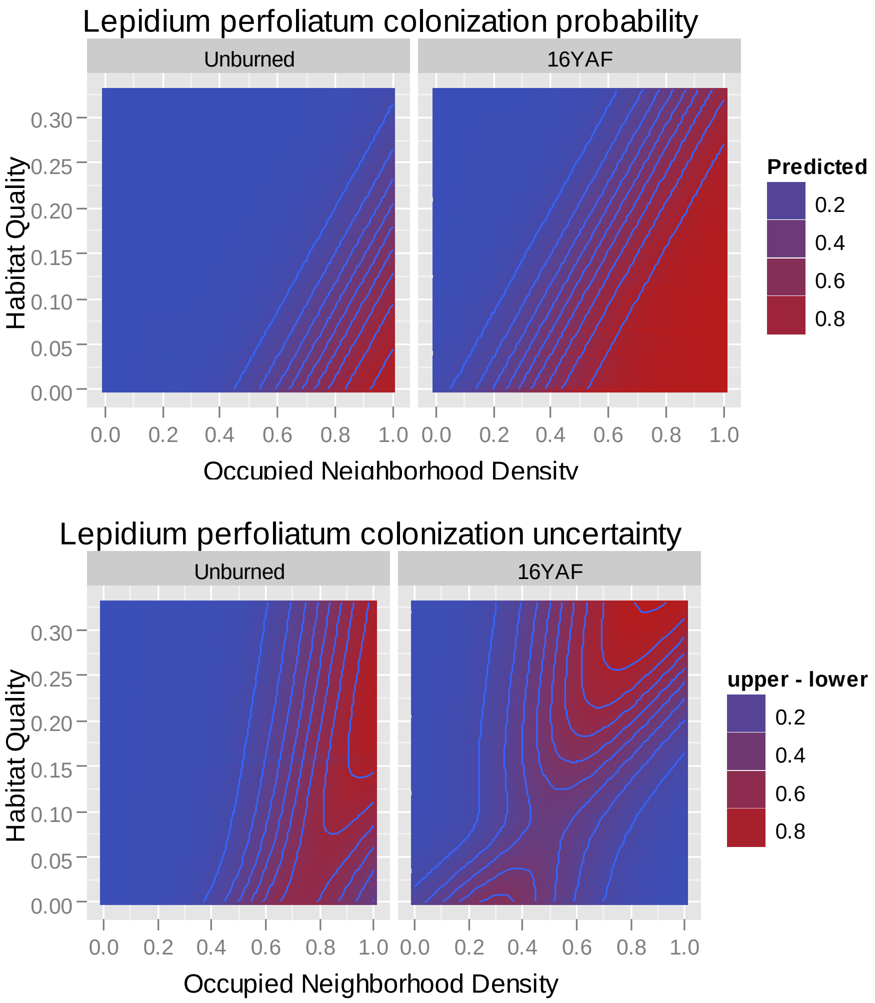


Figure D.7. Predicted colonization probability for *Lepidium perfoliatum* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, habitat suitability on the vertical axis and occupied neighborhood density on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The two panels indicate unburned (left) and inside the 16 year old fire.

