PREDICTIVE INVASION ECOLOGY AND DECISIONS UNDER UNCERTAINTY

Corey Chivers

Doctor of Philosophy

Department of Biology

McGill University

Montréal,Québec

2014-07-15

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

DEDICATION

This work is dedicated to my loving family. For my parents, Barb and Barry Chivers, who it seems to me that being a nurse and an accountant, respectively, have brought me up destined to have an interest in both biology and numbers. For my sisters Ashley and Brittaney who always hold me to finishing what I've started. And for my amazing spouse, best friend, and life partner Kate Cairns, without whom I would not have believed this to be possible.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Brian Leung, for giving me his confidence and for treating me as much as a colleague as a student right from the very beginning. I owe so much of what I have learned during the course of this research to Brian, but even more importantly he has fundamentally changed the way that I think about the world, and this way of thinking will undoubtedly continue to impact the way I learn, grow, and approach new problems for the rest of my life.

I am grateful to my academic supervisors Elena Bennett, Gregor Fussmann, and Claire De Mazancourt for their expert guidance throughout this endeavour. My co-author on one of the chapters of this thesis, Dr. Norman Yan, also deserves thanks here for his help and his considerable knowledge of freshwater flora and fauna.

My lab mates Erin Gertzen, Dominique Roche, Ayaz Hyder, Lidia Della Venezia, Kristina Marie Enciso, Johanna Bradie, and Dylan Schneider have all been both excellent colleagues and dependable friends, often providing perspectives on the work that I would never have had access to without them. A very special thank you is reserved for Paul Edwards. Paul has been my scholastic comrade since our undergraduate days and I couldn't have asked for a better friend,colleague, and geek-out partner. I'd like to thank lab assistant Jeremy Proville for his hard work with survey design and data processing.

The administrative staff in the Biology department are deserving of many thanks for their tireless behind the scenes support of the research and teaching operations on a day to day basis. In particular I would like to thank Linda Morai, who was always available to help navigate any and every administrative need. Linda passed away before this thesis was completed; she will be missed.

I am grateful for the generous funding provided by the Natural Sciences and Engineering Research Council (NSERC) through the Canadian Aquatic Invasive Species Network (CAISN), the Quebec Center for Biodiversity Science (CSBQ) and the Biology department of McGill University for further funding. Many of the more computationally heavy simulations presented herein were made possible by the high performance computing resources at Compute Canada.

I would also like to thank Dr. Martin Krkosek for graciously hosting me as a visiting scholar in his lab at the University of Toronto during my final semesters. Members of the Krkosek lab Luke Rogers, Andrew Bateman, Lindsay Ogston, and Pepijn Luijckx provided helpful support and made for particularly good company during my time there.

Additional thanks are due to many anonymous reviewers for their insightful comments that have strengthened the manuscripts contained herein.

ABSTRACT

This thesis focuses on the development of methods for forecasting and managing the spread of non-native species. By combining statistical modelling and computational simulations with both biological and sociological data, this research aims to provide decision support tools to resource managers and policy makers. With an emphasis on the quantification and propagation of uncertainty through the construction of both classical and Bayesian models, I analyse the implications of various human and biological factors on forecasting the spread of fresh water invasive species. These include: 1) dispersal network structure, 2) population dynamics, 3) environmental suitability, and 4) human behavioural feedbacks to policy interventions.

The first section compares two current approaches to predicting the secondary spread of aquatic invasive species and introduces a novel methodology for the quantitative validation of such predictions. Chapter 2 compares alternative models of human-mediated dispersal and assesses the consequences of the resulting dispersal network structures for predictions of invasion risk at both the local and landscape level. A new approach to validating the predictions made by models of spreading species is developed in Chapter 3. The new approach accommodates both stochastic and epistemic prediction uncertainty and I demonstrate that it has both the appropriate expected error rates as well as increased power compared to existing methods. Application to a published forecast model of *Bythotrephes longimanus* in central Ontario confirms the predicted invasion pattern.

The second section deals with the development and application of new forecasting and management models which are applicable in common situations of limited data availability and limited management resources. Chapter 4 solves problems posed by presence-only data by extending current approaches to species distribution modelling using an observation model of the detection process. Application of this approach to 10 aquatic invasive species in Ontario revealed that the number of sites at which species are detected is not alone predictive of their current and potential range. By quantifying between-species differences in prevalence and detectability, this approach can provide guidance for sampling efforts and management interventions. Finally, Chapter 5 addresses the predicted efficacy of specific management interventions by modelling the human behavioural responses to such interventions. By integrating behavioural responses into a gravity model formulation, the predicted consequences of various policy scenarios on the future spread of aquatic invasives in Ontario is compared. Together, this research provides novel insights into both ecological processes and environmental policy.

ABRÉGÉ

Cette thèse repose sur le développement de méthodes visant à prédire et gérer la propagation des espèces non-indigènes. Combinant modélisation statistique et simulations informatiques sur des données biologiques et sociologiques, cette recherche vise à fournir une aide à la décision aux décideurs et gestionnaires de ressources. En mettant l'accent sur la quantication et la propagation de l'incertitude par la construction de deux modèles classique et Bayesian, j'analyse les implications de différents facteurs humains et biologiques pour la prédiction de la propagation des espèces envahissantes d'eau douce. Il s'agit notamment de: 1) la structure du réseau de dispersion, 2) les dynamiques de population, 3)la qualité de l'environnement, et 4) la réponse aux interventions politiques.

La première section compare deux approches actuellement employées pour prédire la propagation secondaire d'espèces aquatiques envahissantes et introduit une nouvelle méthodologie de validation quantitative de ces prédictions. Le chapitre 2 compare d'autres modèles de dispersion par médiation humaine et évalue les conséquences des structures de réseau de dispersion obtenues pour les prédictions de risques d'invasion, à la fois au niveau local et à l'échelle du paysage. Le chapitre 3 développe une nouvelle approche de validation des prédictions issues de modèles de propagation. Cette nouvelle approche intègre les incertitudes de prédiction stochastique et épistémique et je démontre qu'elle conduit aux taux d'erreur attendus et est plus puissante que les méthodes existantes. Appliquée à un modèle de prévision publié de *Bythoterphes*

longimanus dans le Centre de l'Ontario, cette approche confirme le schéma d'invasion prédit.

La deuxième section porte sur le développement et l'application de nouveaux modèles de prévision et de gestion utilisables dans les limites ordinaires de disponibilité de données et de ressources de gestion. Le chapitre 4 résout les problèmes associés aux données signalant seulement la présence d'une espèce et élargit le champ des approches actuelles de modélisation de distribution d'espèces parl'utilisation d'un modèle d'observation du processus de détection. L'application de cette approche sur 10 espèces aquatiques envahissantes en Ontario révèle que l'usage seul du nombre de sites où les espèces sont détectées ne permet pas de prédire leur distribution actuelle et potentielle. En quantiant les différences inter-spécifiques de prévalence et de détection, cette approche peut aider au développement de méthodes d'échantillonnage et à la mise en place d'interventions de gestion. En dernier lieu, le chapitre 5 traite de l'efficacité prédite d'interventions de gestion spécifiques en modélisant les réponses comportementales des individus à ces interventions. Intégrant les réponses comportementales dans une formule de modèle de gravité, je compare les effets prédits de différents scenarios d'intervention sur la future propagation des espèces envahissantes aquatiques en Ontario. Cette recherche ouvre de nouvelles perspectives aussi bien sur les processus écologiques que les politiques de gestion de l'environnement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Preface

0.1 Thesis Format and Style

This thesis is presented in a manuscript-based format and consists of four papers, each addressing a different aspect of the prediction and management of spreading invasive species.

Chapter 2: Chivers, C. & Leung, B (2012) Predicting invasions: Alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects. *Journal of Applied Ecology* **49**:1113-1123.

Chapter 3: Chivers, C., Leung, B., Yan, N. (in revision to be published in *Methods in Ecology and Evolution*) Validation and calibration of probabilistic predictions in ecology.

Chapter 4: Chivers, C. & Leung, B. (submitted to *Ecology Letters*) The probability of establishment and spread of biological organisms: Presence-only data, hidden states, and the observation process.

Chapter 5: Chivers, C. & Leung, B. (in prep.) Modelling responses to management intervention for controlling the spread of freshwater invasives.

0.2 Contribution of Co-Authors

This thesis is comprised of my original work. I am the primary author for all, in close collaboration with my supervisor, Brian Leung. I constructed the models, analysed the data, and wrote the manuscripts herein. One of the manuscripts has greatly benefited from the contribution of one additional co-author, which is stated below:

Chapter 2

Authors: *Corey Chivers and Brian Leung.* Corey Chivers designed and conducted data collection, built and analysed the models. Corey Chivers wrote the manuscript with input from Brian Leung.

Chapter 3

Authors: *Corey Chivers, Brian Leung, and Norman Yan.* Corey Chivers and Brian Leung conceived the model, Corey Chivers built and analysed the model, analysed the case study and wrote co-wrote the manuscript with Brian Leung and Norman Yan, who provided the data analysed in the case study.

Chapter 4

Authors: *Corey Chivers and Brian Leung.* Corey Chivers and Brian Leung conceived the methodology. Corey Chivers built and tested the theoretical model. Corey Chivers collected the data, built, tested, and analysed the applied model. Corey Chivers wrote the manuscript with input from Brian Leung.

Chapter 5

Authors: *Corey Chivers and Brian Leung.* Corey Chivers and Brian Leung conceived the methodology. Corey Chivers built and tested the theoretical model. Corey

Chivers collected the data, built, tested, and analysed the applied model. Corey Chivers wrote the manuscript with input from Brian Leung.

0.3 Original Contributions to Knowledge

In Chapter 2, I evaluate two alternative formulations of human-mediated dispersal of aquatic invasive species and investigate the interaction between the resulting dispersal networks and the population dynamics of the invading species. The chapter adds to the field of secondary spread prediction by describing the consequences of alternative model formulations on predictions of invasion risks at both the landscape and local levels. This study represents the first comparison of these alternative human-mediated dispersal models using empirical data and is the first to demonstrate the importance of the interaction between the resulting dispersal networks and Allee effects.

I present a novel approach to evaluating and validating the calibration of binary prediction models in Chapter 3. While applicable to the present focus of spreading invasive species, the Validation Metric Applied to Probabilistic Predictions (VMAPP) presented in this chapter represents a general advancement in the field of predictive model validation and is applicable to a wide array of ecological prediction problems including Population Viability Analysis, general Species Distribution Modelling (SDM), metapolulation analysis, and any other realm of ecology in which predictions about binary outcomes is prevalent. The results in this chapter suggest that in many realistic situations, VMAPP is more powerful and less prone to error than previously available methods.

In Chapter 4, I develop a novel modelling framework for building Species Distribution Models (SDMs) when the standard assumptions made by the most common currently employed methods are violated and using this framework I provide a solution to the common problem of presence-only data. This chapter not only represents a significant contribution to the invasion literature, but also yields important results for the field of Biogeography more generally.

Finally, Chapter 5 contributes important insights into the consequences of behavioural feedbacks which may result due to management interventions aiming to mitigate the spread of aquatic invasive species. A major contribution of this chapter is a new decision support tool for managers and policy makers dealing with limited resources available for the management of invasive species.

CHAPTER 1
Introduction

1.1 Introduction

Human activities involving the large-scale movement of people and goods have brought massive benefits to human societies around the globe. As we transport ever more cargo, however, we move with it an unprecedented number of species both purposefully and inadvertently. The geographic extent of these biotic displacements are far beyond that which was previously possible, and are occurring at rates which far outstrip the ecological and evolutionary time-scales on which the processes of natural dispersal operate (Ricciardi, 2007). This effective removal of long-standing biogeographical barriers to dispersal represents a major driver of global change (Vitousek et al., 1997). The resulting global-scale 'biotic homogenization' (Olden et al., 2004; Rahel, 2007) has had considerable ecological (Parker et al., 1999; Occhipinti-Ambrogi and Savini, 2003; Landis, 2004; Dextrase and Mandrak, 2006) and economic consequences (Pimentel et al., 2005; Pejchar and Mooney, 2009; Aukema et al., 2011).

In one instance – freshwater invasives in the North American Great Lakes region – new species have been arriving and establishing reproducing populations at an accelerating rate over the last century (Ricciardi, 2006). Many of the established invasives which have arrived over that last 50 years are believed to have been introduced via ballast water exchange undertaken by international cargo ships (Keller et al., 2011), although the opening of a major canal (Ricciardi, 2006), as well as the aquarium (Padilla and Williams, 2004; Semmens et al., 2004; Duggan et al., 2006; Gertzen et al., 2008; Strecker et al., 2011; Bradie et al., 2013) and ornamental garden trades (Kay and Hoyle, 2001; Dehnen-Schmutz et al., 2007) have also been implicated.

While many of these introduced freshwater species have been relatively innocuous, some have had substantial economic and ecological impacts. These impacts may arise in the form of direct detrimental effects due to increased competition and predation pressure on native flora and fauna (Dextrase and Mandrak, 2006). For example the invasive predatory zooplankton *Bythotrephes longimanus*, native to the Ponto-Caspian region and introduced to the Great Lakes in the early 1980's, induces major changes in the structure of existing food webs, and may have caused multiple extirpations of native zooplankton species in inland lakes in and around the Great Lakes basin (Bourdeau et al., 2011; Yan et al., 2011). Other ecological impacts may arise due to habitat alterations induced by the introduced species. For instance, the invasive carp *Cyprinus carpio* is known to increase turbidity and degrade water quality, reducing macrophyte and benthic invertebrate densities (Kulhanek et al., 2011). In addition to – and sometimes resulting from – the ecological impacts, some introduced freshwater species have been documented to have adverse impacts on ecosystem services. The largest impacts are likely borne by commercial and recreational fisheries, as species such as the parasitic sea lamprey (*Petromyzon marinus*) reduce the population growth rates of many game fish species (Lupi et al., 2003; Rothlisberger et al., 2012). In addition to the considerable realised and potential economic impacts of freshwater invasives on the fisheries industry, other segments of the economy are also affected. Two examples include zebra mussels (*Dreissena polymorpha*) which have direct impacts on power plants and other shoreline industrial operations by fouling water intake pipes (Leung et al., 2002), and the diminishing by up to 16% of lakefront residential property values as a result of Eurasian watermilfoil (*Myriophyllum Spicatum*) infestation (Zhang and Boyle, 2010).

As the potential for harm is quite high, much research has been devoted to developing forecast models and risk assessments for non-native species (Stohlgren and Schnase, 2006; Keller and Lodge, 2007). While some assessments focus on evaluating risks posed by individual target species or specific areas of potential impact (Schneider et al., 1998; Landis, 2004; Chan et al., 2012), there has been a growing recognition that integrative approaches to assessing risk for multiple species at each level of the invasion process can provide more comprehensive, and actionable insights (Leung et al., 2012; Ibáñez et al., 2014). Ultimately, the process of undertaking a risk assessment is to inform management decisions. Possible management actions for invasive species will depend on the stage of the invasion in question and can include measures to prevent the introduction of potential invaders in the first place (Jerde and Lewis, 2007; Meyerson and Mooney, 2007; Bailey et al., 2011), monitoring to improve early detection and eradication of newly established populations (Edwards and Leung, 2009; Keith and Spring, 2013), and/or efforts to limit the impact of established species by halting or reducing the rate at which species spread in the introduced range (Vander Zanden and Olden, 2008; Rothlisberger et al., 2010; Epanchin-Niell and Hastings, 2010). In this thesis, I focus on the forecasting and management of biological invasions in this latter stage, referred to as 'secondary spread'.

1.2 Secondary spread

Conceptually, the process of biological invasion can be divided into a series of stages. At each stage there are barriers, both physical and biological, that need to be overcome for a species or population to continue to the next stage (Colautti and MacIsaac, 2004; Blackburn et al., 2011). In the most general formulation of this conceptual framework, the stages are: 1) Transport, 2) Introduction, 3) Establishment, and 4) Spread. At the secondary, or landscape level, where the focus of the research in this thesis is placed, the final stage of this process can itself be conceptualized as a microcosm of the entire framework, operating at a landscape scale.

In freshwater invasions, natural dispersal is typically constrained (on an ecological time-scale) within the boundaries of a given body of water (Gertzen and Leung, 2011), yet there are additional transport vectors operating at the landscape scale, facilitating the spread of species to inland lakes following the initial introduction and establishment at some location(s) within a region. The primary mode of dispersal across a landscape of inland lakes has repeatedly been found to be human-mediated via transient recreational boaters (Bossenbroek et al., 2001; MacIsaac et al., 2004; Muirhead and Macisaac, 2005; Kerr et al., 2005; Timar and Phaneuf, 2009). As such, aquatic organisms spread among hydrologically disconnected lakes via occasional discrete jumps from invaded to uninvaded sites. These newly colonized locations then form 'disjunct foci' which subsequently become a population source emanating propagules which can proceed to colonize additional locations (Moody and Mack, 1988). This pattern stands in contrast to the continuously radiating wavefront associated with some range expansions, where the square root of the area invaded grows linearly with time (Skellam, 1951; Hastings, 1996). Instead, the range expansion of aquatic invasives is characterized the stochastic colonization of discrete patches of suitable habitat (lakes) in a heterogeneous landscape (Hastings et al., 2005).

As human actors are the primary dispersal vector for aquatic invasive species, modelling dispersal requires modelling human behavioural patterns. To this end, approaches adopted from the human demography and transportation literature called Gravity Models (GMs) have been widely employed (Bossenbroek et al., 2001; MacIsaac et al., 2004; Leung et al., 2004, 2006; Gertzen and Leung, 2011). While there are various forms, in general these models characterise the movement of recreational boaters by describing individual lakes as having 'pull' relative to others based on their size, distance from population centres, and any other factors which may influence a boater's decision to visit a given lake. More recently, an alternative family of models called Random Utility Models (RUMs), which originated in the econometrics literature to model recreational demand (Smith and Kaoru, 1986), have also been applied to modelling human-mediated dispersal (MacPherson et al., 2006; Timar and Phaneuf, 2009). RUMs are derived from a different set of assumptions than GMs, but can be reduced to alternative functional forms used to describe the probability distributions of individual boaters' trip choices (see Chapter 2). Once the behaviour of individual boaters has been characterized, the strength of the dispersal corridors between lakes can be easily estimated by aggregating all of the individual trips.

Forecasting the secondary spread of aquatic invasive species requires linking models of dispersal, which describe the the number of individuals arriving per unit time (eg. propagule pressure), with models describing how propagule pressure relates to the probability of a species establishing a reproducing population, and therefore going on to become a source of new propagules dispersing out to new, uncolonized locations. There are several challenges associated with building these joint models and linking them with data. First, dispersal and establishment are both stochastic processes and rare events, making prediction difficult, especially during the early stages of spread (Smith et al., 1999). Second, our understanding of the processes involved is limited. While reality is high dimensional, our models are simplified abstractions of the real world and are therefore necessarily incomplete (Hilborn and Mangel, 1997). And finally, we are often dealing with incomplete data characterised by low and often patchy spatio-temporal resolution (Leung and Delaney, 2006; Ibáñez et al., 2014).

1.3 Why uncertainty matters

When forecasting spread, we are making predictions about future outcomes (states of nature), and any such predictions will inherently have uncertainty associated with them (Clark et al., 2001). In order to inform sound management decisions, it is important to acknowledge, attempt to quantify, and communicate prediction uncertainty in it's various forms (Elith et al., 2002; Beale and Lennon, 2012; Yemshanov et al., 2013). As it comprises the basis of decision theory, throughout this thesis I take a probabilistic approach to uncertainty wherever possible by describing uncertainty about potential future outcomes as probability distributions. In resource and ecological management, as in many other decision making domains, outcomes are neither equally likely, nor equally valuable and thus the decision maker wishes to choose an action from the set of alternatives which will maximise the expected value, or net benefit, associated with the set of possible outcomes. The expected value is simply the sum of the net benefits associated with each possible outcome, weighted by the conditional probability of each outcome given the action taken by the decision maker (Polasky et al., 2011). This forms the standard objective of decision theory: to choose the alternative which maximizes the expected (or average) net benefit, given that outcomes themselves are uncertain. By representing potential outcomes as probability distributions, decision theory provides the mechanisms for making optimal decisions subject to imperfect information (Dorazio and Johnson, 2003).

1.3.1 Jensen's inequality

Of particular importance to the incorporation of various sources and types of uncertainty into our models of species spread is Jensen's inequality (Ruel and Ayres, 1999; Bolnick et al., 2011). Jensen's inequality is a mathematical property of non-linear functions. It states simply that for a non-linear function $f(x)$, and a sample of x values, the expected value of the function is not equal to the function evaluated at the expected value of x, or $E[f(x)] \neq f(E[x])$. The direction of this inequality depends on the shape of $f(x)$. Specifically, if the function is concave ($f''(x) > 0$) then $E[f(x)] > f(E[x])$, and if it is convex ($f''(x) < 0$) then $E[f(x)] < f(E[x])$ (See Fig. 1–1).

This inequality can have several consequences for the prediction and management of spreading invasions. As a simple but central example, the species establishment probability as a function of propagule pressure is itself a non-linear function. In the case where establishment is not a density dependent phenomenon and each additional propagule contributes independently to the probability of establishment (e.g. no Allee effect (Stephens et al., 1999)), then the curve is convex and the consequence of uncertainty (variability) in propagule pressure will have a known directional effect. However, in the presence of an Allee effect the probability of establishment curve exhibits both concave *and* convex behaviour across the range of propagule pressures (Fig. 1–2). The resulting dynamics of spreading populations will therefore be inherently dependent on both the strength of the Allee effect and the nature of the variability in the system. While I examine the effect that this has on forecasts generated using alternative models of spread explicitly in Chapter 2, this phenomenon is also at play and taken into account in the models presented in Chapters 4 and 5.

1.4 Overarching methodology

The general methodological framework that I adopt in this thesis approaches disparate problems in a unified way. Extending the concept of the 'Ecological Detective' (Hilborn and Mangel, 1997), the elements of this methodology are laid out in figure 1–3, which describes the process of going from data to decisions, making ample use of *in silico* simulations along the way. A key step in the process focuses on evaluating the inferential and predictive capacity of our models of biological phenomena by repeatedly generating observations under controlled conditions where we know the underlying reality. In so doing, we can assess the theoretical properties of our statistical approaches given the same type, quantity, and structure of data as that which we can observe in the real world. This step can also be critically important as a safeguard against faulty logic. As models of biological phenomena become more complex, incorporating more realism with many component parts, they also become increasingly difficult to reason with directly. Generating hypotheses and experimenting with model behaviour using simulation is an invaluable tool for solidifying our understanding of complex biological processes, as it allows us to reproduce, manipulate, and explore large-scale systems in ways that would be infeasible to do in the real-world (Peck, 2004). While following this methodology does not provide a guarantee that our models are 'correct', it does force us to confront and systematize our assumptions about the underlying ecological processes which have led to the patterns observed in data, and I have taken great care to do throughout this thesis.

1.5 Thesis outline

In the following chapters I investigate ecological and economic modelling approaches to the problem of predicting and managing the secondary spread of aquatic invasive species. The sources, magnitudes, and consequences of uncertainty for our ability to both understand and mitigate biological invasions are addressed throughout. Beginning in Chapter 2, I investigate the consequences of human-mediated dispersal model choice for our predictions of invasion risk at both the landscape and local level. I test two approaches to modelling human-mediated dispersal from the literature (Gravity and Random Utility models) and fit the observed behaviour of recreational boaters in Ontario. I then demonstrate how when using these alternative models, the population dynamics and density dependence effects of the invading species can interact with the structure of dispersal networks to significantly alter projected risks at both the landscape and local level. In Chapter 3, I construct a novel method for evaluating spatio-temporal risk predictions and compare this new method against two currently available approaches. I then look at a specific case study of the spread of *Bythotrephes longimanus* in a watershed in Central Ontario. In the case study I test the validity of a published model of *Bythotrephes* spread to accurately predict future invasion risks across multiple sites.

In Chapter 4, I develop a modelling framework which solves a common problem associated with the type of data which we often have available to us for predicting the current and future ranges of invasive species – presence-only records. This form of data, while ubiquitous, presents a particular challenge to modelling efforts aiming to assess the risks posed to sites across space and time. I present a novel approach for solving the unique problem posed by presence-only data which explicitly treats observation as a stochastic process of detections and non-detections. I first demonstrate the theoretical properties of this approach using simulations, then apply it to 10 species of aquatic invasive species in Ontario for which presence-only data are available. Finally, in Chapter 5 I incorporate management interventions aimed at spread-reduction into a gravity modelling framework. Using data that I collected as a part of a survey of recreational boaters in Ontario, I fit a model of behavioural responses to mandatory boat cleaning stations and then investigate the predicted efficacy of such place-based interventions.

Together, this thesis aims to deliver new techniques and modelling approaches for predicting and managing the spread of invasive species under the ubiquitous dual constraints of limited data and limited resources.

References

Aukema, J. E., B. Leung, K. Kovacs, C. Chivers, K. O. Britton, J. Englin, S. J. Frankel, R. G. Haight, T. P. Holmes, A. M. Liebhold, D. G. McCullough, and B. Von Holle, 2011. Economic Impacts of Non-Native Forest Insects in the Continental United States. *PLoS ONE* **6**:e24587.

Bailey, S. a., M. G. Deneau, L. Jean, C. J. Wiley, B. Leung, and H. J. MacIsaac, 2011. Evaluating efficacy of an environmental policy to prevent biological invasions. *Environmental science & technology* **45**:2554–61.

Beale, C. M. and J. J. Lennon, 2012. Incorporating uncertainty in predictive species distribution modelling. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**:247–58.

Blackburn, T. M., P. Pyšek, S. Bacher, J. T. Carlton, R. P. Duncan, V. Jarošík, J. R. U. Wilson, and D. M. Richardson, 2011. A proposed unified framework for biological invasions. *Trends in ecology & evolution* **26**:333–9.

Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. W. Rudolf, S. J. Schreiber, M. C. Urban, and D. a. Vasseur, 2011. Why intraspecific trait variation matters in community ecology. *Trends in ecology & evolution* **26**:183–92.

Bossenbroek, J. M., C. E. Kraft, and J. C. Nekola, 2001. Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecological Applications* **11**:1778–1788.

Bourdeau, P. E., K. L. Pangle, and S. D. Peacor, 2011. The invasive predator Bythotrephes induces changes in the vertical distribution of native copepods in

Lake Michigan. *Biological Invasions* pages 2533–2545.

Bradie, J., C. Chivers, and B. Leung, 2013. Importing risk: quantifying the propagule pressure-establishment relationship at the pathway level. *Diversity and Distributions* **19**:1020–1030.

Chan, F. T., S. a. Bailey, C. J. Wiley, and H. J. MacIsaac, 2012. Relative risk assessment for ballast-mediated invasions at Canadian Arctic ports. *Biological Invasions* **15**:295–308.

Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. a. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. a. Rose, O. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear, 2001. Ecological forecasts: an emerging imperative. *Science* **293**:657–60.

Colautti, R. I. and H. J. MacIsaac, 2004. A neutral terminology to define invasive species. *Diversity and Distributions* **10**:135–141.

Dehnen-Schmutz, K., J. Touza, C. Perrings, and M. Williamson, 2007. A century of the ornamental plant trade and its impact on invasion success. *Diversity and Distributions* **13**:527–534.

Dextrase, A. J. and N. E. Mandrak, 2006. Impacts of Alien Invasive Species on Freshwater Fauna at Risk in Canada. *Biological Invasions* **8**:13–24.

Dorazio, R. and F. Johnson, 2003. Bayesian inference and decision theory-a framework for decision making in natural resource management. *Ecological Applications* **13**:556–563.

Duggan, I. C., C. a. M. Rixon, and H. J. MacIsaac, 2006. Popularity and Propagule Pressure: Determinants of Introduction and Establishment of Aquarium Fish.

Biological Invasions **8**:377–382.

Edwards, P. K. and B. Leung, 2009. Re-evaluating eradication of nuisance species: invasion of the tunicate, Ciona intestinalis. *Frontiers in Ecology and the Environment* **7**:326–332.

Elith, J., M. Burgman, and H. Regan, 2002. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological modelling* **157**.

Epanchin-Niell, R. S. and A. Hastings, 2010. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. *Ecology letters* **13**:528–41.

Gertzen, E., O. Familiar, and B. Leung, 2008. Quantifying invasion pathways: fish introductions from the aquarium trade. *Canadian Journal of Fisheries and Aquatic Sciences* **65**:1265–1273.

Gertzen, E. L. and B. Leung, 2011. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

Hastings, A., 1996. Models of spatial spread: Is the theory complete? *Ecology* **77**:1675–1679.

Hastings, A., K. Cuddington, C. J. Dugaw, S. Elmendorf, A. Freestone, S. Harrison, M. Holland, J. Lambrinos, and U. Malvadkar, 2005. The spatial spread of invasions : new developments in theory and evidence. *Ecology Letters* **8**:91–101.

Hilborn, R. and M. Mangel, 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, NJ.

Ibáñez, I., J. Diez, L. L. P. Miller, J. D. Olden, C. J. B. Sorte, D. M. Blumenthal, B. A. Bradley, C. D'Antonio, J. S. Dukes, R. Early, E. D. Grosholz, and J. J. Lawler, 2014. Integrated Assessment of Biological Invasions. *Ecological Applications* **24**:25–37.

Jerde, C. L. and M. a. Lewis, 2007. Waiting for invasions: a framework for the arrival of nonindigenous species. *The American naturalist* **170**:1–9.

Kay, S. H. and S. T. Hoyle, 2001. Mail Order , the Internet , and Invasive Aquatic Weeds. *Journal of Aquatic Plant Management* **39**:88–91.

Keith, J. M. and D. Spring, 2013. Agent-based Bayesian approach to monitoring the progress of invasive species eradication programs. *Proceedings of the National Academy of Sciences* **2013**:1–6.

Keller, R. and D. Lodge, 2007. Risk assessment for invasive species produces net bioeconomic benefits. *Proceedings of the National Academy of Sciences of the United States of America* **104**:203–207.

Keller, R. P., J. M. Drake, M. B. Drew, and D. M. Lodge, 2011. Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. *Diversity and Distributions* **17**:93–102.

Kerr, J., C. S. Brousseau, and M. Muschett, 2005. Invasive Aquatic Species in Ontario: A Review and Analysis of Potential Pathways for Introduction. *Fisheries* **30**:21–30.

Kulhanek, S. a., B. Leung, and A. Ricciardi, 2011. Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecological applications* **21**:203–13.

Landis, W. G., 2004. Ecological risk assessment conceptual model formulation for nonindigenous species. *Risk analysis : an official publication of the Society for Risk Analysis* **24**:847–58.

Leung, B., J. M. Bossenbroek, and D. M. Lodge, 2006. Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biological Invasions* **8**:241–254.

Leung, B. and D. G. Delaney, 2006. Managing sparse data in biological invasions: a simulation study. *Ecological Modelling* **198**:229–239.

Leung, B., J. M. Drake, and D. M. Lodge, 2004. Predicting invasions: Propagule presure and the gravity of allee effects. *Ecology* **85**:1651–1660.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti, 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings. Biological sciences / The Royal Society* **269**:2407–13.

Leung, B., N. Roura-Pascual, S. Bacher, J. Heikkilä, L. Brotons, M. a. Burgman, K. Dehnen-Schmutz, F. Essl, P. E. Hulme, D. M. Richardson, D. Sol, M. Vilà, and M. Rejmanek, 2012. TEASIng apart alien species risk assessments: a framework for best practices. *Ecology letters* **15**:1475–93.

Lupi, F., J. Hoehn, and G. Christie, 2003. Using an Economic Model of Recreational Fishing to Evaluate the Benefits of Sea Lamprey (Petromyzon marinus) Control on the St. Marys River. *Journal of Great Lakes Research* **29**:742–754.

MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. a. Graniero, 2004. Backcasting and Forecasting Biological Invasions of Inland Lakes. *Ecological Applications* **14**:773–783.

MacPherson, A. J., R. Moore, and B. Provencher, 2006. A dynamic principal-agent model of human-mediated aquatic species invasions. *Agricultural and Resource Economics Review* **35**:144–154.

Meyerson, L. a. and H. a. Mooney, 2007. Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment* **5**:199–208.

Moody, M. and R. Mack, 1988. Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology* **25**:1009–1021.

Muirhead, J. R. and H. J. Macisaac, 2005. Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology* **42**:80–90.

Occhipinti-Ambrogi, a. and D. Savini, 2003. Biological invasions as a component of global change in stressed marine ecosystems. *Marine pollution bulletin* **46**:542–51.

Olden, J. D., N. Leroy Poff, M. R. Douglas, M. E. Douglas, and K. D. Fausch, 2004. Ecological and evolutionary consequences of biotic homogenization. *Trends in ecology & evolution* **19**:18–24.

Padilla, D. K. and S. L. Williams, 2004. Beyond Ballast Water: Aquarium and Ornamental Trades as Sources of Invasive Species in Aquatic Ecosystems. *Frontiers in Ecology and the Environment* **2**:131–138.

Parker, I., D. Simberloff, W. Lonsdale, K. Goodell, M. Wonham, P. M. Kareiva, M. Williamson, B. Von Holle, P. Moyle, J. Byers, and L. Goldwasser, 1999. Impact: toward a framework for understanding the ecological effects of invaders. *Biological Invasions* **1**:3–19.

Peck, S. L., 2004. Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in ecology & evolution* **19**:530–4.

Pejchar, L. and H. a. Mooney, 2009. Invasive species, ecosystem services and human well-being. *Trends in ecology & evolution* **24**:497–504.

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

Polasky, S., S. R. Carpenter, C. Folke, and B. Keeler, 2011. Decision-making under great uncertainty: environmental management in an era of global change. *Trends in ecology & evolution* **26**:398–404.

Rahel, F. J., 2007. Biogeographic barriers, connectivity and homogenization of freshwater faunas: it's a small world after all. *Freshwater Biology* **52**:696–710.

Ricciardi, A., 2006. Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distributions* **12**:425–433.

Ricciardi, A., 2007. Are modern biological invasions an unprecedented form of global change? *Conservation Biology* **21**:329–36.

Rothlisberger, J. D., W. L. Chadderton, J. McNulty, and D. M. Lodge, 2010. Aquatic Invasive Species Transport via Trailered Boats: What is Being Moved, Who is Moving it, and What Can Be Done. *Fisheries* **35**:121–132.

Rothlisberger, J. D., D. C. Finnoff, R. M. Cooke, and D. M. Lodge, 2012. Ship-borne Nonindigenous Species Diminish Great Lakes Ecosystem Services. *Ecosystems* **15**:1–15.

Ruel, J. J. and M. P. Ayres, 1999. Jensens inequality predicts effects of environmental variation. *Trends in Ecology & Evolution* **14**:361–366.

Schneider, D. W., C. D. Ellis, and K. S. Cummings, 1998. A Transportation Model Assessment of the Risk to Native Mussel Communities from Zebra Mussel Spread. *Conservation Biology* **12**:788–800.

Semmens, B., E. Buhle, A. Salomon, and C. Pattengill-Semmens, 2004. A hotspot of non-native marine fishes: evidence for the aquarium trade as an invasion pathway. *Marine Ecology Progress Series* **266**:239–244.

Skellam, J., 1951. Random dispersal in theoretical populations. *Biometrika* **38**:196–218.

Smith, C. S., W. M. Lonsdale, and J. Fortune, 1999. When to ignore advice: invasion predictions and decision theory. *Biological Invasions* **1**:89–96.

Smith, V. and Y. Kaoru, 1986. Modeling recreation demand within a random utility framework. *Economics Letters* **22**:395–399.

Stephens, P., W. Sutherland, and R. Freckleton, 1999. What is the Allee effect? *Oikos* **87**:185–190.

Stohlgren, T. J. and J. L. Schnase, 2006. Risk analysis for biological hazards: what we need to know about invasive species. *Risk analysis : an official publication of the Society for Risk Analysis* **26**:163–73.

Strecker, A., P. Campbell, and J. D. Olden, 2011. The aquarium trade as an invasion pathway in the Pacific Northwest. *Fisheries* **36**:74–85.

Timar, L. and D. J. Phaneuf, 2009. Modeling the human-induced spread of an aquatic invasive: The case of the zebra mussel. *Ecological Economics* **68**:3060–3071.

Vander Zanden, M. J. and J. D. Olden, 2008. A management framework for preventing the secondary spread of aquatic invasive species. *of Fisheries and Aquatic*

Science **65**:1512–1522.

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

Yan, N. D., B. Leung, M. a. Lewis, and S. D. Peacor, 2011. The spread, establishment and impacts of the spiny water flea, Bythotrephes longimanus, in temperate North America: a synopsis of the special issue. *Biological Invasions* pages 2423–2432.

Yemshanov, D., F. H. Koch, M. Ducey, and K. Koehler, 2013. Mapping ecological risks with a portfolio-based technique: incorporating uncertainty and decision-making preferences. *Diversity and Distributions* **19**:567–579.

Zhang, C. and K. J. Boyle, 2010. The effect of an aquatic invasive species (Eurasian watermilfoil) on lakefront property values. *Ecological Economics* **70**:394–404.

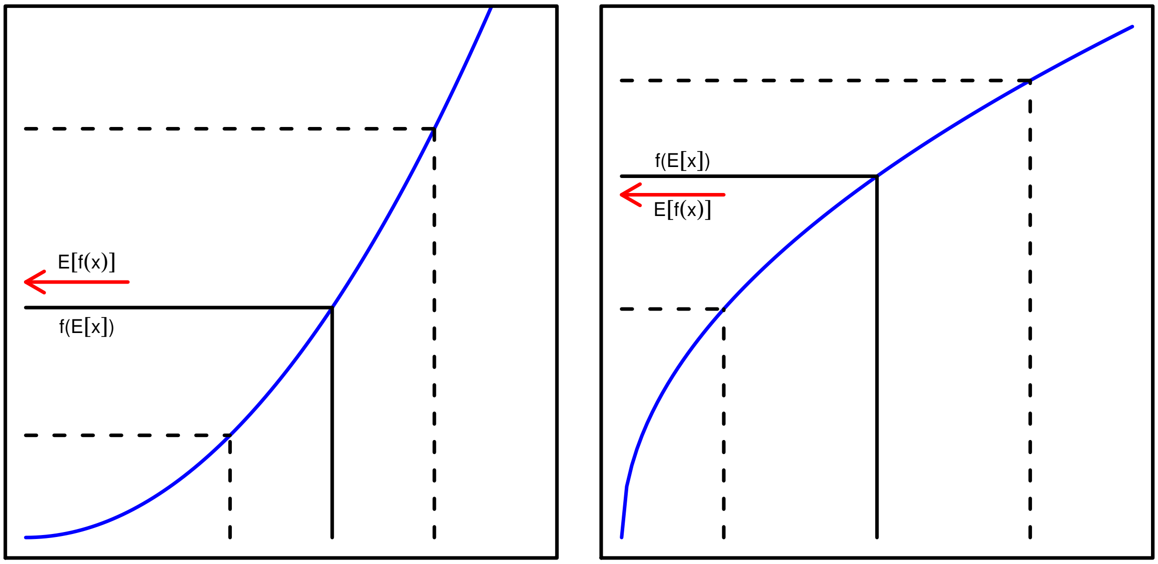


Figure 1–1: **Jensen's inequality.** For any non-linear function, the function evaluated at the mean of a distribution of inputs is not equal to the mean of the function evaluated over the distribution of inputs. A consequence of this is that failure to account for uncertainty not only leads to false precision, but also to systematic biases.

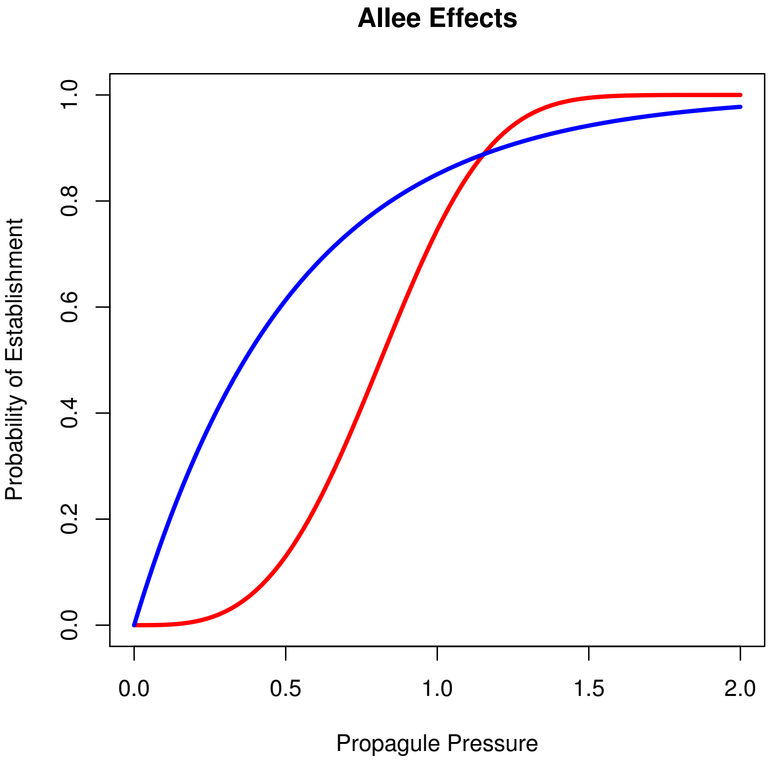


Figure 1–2: **Propagule pressure-probability of establishment relationships.** Blue line is the independence model which is convex over the entire range. Red line is the relationship when an Allee effect is present. In the Allee case, the function is concave for low propagule pressures, and becomes convex for high propagule pressures. The consequence of variability will depend on the strength of the Allee effect, and both the location and dispersion of the distribution of propagule pressures.

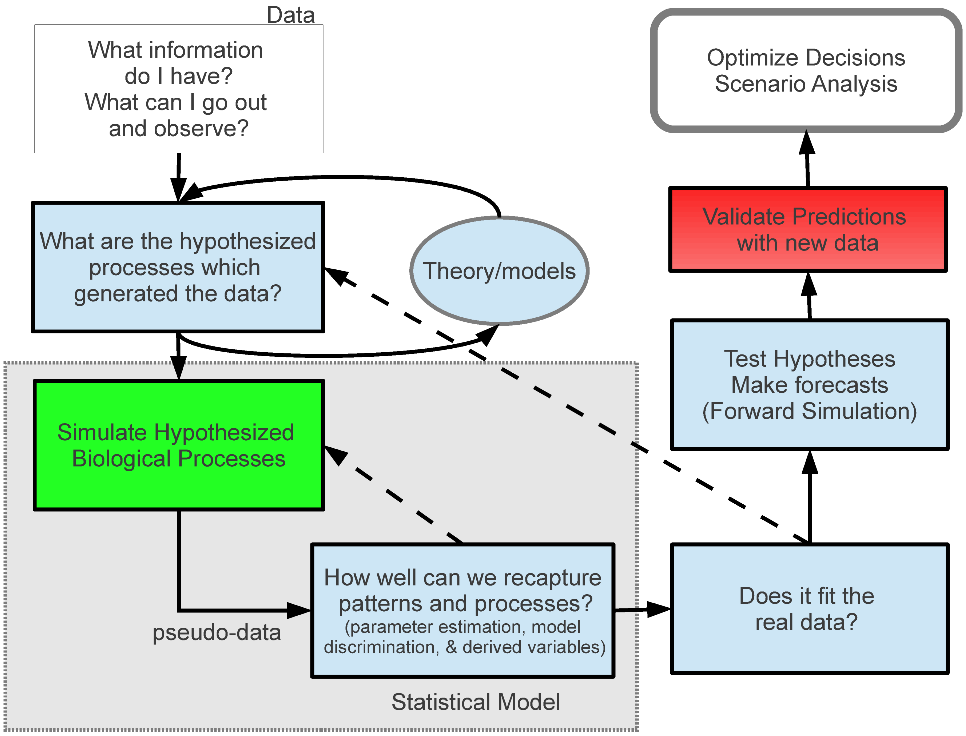


Figure 1–3: **General methodological framework for ecological predictions and decision making under uncertainty.** This framework outlines the approach taken throughout this thesis. I have paid particular attention to the cycle in the grey box, which is often neglected and can lead to erroneous interpretations of data and sub-optimal decisions.

CHAPTER 2

Predicting invasions: Alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects

Corey Chivers[1] , Brian Leung[2] [3]

A version of this chapter has been published in the *Journal of Applied Ecology*.

---

[1] Department of Biology, McGill University.Stewart Biology Building 1205 ave Docteur Penfield Montreal, Quebec. 514-398-1833. corey.chivers@mail.mcgill.ca.

[2] Department of Biology, McGill University

[3] School of Environment, McGill University. brian.leung2@mcgill.ca

2.1 Abstract

1. Human-mediated dispersal has been shown to be the most important vector for the spread of invasive species, yet there has been little evaluation of alternative models of dispersal in terms of differences in their predictions of invasion patterns. Moreover, no analyses have been attempted to elucidate the potential interaction between alternative models of human-mediated dispersal and population dynamical characteristics, such as Allee effects, that are central to the probability of an invasion.
2. Two prominent models in the literature which have previously been employed to predict human movement patterns are explored: a) Gravity models, which use the attractiveness of, and distance to a location to predict travel patterns, and b) Random utility models, which assume that individuals decide where to travel by maximizing the benefits which they receive according to some partially observable function of individual and site characteristics.
3. While distinction is often drawn between them in the literature, we demonstrate that these two approaches can be reduced to alternative functional forms describing the trip taking decisions of individuals.
4. Each model was empirically parameterized using a survey of recreational boaters in Ontario, Canada. Within each model, both boater and site specific characteristics were important and the functional form provided by the gravity model was significantly better at capturing the behaviour of recreational boaters.

5. *Synthesis and applications.* The dispersal and establishment of species into novel habitats are central components of the invasion process and of quantitative risk assessments. However, predictions are dependent on the estimated spatial structure of the dispersal network and its potential interactions with species characteristics. This study demonstrates that Allee effects can interact with dispersal network structure to significantly alter predicted spread rates and that the consequences of these interactions manifest differently at the system and site levels. These insights can be used to inform management interventions aimed at modifying human-mediated dispersal in order to reduce the spread of invasive species.

2.2 Introduction

Invasive species can cause ecological (Parker et al., 1999; Pejchar and Mooney, 2009) and economic impacts (Aukema et al., 2011). In order to prevent or limit the spread of potentially harmful species, management efforts must be informed by reliable estimates of where and when we expect new invasions to occur. The overland dispersal (lake to lake spread) of aquatic invasive species has been shown to be driven primarily by inadvertent human-mediated transportation. Several species have been observed 'hitchhiking' on the hulls, and in the ballasts of recreational water vessels, which are transported on trailers from lake to lake (Johnston et al., 2001; Kraft et al., 2002). A recent study by Gertzen and Leung (2011) comparing human-mediated and fluvial dispersal found that human-mediated dispersal of an invasive species accounted for almost all of the propagules contributing to establishment probability. Understanding where species are likely to spread via this key human-mediated pathway is therefore an important step toward implementing mitigating measures.

There have been two general classes of modelling frameworks developed in the literature to characterize the movement of individuals across a landscape of discrete sites. Gravity models (GM) have been used extensively to characterize human movement patterns, and have been applied successfully in several studies to the spread of invasive species (Leung et al., 2004; Potapov et al., 2011; Muirhead et al., 2011). They work by an analogy to Newtonian gravity, where individuals are attracted to locations proportionally to their mass (which can be any set of measures of desirability of the site) and inversely to the distance between an individual's current location, and the site Schneider et al. (1998); Leung et al. (2004, 2006). An alternative specification, developed in the field of recreational demand econometrics, is the discrete choice random utility model (RUM) (Smith and Kaoru, 1986; Smirnov and Egan, 2012). In this framework, individuals choose a destination from a suite of alternatives by maximizing a utility function based on any set of desirable traits, only part of which is known to the analyst. This model has been used in several recreational demand studies (e.g. Smith and Kaoru (1986), Champ et al. (2003)), but has only recently been applied to the study of spread of invasive species (MacPherson et al., 2006; Timar and Phaneuf, 2009). We show that these models are quite similar and that they can be reduced to simply alternative functional forms to describe an individual's trip decisions.

While it is clear that human vectors are central to the invasion process, the ramifications of employing alternative models of this vector on predicting spread is less clear. Moreover, although it has not been previously examined, one might expect that the consequences of dispersal models may interact with, and be determined by the specific population dynamics of invaders. In particular, stochasticity and Allee effects are both well-known population level factors affecting invasion dynamics (Clark et al., 2003; Drake and Lodge, 2006).

In this study, we address the following three questions: 1) Do the alternative human vector modelling frameworks (gravity and random utility models) differ in their ability to capture actual human behaviour, and therefore characterize dispersal vectors of invasive species? 2) How do these alternative models interact with the population dynamics of invaders? 3) What are the implications of alternative dispersal model specifications on our predictions of invasion risk across space and time (i.e. spread)?

We analysed the predictive ability of competing models of human-mediated dispersal by surveying recreational boaters and examining the ability of each model to recapture the observed trip outcomes. We recognized that differences in model fit are most important if the alternative model formulations lead to human-mediated dispersal networks which yield quantitative differences in our predictions of the spread of invasive species. Given the potential for ecological and economic harm posed by invasive species, predictions of spread across a landscape, as well as invasion risk at specific sites are vital components of informed management policies (Landis, 2004). Thus, we conducted a series of simulation experiments to examine the potential implications of each human-mediated dispersal models for risk assessments, taking into account their interaction with the population dynamics of invaders. We describe how the entropy, or evenness of the predicted connectivity distribution of the dispersal network can interact with population dynamics to hinder spread. Taken together, this work provides new insights into how models of human behaviour affect the predicted structure of discrete dispersal networks, and how the structure of dispersal networks interact with population level processes to influence the spatial spread of invasive species.

2.3 Materials & methods

2.3.1 Survey

We conducted a survey of recreational boaters in Ontario, Canada. We mailed 5,000 invitations to participate in the survey to individuals with registered recreational licenses (boating/fishing) issued by the Ontario Ministry of Natural Resources. Individual names were selected using a spatially stratified random sampling scheme. Approximately 100 invitations were sent to randomly selected individuals in each of 47 major geographic regions of Ontario as defined by the first two digits of their postal code. We developed an online survey instrument using the design approach of Dillman (2000). We employed an interactive map through which participants could quickly and easily identify the lakes which they visited. The advantage of this approach was that we were able to precisely identify lakes which may have been ambiguous due to multiple naming conventions. In this way, we were able to collect more in depth information in a visually intuitive manner. While our survey instrument was only able to capture individuals with access to the internet, 81% of households in Ontario had access to the internet as of 2010 (Statistics Canada, (`http://www40.statcan.gc.ca/l01/cst01/comm36g-eng.htm`). We have no reason to believe that those without internet access would behave differently vis a vis boating behaviour than those with online access. We asked participants to catalogue all of the boating trips which they took and to indicate the primary location where they kept their boat during the 2010 boating season.

Our survey response rate was 11%, with 30% of respondents indicating that they had visited multiple lakes during 2011. Given that we are interested in the behaviour of boaters who transport their boat from lake to lake during the boating season, we retained only those trip outcomes made by multi-lake boaters. This left us with relevant observed source/destination outcomes for 146 individual boaters across Ontario making a total of 2354 boating trips (Fig. 2–1).

2.3.2 Gravity model specification

Gravity models employ an analogy to Newtonian gravity, where the 'pull' of a given site is proportional to some function of desirable lake characteristics (termed 'attractiveness', e.g. size, (Bossenbroek et al., 2001; MacIsaac et al., 2004; Leung and Delaney, 2006)) and inversely related to the distance between a source location and the site. A boater chooses a site to visit according to the degree to which they are 'pulled' to that site, relative to the degree to which they are pulled by all other possible sites. While there are many possible formulations of gravity models, recent comparisons have found that the production-constrained gravity model provides the best estimate of human-mediated dispersal of aquatic invasives (Muirhead, 2007; Muirhead et al., 2011). In the production-constrained formulation, it is assumed that a boater travels from their home location (primary location where they keep their boat), to a destination lake, and then returns to their home location before visiting another lake. Further, the production-constrained gravity model has modest data requirements compared with its alternatives (Muirhead et al., 2011), making it an accessible choice for resource managers. Because we wish to compare models in terms of their ability to capture individual level behaviour, we present a disaggregated formulation of the production-constrained gravity model, in which each individual makes trip destination decisions according to a probability distribution described by the model. The site selection probability distribution, $P(T_{n\bullet})$ for an individual boater n is given as:

$$P^{GM}(T_{nj}) = A_n W_j^e D_{nj}^{-d}, n = 1, \ldots, N, j = 1, \ldots, J. \tag{2.1}$$

Where W_j is the attractiveness of lake j, and D_{nj} is the distance between lake j, and the home location (where they keep their boat) of individual n. Some authors suggest the use of least cost road networks to calculate the effective distance between source and destination (Drake and Mandrak, 2010), however for simplicity here we use the euclidian distance between boater home location and lake centroid. The free parameters d and e describe the shape of the relationship and are fitted to the data (see section 2.3.4). A_n is the 'pull' of all lakes, given by:

$$A_n = 1/\sum_{k=1}^{J} W_k^e D_{nk}^{-d}. \tag{2.2}$$

Such that the probability of a boater n visiting lake j is proportional to the gravitational 'pull' of that lake compared to that of all other lakes. As a simple proxy for lake attractiveness, we used lake surface area in hectares. While other lake characteristics may alter the attractiveness, lake area is most readily available and has been shown to be predictive in previous studies (Leung et al., 2004, 2006; Muirhead et al., 2011; Gertzen and Leung, 2011).

Furthermore, our survey tool provided us with additional boater level information, which we were able to incorporate into the model. Respondents identified which type of boat they owned and we categorized them as large motor boat (> 14'), small motor boat (< 14'), or other. We assumed that boater type would modulate the relationship between lake size (W) and probability of visitation. As such, we incorporated this additional information using dummy variables ($B1$, and $B2$) in the exponent of W:

$$P^{GM}(T_{nj}) = A_n W_j^{e+\beta_1 B1_n+\beta_2 B2_n} D_{nj}^{-d}, n = 1, \ldots, N, j = 1, \ldots, J. \tag{2.3}$$

$$A_n = 1/\sum_{k=1}^{J} W_k^{e+\beta_1 B1_n+\beta_2 B2_n} D_{nk}^{-d}. \tag{2.4}$$

Where $B1$ and $B2$ equal 0 for large motor boat, $B1$ equals 0 and $B2$ equals 1 for small motor boat, and $B1$ equals 1 and $B2$ equals 0 for other. In this way, boat type determines the rate at which each additional hectare of lake area increases the attractiveness of a given lake.

2.3.3 Random utility model specification

The RUM is a discrete choice model used extensively in the econometrics literature to predict the behaviour of recreationalists (Champ et al., 2003). This formulation has recently been applied to predicting the spread of invasive zebra mussels in Wisconsin (Timar and Phaneuf, 2009), and in a simulation study of the spread, and management of Eurasion watermilfoil (MacPherson et al., 2006). In this model, boaters are assumed to behave as rational actors, maximizing their utility. For a given trip, a boater chooses the site which maximizes their utility function U, which is only partially observable by the analyst. We can separate the utility function, therefore, into two parts. The utility which boater n would derive from visiting lake j can be then be written as the sum of the observable part V_{nj}, and an error term ϵ_{nj}.

$$U_{nj} = V_{nj} + \epsilon_{nj}, n = 1, \ldots, N, j = 1, \ldots, J. \tag{2.5}$$

Where V_{nj} is any linear function of the attributes of boater n and site j.

$$V_{nj} = \boldsymbol{\beta} X_{nj} \tag{2.6}$$

We can then re-write the utility that would be derived by boater n by visiting each site in terms of the probability that they will choose that site over all other alternatives.

$$\begin{align} P^{RUM}(T_{nj}) &= Pr(U_{nj} \geq U_{nk} \forall k \neq j) \tag{2.7} \\ &= Pr(V_{nj} + \epsilon_{nj} \geq V_{nk} + \epsilon_{nk} \forall k \neq j) \tag{2.8} \\ &= Pr(\epsilon_{nk} - \epsilon_{nj} \leq V_{nj} - V_{nk} \forall k \neq j), n = 1, \ldots, N, \ j = 1, \ldots, J. \tag{2.9} \end{align}$$

If we model the error term ϵ using the type-I extreme value distribution as is most commonly done, the model reduces to a simple logit, and the distribution describing the probability that boater n will choose to visit lake j is given by:

$$P^{RUM}(T_{nj}) = \frac{exp(V_{nj})}{\sum_{k=1}^{L} exp(V_{nk})}, n = 1, \ldots, N,\ j = 1, \ldots, J. \tag{2.10}$$

For further details of this model, see (Champ et al., 2003). The parameters ($\boldsymbol{\beta}$) are easily fit given the observed trip outcomes using maximum likelihood (see section 2.3.4).