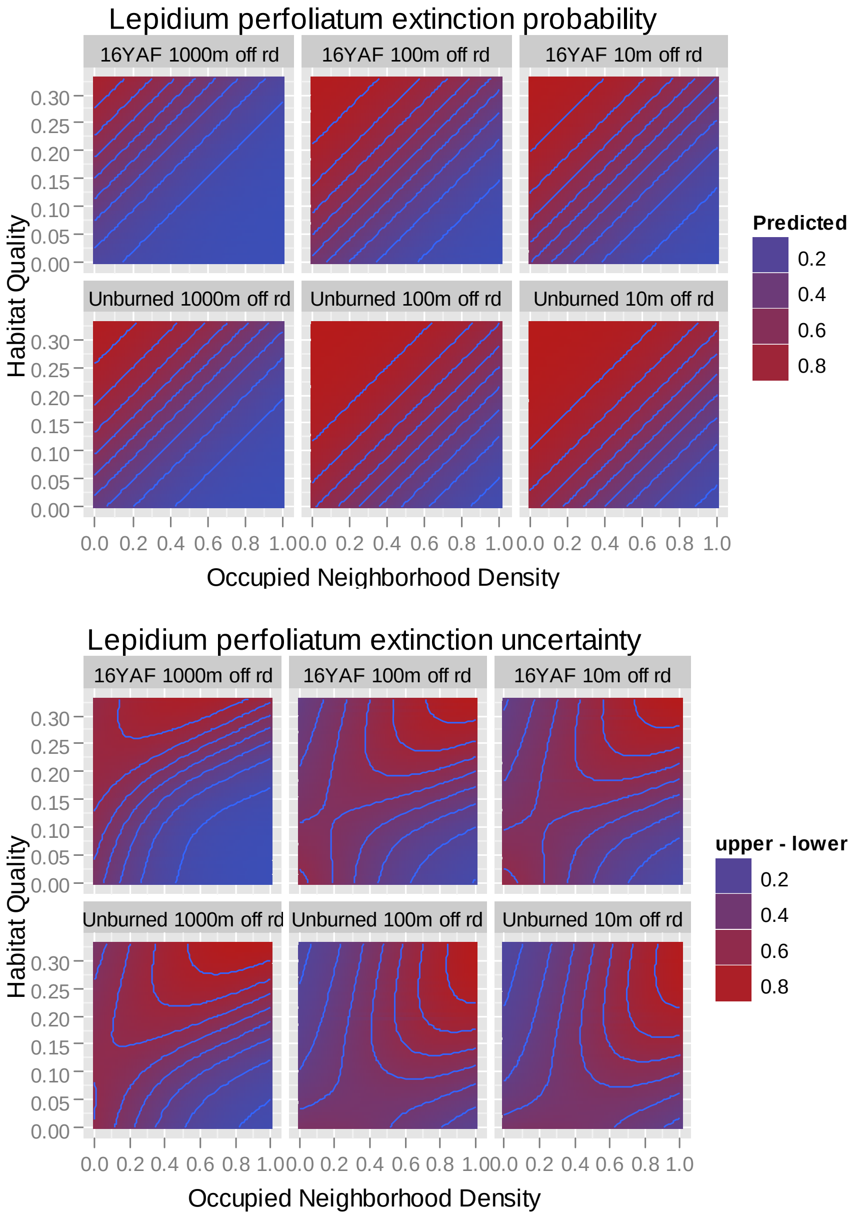


Figure D.8. Predicted extinction probability for *Lepidium perfoliatum* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, habitat suitability on the vertical axis and occupied neighborhood density on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The top three panels are the predictions inside the 16 year old fire at three distances from the paved road and the bottom three panels are the predictions in unburned areas at three distances from the paved road.

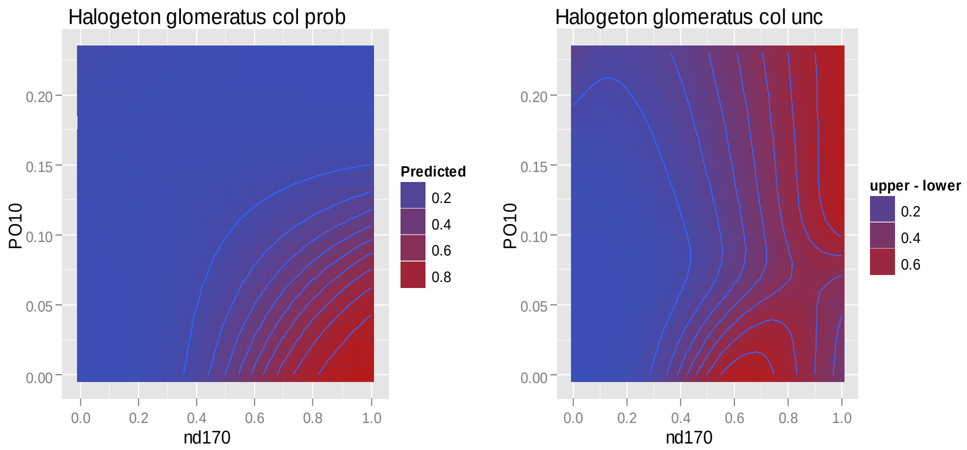


Figure D.9. Predicted colonization probability for *Halogeton glomeratus* from 2010 to 2011 and associated uncertainty. Predictions are arrayed on two independent variables, occupied neighborhood density (nd170) on the horizontal axis and habitat suitability (PO10) on the vertical axis. The colors indicate low (blue) to high (red) probability in the left panel while blue colors in the right panel indicate narrow prediction intervals (greater certainty) and red indicates wider confidence intervals (greater uncertainty). Contours are isolines of either the probability or uncertainty.