As with the gravity model, we incorporated the additional boater level predictor of boat type into the utility function of the RUM. We did this by adding two dummy variables to describe the three categories of boat type, with the same definitions as in the gravity model. Our full (observable) utility function is therefore formulated as:

$$U_{nj} = \beta_1 W_j + \beta_2 D_{nj} + (\beta_3 B1_n + \beta_4 B2_n) W_j. \tag{2.11}$$

Since boat type is a boater level variable, we do not include it into the main effect part of the utility function, as it cancels out when summing across the entire choice set of a given boater. Instead, we model the interaction between boat type and lake size (W).

In both the RUM and GM model formulations, we have made two key assumptions: 1) Boater behaviour is constant across time, and 2) boater trips are distributed independently and identically according to each model.

2.3.4 Fitting and model selection

The parameters ($\boldsymbol{\theta}$) of each model can be fit using maximum likelihood. Our survey data provides us with observations of the number of trip outcomes S_{nj} , for each boater n, to a given lake j. From these observations, we can write the log-likelihood for model M as:

$$LL^{M}(\boldsymbol{\theta} \,|\, \boldsymbol{D}) = \sum_{n=1}^{N} \sum_{j=1}^{J} S_{nj} log[P^{M}(T_{nj} \,|\, \boldsymbol{\theta})] \tag{2.12}$$

We fit the parameters of each model, including reduced models using maximum likelihood implemented in the R statistical programming environment (R Core Team, 2013). Reduced models were those in which we removed the boater level parameters pertaining to boat type. Each model was then compared in terms of its relative performance using two separate metrics. The first metric of model selection we used is the Akaike Information Criterion (AIC) (Burnham and Anderson, 2002). The second metric we used was the simple coefficient of determination (R^2) between the predicted and observed total number of trips taken to each lake in our study system. From this metric we could compare the relative proportions of total variation in the number of visits across all lakes explained by each model.

2.3.5 Spread simulations: Examining theoretic model behaviour and interactions with population demographics

Ultimately, we are constructing our models of human movement patterns between discrete patches to use in making predictions about the spread of species which are being dispersed across this network of patches. While spread is a stochastic process, where introductions lead to viable population establishments at a given site in a non-deterministic manner, we can use repeated simulations to characterize the expected trajectory of a given invasion process (Peck, 2004). By simulating the spread process under each of our competing models, we can compare the predicted trajectories to make inferences about the consequences of model specification on spread prediction. Differences in predicted spread rates, as well as predicted invasion risk at the individual site level may have an effect on management decisions regarding mitigation and control (see Chapter 5).

To conduct these simulations, we followed the procedure outlined in (Leung and Delaney, 2006). We model the stochastic spread process as

$$G(Q_{jt}) = 1 - e^{-(\alpha Q_{jt})^c} \tag{2.13}$$

Where the probability of invasion is given as a function of the number of propagules Q, arriving at time t to site j. The function is described by two shape parameters. The first, α, is a per propagule multiplier proportional to $-ln(1-p)$, where p is the per propagule probability of establishment. The additional parameter c allows us to describe an Allee effect, where the per propagule establishment probability is disproportionately lower at low propagule pressures (Dennis, 2002). The strength of the Allee effect increases as $c > 1$. Non-negligible Allee effects have been observed in some aquatic invasives. This parameter has been estimated as 1.86 ($P < 0.0001; Ho : c = 1$) for zebra mussels using an invasion time series (Leung et al., 2004). Wittmann et al. (2011) also detect an Allee effect using a stage-structured model of the invasive zooplankton Bythotrephes.

To calculate the number of propagules Q arriving at site j, we sum across the probability distribution of each boater having visited an invaded lake before arriving at site j. To do this, we first calculate the proportion of boaters at each source location which have visited an invaded lake as:

$$X_{it} = O_i \sum_{h=1}^{H} P^M(T_{ih}) \tag{2.14}$$

Where O_i is the number of boaters at source location i, and $P^M(T_{ih})$ is the probability of a boater at source location i visiting lake h as given by the model M under which simulations are being carried out. X_{it} is the number of boaters in source location i having visited an invaded lake in time step t. We derived O_i from data obtained from the Ontario Ministry Natural Resources on the number of registered boaters in Ontario in each of 526 postal regions identified by the first three postal code digits. The next step is to calculate the propagule pressure Q arriving at lake j in time t as:

$$Q_{jt} = \sum_{i=1}^{K} P^M(T_{it}) X_{it} \tag{2.15}$$

Which is the total boater traffic from all invaded sources to lake j in time step t. For more details, see (Gertzen and Leung, 2011). While each human vector model predicts a unique trip distribution matrix, the total number of boater trips taken, or the overall magnitude of traffic flow in the system as a whole is constant across both models. Any difference in the observed rates of spread in our simulations therefore is a result of the dispersal network structure, and not the absolute magnitude of between-lake movement.

While there are roughly 250,000 lakes and rivers in Ontario, in order to render our simulations computationally feasible, we simulate spread across only those lakes with a surface area larger than 10 hectares. Additionally, we removed lakes above 52° latitude, as these lakes are not accessible by any roadways connecting them to the southern lakes. This left us with 781 lakes in our simulation set. Each independent simulation began with a seed invasion in Lake Ontario, and was run forward 30 years. By seeding the invasion in Lake Ontario, we recreate the most likely invasion scenario for Ontario inland lakes. As of 2006, the great lakes are known to have been invaded by at least 182 species (Ricciardi, 2006), making it the most likely source location of a novel species spread to inland lakes.

In order to analyse potential interactions between population dynamics and the human vector model, we examined the effect of population establishment parameters and we ran repeated simulations across a range of parameter values of both α (7.5-e05,1.0e-04,1.25e-04,1.5e-04) and c (1,1.5,2,2.5). For each simulation, we used either the best fitting GM or RUM of boater behaviour. As our metrics of invasion progress, for each run, we retained the cumulative number of lakes invaded. An example realization of our simulated spread procedure can be seen in Fig. 2–2. Additionally, we compared the relative invasion risk at each of three specific selected sites. Lakes Simcoe, Nipissing, and Nipigon were selected due to their large size, making them more at risk to invasion, as well as due to their relative distances from the source location of invasion. While these lakes by no means represent a random sample, they provide a convenient gradient of baseline risk along which to observe the rate at which deviations between models occur. For these lakes, we retained the time to invasion across every simulation for every parameter combination. We calculated the risk to a given lake as the proportion of simulation realizations in which the site became invaded before the end of the 30 year time horizon.

2.4 Results

2.4.1 Model fitting and model selection

Formal model selection identified the GM as the most likely, given the data. The GM provided superior fit to the RUM with a ΔAIC value of 3229 between the full GM and the full RUM, for the observed pattern of boater trips. Table 2–1 provides the ΔAIC for each model, sorted in increasing order (decreasing order of goodness of fit). Maximum likelihood parameter estimates and their 95% confidence intervals for each full model are given in Table 2–2. All fitted parameter values have direction and magnitude which we would expect. In the gravity model $0 < e < 1$, indicating a diminishing marginal effect of each additional hectare of lake surface area. The boater specific dummy variables ($\beta_1 > 0, \beta_2 < 0$) indicate that the marginal effect of each additional hectare decreases fastest for small motor boats. That $d > 0$ indicates that closer lakes are more attractive than more distant ones. Similarly, in the RUM $\beta_1 > 0$, and $\beta_2 < 0$ indicate a positive relationship between lake area and utility, and a negative relationship between distance and utility, respectively. As with the gravity model $\beta_3 > 0$, and $\beta_4 > 0$ indicate that the marginal utility of an additional hectare diminishes fastest for mall motor boats. The full GM was able to account for 58% of the variation in trip outcomes across all boaters, compared to only 42% for the full RUM (Fig. 2–3). As a check for bias, we fit a linear regression to the predicted-observed points. The equations of fit were $y = 0.00018[\pm 0.00016] + 0.86[\pm 0.034]x_{GM}$ and $y = -0.00031[\pm 0.00018] + 1.24[\pm 0.058]x_{RUM}$. While neither intercepts deviate significantly from zero, the slope of the GM is less than one, indicating a tendency to overestimate the traffic to high frequency lakes, while the RUM tends to underestimate traffic to high frequency lakes.

There are three main components of the differences between the dispersal networks predicted by each model. 1) Each model generally predicted higher traffic to large lakes which are close to dense population sources, as expected. However, the rank ordering of individual lakes can differ substantially within this broader pattern (Fig. 2–4a). 2) The average predicted distance travelled by boaters was higher in the GM (190 km) compared to the RUM (140 km). 3) How evenly, or unevenly the predicted traffic was spread across different sites differed between models. In order to quantitatively evaluate this characteristic, we calculated the Shannon entropy of the traffic distributions predicted by the two models. Entropy can be thought of as a measure of evenness (Hill, 1973). Probability distributions with higher entropy are more evenly dispersed. As entropy decreases, the distribution becomes more uneven, or more sharply peaked, such that more of the mass of the distribution is concentrated in fewer sites. We calculate the Shannon entropy of the predicted distributions P for model M as:

$$H(P^M) = -\sum_{i=1}^{I} P_i^M log(P_i^M) \tag{2.16}$$

As $H(P^M) \rightarrow 0$, boater traffic is concentrated entirely in one lake. The maximum entropy distribution would be that which assigns $P_i = 1/n$ to all lakes in the system. Comparing the predictions of the two models, we find that the gravity model ($H(P^{GM}) = 5.018$) represents a more uneven predicted distribution than that of the RUM ($H(P^{RUM}) = 5.36$). The stronger concentration of traffic predicted by the GM, compared with the RUM can be seen by comparison of the rank ordered distributions in Fig. 2–4b. The consequences of these differences are analysed in the following section.

2.4.2 Implications for spread and risk assessment

When the per-propagule probability of establishment is low (small α values), and there is no Allee effect present, both dispersal models predict similar rates during the early phases of invasion (Fig. 2–5). The deviations between early rates of spread under the alternative dispersal models increases drastically, however, as the strength of the Allee effect increases. At the extreme end of invasiveness and Allee effect ($\alpha = 1.5e-04$ and $c = 2.5$), we observe an over ten-fold increase in the cumulative total number of sites invaded by the end of the 30 year time horizon. The degree of deviation induced by increased Allee strength is also modulated by the independent population growth, or per-propagule invasiveness parameter α. This can be seen by observing the magnitude of deviation at each row of Fig. 2–5, which increases as the parameter α increases.

In the absence of an Allee effect, the deviations in late-stage rates of spread can be accounted for by the differences in predicted mean distance travelled under each model. Spread in Southern Ontario occurs at similar rates under each model due to high population density, where the distances between population sources and lakes are short. However, as the invasion progresses northward into the more sparsely populated regions, spread under the RUM is slowed substantially due to the increased distances required to reach additional lakes. When an Allee effect is present, the difference in spread rate is apparent throughout the entire time series. The relative entropies of the dispersal distributions (i.e. the variance in total inbound propagules arriving across all uninvaded sites) can account for this further deviation. The expected rate at which the proportion of previously uninvaded sites become invaded can be written as $R = E[G(Q_{\bullet t})]$, where $G(Q_{\bullet t})$ is a vector of invasion probabilities for all uninvaded sites, and is given by Eqn. 2.15. When an Allee effect is present, the function $G(Q_{\bullet t})$ is concave over part of its range. By Jensen's inequality we know that for a convex function $E[G(Q_{\bullet t})] \geq G[E(Q_{\bullet t})]$. From this we can see that as the variance of inbound propagules increases over the concave range as a result of a more uneven dispersal distribution, the rate of new invasions increases as well.

Spread at the landscape level may be of interest to regional managers; however, the risk of invasion posed at specific sites will inform management decisions made at the lake level. To see the differences in the site-level invasion risk predicted by our alternative dispersal models, we also looked at three specific inland lakes (Lakes Simcoe, Nipissing and Nipigon). These three lakes occur at increasing distances from our source location (Lake Ontario), respectively. By observing the differences in invasion risk predicted at these sites, it is possible to see how uncertainty and deviations between model predictions increase as we move further from the source of invasion. Fig. 2–6 shows probability of invasion (risk) as a function of time at each of the three sites across our range of population parameters. The risk of invasion posed at each of these three sites is always higher under the GM, with the exception lakes Nipissing and Nipigon under strong Allee effect where projected risk is very near zero and indistinguishable between models.

2.5 Discussion

In this study, we have shown that a GM can better capture the behaviour of individual boaters in Ontario than an RUM. Ultimately, these two alternative models can be represented simply as different functional forms which we can use to describe a boater's trip-taking probability distribution. In the case of our sample of Ontario boaters, the functional form of the GM provides a better representation of the probabilistic process through which boaters select which lakes to visit from a suite of alternatives.

Both of the behavioural models considered in this study were built using only the distance between the boater source location, lake size (surface area in hectares) and boat type as explanatory factors. We recognize that there may be a suite of additional variables which may add further explanatory power. Previous work has incorporated additional lake predictors, as well as additional interactions between individual level and lake variables. These have included lake clarity (measured as secci depth), cost of access and whether or not a given boater is an angler (Champ et al., 2003; Timar and Phaneuf, 2009). Here, we have used lake size, distance from boater's home location, and boat type as these are the most readily available data with which to build a model of boater behaviour for the purposes of assessing invasion risk. Both the gravity and RUMs can easily be extended to incorporate any number of additional lake and boater specific variables. In the RUM, one would need simply to include additional linear, or higher order, predictors ($\boldsymbol{\beta}$) in V_{nj} (see eqn. 2.6). Within a GM, lake level predictors could be modelled by expansion of W_j into:

$$W_j^{e+\beta_1 X_{1j}+\ldots+\beta_i X_{ij}} \tag{2.17}$$

With additional explanatory variables of lake attractiveness each requiring the fitting of an additional free parameter (see section 2.4.1).

We have also demonstrated that the choice of modelling framework used to characterize the human-mediated vector can have important consequences for predicting the future spread of invasive species. The deviation between spread predictions under the two models analysed here interact with population level factors of the invading species. In the presence of a strong Allee effect, boaters behaving according to an RUM do not act in such a way as to generate propagule pressures high enough to overcome the demographic barriers to establishment. The inability to overcome these barriers is a consequence of the evenness of the predicted trip distribution of boaters under the RUM, as measured using the Shannon entropy of the predicted trip distribution. Without the centralized "hub" lakes (those highly connected lakes with very high visitation frequency) predicted by the GM, a situation can arise where there is not a sufficiently concentrated flow of individuals from invaded lakes to uninvaded lakes. The data suggests rather, that individuals do act in such a way as to concentrate traffic to a small number of "hub" lakes, as predicted by a GM, and that this level of concentration is sufficient to overcome even very strong demographic barriers. The existence of such hub lakes and their importance to the spread of aquatic invasive species has been noted in the literature (MacIsaac et al., 2004; Muirhead and Macisaac, 2005). A misspecified behavioural model of human-mediated dispersal may underestimate the importance of these sites, leading to potentially overoptimistic projections of lake to lake spread.

We also know, however, that just predicting the rate of spread may not be the most relevant metric of interest to a resource manager who is making decisions at the local level. A more relevant measure at the lake level is the risk of invasion posed at particular sites. We analysed how our dispersal models affect site level predictions of risk by pulling out three of the larger, more important sites in Ontario: Lakes Simcoe, Nipissing, and Nipigon. For these specific lakes, the predicted probability of establishment over time differed dramatically between the GMs and RUMs, even in the absence of Allee effects. Indeed, these are three of the largest inland lakes in Ontario, all of which are probably receiving sufficient propagule pressure rather early on to overcome the demographic barrier of an Allee effect. From this result we can see that the way in which the underlying behavioural model interacts with the population dynamics of the invading species manifests differently at the site level, than at the landscape level.

When making policy decisions regarding invasive species, managers need informed estimates of invasion risk across space and time (Epanchin-Niell and Hastings, 2010). This study suggests that for boaters in Ontario, a GM of individual behaviour most accurately characterizes this single most important vector of overland invasive spread, and that alternative formulations of human vector dispersal models can interact with the population dynamics of the invading species to produce large deviations in predicted spread. These deviations manifest differently at the system level than at the level of individual lakes. While managers of inland freshwater resources should be aware of how these interactions impact assessments of risk, our results are general and hold for any species spreading across a network of discrete patches.

Future work could look at the utility of implementing GMs in the context of management interventions. Were managers to implement policies aimed at limiting the spread of an aquatic invasive species by levying a launching fee, or requiring hull sanitation procedures at either at-risk or currently infested lakes, boaters may change their behaviour. Changes in boater behaviours resulting from management interventions could potentially alter the structure of the human mediated dispersal network. This, as we have shown, will have consequences for our predictions of spread. Previous work has employed RUMs to incorporate these behavioural feedbacks (MacPherson et al., 2006; Timar and Phaneuf, 2009), however in light of our current results, it may be appropriate to include these behaviours directly in a GM formulation.

While we have shown here that the interaction between intra-patch dispersal connectivity and stochastic establishment dynamics within patches interact to determine rates of species spread across a landscape, there will no doubt be effects of other factors, such as spatial and temporal environmental heterogeneity (Melbourne et al., 2007), biotic interactions (Hunt and Yamada, 2003), as well as temporal variation in the dispersal network structure itself which should be considered in future studies.

2.6 Acknowledgments

The authors would like to thank all of our survey respondents, as well as Jason Cologna at the Ontario Ministry of Natural Resources for help registered boater mailing list data. Funding for this study was provided by the Canadian Aquatic Invasive Species Network.

References

Aukema, J. E., B. Leung, K. Kovacs, C. Chivers, K. O. Britton, J. Englin, S. J. Frankel, R. G. Haight, T. P. Holmes, A. M. Liebhold, D. G. McCullough, and B. Von Holle, 2011. Economic Impacts of Non-Native Forest Insects in the Continental United States. *PLoS ONE* **6**:e24587.

Bossenbroek, J. M., C. E. Kraft, and J. C. Nekola, 2001. Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecological Applications* **11**:1778–1788.

Burnham, K. P. and D. R. Anderson, 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer.

Champ, P. A., K. J. Boyle, and T. C. Brown, 2003. A Primer on Nonmarket Valuation. The Economics of Non-Market Goods and Resources. Springer.

Clark, J., M. Lewis, J. McLachlan, and J. HilleRisLambers, 2003. Estimating population spread: What can we forecast and how well? *Ecology* **84**:1979–1988.

Dennis, B., 2002. Allee effects in stochastic populations. *Oikos* **96**:389–401.

Dillman, D. A., 2000. Mail and Internet surveys: the tailored design method. Psychology/Sociology/Marketing. J. Wiley, second edition.

Drake, D. A. R. and N. E. Mandrak, 2010. Least-cost transportation networks predict spatial interaction of invasion vectors. *Ecological Applications* **20**:2286–2299.

Drake, J. M. and D. M. Lodge, 2006. Allee Effects, Propagule Pressure and the Probability of Establishment: Risk Analysis for Biological Invasions. *Biological Invasions* **8**:365–375.

Epanchin-Niell, R. S. and A. Hastings, 2010. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. *Ecology letters* **13**:528–41.

Gertzen, E. L. and B. Leung, 2011. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

Hill, M., 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**:427–432.

Hunt, C. and S. Yamada, 2003. Biotic resistance experienced by an invasive crustacean in a temperate estuary. *Biological Invasions* **5**:33–43.

Johnston, L. E., A. Ricciardi, and J. T. Carlton, 2001. Overland dispersal of aquatic invasive species: a risk assessment of transient recreational boating. *Ecological Applications* **11**:1789–1799.

Kraft, C. E., P. J. Sullivan, A. Y. Karatayev, L. E. Burlakova, J. C. Nekola, L. E. Johnson, and D. K. Padilla, 2002. Landscape patterns of an aquatic invader: assessing dispersal extent from spatial distributions. *Ecological Applications* **12**:749–759.

Landis, W. G., 2004. Ecological risk assessment conceptual model formulation for nonindigenous species. *Risk analysis : an official publication of the Society for Risk Analysis* **24**:847–58.

Leung, B., J. M. Bossenbroek, and D. M. Lodge, 2006. Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biological Invasions* **8**:241–254.

Leung, B. and D. G. Delaney, 2006. Managing sparse data in biological invasions: a simulation study. *Ecological Modelling* **198**:229–239.

Leung, B., J. M. Drake, and D. M. Lodge, 2004. Predicting invasions: Propagule presure and the gravity of allee effects. *Ecology* **85**:1651–1660.

MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. a. Graniero, 2004. Backcasting and Forecasting Biological Invasions of Inland Lakes. *Ecological Applications* **14**:773–783.

MacPherson, A. J., R. Moore, and B. Provencher, 2006. A dynamic principal-agent model of human-mediated aquatic species invasions. *Agricultural and Resource Economics Review* **35**:144–154.

Melbourne, B. a., H. V. Cornell, K. F. Davies, C. J. Dugaw, S. Elmendorf, A. L. Freestone, R. J. Hall, S. Harrison, A. Hastings, M. Holland, M. Holyoak, J. Lambrinos, K. Moore, and H. Yokomizo, 2007. Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? *Ecology letters* **10**:77–94.

Muirhead, J. R., 2007. Forecasting dispersal of nonindigenous species. Phd, University of Windsor.

Muirhead, J. R., M. A. Lewis, and H. J. MacIsaac, 2011. Prediction and error in multi-stage models for spread of aquatic non-indigenous species. *Diversity and Distributions* **17**:323–337.

Muirhead, J. R. and H. J. Macisaac, 2005. Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology* **42**:80–90.

Parker, I., D. Simberloff, W. Lonsdale, K. Goodell, M. Wonham, P. M. Kareiva, M. Williamson, B. Von Holle, P. Moyle, J. Byers, and L. Goldwasser, 1999. Impact:

toward a framework for understanding the ecological effects of invaders. *Biological Invasions* **1**:3–19.

Peck, S. L., 2004. Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in ecology & evolution* **19**:530–4.

Pejchar, L. and H. a. Mooney, 2009. Invasive species, ecosystem services and human well-being. *Trends in ecology & evolution* **24**:497–504.

Potapov, A., J. Muirhead, N. Yan, S. Lele, and M. Lewis, 2011. Models of lake invasibility by Bythotrephes longimanus, a non-indigenous zooplankton. *Biological Invasions* pages 2459–2476.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ricciardi, A., 2006. Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distributions* **12**:425–433.

Schneider, D. W., C. D. Ellis, and K. S. Cummings, 1998. A Transportation Model Assessment of the Risk to Native Mussel Communities from Zebra Mussel Spread. *Conservation Biology* **12**:788–800.

Smirnov, O. a. and K. J. Egan, 2012. Spatial random utility model with an application to recreation demand. *Economic Modelling* **29**:72–78.

Smith, V. and Y. Kaoru, 1986. Modeling recreation demand within a random utility framework. *Economics Letters* **22**:395–399.

Timar, L. and D. J. Phaneuf, 2009. Modeling the human-induced spread of an aquatic invasive: The case of the zebra mussel. *Ecological Economics* **68**:3060–3071.

Wittmann, M. J., M. a. Lewis, J. D. Young, and N. D. Yan, 2011. Temperature-dependent Allee effects in a stage-structured model for Bythotrephes establishment. *Biological Invasions* **13**:2477–2497.

Model	ΔAIC
GM	
$P^{GM}_{nj} = A_n W_j^{e+\beta_1 B1_n + \beta_2 B2_n} D^{-1}_{nj}$ *	0
$P^{GM}_{nj} = A_n W_j^{e} D^{-1}_{nj}$	62
RUM	
$U_{nj} = \beta_1 W_j + \beta_2 D_{nj} + (\beta_3 B1_n + \beta_4 B2_n) W_j + \epsilon$ §	3139
$U_{nj} = \beta_1 W_j + \beta_2 D_{nj} + \epsilon$	3187
NULL	
$P_j = 1/J$	14732

Table 2–1: **Model comparison by ΔAIC.** *A_n follows the same form as the parenthetical part shown in the table. §Shown are only the utility (both the observable and random) components of the random utility model. See equations 2.6-2.11 for full specification.

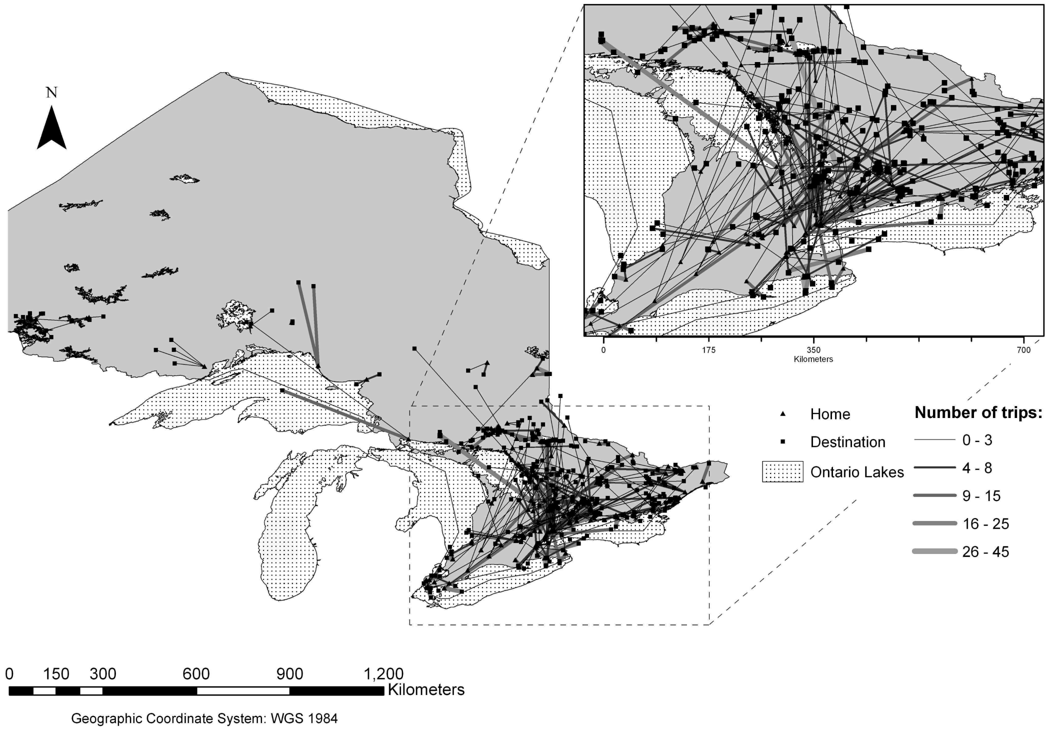


Figure 2–1: **Trips reported by survey respondents.** The location where boaters stored their boat during the boating season is indicated with closed squares. The thickness of the line home between location and the destination lake is proportional to the number of trips taken. The inset panel shows a zoomed in section of Southern Ontario for better visualization.

Parameter	MLE($\hat{\theta}$)	95% CI
Gravity Model (GM)		
e	0.51	[0.48, 0.52]
β_1	0.14	[0.083, 0.202]
β_2	-0.13	[-0.179, 0.0883]
d	1.86	[1.82, 1.89]
Random Utitlity Model (RUM)		
β_1	0.0011	[0.00106, 0.0012]
β_2	-1.40	[-1.447, -1.344]
β_3	0.00043	[0.000273, 0.000578]
β_4	-0.00044	[-0.000614, -0.000276]

Table 2–2: **Maximum likelihood parameter values and 95% confidence intervals for each human vector model.**

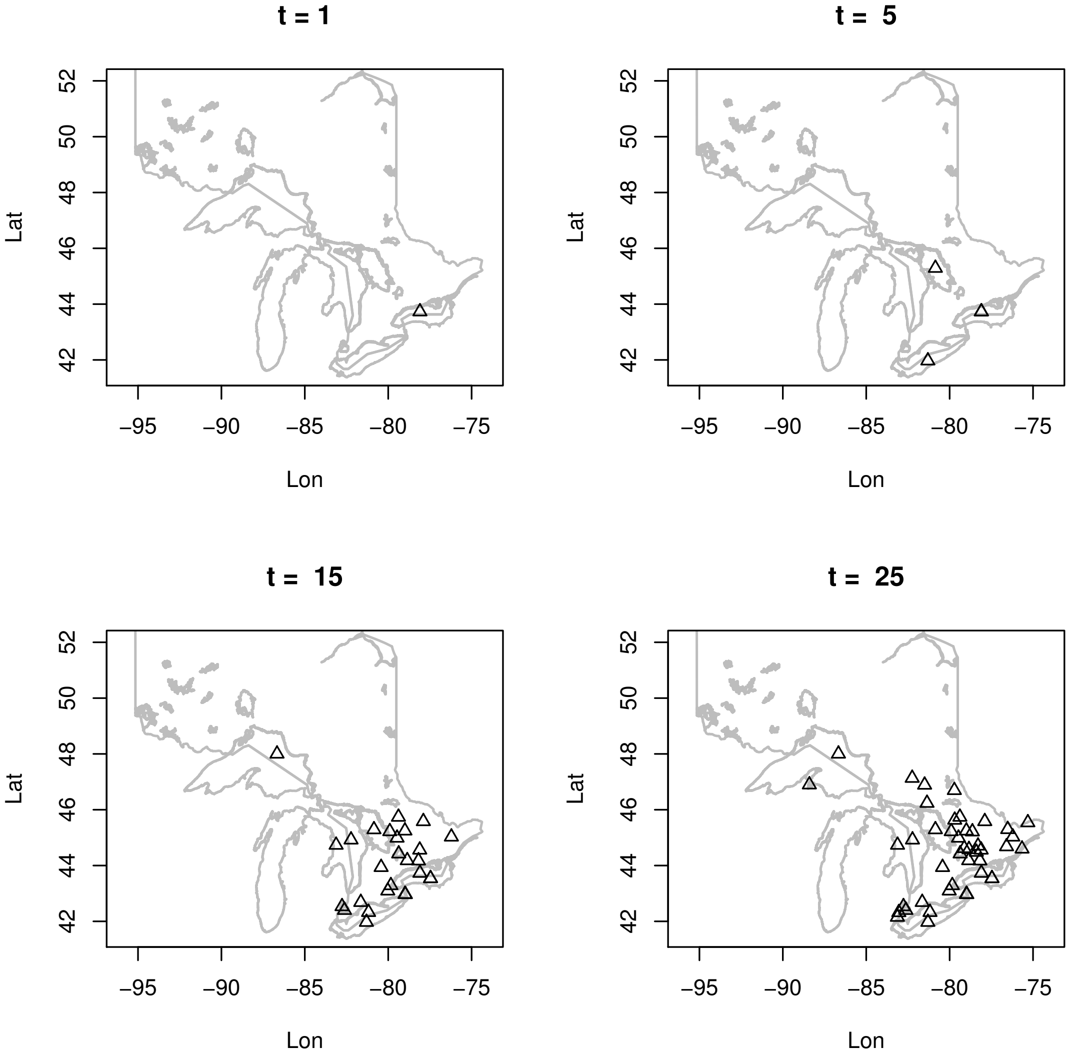


Figure 2–2: **Map of an example outcome of a spread simulation.** Triangles indicate lakes which have become invaded as of time t. Shown is a single realization of the spread process under the gravity model with parameters $\alpha = 1.25e-04, c = 2$.

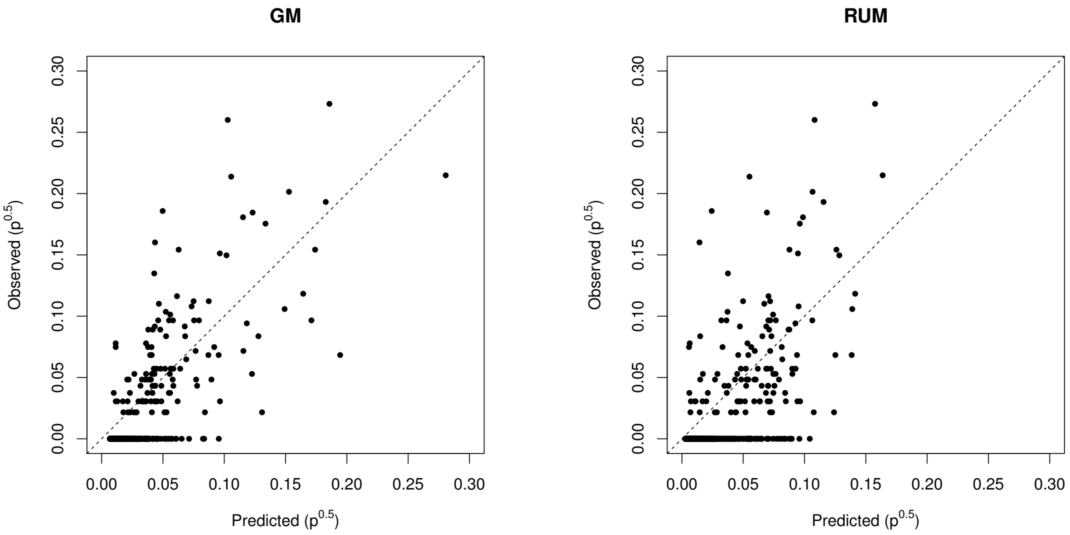


Figure 2–3: **Predicted vs. observed relative lake visitation frequency (p) for both models.** Both axes have been square-root transformed to better visualize low values. Coefficients of correlation (R^2) are 0.58, and 0.42 for the gravity (GM) and random utility (RUM) models respectively. The dotted line is the 1:1 line.

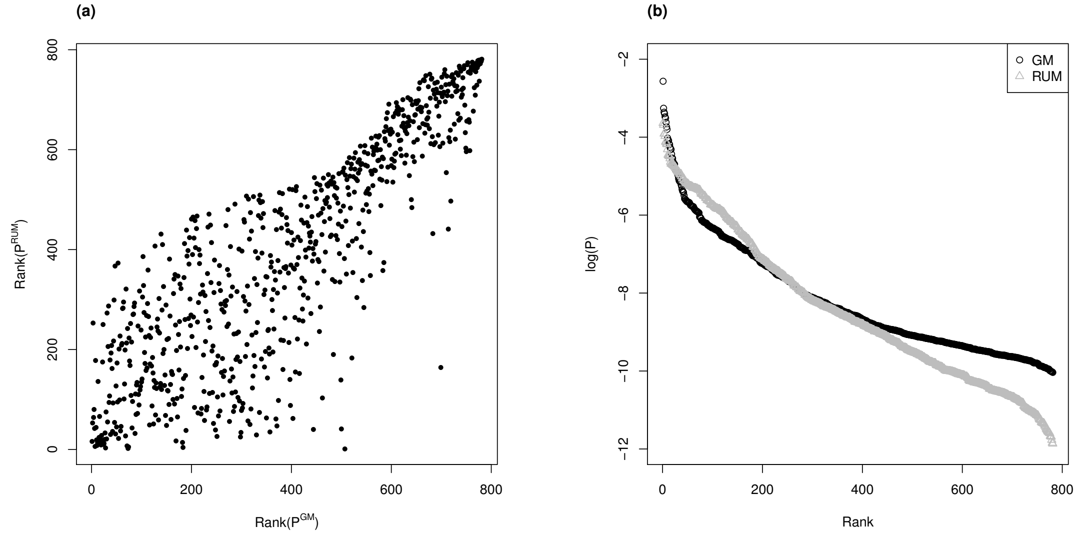


Figure 2–4: **Evenness of dispersal networks predicted by RUM vs GM.** Panel (a): Comparison of the ranking of predicted traffic to individual lakes for each model. Panel (b): Ranked predicted boater traffic distributions for the gravity (black circles) and random utility (grey triangles) models.

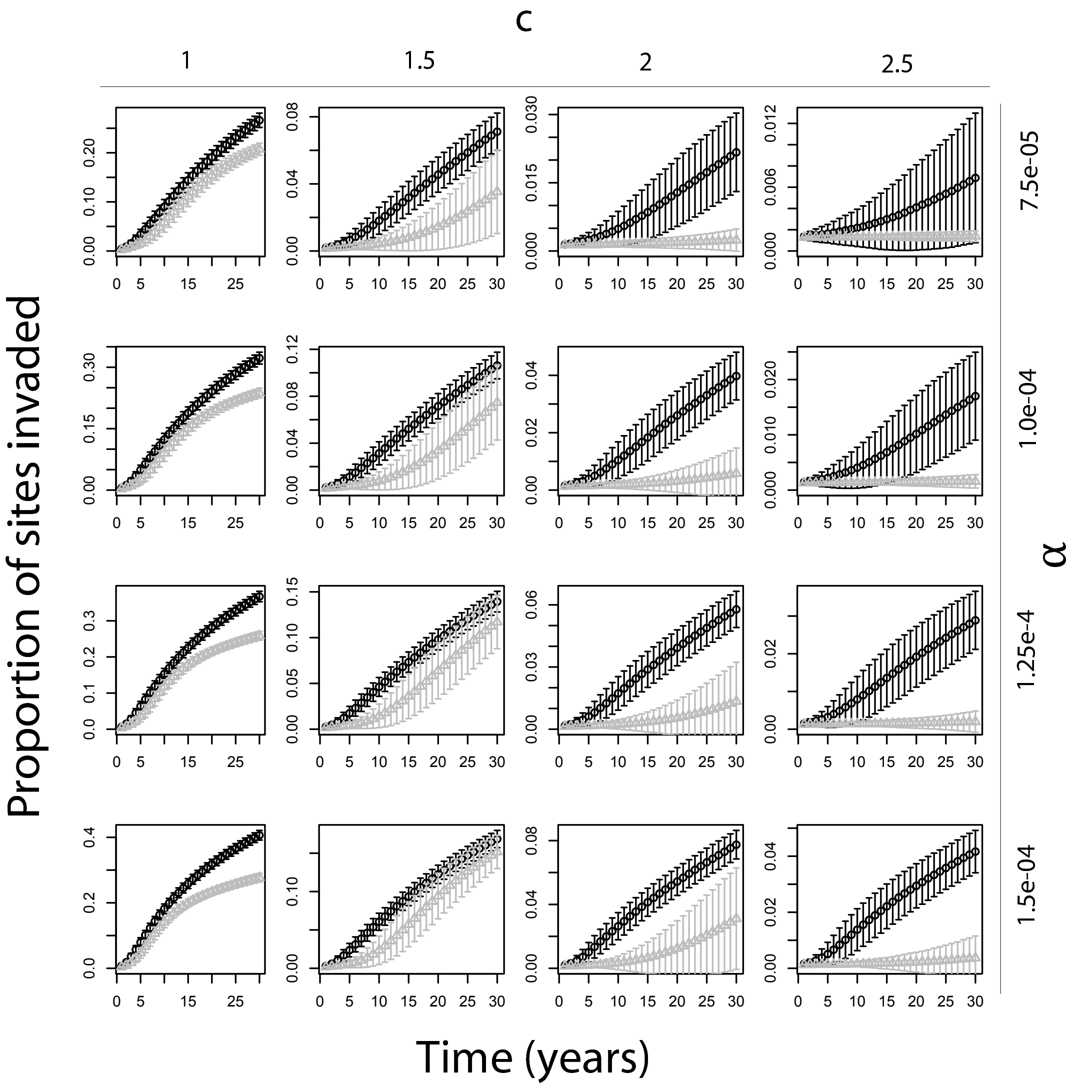


Figure 2–5: **Invasion trajectories.** Proportion of total number of lakes invaded predicted using the gravity model (black circles) and the random utility model (grey triangles). Panels show factorial parameter combinations of $\alpha = 7.5 - e05, 1.0e - 04, 1.25e - 04, 1.5e - 04$ for no Allee effect ($c = 1$), and increasing Allee effects of $c = 1.5, 2, 2.5$. Each model and parameter set was run for 1000 replicates. Error bars show the range encompassed by 95% of invasion simulations.

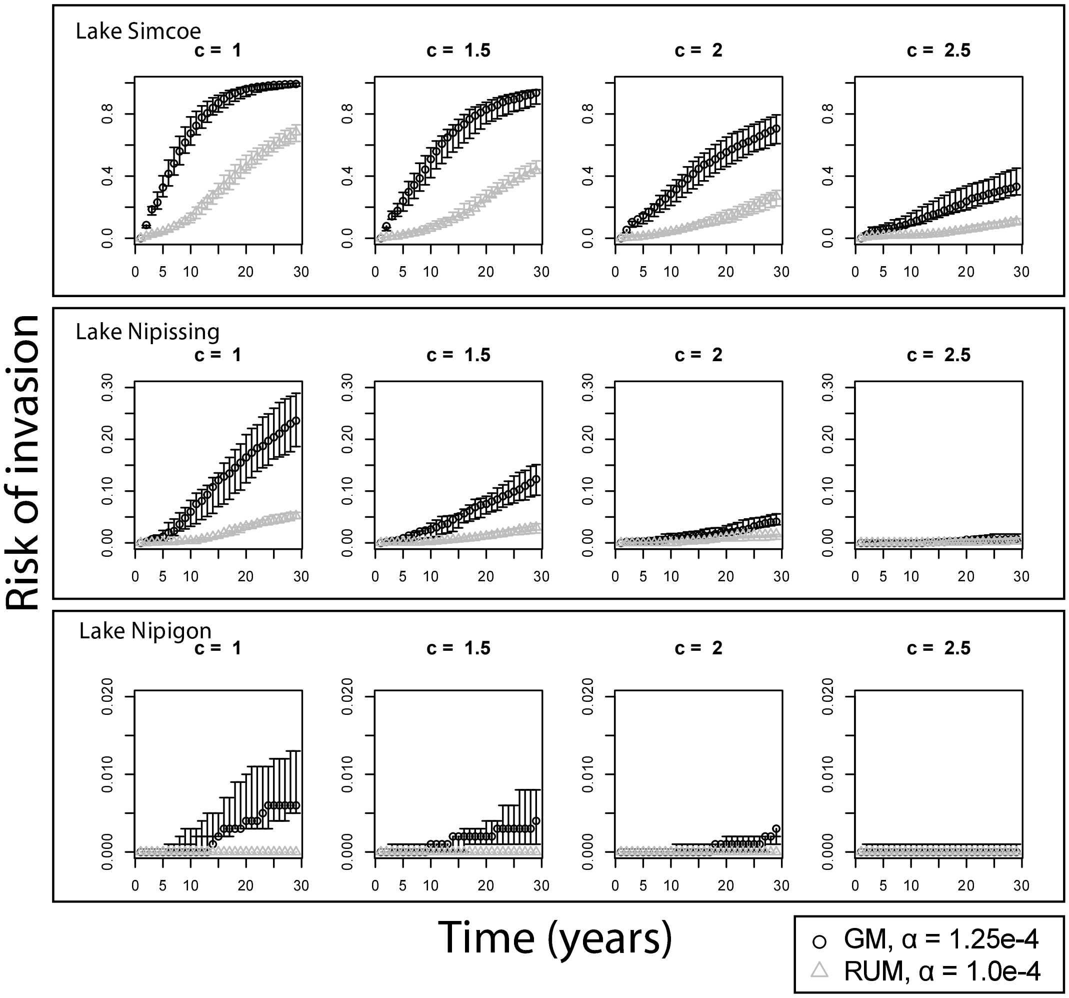


Figure 2–6: **Invasion risk to selected locations (Lakes Simcoe, Nipissing, and Nipigon).** Plots show the invasion risk (proportion of times invaded across 1000 spread simulations) across time with 95% confidence intervals. Panel columns show Allee effect increasing to the right. Alternative human vector models are shown in black circles (GM), and grey triangles (RUM).

Connecting statement

In the previous chapter, I investigated the consequences of employing alternative models of human-mediated dispersal on predictions of invasion risk at both the landscape and local level. I showed how the population dynamics, and particularly density dependent effects of the invading species can interact with the structure of the resulting dispersal network to significantly alter the projected risks. While the implications for management were suggested in this chapter, I will return to this aspect, and in particular spread mitigation efforts, in Chapter 5. The general models of human-mediated dispersal presented in the previous chapter provide a species-independent framework for generating projections of invasion risk. In the next chapter, I develop a novel method for evaluating predicted risks generated by such a model and look at a specific case study of the spread of *Bythotrephes longimanus* in a watershed in Central Ontario. The case study highlights both the applicability of the human-mediated spread risk projection framework and the validity of the underlying model for accurately assessing these risks. The novel methodology for evaluating the performance, and specifically the calibration, of risk projection models is developed in detail and is compared against two currently available approaches.

CHAPTER 3
Validation and calibration of probabilistic predictions in ecology

Corey Chivers[1] , Brian Leung[2] [3] , & Norman Yan[4]

A version of this chapter is currently in revision for publication in *Methods in Ecology and Evolution*.

---

[1] Department of Biology, McGill University.Stewart Biology Building 1205 ave Docteur Penfield Montreal, Quebec. 514-398-1833. corey.chivers@mail.mcgill.ca.

[2] Department of Biology, McGill University

[3] School of Environment, McGill University. brian.leung2@mcgill.ca

[4] Department of Biology, York University, Toronto, ON. nyan@yorku.ca

3.1 Abstract

1. Predictive models in ecology are important for guiding policy and management. However, they are necessarily abstractions of natural systems, making predictive validation imperative. Models which make predictions about binary outcomes (eg. Species distribution models, population viability analysis, predictive disease/invasion models) are widespread in the ecological literature. When supporting probability-based management decisions, these predictions need to be assessed with respect to the degree to which predicted probabilities agree with future outcomes. Many predictive models are not validated using external data, and are often only assessed in terms of their ability to discriminate between outcomes rather than the degree to which they predicted the correct probabilities, and by extension the correct risk.
2. We develop a novel Validation Metric Applied to Probabilistic Predictions (VMAPP) which provides a goodness-of-fit test for binary probabilistic prediction models. We analyse the theoretic properties of this test, and compare its performance against existing methods. We demonstrate the utility of VMAPP by application to a published model in invasion biology which forecasts the establishment probability of the zooplanktivorous spiny water flea (*Bythotrephes longimanus*). We selected 102 additional sites to sample four years after the training data was collected and use this independently collected data to assess predictive reliability using VMAPP.
3. Theoretic simulation analysis shows that VMAPP outperforms existing metrics (Cox's regression technique and Hosmer & Lemeshow's χ^2 test) in terms

of statistical power to identify model miscalibration. Further, we find that under realistic conditions where model parameters are estimated (and have associated uncertainty) that VMAPP is more robust, retaining the appropriate Type-I error rates (5%) where previous metrics fail ($\leq$17%) . Application of VMAPP to a published invasion model using empirical validation data show that in addition to having high discriminative power, the model's probabilistic predictions agree with the observed outcomes as measured by VMAPP.

4. We argue that quantifying ecological predictions as probabilities with associated uncertainty provides the most useful information to support management decisions. Ecological predictions, while uncertain, should still be rigorously validated. Identifying the circumstances in which our predictions deviate from observation can further inform the next generation of the model, bringing prediction and reality ever closer.

3.2 Introduction

From identifying the risk of species loss (Lee and Jetz, 2011), to predicting range shifts under climate change (Chen et al., 2011), or forecasting the spread of invasive species (Gertzen and Leung, 2011a), ecology has the potential to contribute greatly as a predictive science. Using ecological models to make predictions about unknown or future states of nature is an essential component of informing policy related to global change (Clark et al., 2001; Burgman and Yemshanov, 2013). In order for the predictions of ecological models to remain relevant to decision makers, however, we need to rigorously assess the validity of the predictions generated from our models (Schmolke et al., 2010). This is most readily apparent when ecological predictions aim to inform risk management. Since risk can be defined as the probability of an event times the severity of its impact (Suter, 2006), we need to know whether our models are providing us not with some exact single prediction of a given outcome (eg. definitive presence or absence, alive or dead, extant or extinct, etc.), but rather with the correct *probabilities* associated with each. Thus, using a simple threshold approach to separate predictions into binary classifications, one would rule all cases with low predicted probability (e.g. $\hat{p} = 0.1$) as "absent", however if the model is correct, we should observe 10% of these cases as actually being present.

Ideally, a procedure to assess predictive validity should have three key aspects. First, the type of predictions on which a model is validated should be in line with it's stated purpose. If a model is built with the goal of being able to predict outcomes beyond the spatial or temporal domain of the data on which it was built, then validation should be carried out on outcomes similarly beyond these ranges (Guisan and Zimmermann, 2000). Secondly, the validation procedure should examine the range of model predictions (from low to high predicted probabilities), and be able to identify regions of significant departures of model predictions from observations (Pearce and Ferrier, 2000; Phillips and Elith, 2010). Third, the validation metric should provide an absolute measure of predictive performance, as opposed to the relative measures commonly used for model selection (Mason, 2008). Finally, it should be powerful. Ecology, like all endeavours, is subject to the ubiquitous constraint of limited resources. A validation procedure should therefore make efficient use of the limited data available to most powerfully detect discrepancies between model predictions and reality, when they exist.

In this paper, we focus on evaluating models that make predictions on binary outcomes (i.e. the outcomes of processes which lead to one of two possible states). Such models are widely used in the ecological literature, yet the evaluation of these models has been limited with respect to the validation discussed above. Examples of this class of ecological models include species distribution models (Elith et al., 2011; Wenger et al., 2013; Broennimann et al., 2012), population viability analyses (Brook et al., 2000; Staples et al., 2004; Heard et al., 2013), and establishment of invasive species (Gertzen and Leung, 2011a; Seebens et al., 2013; Bradie et al., 2013). In many cases, making reliable probabilistic predictions is essential from a decision making perspective (but see Lawson et al. (2014) and Kuk et al. (2014) for discussion of cases where correct ranking of probabilities is sufficient). Whether we are employing expected value theory in a cost-benefit analysis or defining acceptable levels of risk, we are often relying on the estimates of probability from our models to be accurate.

Validation of probabilistic models involves assessing the quality, or goodnesss-of-fit, of generated predicted probabilities in comparison with actual observations. While there are nine attributes that define the quality of a probabilistic model (Murphy, 1993), often we focus on examining discriminative ability, and/or calibrartion (also termed reliability) (Lawson et al., 2014). The goal when building binary outcome models is often viewed as being able to maximally discriminate between the two possible states of nature across a variety of scenarios or conditions. When discrimination is the primary goal, we say that the problem is one of binary classification (Freeman and Moisen, 2008). In this view, the modeller's objective is to place very high scores on positive cases, and conversely, low scores negative cases. This objective is conceptually similar to the concept of sharpness, in which the predicted probabilities are concentrated around 0 and 1 (in the case of binary outcomes), however differs in that while sharpness is a property solely of the prediction, discrimination is a joint property of the prediction and the outcomes (Gneiting and Katzfuss, 2014). Alternatively, the prediction problem can be viewed as an assessment of risk, with the goal being to have predicted probabilities match the actual probabilities associated with each case, given the available data. That is, we want the predicted probabilities, based on the available predictor variables and a given model structure, to be in agreement with the true probabilities (ie. a well calibrated model, also referred to as a reliable model) (Phillips and Elith, 2010; Pearce and Ferrier, 2000). As the simplest example, a well calibrated prediction on the outcome of a fair coin landing heads-up would be 0.5, and we could statistically compare this prediction against observed outcomes. While prediction calibration is clearly an important aspect of model validation, a recent review found that a majority of species distribution modelling studies are only evaluated in terms of discrimination, in part because common methods of modelling species distributions use presence-background data and do not purport to make well calibrated predictions (Ward et al., 2009; Lawson et al., 2014).

In this paper, we examine two existing goodness-of-fit tests for predictive calibration used in Ecology and demonstrate that these approaches may be flawed under realistic conditions. We provide a solution by developing a novel approach: the Validation Metric Applied to Probabilistic Predictions (VMAPP), which compares probabilistic model predictions to observed outcomes from an independent data set. Our metric provides a goodness-of-fit test against four forms of model miscalibration which is more powerful than previously available techniques. The goodness-of-fit measures assess whether model predictions are biased overall, as well as whether the magnitude and direction of bias changes across the prediction range. We demonstrate that VMAPP is able to incorporate both stochastic and parametric uncertainties inherent to predictive models when assessing model fit and show that that previous methods suffer from inflated type-I errors under these conditions. In addition to the goodness-of-fit tests, VMAPP also provides an estimator of the magnitude and direction of deviation between predicted and actual risks for use as a diagnostic tool similar to a residual plot in classical statistical models. All of these outputs are easily obtained using an R package (R Core Team, 2013) which we have developed to accompany this paper.

3.3 Materials & methods

We first outline the logic of our metric and describe a recipe for how it is computed. We then assess the properties of VMAPP theoretically, testing its performance in terms of statistical power to identify miscalibrated predictions, as well as evaluating the accuracy of the estimator of model deviations from reality across the range of model predictions. We then demonstrate the utility of VMAPP via application to a published model in invasion ecology for which validation data has since been collected (Gertzen and Leung, 2011b).

3.3.1 Validation metric

The different ways in which a model may perform well in terms of its ability to discriminate between positive and negative events, yet place erroneous probabilities on individual predicted outcomes can be categorized into a discrete set of miscalibration pathologies as described in Pearce and Ferrier (2000) and revisited in Phillips and Elith (2010). First, predicted probabilities may be biased overall higher (or lower) than the observed rates of presence events. These cases are represented graphically using stylized examples in Fig. 3–1, panels D) and G). In these situations, discrimination may indeed be quite high, yet the predicted rates are everywhere biased in the same direction. Second, the spread of the predictions may be biased relative to the spread of the generating probabilities. The model may predict overly low probabilities in the low range of predicted probabilities, and overly high probabilities in the high range (Fig 3–1 M). This situation can occur, for instance, in simple logistic regression when a covariate in the model is only loosely correlated with the true predictor. Finally, model predictions may be higher than the true probabilities in the low range and lower than true probabilities in the high range (Fig 3–1 E). In this situation, model discrimination may also be low, as the distribution of predicted probabilities is narrower than the true probabilities. While model predictions may be miscalibrated in an infinite variety of ways (for instance linearity error (Phillips and Elith, 2010)), we focus on this discrete set in order to build an inferential mechanism for identifying these common pathologies.

To explain how VMAPP assesses these miscalibration pathologies, and estimates the degree of model miscalibration at each point in the predictive range we first define a few terms: For every outcome in a validation set consisting of n observations ($R_i \in \{0,1\}, i = 1, \ldots, n$), the true (but unknown) probability is denoted by $p_i = P(R_i = 1)$, and predicted probabilities given a model (or ensemble of models) M and some data D by $\hat{p}_i = P(R_i = 1 \mid M, D)$ for each $i = 1, \ldots, n$. We begin with the case where our model makes point estimates of each probability, but will generalize this to consider uncertainty in each $\hat{p}_i$ below. Of course, we do not observe the real-valued p_i directly, but rather observe binary outcomes ($R_i \in \{0,1\}$) in our validation data. To compare our predictions with these binary observations we can simulate an arbitrarily large number (J) of predicted outcomes ($S_{ij} \sim Bernoulli(\hat{p}_i),\ j = 1, \ldots, J$) using probabilities from the model. The degree of miscalibration (or deviation) is the difference between model and reality for each validation case and is defined as $\delta_i = \hat{p}_i - p_i$.

In order to assess miscalibration pathologies, VMAPP compares each vector $\boldsymbol{S}_j, j = 1 \ldots, J$ of simulated outcomes from the predictive model to the vector of real observations $\boldsymbol{R}$ in the validation data. Of course, because we are dealing with stochastic events, we do not necessarily expect the simulated and observed occurrences to always match even when the model is valid, and VMAPP focuses on these discrepancies ($S_{ij} \neq R_i$) when they occur. Under the null hypothesis that our model perfectly characterizes the underlying probabilities ($\hat{p}_i = p_i,\ i = 1, \ldots, n$), we would expect that when our simulated outcomes differ from the data, they will do so as frequently in one direction as the other. In other words, $S_{ij} > R_i$ ($S_{ij} = 1$, $R_i = 0$) is as likely as $S_{ij} < R_i$ ($S_{ij} = 0$, $R_i = 1$). In contrast, if model predictions overestimate the probability ($\hat{p}_i > p_i$), then we would expect $S_{ij} > R_i$ to occur more frequently than $S_{ij} < R_i$ (and the reverse scenario for underestimation $\hat{p}_i < p_i$) (Fig.3–1 F,I).

Next, if the direction of deviation changes in sign over the range of model predictions, $\boldsymbol{\delta}$ will either be an increasing or decreasing function of $\hat{\boldsymbol{p}}$ (Fig. 3–1. J,M). Specifically, if $\hat{p}_i < p_i$ for all i where $\hat{p}_i$ is small, but switches to $\hat{p}_i > p_i$ for all i where $\hat{p}_i$ is large, then the proportion of $S_{ij} > R_i$ to $S_{ij} < R_i$ will likewise increase over the range of $\hat{p}$ (Fig. 3–1. L), and *vice versa* in the other direction (Fig. 3–1. O).