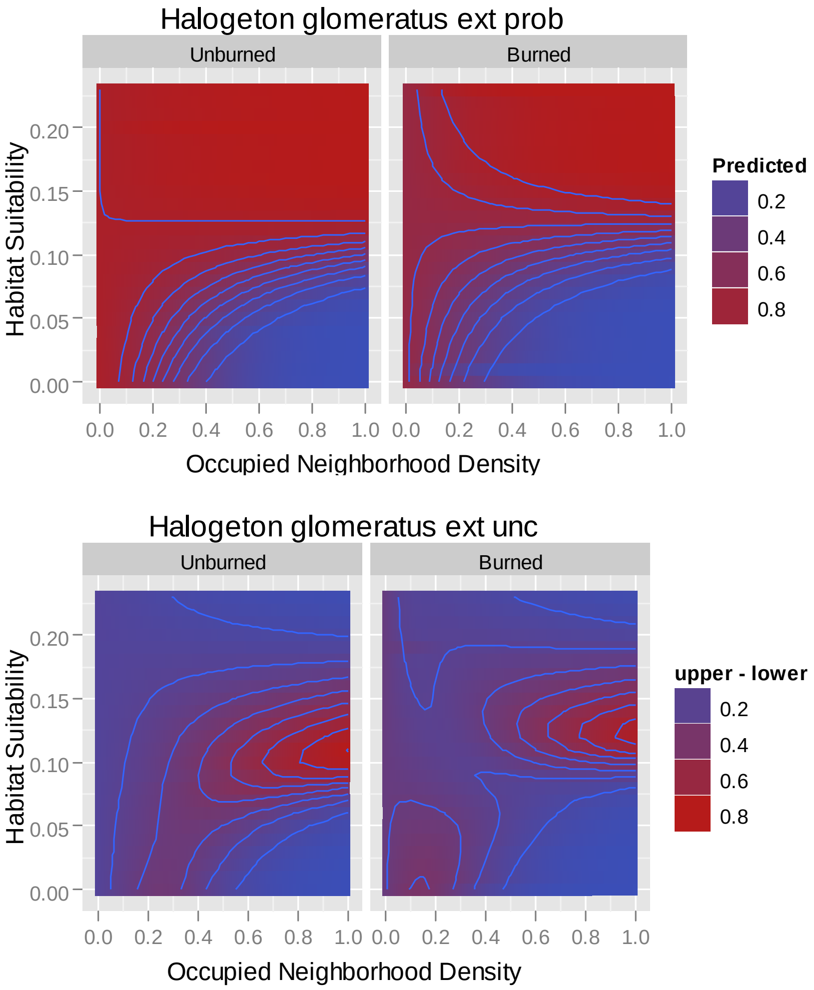


Figure D.10. Predicted extinction probability for *Halogeton glomeratus* from 2010 to 2011 and associated uncertainty. Predictions are arrayed on two independent variables, occupied neighborhood density on the horizontal axis and habitat suitability on the vertical axis. The colors indicate low (blue) to high (red) probability in the top two panels while blue colors in the bottom two panels indicate narrow prediction intervals (greater certainty) and red indicates wider confidence intervals (greater uncertainty). Contours are isolines of either the probability or uncertainty. The left panels are the predictions for unburned sites while the right panels are predictions for burned sites.

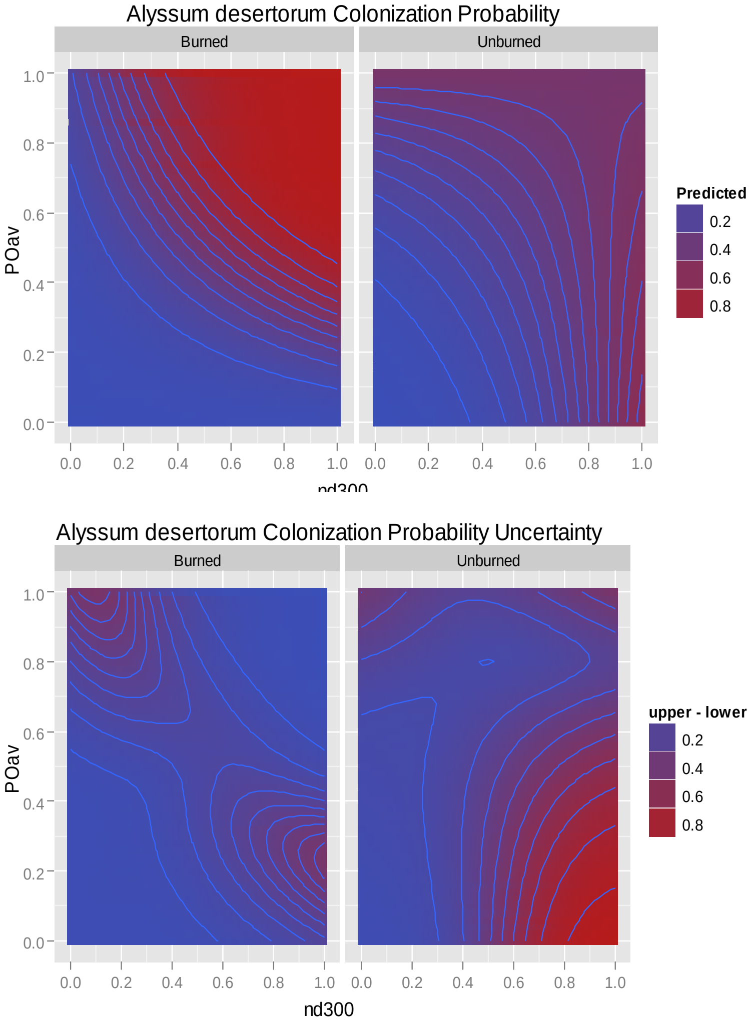


Figure D.11. Predicted colonization probability for *Alyssum desertorum* from 2010 to 2011 and associated uncertainty. Predictions are arrayed on two independent variables, occupied neighborhood density (nd300) on the horizontal axis and habitat suitability (POav) on the vertical axis. The colors indicate low (blue) to high (red) probability in the top two panels while blue colors in the bottom two panels indicate narrow prediction intervals (greater certainty) and red indicates wider confidence intervals (greater uncertainty). Contours are isolines of either the probability or uncertainty. The left panels are the predictions for unburned sites while the right panels are predictions for burned sites.

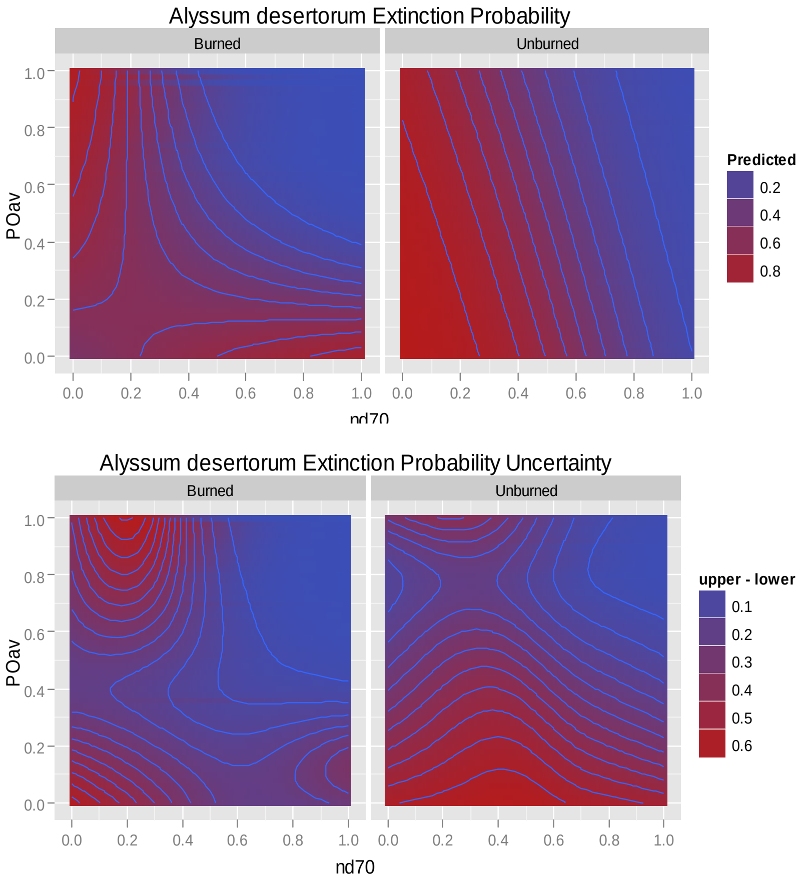


Figure D.12. Predicted extinction probability for *Alyssum desertorum* from 2010 to 2011 and associated uncertainty. Predictions are arrayed on two independent variables, occupied neighborhood density (nd300) on the horizontal axis and habitat suitability (POav) on the vertical axis. The colors indicate low (blue) to high (red) probability in the top two panels while blue colors in the bottom two panels indicate narrow prediction intervals (greater certainty) and red indicates wider confidence intervals (greater uncertainty). Contours are isolines of either the probability or uncertainty. The left panels are the predictions for unburned sites while the right panels are predictions for burned sites.