Using the above logic, we can assay for model miscalibration by testing two null hypotheses. First, when $\hat{p}_i = p_i$ for all $i = 1, \ldots, n$ then $P(S_{ij} > R_i \,|\, S_{ij} \neq R_i) = 1/2$. Thus, the expected proportion of positive discrepancies to negative ones is equal to 1/2 for every outcome in the validation set. By repeatedly simulating outcome vectors $\boldsymbol{S}_j, j = 1, \ldots, J$ from each element $i = 1, \ldots, n$ of the prediction vector $\hat{\boldsymbol{p}}$, each time calculating this proportion, we can assess how often it is greater than, or less than the null hypothesis of 1/2. If it is greater(less) than 1/2 most (ie $1 - \alpha$) of the time, then the model predictions overall are significantly biased high(low). Second, if there is no change in the direction of bias over the range of $\hat{\boldsymbol{p}}$, we would expect $P(S_{ij} > R_i \,|\, S_{ij} \neq R_i)$ to be independent of $\hat{p}$. The alternative hypothesis being that $P(S_{ij} > R_i \,|\, S_{ij} \neq R_i)$ is either an increasing, or decreasing function of $\hat{\boldsymbol{p}}$. We accomplish this test by fitting a logistic function to each of the J vectors of discrepancies between simulated and observed outcomes. If the slope is greater(less) than zero more than $1 - \alpha$ of the time, then we reject the null hypothesis that $\boldsymbol{\delta}$ is independent of $\hat{\boldsymbol{p}}$, and conclude that model predictions switch from being significantly biased low(high) for small values of $\hat{\boldsymbol{p}}$, to being high(low) for larger values of $\hat{\boldsymbol{p}}$. An algorithm for conducting these VMAPP goodness-of-fit tests is given in box 1.

Accounting for uncertainty

Up to this point, we have developed the logic behind VMAPP as a tool for identifying whether model predictions are in concordance with the probabilities underlying the system being modelled. In doing so, we have taken explicit account of the stochastic uncertainty underlying the conversion of continuous probabilities into the observed binary outcomes in our validation data. Missing however, is the acknowledgement that even if we have constructed a valid model of our system, we have typically estimated the parameters of that model using a sample data. The resulting estimates, of course, have associated uncertainty as a result of sampling variation. The consequence of which is that even when our model is valid, we do not expect $\hat{p}_i$ to exactly equal $p_i,\ i = 1, \ldots, n$. Since what we wish to know is whether our *model* is valid, failure to account for this sampling uncertainty may result in the erroneous rejection of a valid model due to errors in parameter values.

There are several techniques to account for parameter uncertainty due to sampling variation in model predictions (Cressie et al., 2009; Meyer et al., 1986; Wenger et al., 2013), but the underlying concept is the same. Rather than predicting a single probability for each new case, propagation of uncertainty due to sampling variation entails predicting a *distribution* of probabilities. In this study, we use the method of bootstrap re-sampling to build uncertainty distributions of predicted probabilities for each case in the validation data. Now, instead of having a single prediction vector $\hat{p}_i,\ i = 1, \ldots, n$, we have a distribution of J prediction vectors denoted $\hat{p}_{ij} \sim P(\hat{p}_i \,|\, M, D),\ i = 1, \ldots, n\ j = 1, \ldots, J$.

Box 1. Algorithm steps for VMAPP miscalibration tests:

1. FOR j in $j = 1, \ldots, J$, DO:
2. Simulate a predicted outcome for each point in the validation set ($S_{ij} \sim Bernoulli(\hat{p}_{ij}),\ i = 1, \ldots, n$). If there is no parameter uncertainty accounted for in the model, $\hat{p}_{i1} = \hat{p}_{i2} = \ldots = \hat{p}_{iJ}$.
3. Compute the difference between the simulated and observed outcomes ($\boldsymbol{S}_j - \boldsymbol{R}$). Map $S_{ij} > R_i \rightarrow \Delta_{ij} = 1$, and $S_{ij} < R_i \rightarrow \Delta_{ij} = 0$. Exclude any cases where $S_{ij} = R_i$ (Fig. 3–1, row 3).
4. Calculate the mean discrepancy direction $\bar{\Delta}_j = \sum_{i=1}^{n} \Delta_{ij}/n$.
5. Fit a logistic regression $ln\frac{P(\Delta_{ij})}{1-P(\Delta_{ij})} = \beta_{0j} + \beta_{1j}\hat{p}_{ij}$, and retain the fitted slope value β_{1j}.
6. END FOR LOOP.
7. Compare the distribution of $\bar{\Delta}$ to $H_o : \bar{\Delta} = 0.5$. If $\bar{\Delta}_j$ greater(less) than 1/2 more than $1 - \alpha$ of the time ($j = 1, \ldots, J$), model predictions are significantly biased high(low) overall (See Fig. 1, columns 2 & 3).

8. Compare the distribution of β_1 to $H_o : \beta_1 = 0$. If miscalibration changes direction from under- to over-prediction across the range of prediction, $H_A : \beta_1 > 0$ (Fig. 3–1, column 4). Similarly, a change from over to under is $H_A : \beta_1 < 0$ (Fig. 3–1, column 5).

3.3.2 Theoretic analysis of VMAPP

Power and type-I error rates

We used simulation tests to evaluate the theoretic behaviour of VMAPP, and compare its performance with both the the Hosmer-Lemeshow (HL) test (Hosmer and Lemeshow, 2000) and Cox's method (Pearce and Ferrier, 2000) (see appendix B.1.2 for a description of these methods). First, we establish that the type-I error rates for identifying model miscalibration pathologies match with the expected rates. To do this, we first assumed a model which generates predictions perfectly concordant with the probabilities underlying the validation data (ie. $\hat{p_i} = p_i,\ i = 1, \ldots, n$). This represents the special case where the model is not only valid, but its parameters are known with complete certainty. We simulated a validation procedure, drawing 100 points from the underlying probabilities and calculated each of the validation metrics. We repeated this 1000 times and calculated the proportion of times each metric resulted in the rejection of the null hypothesis of a valid model at the $\alpha = 0.05$ level. To assess power to detect model miscalibration, we repeated the procedure using systematically biased model predictions. We generated predictions following each of the four miscalibration pathologies outlined in section 3.3.1 at medium and high levels of absolute deviation (Fig. 3–3).

To test the performance of each metric under the more realistic scenario of a model parameterized using empirical observations, we repeated the simulation experiment, this time introducing a model fitting step to sample data. We simulated 100 data points at a time from a simple logistic model with a single predictor variable. Using this data, we fit the parameters of the correct model formulation using maximum likelihood and used the fitted model to make predictions on another set of 100 validation data points, also simulated from the true underlying model. Repeating this 1000 times, We then assess the type-I error rates by calculating the proportion of times that each metric detected miscalibration.

To incorporate uncertainty caused by sampling variation in the data, we also considered a bootstrapping procedure to generate a distribution of parameter estimates (Manly, 2001). In each replicate of "reality", the model was refit 1000 times, each time using 100 samples (with replacement) from the 100 data points. The resulting distribution of parameter values was then used to predict a distribution of probabilities for each case in the validation set. We repeated this procedure 1000 times to test the type-I error rates of VMAPP and Cox's method. The HL test, however, has no obvious mechanism for testing predictions which include parameter uncertainty and so was not included in this final comparison.

Estimation of $\boldsymbol{\delta}$

In addition to the goodness-of-fit tests, VMAPP also provides an estimation of the magnitude and direction of deviations ($\boldsymbol{\delta}$) between model predictions and observed frequencies in the validation data. This procedure is analogous to calibration plots by Phillips and Elith (2010) for presence-only data, based on Cox's method

(Pearce and Ferrier, 2000), where the frequency of positive validation outcomes is plotted against the predicted probability from a model. By estimating $\boldsymbol{\delta}$, predictive calibration can be inspected over the range of predictions, potentially identifying miscalibration behaviour not explicitly accounted for in the goodness-of-fit tests. To accomplish this we need to estimate two functions using our Monte Carlo outcomes. First, $f_1(\hat{\boldsymbol{p}})$ estimates the probability of discrepancies irrespective of direction, or $P(\boldsymbol{S} \neq \boldsymbol{R})$ (Fig. 3–1 2nd row). Next, $f_2(\hat{\boldsymbol{p}})$ estimates $P(\boldsymbol{S} > \boldsymbol{R} \,|\, \boldsymbol{S} \neq \boldsymbol{R})$ as a function of $\hat{\boldsymbol{p}}$ (Fig. 3–1 3rd row). We use parametric functional forms which capture the expected range of shapes of $f_1(\hat{\boldsymbol{p}})$ and $f_2(\hat{\boldsymbol{p}})$ (see Appendix B.1.1) and fit using maximum likelihood to each set of discrepancies between the simulated and observed outcomes. With these two function estimations in hand, we can calculate the estimated deviation as a function of the model predictions:

$$\hat{\boldsymbol{\delta}}(\hat{\boldsymbol{p}}) = 2(f_2(\hat{\boldsymbol{p}}) - 1/2)f_1(\hat{\boldsymbol{p}}). \tag{3.1}$$

The proof of which is given and an algorithm for estimating f_1 and f_2 is given in Appendix B.1.1. Code is provided as an R package (R Core Team, 2013) available at `https://github.com/cjbayesian/rvmapp`.

To assess the ability of VMAPP to estimate the magnitude and direction of deviation between model predictions and reality (which we simulated and therefore knew), we simulated random scenarios representing arbitrary combinations of each of the four miscalibration pathologies. For each simulation, we applied VMAPP to the miscalibrated predictions on an independent validation dataset. Each time, VMAPP generates a distribution of the estimated deviation $\hat{\boldsymbol{\delta}}$. Since the estimate is not a single value, but rather a distribution representing uncertainty, we evaluate whether this distribution is predictive of the true deviation using the approach outlined in Leung and Steele (2013). This approach tests for bias in both location and dispersion using a probability-probability ($P - P$) plot (Leung and Steele, 2013). If $\hat{\boldsymbol{\delta}}$ is an unbiased estimator, we expect the true value $\boldsymbol{\delta}$ to fall below 1% of the cumulative distribution function (cdf) of $\hat{\boldsymbol{\delta}}$ 1% of the time, below 10% of the cdf 10% of the time, and so on. The expected result if no bias exists is a P-P plot following a 1:1 line between the expected percentile and the proportion times, under repeated simulation, that the real values of $\boldsymbol{\delta}$ fall into these percentiles.

3.3.3 Application: Spread of the invasive planktivore *Bythotrephes longimanus*

Watershed Description and background

Watershed 2EB comprises the District of Muskoka and parts of the counties of Haliburton, Nipissing and Parry Sound in south-central Ontario, Canada (Fig. 3–4). There are 1636 lakes that are > 1 ha in the watershed. The lakes in the watershed have the longest history of secondary spread of *Bythotrephes* in North America (Yan et al., 1992; MacIsaac et al., 2004).

Data collection

Prior to 2005, 23 lakes in the watershed were known to be invaded with *Bythotrephes.* In 2005 and 2006, 311 spatially and size stratified lakes were sampled for the presence/absence of *Bythotrephes* (see appendix B.1.3). This sampling effort increased the number of known invaded sites from 23 to 46. The predictive model outlined in the next section was built using the data collected during 2005/2006 as well as the historical record of confirmed invasions dating back to 1989. For validation, an additional 102 lakes were selected to be compared against the predictions from from the model. The validation lakes were chosen to cover the full extent of the prediction range (low, medium, and high predicted probabilities) and were sampled in 2010.

Predictive modelling framework

The modelling framework for forecasting the spatial spread of *Bythotrephes* is described in detail in Gertzen and Leung (2011b). The methodology incorporates several levels of real world uncertainty involved in predicting the trajectory of the invasion process. First, while the data consists of presence/absence records, only a small fraction of total possible sites have been sampled, with many unsampled sites representing unknown potential sources or sinks of propagules to the system. Second, the sites which have been sampled (one or more times) represent observations at relatively few, non-evenly distributed time intervals. A site at which a species is observed to be present when sampled at time t may have become colonized at any time between when it was last observed to be free of invasion and t. Third, while the probability of establishment will increase with the inbound propagule pressure to a given site, the success or failure to establish a persistent population is subject to demographic and environmental stochasticity. By using maximum likelihood estimation in concert with stochastic simulation, this model was fit to the *Bythotrephes* data collected up to 2006.

3.4 Results

3.4.1 Theoretic properties of the validation metric

Type-I error rates and Power - For the null case where we have the true model structure and know the true parameters (Fig. 3–3), our simulation results show both VMAPP and Cox's method to have the expected type-I error rate for detecting model miscalibration (5% for $\alpha = 0.05$). The rate of type-I errors for the HL test was an inflated 11%. When model parameters were fit, but parameter uncertainty (sampling variation) not accounted for, all tests resulted in inflated type-I errors (HL: 27%, Cox: 23%, and VMAPP: 16%). When parametric uncertainty was accounted for, VMAPP yielded the appropriate type-I errors yet Cox's method still had an inflated type-I rate of 17% (Table 3–1).

The power of each metric to detect miscalibration increased as the size of deviation increased for all four types of model miscalibration. VMAPP outperformed both Cox and HL. VMAPP obtained a power of 94% in the high category of underestimation compared with 84% for Cox's method and 72% for HL (Fig. 3–3 A & B). Power comparisons were similar for cases where the direction of bias switched from under to overestimation across the predictive range (at high deviation, the power of the tests was 94%, 91%, and 90%, respectively) (Fig. 3–3 C). Both VMAPP and Cox's method outperformed HL by the widest margin when the direction of bias switched from over to underestimation across the predictive range (94% and 40%, respectively) (Fig. 3–3 D).

Estimation of δ - P-P plots showed that the estimator is not biased in either spread or location, with observed percentiles following closely with expected along the 1:1 line (Fig. 3–2).

3.4.2 Empirical results

The model predicted probabilities of *Bythotrephes* establishment at the validation lakes were consistent with the observed establishments. Discrimination was also high, with AUC values for the bootstrapped predictions of 0.82 ± 0.3. Application of VMAPP to the model predictions and validation data did not detect any significant overall bias ($p = 0.50$) nor any changing bias across the predictive range ($p = 0.37$).

Fig. 3–4 shows the predicted probabilities from the model on the 102 validation lakes. While the estimated deviation between model predictions and reality are small, Fig. 3–4 B) demonstrates how this VMAPP product can be mapped and visually inspected for spatial patterns in model deviation. For example, we can see that the maximum expected deviation between our model predictions and actual risk (3.4%) occurs at the lake which the model predicted most likely to be invaded. Inspecting the 95% CIs, we see that the model could be underestimating the risk by as much as 17% or overestimating by as much as 16%. Again, we see that the uncertainty in this value is greatest at predicted risks near 50%.

3.5 Discussion

If ecology is to provide support for management and policy, the credibility of ecological predictions is essential. This requires rigorous assessments of our predictive models using independent validation data (Manel et al., 2002; Schmolke et al., 2010). For some predictions, the goal is to accurately capture a 'snapshot' in time, requiring validation data collected concurrently with the data on which the model was fit. In cases where our models are being used to forecast systems which are in flux (non-equilibrium - eg. invasions, climate change driven range shifts) into the future, it is particularly important to have temporal separation between fitted and validation data sets (Václavík and Meentemeyer, 2012; Márcia Barbosa et al., 2013). Further, predictive models which generate probability estimates need to be evaluated in terms of the degree to which the probabilities themselves are accurate, if they are going to provide meaningful inputs into calculations involving risk.

There is a paucity of studies which have validated whether the quantitative probabilities from models matches the empirical binary outcomes (Lawson et al., 2014), and we advocate that this should become common practice. Our work in this manuscript will aid such validation by developing a novel approach (VMAPP) to evaluate any binary probabilistic model. We have demonstrated via theoretical simulations that it is simultaneously more robust (appropriate type-I errors) and more powerful than two existing approaches. We have also shown that our method can be used to estimate the magnitude of deviation of probabilistic predictions from reality using empirical observations, along with the associated uncertainty in these estimates (Jolliffe, 2007). Moreover, we applied it to a real-world non-trivial ecological model (Gertzen and Leung, 2011b), comparing *a priori* model predictions to >100 independently sampled locations for validation. A complete validation of binary prediction models in ecology using data independent of that used to construct a model is rarely carried out in practice (Guisan and Thuiller, 2005). Most studies, at best, use a data partitioning scheme to conduct internal validation on a single data set. That is, validation is conducted by fitting a model to some subset of the data and holding out the rest on which to assess model predictions (i.e., cross-validation). While this is the best that can be achieved in many situations where only a single dataset is available and VMAPP can be applied in this setting, the possibility of bias resulting from building models to ”chase” internal validation performance is problematic (Hirsch, 1991) and has been observed to lead to inflated performance when compared against independently collected data (Randin et al., 2006; Araújo et al., 2005). Here, we assessed a previously published model using independent data collected four years after the data on which the model was built and demonstrated that the model yielded probabilistic predictions which matched observed outcomes. We thereby provided an account of model performance which is not subject to unintended biases which can occur when the analyst has access to all of the data at once (ie. researcher degrees of freedom). Additionally, the team which carried out the validation sampling was 'blind', with guidance on which sites to sample but no knowledge of model predictions (Cairns and Yan, 2011).

3.5.1 Comparison with other approaches to validation

Some previous validation studies have employed non-probabilistic measures of agreement between model predictions and observations. These have included measures of spatial agreement between predictions and observations or summary statistics such as the net rate of spread in the case of biological invasions (e.g., Higgins et al. (2001)). Alternatively, depending on the objectives of a modelling exercise, various other validation metrics may be appropriate (Guisan and Thuiller, 2005). For instance, some management plans seek to simply rank sites by the likelihood of establishment in order to focus efforts and resources on the top n sites. In such cases, commonly used metrics such as Area Under the receiver operating charactaristic Curve (AUC), which indicates whether ranking of predictions fits with observed outcomes, is most useful (Peterson et al., 2008). However, ranking alone may be undesirable if the absolute magnitude of the difference in risks between sites is small. More generally, if one wished to determine whether an acceptable environmental risk threshold were exceeded, weigh costs versus benefits of management, or any other analyses of expected utility, it would be critical that the predicted probabilities from models are quantitatively correct. The validation of probabilistic predictions provided by VMAPP will be most applicable in these cases, and it is therefore complementary to measures such as AUC. Thus, for instance, in our real-world application to *Bythotrephes* spread in Ontario (Gertzen and Leung, 2011b), management options involving costly interventions could be evaluated using cost-benefit analysis, where the risk posed by *Bythotrephes* invasions requires estimates of the probability of their establishment over time.

VMAPP improves upon the few existing inferential methods to evaluate probabilistic predictions, namely the Hosmer-Lemeshow (HL) (Hosmer and Lemeshow, 2000) test and Cox's method (Miller et al., 1991; Pearce and Ferrier, 2000). The HL test is simple in its conception and implementation, however it suffers from several drawbacks. First, the process of binning data represents a loss of information about the differences between cases contained in each bin. Second, the choice of bin sizes introduces an arbitrary decision imposed on the analyst. Third, the HL test has inflated type-I errors, and will falsely indicate model miscalibration more often than theoretically expected. Despite this, the HL test is underpowered to detect miscalibrated probabilistic models compared to other approaches, including VMAPP. Finally, HL provides a single measure of calibration without any reference to *how* a model may be miscalibrated, and does not identify where a model is failing and in what direction.

In contrast, both VMAPP and Cox's method do not require the analyst to impose any arbitrarily chosen binning scheme on the validation data, and are able to directly identify different miscalibration pathologies. However, VMAPP improves upon Cox's method in several ways. Although both Cox's method and VMAPP are valid when the true model parameters are used, the underlying parameters are typically unknown and need to be fit from empirical data. In these cases, Cox's method has inflated Type-I error rates, whereas VMAPP does not. Second, VMAPP has demonstrably more power to detect the miscalibrated predictions of invalid models than Cox. Finally, VMAPP provides techniques which go beyond previous metrics by providing an estimate of the deviation between model predicted and actual probabilities across the range of predictions.

3.5.2 VMAPP extensions

Currently, VMAPP provides a method for estimating deviation between predicted and actual probabilities as a function of the predicted probabilities themselves. From this, one can easily produce calibration plots similar to those in Phillips and Elith (2010) by plotting predicted probability against probability estimates given the validation data ($\hat{p}\ vs\ \hat{p} - \hat{\delta}$). Our method, however, is more general and can be extended to allow for the estimation of δ as a function of any covariate of interest (included in the model or not). In this way, VMAPP could be used to suggest structural areas in which a model may be improved (Bolker, 2008). For example, VMAPP could be extended to assess the effects of spatial autocorrelation on model predictions by identifying systematic deviations (biases) as a function of space. Alternatively, by estimating δ as a function of one or more covariates used in the model's development, the structural contribution of those covariates could be assessed. For instance, if a covariate entered the model as a linear term, yet was actually exerting a nonlinear effect, the δ estimation would provide insights into this model error. Deriving the appropriate functional forms of f_1 and f_2 in these contexts would require further research, but should otherwise be possible with the technology described in this paper.

3.5.3 Concluding remarks

A complete assessment of any probabilistic ecological model requires a rigorous, quantitative means of comparing independently observed outcomes to model predictions. The method we have described here provides not only a powerful inferential technique for diagnosing miscalibrated predictions, but also a measure of the degree to which predicted probabilities deviate from the probabilities underlying the system in question. By quantifying where, and to what degree, our predictions deviate from reality, we produce clues which can help to improve our models, thereby narrowing the gap between prediction and reality.

3.6 Acknowledgements

This research was funded in part by the Canadian Aquatic Invasive Species Network NSERC grant and NSERC discovery grant to BL. The authors thank Johanna Bradie for her helpful comments on the manuscript.

References

Araújo, M., R. Pearson, W. Thuiller, and M. Erhard, 2005. Validation of species-climate impact models under climate change. *Global Change Biology* **11**:1504–1513.

Bolker, B. M., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ.

Bradie, J., C. Chivers, and B. Leung, 2013. Importing risk: quantifying the propagule pressure-establishment relationship at the pathway level. *Diversity and Distributions* **19**:1020–1030.

Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G. Yoccoz, W. Thuiller, M.-J. Fortin, C. Randin, N. E. Zimmermann, C. H. Graham, and A. Guisan, 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* **21**:481–497.

Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, H. R. Akçakaya, and R. Frankham, 2000. Predictive accuracy of population viability analysis in conservation biology. *Nature* **404**:385–7.

Burgman, M. A. and D. Yemshanov, 2013. Risks, decisions and biological conservation. *Diversity and Distributions* **19**:485–489.

Cairns, A. and N. D. Yan, 2011. New reports of Bythotrephes, summary of methodology and database design. Technical Report January, Canadian Aquatic Invading Species Network, Toronto, ON.

Chen, I.-C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas, 2011. Rapid range shifts of species associated with high levels of climate warming. *Science*

333:1024–6.

Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. a. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. a. Rose, O. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear, 2001. Ecological forecasts: an emerging imperative. *Science* **293**:657–60.

Cressie, N., C. a. Calder, J. S. Clark, J. M. Ver Hoef, and C. K. Wikle, 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological applications : a publication of the Ecological Society of America* **19**:553–70.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates, 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**:43–57.

Freeman, E. A. and G. G. Moisen, 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling* **217**:48–58.

Gertzen, E. and B. Leung, 2011a. Propagule pressure, Allee effects and the probability of establishment of an invasive species (Bythotrephes longimanus). *Ecosphere* **2**.

Gertzen, E. L. and B. Leung, 2011b. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

Gneiting, T. and M. Katzfuss, 2014. Probabilistic Forecasting. *Annual Review of Statistics and Its Application* **1**:125–151.

Guisan, A. and W. Thuiller, 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**:993–1009.

Guisan, A. and N. E. Zimmermann, 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**:147–186.

Heard, G. W., M. a. McCarthy, M. P. Scroggie, J. B. Baumgartner, and K. M. Parris, 2013. A Bayesian model of metapopulation viability, with application to an endangered amphibian. *Diversity and Distributions* **19**:555–566.

Higgins, S. I., D. M. Richardson, and R. M. Cowling, 2001. Validation of a spatial simulation model of a spreading alien plant population. *Journal of Applied Ecology* **38**:571–584.

Hirsch, R., 1991. Validation samples. *Biometrics* **47**:1193–1194.

Hosmer, D. and S. Lemeshow, 2000. Assessing the Fit of the Model. In *Applied Logistic Regression*, chapter 5, pages 143–202. John Wiley & Sons, Inc., second edition.

Jolliffe, I. T., 2007. Uncertainty and Inference for Verification Measures. *Weather and Forecasting* **22**:637–650.

Kuk, A. Y. C., J. Li, and A. John Rush, 2014. Variable and threshold selection to control predictive accuracy in logistic regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* .

Lawson, C. R., J. a. Hodgson, R. J. Wilson, and S. a. Richards, 2014. Prevalence, thresholds and the performance of presence-absence models. *Methods in Ecology and Evolution* **5**:54–64.

Lee, T. M. and W. Jetz, 2011. Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings. Biological sciences / The Royal Society* **278**:1329–38.

Leung, B. and R. J. Steele, 2013. The value of a datum - how little data do we need for a quantitative risk analysis? *Diversity and Distributions* **19**:617–628.

MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. a. Graniero, 2004. Backcasting and Forecasting Biological Invasions of Inland Lakes. *Ecological Applications* **14**:773–783.

Manel, S., H. C. Williams, and S. Ormerod, 2002. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**:921–931.

Manly, B. F., 2001. Randomization, Bootstrap and Monte Carlo Methods in Biology. CRC Press, Chapman & Hall, Boca Raton, FL, second edi edition.

Márcia Barbosa, A., R. Real, A.-R. Muñoz, and J. a. Brown, 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions* **19**:1333–1338.

Mason, S., 2008. Understanding forecast verification statistics. *Meteorological Applications* **15**:31–40.

Meyer, J., C. Ingersoll, L. McDonald, and M. Boyce, 1986. Estimating uncertainty in population growth rates: jackknife vs. bootstrap techniques. *Ecology* **67**:1156–1166.

Miller, M. E., S. L. Hui, and W. M. Tierney, 1991. Validation techniques for logistic regression models. *Statistics in medicine* **10**:1213–26.

Murphy, A., 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and forecasting* **8**:281–293.

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

Peterson, A. T., M. Papes, and J. Soberón, 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* **213**:63–72.

Phillips, S. J. and J. Elith, 2010. POC plots: calibrating species distribution models with presence-only data. *Ecology* **91**:2476–84.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Randin, C. F., T. Dirnböck, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan, 2006. Are niche-based species distribution models transferable in space? *Journal of Biogeography* **33**:1689–1703.

Schmolke, A., P. Thorbek, D. L. DeAngelis, and V. Grimm, 2010. Ecological models supporting environmental decision making: a strategy for the future. *Trends in Ecology & Evolution* **25**:479–86.

Seebens, H., M. T. Gastner, and B. Blasius, 2013. The risk of marine bioinvasion caused by global shipping. *Ecology Letters* **16**:782–90.

Staples, D., M. Taper, and B. Dennis, 2004. Estimating population trend and process variation for PVA in the presence of sampling error. *Ecology* **85**:923–929.

Suter, G. W., 2006. Ecological Risk Assessment. CRC Press, Taylor & Francis Group, Boca Raton, FL, second edition.

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

Ward, G., T. Hastie, S. Barry, J. Elith, and J. R. Leathwick, 2009. Presence-only data and the em algorithm. *Biometrics* **65**:554–63.

Wenger, S. J., N. a. Som, D. C. Dauwalter, D. J. Isaak, H. M. Neville, C. H. Luce, J. B. Dunham, M. K. Young, K. D. Fausch, and B. E. Rieman, 2013. Probabilistic accounting of uncertainty in forecasts of species distributions under climate change. *Global Change Biology* **19**:3343–3354.

Yan, N., W. Dunlop, T. Pawson, and L. Mackay, 1992. Bythotrephes cederstroemi (Schoedler) in Muskoka lakes: first records of the European invader in inland lakes in Canada. *Canadian Journal of Fisheries and Aquatic Sciences* **49**:422–426.

Table 3–1: **Type-I error rates of each test under three conditions.**

	Type-I error rate		
Test	No Uncertainty	Unaccounted Uncertainty	Accounted Uncertainty
Hosmer-Lemeshow	0.11	0.27	N/A
Cox's Method	0.05	0.23	0.17
VMAPP	0.05	0.16	0.05

Figure 3–1: **Conceptual outline of VMAPP.** Each column shows a possible miscalibration scenario. The first row shows the underlying *true* probabilities p generating success/failures (blue dashed lines) and the predicted probabilities $\hat{p}$ from a hypothetical model (solid black lines). The difference between p and $\hat{p}$ across the predicted range is indicated by δ. The second row shows discrepancies between a single set of simulated outcomes from the prediction model (S) and empirical observations (R). Successes (1s) in this row represent discrepancies that have occurred ($S \neq R$), failures (0s) represent no discrepancy ($S = R$). The third row shows the direction of discrepancy ($S > R \,|\, S \neq R \rightarrow 1, S < R \,|\, S \neq R \rightarrow 0$). The red lines are stylized fitted curves. By combining these curves over repeated simulations, VMAPP estimates δ and provides an inferential mechanism for identifying model miscalibration.