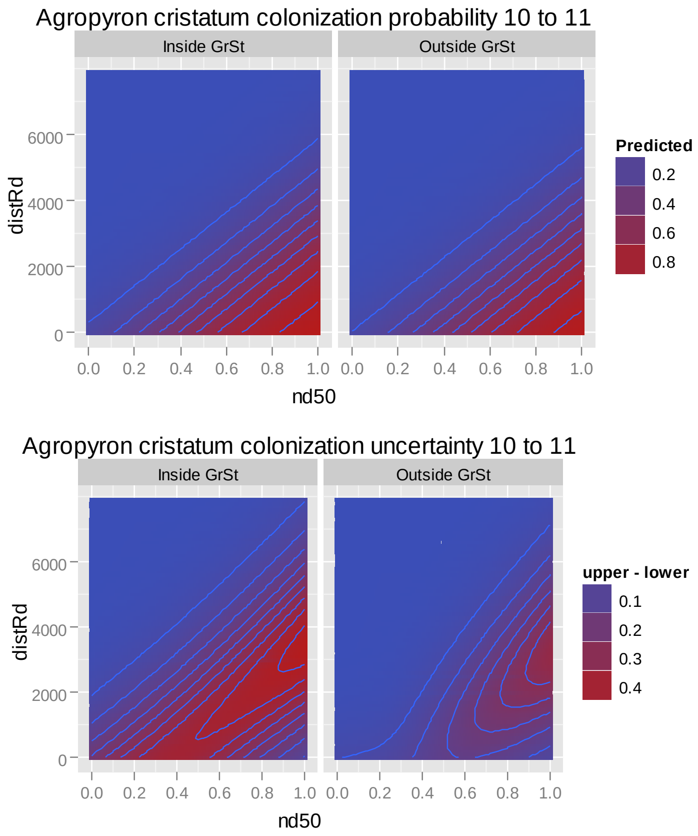


Figure D.13. Predicted colonization probability for *Agropyron cristatum* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, distance to paved road (distRd) on the vertical axis and occupied neighborhood density (nd50) on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The two panels indicate differences between inside the greens strip (GrSt) and outside the green strip. A green strip is an area where this species was drill seeded as a management intervention.

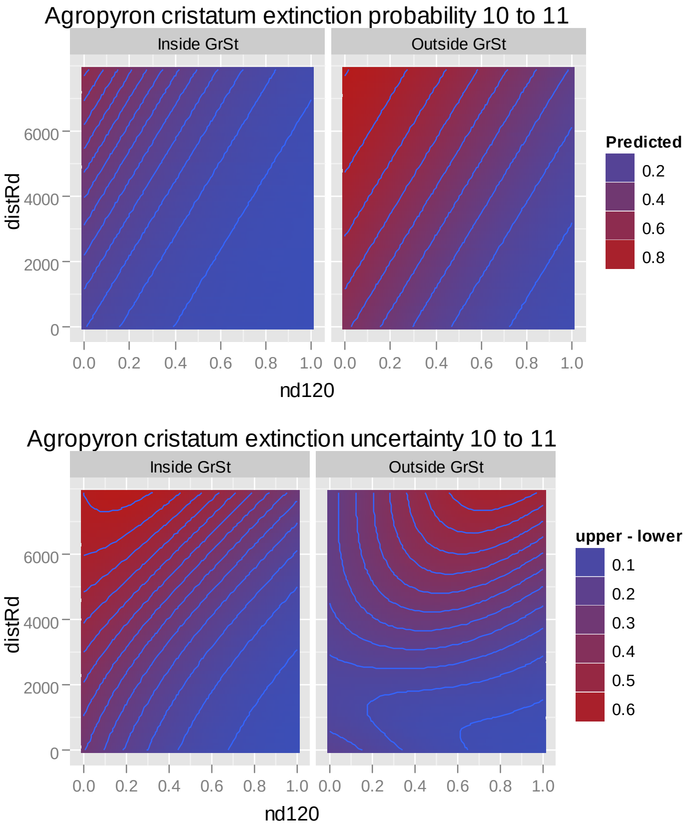


Figure D.14. Predicted extinction probability for *Agropyron cristatum* from 2010 to 2011 and associated uncertainty. Predictions and uncertainty are arrayed on two independent variables, distance to paved road (distRd) on the vertical axis and occupied neighborhood density (120) on the horizontal axis. Probabilities and uncertainty range from low (blue) to high (red) and the light blue lines are isolines of the probabilities or confidence interval widths. Uncertainty was defined as the width of the 95% prediction confidence interval. The two panels indicate differences between inside the greens strip (GrSt) and outside the green strip. A green strip is an area where this species was drill seeded as a management intervention.