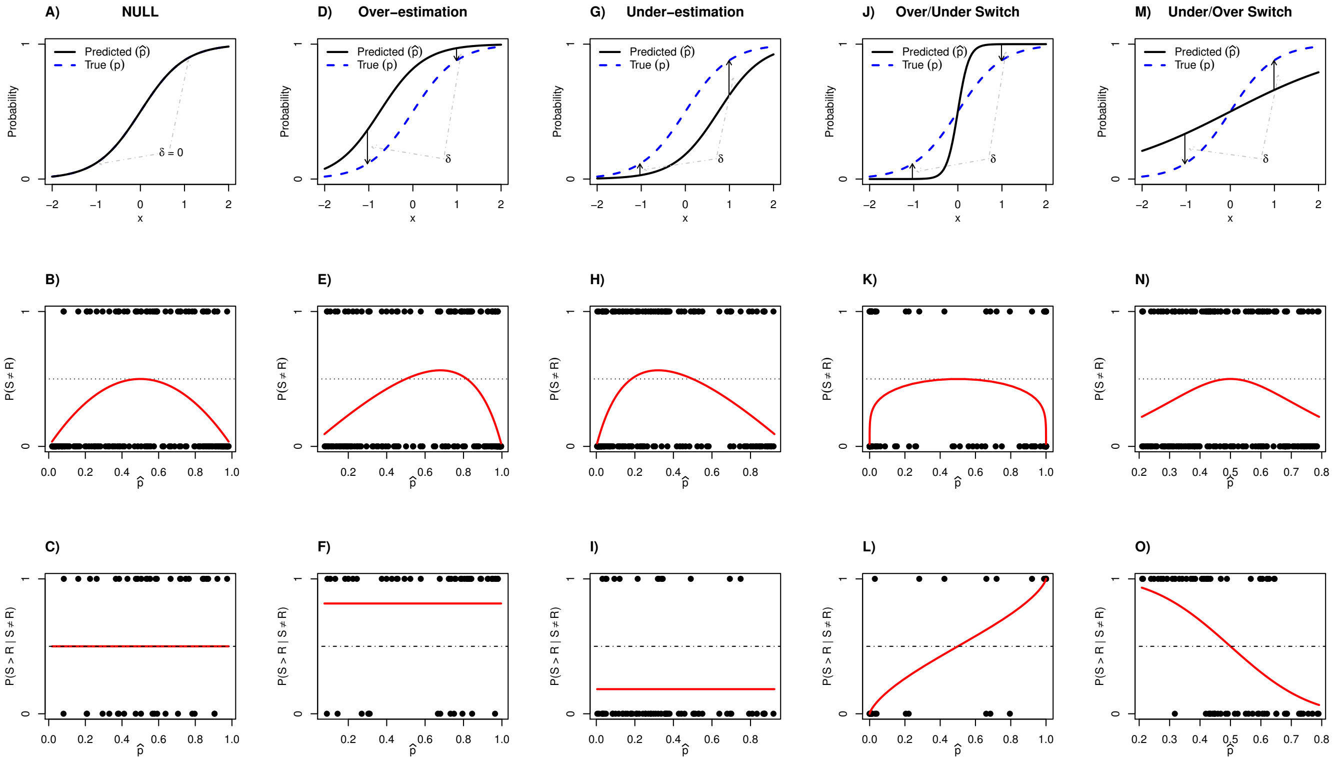

A) NULL
D) Over−estimation
G) Under−estimation
J) Over/Under Switch
M) Under/Over Switch
Predicted ($\hat{p}$)
True (p)
Probability
x
δ = 0
δ
B)
E)
H)
K)
N)
P(S ≠ R)
$\hat{p}$
C)
F)
I)
L)
O)
P(S > R | S ≠ R)

Figure 3–1

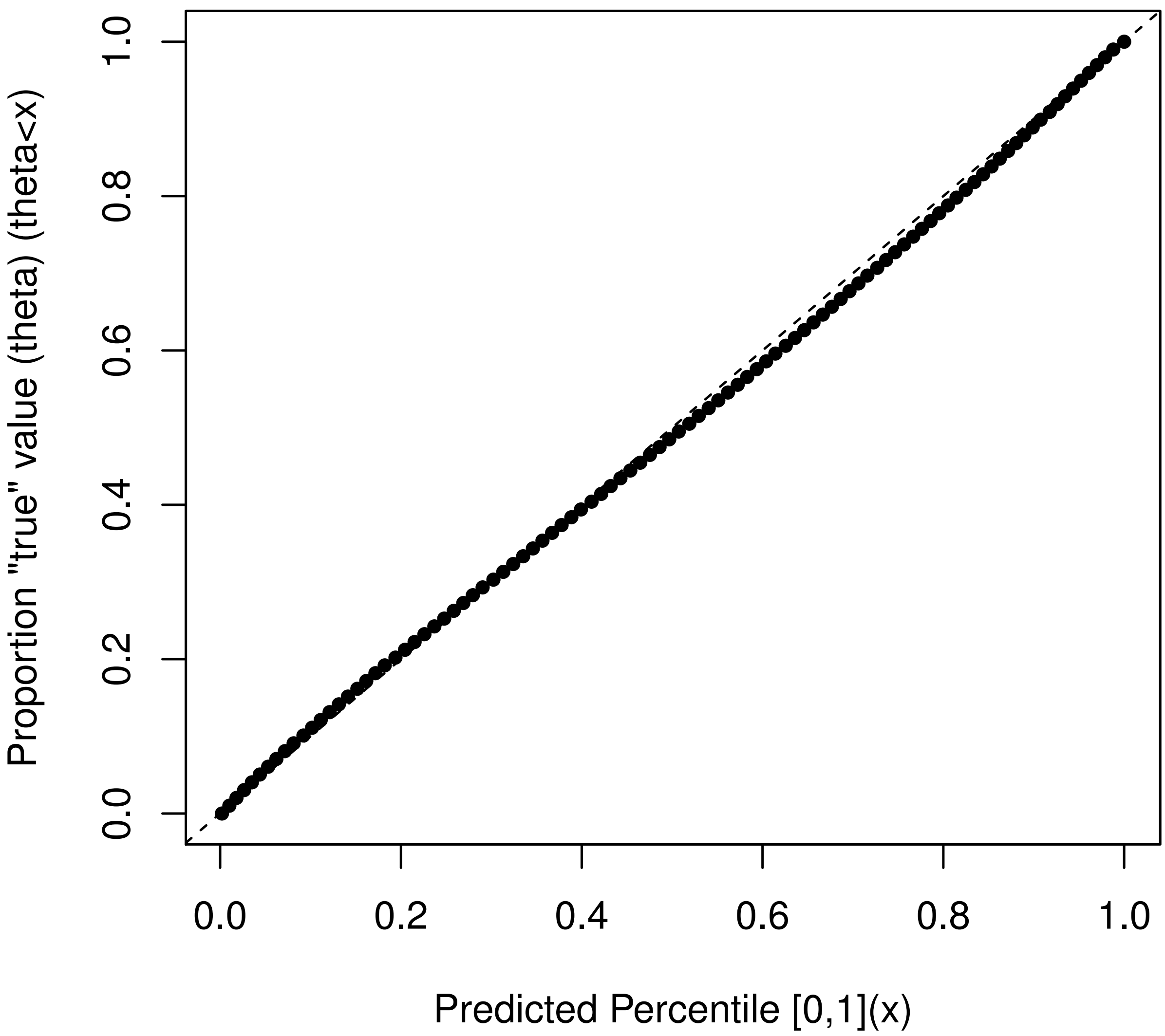


Figure 3–2: **VMAPP theoretical performance.** VMAPP performance in identifying deviations between model predicted probabilities and actual probabilities ($\hat{\delta}$) for 1000 simulated model-validation set pairs ($n = 100$ for each validation set). Under perfect predictions of deviation, the expectation is a 1:1 line.

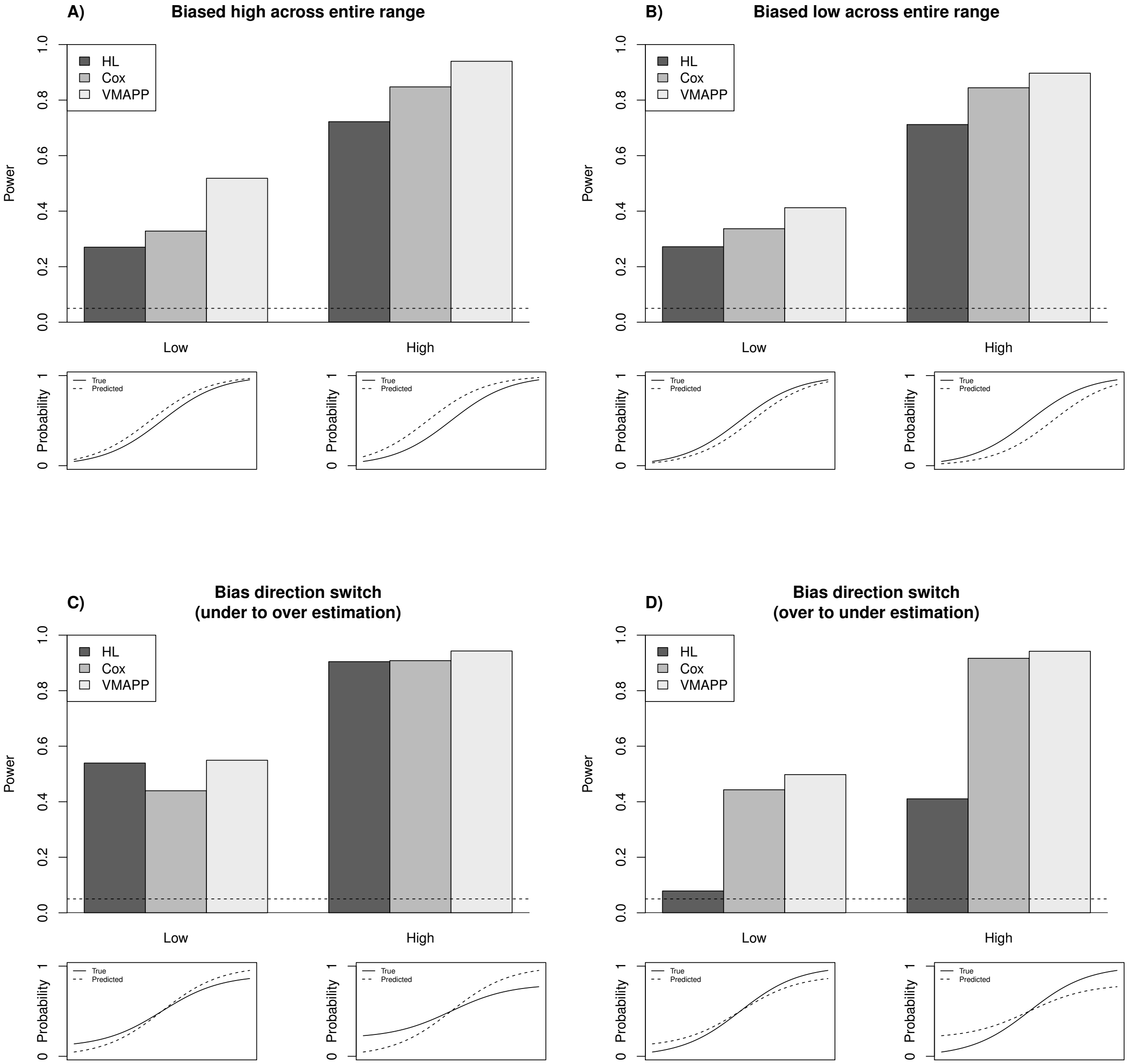


Figure 3–3: **Power $(1-\beta)$ comparisons between Hosmer-Lemeshow, Cox's method, and VMAPP.** Panels A-D correspond in order with the four miscalibration pathologies in Figure 3–1. Results are based on 1000 simulated validation procedures at two levels of deviation between model and reality. The sub-panels in each show the shape of the model predictions and underlying generating probabilities for each case. $n = 100$ for each validation set. Note that the power of all tests goes to 100% as $n \to \infty$.

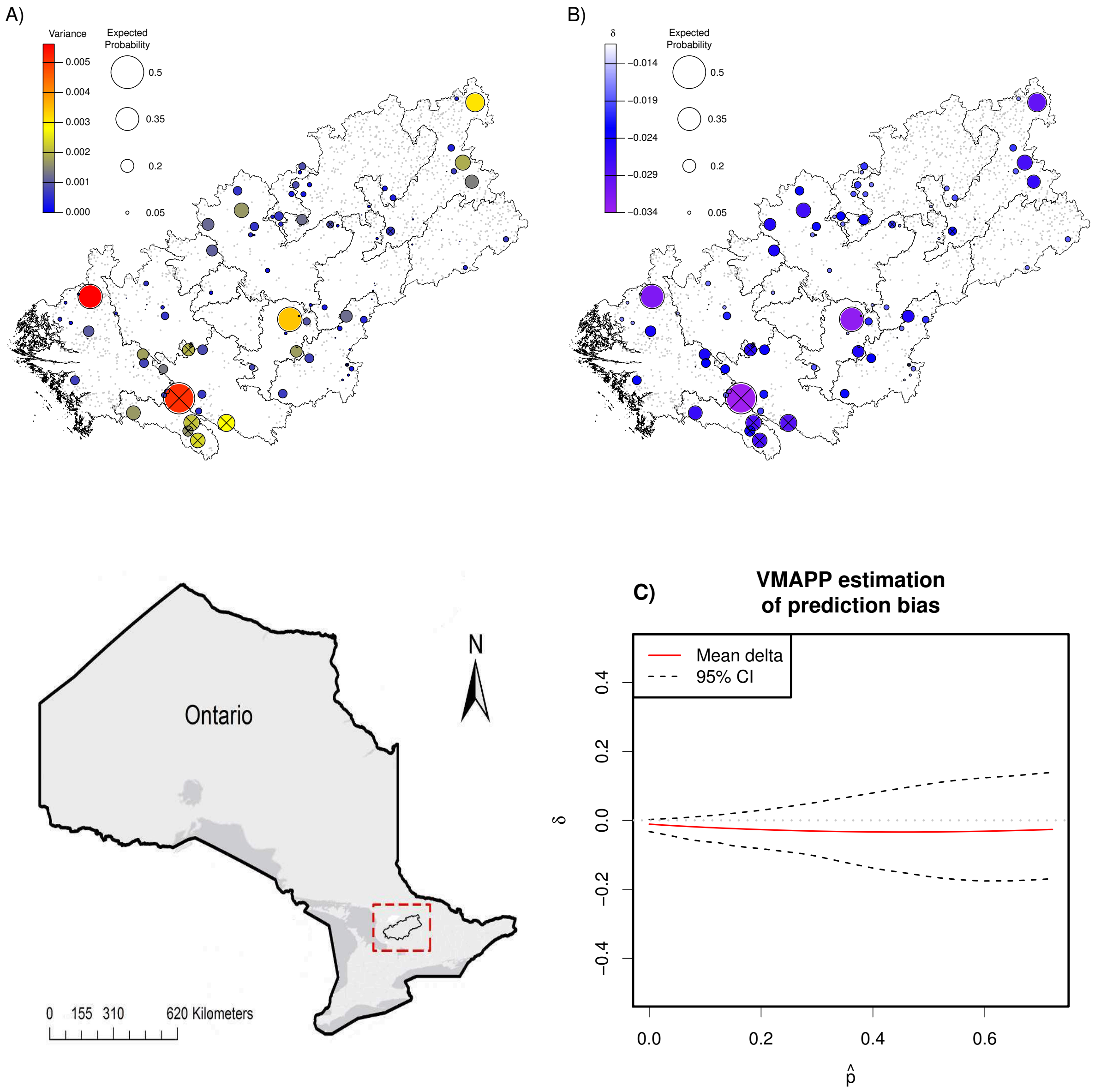


Figure 3–4: **Predictions of Bythoterphes occurances in 2010.** Panel A shows mean predicted probabilities of occurrence of *Bythoterphes* at each of the 102 validation lakes (shown as size of circles) and the variance in the predicted probabilities (in heat colours). Panel B shows the estimated deviation between model predictions and actual probability ($\hat{\delta}$) at each of the validation lakes using VMAPP. In both A & B, lakes which were observed to be invaded in 2010 are marked with an X. Panel C shows mean and 95% confidence intervals of $\hat{\delta}$ as a function of $\hat{p}$ for the *Bythotrephes* validation data.

Connecting statement

In the previous chapter, I developed a novel methodology for validating projected risk models and showed it's utility in a case study of *Bythotrephes longimanus* spread in a watershed in Central Ontario. The case study was an evaluation of a published risk model which incorporates various aspects of uncertainty including stochasticity and sparse data. While the data which went into both the building and validation of this model were sparse, they were the result of a very large and resource-intensive sampling effort to identify the presence and absence of *Bythotrephes* at many locations in the watershed. More common, however, is that we only have records of presences, and these records exist only for a limited subset of locations. Rather than being the result of systematic sampling, these presence records often exist as a result of either opportunistic sampling or as reported sightings from citizen scientists. This form of data, while ubiquitous, presents a particular challenge to modelling efforts aiming to assess the risks posed to sites across space and time. Specifically, without data on species absences, currently available techniques are not able to estimate species prevalence or site-specific probabilities of presence, and are therefore only able to estimate relative, as opposed to absolute risks. In the next chapter, I present a novel modelling framework for solving this problem in the context of spreading invasive species by modelling the spatio-temporal pattern of establishments as hidden states and the detection of presences as a stochastic observation process. I apply the framework to 10 aquatic invasive species in Ontario for which presence-only data were available. By providing a solution to the presence-only problem, this framework represents a major contribution to both the invasive species literature as well as the field of Biogeography more generally.

CHAPTER 4
Estimating the probability of establishment and spread of biological organisms: issues of uncertainty, detection and presence-only data

Corey Chivers[1] , Brian Leung[2] [3]

A version of this chapter has been submitted to *Ecology Letters*.

---

[1] Department of Biology, McGill University.Stewart Biology Building 1205 ave Docteur Penfield Montreal, Quebec. 514-398-1833. corey.chivers@mail.mcgill.ca.

[2] Department of Biology, McGill University

[3] School of Environment, McGill University. brian.leung2@mcgill.ca

4.1 Abstract

Estimating the current and future spatial extent of invasive species is of paramount importance to both scientific understanding and management decisions. Given their ubiquity, presence-only data have become increasingly employed for these purposes. However, absence of species may occur either because they are truly absent, or have not yet been detected; for presences, only date of discovery is known rather than true time of establishment. In this study, we model establishments as hidden states and detection as a stochastic observation process to solve the problems inherent to presence-only data. We show theoretically that we can estimate probability of detection, which sites are likely invaded but not yet detected, as well as the past, current, and future extent of invasions. We apply the model to 10 aquatic invasive species in Ontario. We estimate that for zebra and quagga mussels, the majority of invaded sites have already been detected (76% and 95%, respectively), but in contrast, for two species of mysterysnail and the Eurasian water-milfoil the extent of invasion is likely 10 times greater than currently detected.

4.2 Introduction

Estimating the current and potential future extent of a species' geographic range is of obvious value for both scientific understanding and management of biodiversity. There is a large and growing literature on techniques and methodologies for estimating a species' geographic range (Elith and Leathwick, 2009). Most of the common approaches, however, tend to assume that the species of interest is at equilibrium with the environment, and do not take into account dispersal limitation and other spatio-temporal processes. Modified applications of these modelling frameworks to invasive species (termed invasive Species Distribution Models, iSDMs) on the other hand, consider cases which violate the assumptions underlying most SDM methodologies (Václavík and Meentemeyer, 2009). A minority of studies explicitly incorporate dynamic spatio-temporal processes by joining dispersal and physiological habitat suitability (Leung and Mandrak, 2007; Catterall et al., 2012; Ibáñez et al., 2014). However, current joint models are limited because they assume spatially and temporally explicit presence and absence data, which are unavailable for most species. Instead, data typically consist of the locations of species presences, while the vast majority of locations have unknown status. As such, many SDM approaches rely on randomly selected background sites as stand-ins for absences, or so called pseudo-absences (Elith et al., 2006). The use of pseudo-absences in conjunction with correlative statistical modelling approaches are becoming increasingly prevalent in the literature, and are often used to produce predictive maps of species distributions (Guisan and Thuiller, 2005; Hijmans, 2012).

These correlative statistical approaches model the joint probability distribution of environmental covariates and species occurrences, yet do not directly predict the probability of a species occurring at a given site (Elith et al., 2011). While suitability indices or a ranking of the invasion risk across sites can act as a management aid, non-temporally explicit, non-probabilistic suitability maps can only provide limited decision support. Particularly in the case of invasive species, static suitability maps do not directly assess the risks posed at different sites across the landscape over a given time horizon. Further, knowledge of where species are likely to be present, and yet have gone undetected can guide resources toward more efficient search and monitoring strategies. To be maximally useful, a decision support tool would provide quantities relevant to decision theoretical frameworks such as utility maximization or cost-benefit analysis (Polasky et al., 2011). While some recent papers have attempted to provide solutions to determining probabilities using presence-only data, they do so in the context of species which are at equilibrium (not spreading) with their environment (Li et al. 2011; Royle et al. 2012, but see critique from Hastie and Fithian 2013). When projecting the spatial extent of a species which is not at equilibrium with the environment over time, we need to assess the probability, at each point in time, that a species is, or will be, present at a given location. Beyond being useful at the local scale, such predictions are particularly important for range expanding species, where colonized locations become focal points, providing new propagules on which continuing spread will depend (Muirhead and Macisaac, 2005).

Methods to combine presence-only and presence/absence data to forecast invasive species have been developed recently. In one approach, the problem was broken into a two step procedure, first building a Maxent model then directing field sampling for presence/absence data collection based on the suitability maps (Gormley et al., 2011). Some have suggested that when modelling an invasive species, presence data from the native range should be used under the assumption that the species had sufficient time to reach a kind of equilibrium, not limited by dispersal over the range of environmental conditions present (Guisan and Thuiller, 2005). Other efforts go beyond static assumptions and explicitly model species spread via time-dependent dispersal and colonization processes. Leung and Mandrak (2007) develop and apply a method for joining dispersal and habitat suitability to the spread of zebra mussels using temporal presence/absence data. Accounting for uncertainty in colonization time by using a methodology for incorporating sparse sampling records and explicitly modelling the dynamics at unsampled locations was further developed by Gertzen and Leung (2011). Several recent studies of plant invasions have described general Bayesian approaches which can also simultaneously estimate dispersal and habitat suitability (Catterall et al., 2012; Ibáñez et al., 2014). These approaches require species distribution (presence/absence) data at multiple points in time (snapshots) and can account for various uncertainties including the unknown, but bounded colonization times. In each of these studies, explicit knowledge of at least a subset of sites representing true absences is required.

In this study, we tackle two of the main challenges facing efforts to model current and future ranges of invasive species, specifically dispersal limitation and presence-only data. We describe the rational behind a novel approach which uses a probability of detection (observation) model and presence-only data to predict the current and future range of invasive species. Our approach demonstrates how these two problems can, in fact, help to solve each other. By modelling spatio-temporal establishment as a hidden process combined with an observation process, we simultaneously estimate population dynamical parameters (rates of spread), current extent of invasion, environmental tolerances (habitat suitability), and probabilities of establishment and detection.

We begin by outlining the generalized form and rational of our modelling framework. We highlight the modular nature of this approach, indicating where alternative sub-models may be employed. We then theoretically examine our approach using simulated invasions, validating the ability of the model to recapture the underlying parameters of each sub-model, as well as the unknown colonization timings represented as hidden states in the model. We then apply our approach to presence-only data reported for 10 aquatic invasive species (4 plants, 6 invertebrates) in Ontario. Finally we discuss the benefits of using this approach to modelling the spread of invasive species from both a management and scientific perspective.

4.3 Materials & methods

The modelling framework consists of the joining together of three subcomponent models: 1) Dispersal, 2) Establishment (colonization), and 3) Detection (observation). Since our approach is applicable at various spatial scales, dispersal vectors, and for many types of species undergoing range expansion, we first outline the model in its generalized formulation. The specific modelling choices and functional forms pertaining to each of the three submodels are left to the case-study application of our framework to 10 aquatic invasives in Ontario.

4.3.1 The generic model formulation

Consider a heterogeneous network of habitat patches (referred to here as sites) connected by potential dispersal corridors over which propagules may travel. At some starting time $t = 0$, only a small subset of the sites contain established populations of a species of interest. Sites are heterogeneous in environmental characteristics, and hence in their suitability for the target species. The environmental characteristics are therefore predictors of the relationship between the number of inbound propagules and the probability that the species will establish a population at (invade/colonize) a given previously uninvaded site. We can write this probability for a given time-step t as:

$$\rho_{E_i}(t) = P(S_i = t \,|\, \boldsymbol{\theta}_E, \zeta_i, Q_{i,t}) \tag{4.1}$$

Where S_i defines the time-step in which the target species establishes a reproducing population at site i. This probability is conditional on the parameters $\boldsymbol{\theta}_E$ describing how environmental characteristics of the site ($\boldsymbol{\zeta}_i$) affect the probability of establishment as a function of the inbound propagule pressure $Q_{i,t}$. We assume this to be some increasing function of propagule pressure, such that the probability of establishment is positively related to the number of inbound propagules, but the strength and shape of that relationship is determined by the environmental conditions at a given location. The propagule pressure itself depends on the invasion status of all other sites which are potential sources of inbound propagules as well as the strength of the dispersal corridor between them. That is, every site j where

$S_j < t$ contributes propagules dependent upon synonymously the connection, corridor, or dispersal potential (distance/quality/strength/connectivity) between sites i and j. The connections themselves are determined by an underlying model of potential dispersal with parameters $\boldsymbol{\theta}_\lambda$. For invasive species, we assume the dispersal model to be a network based model of human-mediated dispersal between discrete habitat patches, but even simpler distance-based kernels are also possible. Thus, in general form, the propagule pressure to site i is a function of both the invasion status of all other sites and the dispersal potential connecting them to site i:

$$Q_{i,t} = f(\boldsymbol{\theta}_\lambda, S_j \forall j \neq i). \tag{4.2}$$

Once a site becomes invaded, we assume there is some probability ρ_{D_i} that the species will be detected at that location during each subsequent time step $t \geq S_i$. This probability may itself be a function of some site specific characteristics $\boldsymbol{\eta}_i$ with parameters $\boldsymbol{\theta}_D$. The probability of a site being detected at time $t = D_i$ can then be modelled as the joint probability of that site remaining uninvaded up to time $S_i - 1$, becoming invaded in time S_i, subsequently remaining undetected from time S_i to time $D_i - 1$, then finally being detected at time D_i. We can write this probability using a joint survival equation:

$$g(S_i, D_i, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_i, \boldsymbol{\eta}_i) = \underbrace{\left[\prod_{t=1}^{S_i-1}(1-\rho_{E_i}(t))\right] \rho_{E_i}(S_i)}_{\text{Hidden state space}} \underbrace{\left[\prod_{t=S_i}^{D_i-1}(1-\rho_{D_i})\right]\rho_{D_i}}_{\text{Detection Process}} \tag{4.3}$$

Since our only data for a given site is the time of detection, we do not observe S_i directly but rather know only that it occurred at or prior to D_i. Therefore, in order to calculate the overall likelihood of observing a detection at time D_i, we consider the likelihoods over all possible hidden states $S_i \leq D_i$. The possible hidden states (time of invasion) which may have resulted in a detection at $t = D_i$ are: invasion in the same year as the detection ($S_i = D_i$), *or* invasion in the previous year ($S_i = D_i - 1$), *or* invasion in the year before that ($S_i = D_i - 2$), *or* the year before that, and so on. For example, for a site detected at the third time-step, it may have become established in the first time-step *or* the second *or* the third time-step. We therefore maginalize over the collection of possible hidden state space to obtain:

$$h(D_i, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_i, \boldsymbol{\eta}_i) = \sum_{s=1}^{D_i} g(S_i = s, D_i, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_i, \boldsymbol{\eta}_i). \tag{4.4}$$

Which is the likelihood of first detecting the species to be present at $t = D_i$, given the model and parameters. The overall likelihood for the N sites which have been detected as of the current time T_{now} is therefore the product of their individual likelihoods,

$$\mathcal{L}_1(\boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda \,|\, \boldsymbol{D}, \boldsymbol{\zeta}, \boldsymbol{\eta}) \;=\; \prod_{i=1}^{N} h(D_i, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_i, \boldsymbol{\eta}_i). \tag{4.5}$$

However, eqn 4.5 only describes the likelihood for those sites which were detected to be invaded as of the current time T_{now}. To incorporate those sites at which we have not detected our target species either because the site is not invaded or because it is invaded and has yet to be detected, we need to modify eqn 4.3, removing the ρ_{D_i} term and allowing for the state where the site has not yet been invaded (ie $S_i > T_{now}$).

$$m(S_i, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_i, \boldsymbol{\eta}_i) = \begin{cases} \left[\prod_{t=1}^{T_{now}} (1-\rho_{E_i}(t))\right] & \text{if } S_i > T_{now}, \\ \left[\prod_{t=1}^{S_i-1} (1-\rho_{E_i}(t))\right] \; \rho_{E_i}(S_i) \; \left[\prod_{t=S_i}^{T_{now}} (1-\rho_{D_i})\right] & \text{if } S_i \leq T_{now}. \end{cases} \tag{4.6}$$

The likelihood across each of the K lakes in which the target species has not been detected to date is now given (similarly to eqn 4.5) as:

$$\mathcal{L}_2(\boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda \,|\, \boldsymbol{\zeta}, \boldsymbol{\eta}) \;=\; \prod_{k=1}^{K} \left[\sum_{s=1}^{T_{now}+1} m(S_i = s, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda, \boldsymbol{\zeta}_k, \boldsymbol{\eta}_k) \right] \tag{4.7}$$

And the complete likelihood is now the product of eqns 4.5 and 4.7:

$$\mathcal{L}(\boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda \,|\, \boldsymbol{D}, \boldsymbol{\zeta}, \boldsymbol{\eta}) \;=\; \mathcal{L}_1\mathcal{L}_2 \tag{4.8}$$

With this likelihood function in place, all model parameters associated with the invasion and the detection processes along with their associated uncertainty can then be estimated using a Bayesian approach. We employ MCMC using an adaptive Metropolis-Hastings algorithm to obtain samples from the posterior predictive distributions of the model parameters via the familiar Bayesian inversion:

$$P(\boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda \,|\, D, \boldsymbol{\zeta}, \boldsymbol{\eta}) \propto P(\boldsymbol{D}, \boldsymbol{\zeta}, \boldsymbol{\eta} \,|\, \boldsymbol{\theta}_E, \boldsymbol{\theta}_D, \boldsymbol{\theta}_\lambda) P(\boldsymbol{\theta}_E) P(\boldsymbol{\theta}_D) P(\boldsymbol{\theta}_\lambda). \tag{4.9}$$

In addition to providing uncertainty estimations for each of the model parameters, the use of MCMC provides us with a mechanism for probabilistically calculating eqn 4.1 for every site at each time-step $t \leq D_i$ for the detected sites, and $t \leq T_{now}$ for those sites at which the target species has not yet been detected. This is particularly useful since that in order to calculate the propagule pressure to each site ($Q_{i,t}$), the invasion state of all sites $j \neq i$ is required since those sites which have become invaded prior to t are those which will be providing propagules, as determined via the dispersal model (Eqn. 4.2). We obtain these hidden states by stochastically simulating spread from $t = 0$ to $t = T_{now}$ at every step of the MCMC chain. For reference, the parameters of the model are described in Table 4–1.

4.3.2 Prediction and inference

Once model parameters have been estimated, there are several quantities which can be estimated using this formulation. First, our parameter estimates from the habitat suitability sub-model allow us to make inferences about species' tolerance relationships with the environment. Similarly inference about the strength of dispersal corridors is possible using parameter estimates from the dispersal model. Perhaps of most immediate use to managers, a predictive distribution of the state vector $\boldsymbol{S}$

can easily be generated via stochastic simulation from the joint posterior predictive distribution of parameters. This allows us to quantify both site-specific and system-level risk of invasion over time. These can be generated both looking backwards, predicting which sites are likely already to be invaded, as well as projecting forward, making forecasts about a species' future spatial distributions (MacIsaac et al., 2004). Since the forecasts are probabilistic, they provide natural inputs to both risk assessments and decision-theoretic management plans (eg. maximum expected utility calculations (Polasky et al., 2011)). Similarly, traditional potential distribution maps can be generated, indicating locations for long-term monitoring as well as for comparative studies between the projected invaded and native ranges (Beaumont et al., 2009). Additionally, since we explicitly model the observation process, we can make predictions and ask questions about differences in detectability and reporting between species. By comparing the predicted current extent of a given species with the date of first detection a every location, our framework can provide guidance on which species are likely going unreported as well as where to focus future monitoring efforts (Hui et al., 2011).

4.3.3 Theoretical analysis

The theoretical ability of the general approach which we have outlined above to recapture underlying processes and parameters was tested using simulated data for which the true parameters were known. To make the simulated scenarios as realistic as possible, we used the system of Ontario Lakes and the same environmental variables which we use in our real-world application (section 4.3.4). Further, we employ the same general sub-model structures for the dispersal, suitability, and detection components but simplify by only considering a single environmental predictor (we use 5 in the application). We simulated 1000 *in silico* invasion scenarios using randomly selected model parameters. For each simulated invasion, we collected virtual detection data on which to fit the model (see example simulation in Fig. 4–1). Using this procedure, we were able to determine whether our statistical formulation was able to correctly characterise the generating parameters for all three sub-models using only information about where and when species were detected. Following Leung and Steele (2013), we evaluated the resulting posterior predictive distributions by producing P-P plots which compare the frequency with which each generating parameter falls into different quantiles of the posterior. P-P plots with a 1:1 relationship indicate that the posterior distributions are unbiased estimators of the true parameter values.

4.3.4 Application

We demonstrated the utility and versatility of our approach by applying our model to presence-only records of 4 invasive aquatic plants and 6 invasive aquatic invertebrates from the Early Detection & Distribution Mapping System for Ontario (EDDMapS, 2013). The data consist of time-stamped, geo-referenced observations of each species from a variety of sources including monitoring programs as well as citizen submitted sightings. All reported presences have been verified by the Ontario Federation of Anglers and Hunters. The spatial distributions of the detections data are shown in Fig. 4–2.

4.3.5 Ontario dispersal model

Aquatic invasive species have been found to be primarily dispersed via human-mediated vectors (Johnston et al., 2001; Timar and Phaneuf, 2009; Gertzen and Leung, 2011). In order to characterise the potential dispersal corridors between inland lakes, therefore, we need to characterise the behaviour of the human actors (recreational boaters). Past studies have found so-called Gravity Models to be predictive of human behaviour (Chivers and Leung, 2012) as well as predictive of the spread of invasives mediated by their behaviour (Leung et al., 2004, 2006; Muirhead and MacIsaac, 2011). The dispersal model which we applied to the 10 Ontario species is a production-constrained gravity model. This formulation is not data intensive, requiring at a minimum the home locations of boaters in the system to be fit (Muirhead and MacIsaac, 2011). This data was obtained from the Ontario Ministry of Natural Resources as the number of registered boaters in each of 526 postal regions as identified by the first three digits. For details of the dispersal model formulation see Chivers and Leung (2012).

4.3.6 Establishment & suitability model

Our establishment and suitability model links the number of propagules arriving at a given site at each time step with the probability that the site will become invaded. Following the formulation described by Bradie et al. (2013), we modelled the establishment process as a survival function, where the probability of establishment is the complement of the cumulative probability of each propagule failing to result in establishment. In order to incorporate environmental heterogeneity between sites, we extended this formulation to include environmental predictors of habitat suitability such that the suitability of a site changes the shape of the propagule pressure-establishment probability curve. Specifically,

$$\rho_E(t) = 1 - q(z_{Ei})^{Q_{i,t}^{c}} \tag{4.10}$$

where $Q_{i,t}$ is the number of propagules arriving at site i during time step t as determined by the colonization state $\boldsymbol{S}$ of all other sites and the strength of the dispersal corridor between the currently established sites and site i. The value of $q(z_{Ei})$ is unique to each site and determines the ease with which a site may become invaded by changing the strength of the relationship between propagule pressure and probability of establishment. As $q(z_{Ei})$ defines the probability of a propagule *not* resulting in an establishment, sites with lower $q(z_{Ei})$ values are more susceptible to invasion. In order to make the interpretation of the model coefficients intuitive, we estimate instead the complement of $q(z_{Ei})$, and we do so using a simple logistic function with linear terms for the environmental characteristics of a given site i:

$$q(z_{Ei}) = 1 - \frac{1}{1 + exp(-z_{Ei})}, \tag{4.11}$$

$$z_{Ei} = \beta_{E0} + \beta_{E1}\zeta_{i1} + \ldots + \beta_{EJ}\zeta_{iJ} \tag{4.12}$$

Where $\boldsymbol{\beta}_E$ is a vector of $J + 1$ parameters to be fit and $\boldsymbol{\zeta}_i$ is a vector of J environmental characteristics at site i. The establishment model is a discrete-time stochastic process. Sites with smaller values of $q(z_{Ei})$ are more easily colonized, requiring fewer propagules on average to become invaded, while sites with large values of $q(z_{Ei})$ will remain uncolonised (probabilistically) even when receiving a large number of inbound propagules. The parameter c allows for the possibility of density-dependent Allee effects, which have interactive effects with the network structures (Chivers and Leung, 2012).

4.3.7 Environmental variables

For the habitat suitability component of our model, we obtained environmental predictors from the BIOCLIM database (Hijmans et al., 2005). Environmental attributes of each lake with a surface area > 10 hectares ($n = 781$) were obtained by computing the average value of each BIOCLIM variable over lakes' the surface area. As many of the variables were highly correlated, we performed dimensionality reduction using PCA (Jackson, 2005), retaining the first 5 principal components which together contained 95% of the variance within the study area. Some previous studies have dealt with the multi-colinearity of environmental predictors by choosing only a subset of relatively uncorrelated variables which are thought *a priori* to be relevant to species tolerances (Ibáñez et al., 2009; Kulhanek et al., 2011). Such an approach has the advantage that it can, in some cases, yield biologically interpretable fitted parameters. In our case, however, we chose to use the dimensionally reduced representation of the entire set of predictors with the aim of accounting for most of the variance in environmental conditions while being general enough to be applicable across our variety of plants and invertebrates.

4.3.8 Detection model

We assume that in each time step, if a site is currently invaded, a detection can either occur or not, and that this is a probabilistic process. The probability of detection may be a function of any number of site characteristics $\boldsymbol{\eta}_i$. We modelled this relationship again using a simple logistic form:

$$\rho_{Di} = \frac{1}{1 + exp(-z_{Di})}, \tag{4.13}$$

$$z_{Di} = \beta_{D0} + \beta_{D1}\eta_{i1} + \ldots + \beta_{DJ}\eta_{iJ} \tag{4.14}$$

The underlying observation process may be heterogeneous across time as well as space. For instance, as public awareness of a given species changes over time, so too would we expect the probability of invaded sites being detected. For simplicity and generality, however, here we assume that each time step is independent with equal probability of detection. This assumption could easily be relaxed, adding an additional time-dependent predictor of probability of detection and incorporating additional information about public awareness and other factors affecting search effort over time. Additionally, while we present here the general form of the detection model, we did not have any *a priori* site-specific predictors of search effort.

4.3.9 Prediction, forecasting, and validation

In order to forecast the future distributions of each species, we ran repeated stochastic realisations of spread out 10 years past the last detection data points (2013-2023). While the distribution of spread trajectories resulting from these repeated simulations incorporates the stochastic uncertainty inherent in the dynamics of future spread, we also integrate parameter uncertainty by drawing parameter values randomly from the posterior distribution of the underlying dispersal and suitability models for each realisation. Using this approach, we are able to estimate the past, current, and future distributions of each species. First, by counting the fraction of $R = 10,000$ simulated realisations in which each site becomes invaded by time t, we can estimate the probability of presence at every site in space and time as $P(S_i \leq t) = \frac{1}{R}\sum_{r=1}^{R}(S_i \leq t)$; $i = 1, \ldots, i = n$; $t = T_{firstdetection}, \ldots, t = 2023$. Second, we can separate out the predicted current versus potential distribution of each species (Václavík and Meentemeyer, 2009). The cumulative future risk of invasion posed at each site can be assessed by computing the difference between the probability of presence by 2023 and the probability of presence as of 2013. Maps of such cumulative future risk can then be produced to identify sites which may benefit most from localized increases in early detection efforts (Vander Zanden et al., 2010) and preventative management actions (Leung et al., 2002). Third, the distribution of the number, or proportion, of sites invaded in any given year is easily obtained by keeping track of this count for each realisation. And finally, the predicted number and location of detections can be obtained by additionally simulating the detection process itself, also by drawing parameters from the posterior distribution.

When presence/absence data are available, the outcome of interest on which to perform predictive validation is whether or not the species is present at the validation sites (Hijmans, 2012). Model predictions are then typically compared against these observations and some measure of predictive accuracy is computed (eg AUC). Since we do not have presence/absence data, rather detections and non-detections, we assess model performance based on these same outcomes. To assess model performance for each species, we fit our model using only those detections dated prior to 2006. We then project the spread of each species up to 2013. From this, we calculated the probability of making a new detection at each site during this period, which we then compared against actual outcomes of detection or no detection at each site reported between 2006-2013.

4.4 Results

4.4.1 Theoretical simulations

The dispersal and detection model parameters were predicted without any detectable theoretical bias as illustrated using P-P plots comparing expected and observed quantiles of the true parameters within the posterior distributions (Fig. 4–3). The only slight deviation occurred in the estimates of the environmental suitability parameter β_{E1}. The model showed a very slight tendency to underestimate this coefficient as indicated by the P-P plot curving slightly below the 1:1 line. This effect is likely caused by the spatial structure of the environment and in particular spatial autocorrelation of the predictor which is known to introduce systematic biases in models of metapopulation dynamics (González-Megías et al., 2005). The resulting predictions of the hidden state variables (colonization timing) followed a 1:1 predicted/observed pattern ($R^2 = 0.96$) and were not detectably affected by the slight β_{E1} bias.

4.4.2 Ontario aquatic invasives

The probability of species presences are mapped and overlaid with the location of all detections as of 2013 (Fig. 4–4). The Spiny waterflea is predicted to be the most wide-spread species with 25% ([19-31%] 95%BCI) of sites predicted to be invaded as of 2013, followed by Purple loosestrife (20% [15-25%]) and Rusty crayfish (16% [10-20%]) (Fig. 4–5). The predicted extent of the Flowering rush in 2013 is fairly low at 8% of sites, however, it is also characterised by the highest level of uncertainty in terms of colonization over time, with a lower credible interval suggesting it could be as low as 3.0% and as high as 14%. The Zebra mussel and Banded mystery snail, on the other hand, were both predicted with less uncertainty having credible intervals spanning the relatively small ranges of 3% and 0.8%, respectively (Fig. 4–5). One trivial possibility is that the number of sites currently invaded is simply a function of the number of presences detected. The predicted number of sites currently invaded, however, was not correlated with the number of detections ($r = 0.22, p = 0.55$), indicating that there is a large amount of variance in detectability between species. The ratio of the number of predicted invaded sites to the number of detections as of 2013 is highest for the two snail species in our dataset. The Chinese mystery snail has not yet been detected at 92% of the sites for which it is predicted to be invaded (6 detected out of 68.6 predicted invaded), and 87% of invaded sites similarly remain undetected of Banded mysterysnail (10 out of 68.5). By comparison, the zebra mussel is predicted to be present but undetected at only 24% of the sites in its current range, while the quagga mussel appears to have been detected in all but 1% of the sites where it is currently established (Fig. 4–6). The complete spatio-temporal record of detections shown with the predicted probabilities of presence are available as an online supplement (`https://cjbayesian.github.io/detection_model_spread_animations`).

Three of the 10 species had mean AUC values greater than 0.9, indicating a very high degree of predictive discrimination between the sites which were detected in the 2006-2013 period and those that were not (Table 4–2). Only the Banded

Mystery snail had a mean AUC < 0.7, with an AUC ranging from 0.43 (worse than random) to 0.66. Some of this poor performance may be attributable to low number of detections (4) recorded prior to 2006 on which the model was fit. However, there were only 2 detections recorded prior to 2006 for the Chinese Mysterysnail, for which the validation AUC was 0.76 [0.5,0.89].

4.5 Discussion

Recently, there has been an increasing number of papers attempting to predict the spatial distribution of invasive species in their introduced ranges (Kulhanek et al., 2011; Gallien et al., 2012; Hallstan et al., 2013). While these studies use presence/absence data to estimate current and future distributions, in only a minority of instances are data on both presences and absences actually available. The far more common situation is that we have information on where a species has been detected (eg. presences) but lack any information about which locations are now, or were known at one point in time to have been free of a given species (Ward et al., 2009). The fundamental asymmetry between the knowledge of, and certainty about, presences vs absences makes gathering information about the latter particularly problematic (Lobo et al., 2010). Further, invasive species pose a special challenge to attempts at characterizing their current and potential distributions (Václavík and Meentemeyer, 2009). Specifically, invasive species may not have had sufficient time to have reached every site in the introduced range in which they may thrive once reached, and yet many of the standard methodologies for modelling species geographic distributions are grounded in the assumption that species are in spatial equilibrium with their environment (Václavík and Meentemeyer, 2012). In this paper we have presented an approach to modelling the past, current, and future distributions of invasive species using presence-only data which solves these two main problems with current approaches. We have shown here that by joining models of the spatio-temporal patterns of the dispersal and establishment of invasive species with the observation process by which they are detected, we can simultaneously solve the problem of non-equilibrium spatial distributions while accounting for the uncertainty posed by presence-only data.

Our predictions revealed patterns which would not be readily discernible using current presence-only methods. First, the proportion of sites invaded and their spatial configuration does not follow exactly the pattern of detections. Instead, the unique spatio-temporal distribution of detections inform the model not only about species-specific environmental tolerances, but also about which trajectories of invasions were more likely to lead to the observed detections. For instance, when the time between detections is several years apart, and the locations of those detections are not close together in space, the model 'fills in the gaps' with likely scenarios, identifying the location and timing of stepping stone invasions that were most likely to have occurred in between. Second, we found that the number of detections were not on their own predictive of the number of sites invaded. As of 2013, the conspicuous and publicly well known zebra mussel had been detected at 85 sites and is predicted to be currently established in 113 locations, while the Chinese and banded mysterysnails are predicted to be invaded at 75 and 79 sites, while only having been detected at 6 and 10 sites, respectively (Fig. 4–6). A similar fraction of sites predicted to be invaded have been detected with Eurasian water-milfoil (17 detected out of 111 invaded). While the mysterysnails and water-milfoil may go unidentified and otherwise under-reported at many locations, the quagga mussel, thought to have a fairly restricted range seems to have been detected at nearly all of the sites at which it is currently expected to be established. Failure to account for these differences in detection rates between species, therefore, could lead in some cases to greatly underestimated predictions of both the current extent and future risk of invasion across sites.

In addition to being predicted to be the most currently widespread species in our set, Spiny waterflea and Purple loosestrife also present the highest cumulative future risk of invasion to additional sites over the next ten years (Fig. 4–7). While partly a consequence of the force of invasion (more currently invaded sites represent a larger total flow of propagules to uninvaded sites), this is only one of the driving factors of future invasion risk. Since our model estimates environmental tolerances, we also observe species for which the majority of suitable sites have already been invaded, effectively saturating available habitat and limiting the rate of future spread. For instance, our model predicts Eurasian water-milfoil to have quickly established populations at as many as 20% of sites in Ontario within the first 3 years following its initial discovery in Lake Erie in 1992. Since then, however, additional spread has been limited to a rate of a little more than one new invasion per year (Fig. 4–5). Both of the snail species in our set show similar profiles of quickly saturating suitable locations, suggesting that most of the currently uninvaded locations are likely to remain uninvaded over the next 10 years (Fig. 4–5 and 4–7). The spatial distribution of future risks is also informative, as it tells us which species present risks over a broad spatial range and which are confined to a specific region. This is once again different from purely correlative suitability models as we are assessing future probabilities of establishment resulting from ongoing dispersal, as opposed to changes in the environment. For instance, the projected future risk of Zebra mussel invasion is confined to central Ontario, while Rusty crayfish, Flowering rush, and European common-reed present risks in both south-eastern and north-western Ontario.

The general approach developed here is widely applicable to any species undergoing range expansion for which temporally explicit presence-only data exist. We built the specific model components for the application with the goal of simplicity in mind. However, the specific forms of the dispersal, establishment, and detection sub-components of the model can readily be swapped out for alternative formulations. While we demonstrated the utility of the approach using a gravity model of human-mediated dispersal, there is nothing stopping any other form of dispersal model (eg coupled map lattice, cellular automata, dispersal kernels - fat tailed, or otherwise, etc) from being employed instead. Similarly, we assumed a simple suitability and establishment model using a first order logistic relationship and where within-site establishment was represented as a dichotomous variable, ignoring local scale population proliferation (Ibáñez et al., 2014). It should be possible, however, to include more complex intra-patch population dynamics with only minor adjustments to the general approach outlined in this paper. Finally, our observation model was the simplest possible, assuming constant probability of detection across both space and time. We have validated the theoretical behaviour of the model when incorporating site-level predictors of detection probability, but future work could extend this approach to allow probability of detection to vary over time as well. The advantage of such a general, modular system is that it can be readily extended and applied to additional species as a part of integrated risk assessments (Ibáñez et al., 2014) and management decision making, increasing its attractiveness for policy applications (Leung et al., 2012).

4.5.1 Concluding remarks

In summary, the framework we have presented here solves two pressing challenges facing the modelling of invasive species distributions using presence-only data. First, our framework explicitly incorporates the spatio-temporal dispersal and establishment processes, going beyond the assumption of species-environment equilibrium made by most presence-only methods. Second, by modelling the data generating process as the joint outcome of hidden dispersal and establishment events, as well as a stochastic observation process leading to the detected presences, we provide a solution to the problem of presence-only data in non-equilibrium situations. Together, this framework provides both a powerful tool for informing management decisions, as well as for facilitating scientific inquiry into how the native ranges of invasive species compare with those in the introduced environment.

4.6 Acknowledgements

We thank J. Bradie for helpful comments on the manuscript. This work was funded by the Canadian Aquatic Invasive Species Network, through grants awarded to Dr. B. Leung.

References

Beaumont, L. J., R. V. Gallagher, W. Thuiller, P. O. Downey, M. R. Leishman, and L. Hughes, 2009. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions* **15**:409–420.

Bradie, J., C. Chivers, and B. Leung, 2013. Importing risk: quantifying the propagule pressure-establishment relationship at the pathway level. *Diversity and Distributions* **19**:1020–1030.

Catterall, S., A. R. Cook, G. Marion, A. Butler, and P. E. Hulme, 2012. Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography* **35**:001–011.

Chivers, C. and B. Leung, 2012. Predicting invasions: alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects. *Journal of Applied Ecology* **49**:1113–1123.

EDDMapS, 2013. Early Detection & Distribution Mapping System for Ontario. The University of Georgia - Center for Invasive Species and Ecosystem Health.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M. Overton, A. Townsend Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann, 2006. Novel methods improve prediction of species' distributions from occurrence data.

Ecography **29**:129–151.

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

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates, 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**:43–57.

Gallien, L., R. Douzet, S. Pratte, N. E. Zimmermann, and W. Thuiller, 2012. Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography* **21**:1126–1136.

Gertzen, E. L. and B. Leung, 2011. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

González-Megías, A., J. Gómez, and F. Sánchez-Piñero, 2005. Consequences of spatial autocorrelation for the analysis of metapopulation dynamics. *Ecology* **86**:3264–3271.

Gormley, A. M., D. M. Forsyth, P. Griffioen, M. Lindeman, D. S. Ramsey, M. P. Scroggie, and L. Woodford, 2011. Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species. *The Journal of applied ecology* **48**:25–34.

Guisan, A. and W. Thuiller, 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**:993–1009.

Hallstan, S., R. Johnson, and L. Sandin, 2013. Effects of Dispersal-Related Factors on Species Distribution Model Accuracy for Boreal Lake Ecosystems. *Diversity* **5**:393–408.

Hastie, T. and W. Fithian, 2013. Inference from presence-only data; the ongoing controversy. *Ecography* **36**:864–867.

Hijmans, R. J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* **93**:679–88.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**:1965–1978.

Hui, C., L. C. Foxcroft, D. M. Richardson, and S. MacFadyen, 2011. Defining optimal sampling effort for large-scale monitoring of invasive alien plants: a Bayesian method for estimating abundance and distribution. *Journal of Applied Ecology* **48**:768–776.

Ibáñez, I., J. Diez, L. L. P. Miller, J. D. Olden, C. J. B. Sorte, D. M. Blumenthal, B. A. Bradley, C. D'Antonio, J. S. Dukes, R. Early, E. D. Grosholz, and J. J. Lawler, 2014. Integrated Assessment of Biological Invasions. *Ecological Applications* **24**:25–37.

Ibáñez, I., J. a. Silander, A. M. Wilson, N. LaFleur, N. Tanaka, and I. Tsuyama, 2009. Multivariate forecasts of potential distributions of invasive plant species. *Ecological applications : a publication of the Ecological Society of America* **19**:359–75.

Jackson, J. E., 2005. A User's Guide to Principal Components. Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ.

Johnston, L. E., A. Ricciardi, and J. T. Carlton, 2001. Overland dispersal of aquatic invasive species: a risk assessment of transient recreational boating. *Ecological Applications* **11**:1789–1799.

Kulhanek, S. a., B. Leung, and A. Ricciardi, 2011. Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecological applications* **21**:203–13.

Leung, B., J. M. Bossenbroek, and D. M. Lodge, 2006. Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biological Invasions* **8**:241–254.

Leung, B., J. M. Drake, and D. M. Lodge, 2004. Predicting invasions: Propagule presure and the gravity of allee effects. *Ecology* **85**:1651–1660.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti, 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings. Biological sciences / The Royal Society* **269**:2407–13.

Leung, B. and N. E. Mandrak, 2007. The risk of establishment of aquatic invasive species: joining invasibility and propagule pressure. *Proceedings. Biological sciences / The Royal Society* **274**:2603–9.

Leung, B., N. Roura-Pascual, S. Bacher, J. Heikkilä, L. Brotons, M. a. Burgman, K. Dehnen-Schmutz, F. Essl, P. E. Hulme, D. M. Richardson, D. Sol, M. Vilà, and M. Rejmanek, 2012. TEASIng apart alien species risk assessments: a framework for best practices. *Ecology letters* **15**:1475–93.

Leung, B. and R. J. Steele, 2013. The value of a datum - how little data do we need for a quantitative risk analysis? *Diversity and Distributions* **19**:617–628.

Li, W., Q. Guo, and C. Elkan, 2011. Can we model the probability of presence of species without absence data? *Ecography* **34**:1096–1105.

Lobo, J. M., A. Jiménez-Valverde, and J. Hortal, 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* **33**:103–114.

MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. a. Graniero, 2004. Backcasting and Forecasting Biological Invasions of Inland Lakes. *Ecological Applications* **14**:773–783.

Muirhead, J. R. and H. J. Macisaac, 2005. Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology* **42**:80–90.

Muirhead, J. R. and H. J. MacIsaac, 2011. Evaluation of stochastic gravity model selection for use in estimating non-indigenous species dispersal and establishment. *Biological Invasions* **13**:2445–2458.

Polasky, S., S. R. Carpenter, C. Folke, and B. Keeler, 2011. Decision-making under great uncertainty: environmental management in an era of global change. *Trends in ecology & evolution* **26**:398–404.

Royle, J. A., R. B. Chandler, C. Yackulic, and J. D. Nichols, 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution* **3**:545–554.

Timar, L. and D. J. Phaneuf, 2009. Modeling the human-induced spread of an aquatic invasive: The case of the zebra mussel. *Ecological Economics* **68**:3060–3071.

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

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

Vander Zanden, M. J., G. J. Hansen, S. N. Higgins, and M. Kornis, 2010. A pound of prevention, plus a pound of cure: early detection and eradication of invasive species in the Laurentian Great Lakes. *Journal of Great Lakes Research* **36**:199–205.

Ward, G., T. Hastie, S. Barry, J. Elith, and J. R. Leathwick, 2009. Presence-only data and the em algorithm. *Biometrics* **65**:554–63.

Symbol	Description
$\boldsymbol{\mathcal{L}}$	Likelihood of the model parameters given the observed detections.
S_i	Year of establishment of the i^{th} site (unknown state).
D_i	Year of first detection at the i^{th} site.
$\boldsymbol{\zeta}_i$	Vector of J environmental characteristics of the i^{th} site
$\boldsymbol{\eta}_i$	Vector of characteristics of the i^{th} site which may predict ρ_D
$\boldsymbol{\theta}_\lambda$	Parameters of the dispersal model (d, e in the case of our gravity model).
$\boldsymbol{\theta}_E$	Parameters of habitat suitability model ($c, \beta_{E0}, \ldots, \beta_{EJ}$). These define the relationship between propagule pressure and probability of establishment.
$\boldsymbol{\theta}_D$	Parameters of probability of detection model ($\beta_{D0}, \ldots, \beta_{DJ}$).
$\rho_{E_i}(t)$	Probability of establishment at the i^{th} site at time t. This is a function of propagule pressure, population dynamical parameters, and habitat suitability (environment).
ρ_{D_i}	Probability of detection at the i^{th} site. This is the probability, given that the species is present, that it will be detected in a given unit time. Can be a function of site characteristics.

Table 4–1: **Definitions of symbols**

Species	Year first detected	No. Detections prior to 2006	No. Detections 2006-2013	AUC [95%]
Banded Mysterysnail *Viviparus georgianus*	1998	4	6	0.56 [0.43-0.66]
Chinese Mysterysnail *Cipangopaludina chinensis*	1996	2	4	0.76 [0.50-0.89]
Eurasian Water-milfoil *Myriophyllum spicatum L.*	1992	7	10	0.70 [0.61-0.83]
European Common Reed *Phragmites australis*	1948	31	7	0.92 [0.63-0.98]
Flowering Rush *Butomus umbellatus L.*	1949	8	9	0.82 [0.57-0.94]
Purple Loosestrife *Lythrum salicaria L.*	1991	71	13	0.81 [0.76-0.85]
Quagga Mussel * *Dreissena bugensis*	1992	6	1*	0.94 [0.72-0.99]
Rusty Crayfish *Orconectes rusticus*	1964	51	9	0.83 [0.78-0.87]
Spiny Waterflea *Bythotrephes longimanus*	1982	58	27	0.76 [0.71-0.81]
Zebra Mussel *Dreissena polymorpha*	1988	70	15	0.96 [0.94-0.97]

Table 4–2: **Results of validation of detection predictions.** The number of detections used for fitting (prior to 2006) and validation (2006-2013) are given for comparison. *Note that while Quagga Mussel seems to perform well in terms of AUC, care should be taken in interpreting this result as there is only one detection in the validation period.

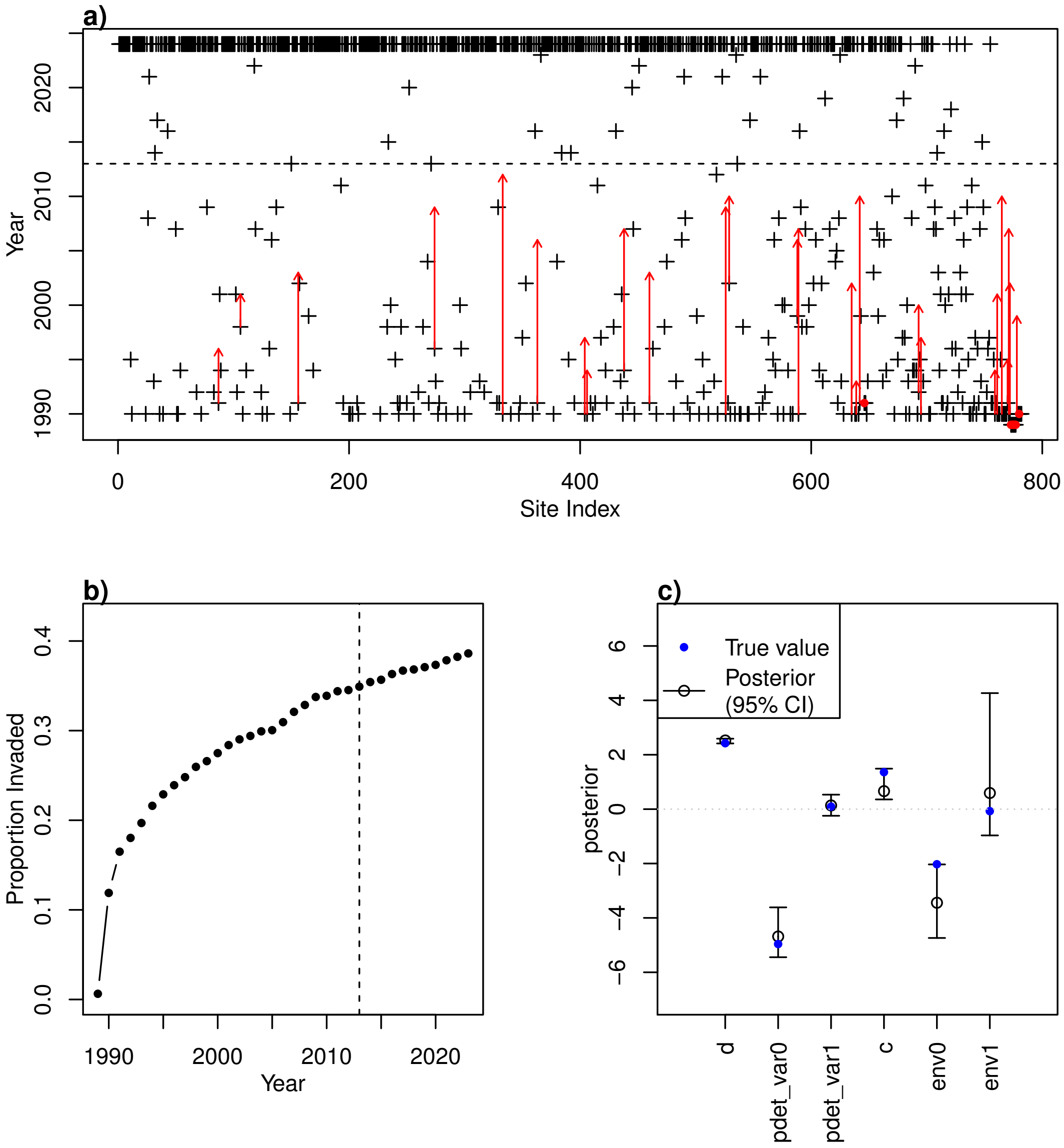


Figure 4–1: **An example simulated spread and detection outcome.** (a) Black crosses indicate the year of invasion for each of the 781 sites. Crosses at 2024 indicate sites which did not become invaded as of 2023. Red arrows indicate sites which had detections. The base of the arrow is at the year of invasion and the tip indicates the year of detection. (b) The proportion of sites invaded over time, including projections 10 years into the future. (c) Generating (true) values of the model parameters in blue with the posterior predictive distribution shown with open black circles (95% CI error bars).

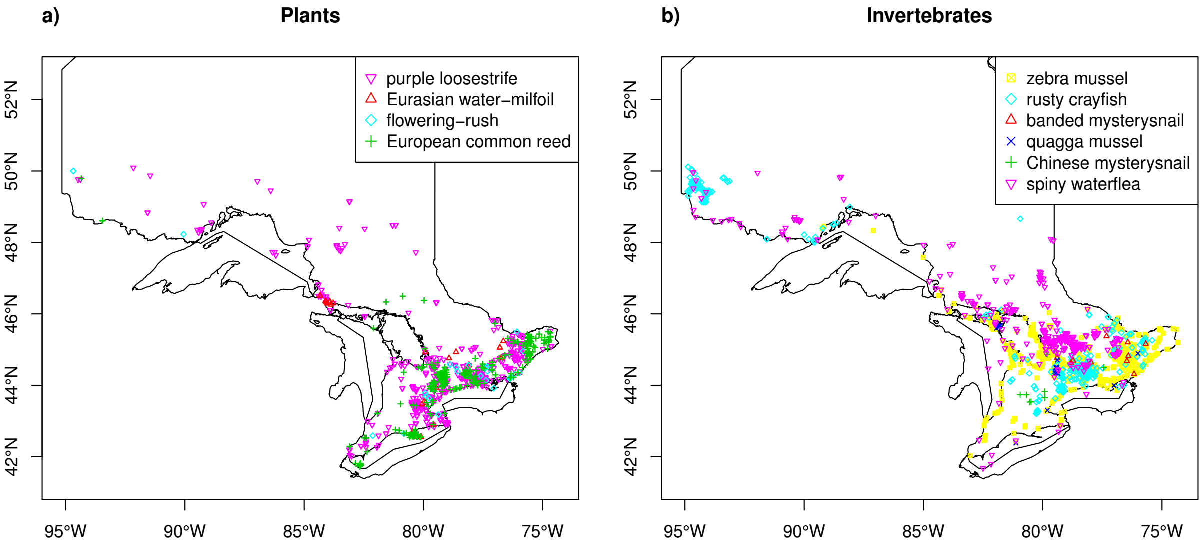


Figure 4–2: **All reported sightings as of 2013.** (a) 4 species of aquatic invasive plants and (b) 6 aquatic invasive invertebrate species in Ontario. Data from EDDMapS (2013).

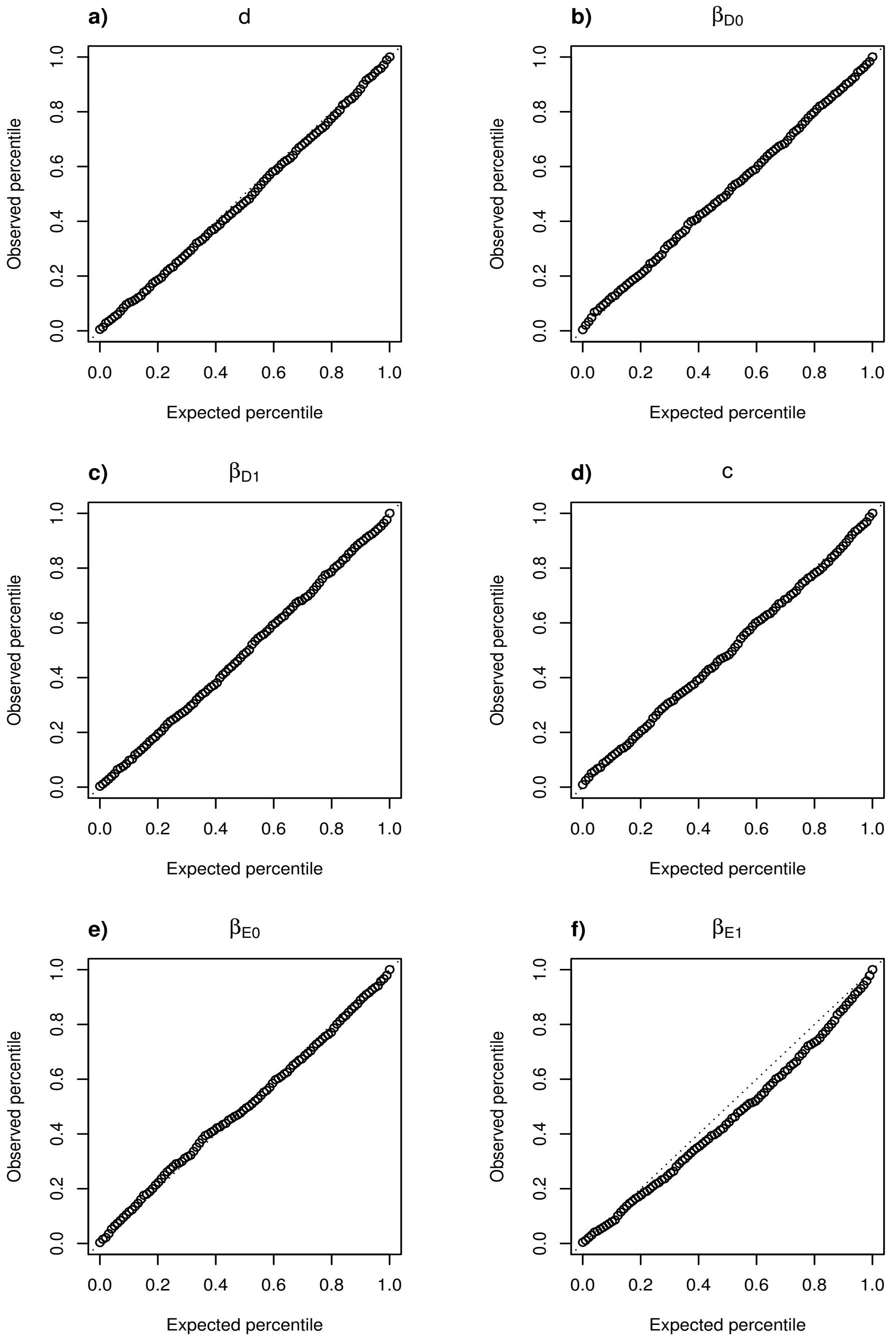


Figure 4–3: **P-P plots of posterior predictive distributions in model fits on simulated data.** A 1:1 relationship indicates that the true (generating) parameter falls in every quantile of the predictive distribution with the expected frequency meaning that the posterior distribution is an unbiased estimator of the underlying model parameter.

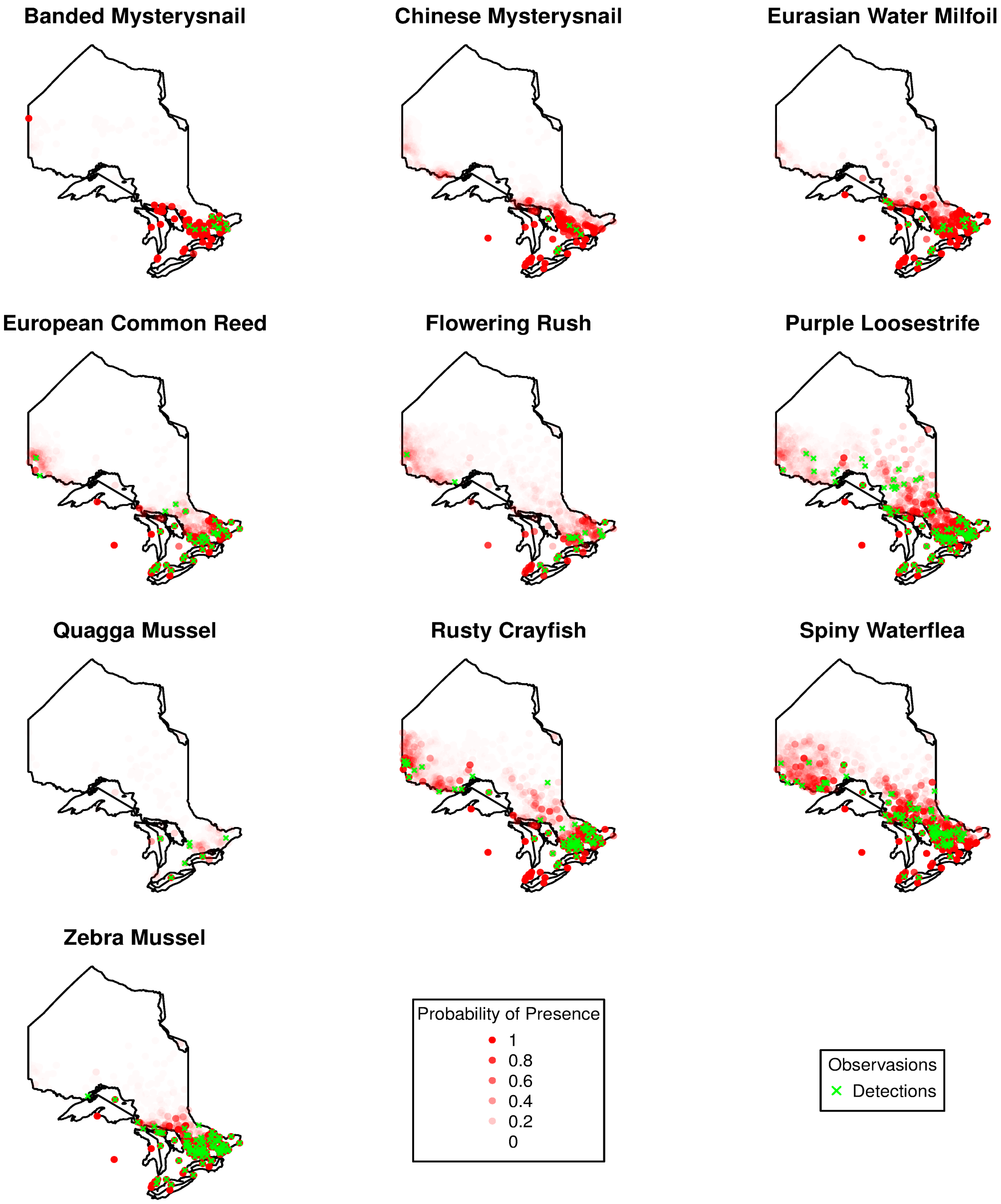


Figure 4–4: **Spatial distribution of probability of presence as of 2013 for 10 aquatic invasive species in Ontario.** Locations at which each species have been detected are indicated with green crosses.

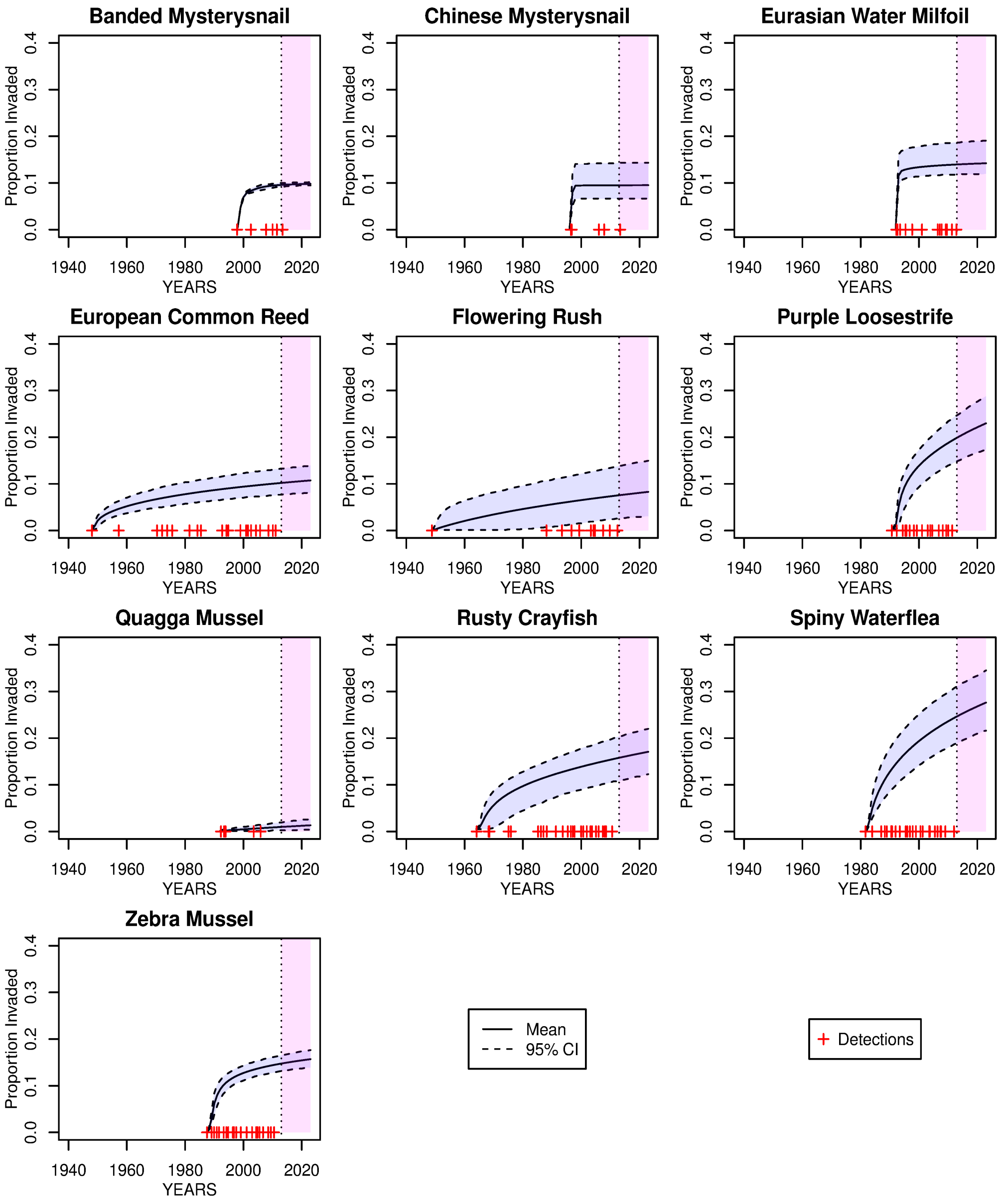


Figure 4–5: **Hind-casts and forecasts of the proportion of sites invaded for each of the 10 aquatic invasive species in Ontario.** The vertical line demarcates the boundary (2013) between hindcasts and projected trajectories over a 10 year time horizon. Red crosses indicate the temporal distribution of new detections for each species.

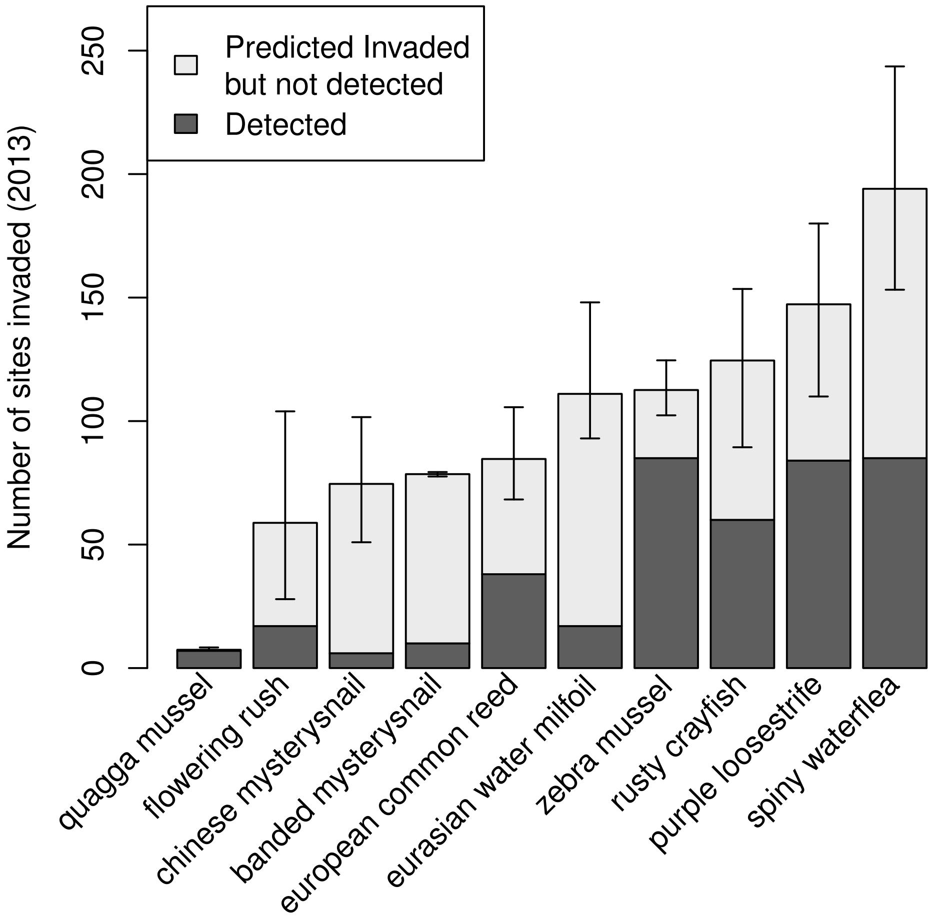


Figure 4–6: **Number of sites predicted to be invaded compared with the number of sites at which each species have been detected**. Total number of invasions stacked by those which have been detected and those which are predicted to have been invaded by 2013 and yet remain undetected.

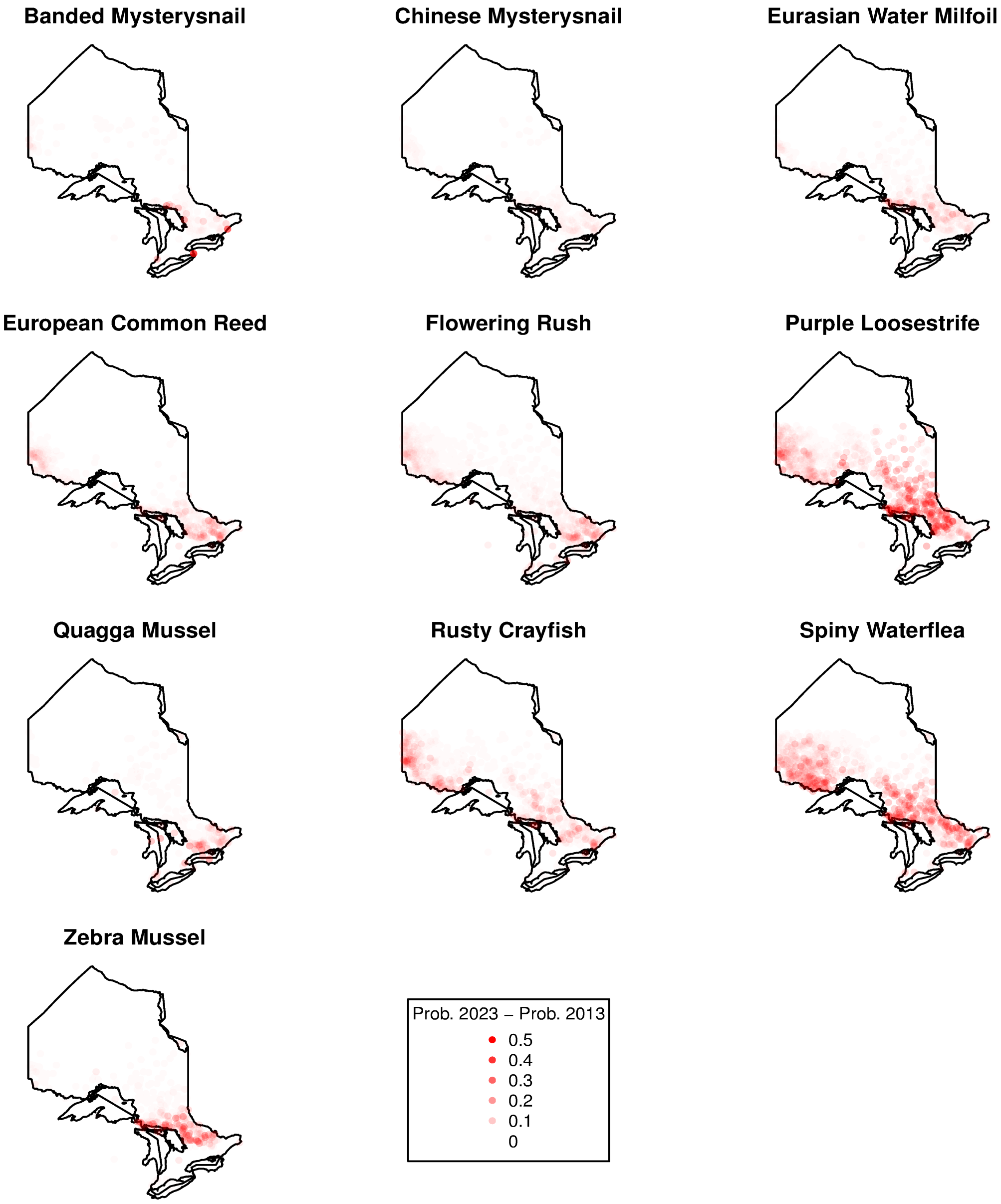


Figure 4–7: **Cumulative additional risk of invasion over the time horizon spanning from 2013-2023.** Of use to management decisions is information not only about which sites are likely to be currently invaded, but also which sites are at highest risk of future invasion and hence would benefit most from mitigation efforts.

Connecting statement

In the previous chapter, I developed a novel approach to predicting the spread of invasive species when presence-only data is available. This methodology works by modelling establishments as hidden states and detections as a stochastic observation process. The application to 10 aquatic invasive species revealed that while for some species, the detected sites represent a large proportion of the current extent of the invasion, others are predicted to be currently established at as many as 10 times more sites those at which they have been detected. Further, while some species have already colonized the majority of sites for which they are environmentally suited, others (Spiny waterflea and Purple loosestrife) are predicted to continue to spread and become established at many more sites over the next decade. Management efforts to reduce spread would be best targeted at species who are still in the early stages of invasion with many suitable sites into which to expand where such efforts would be most beneficial. In the next chapter, I present a framework for modelling the efficacy of management interventions aimed at spread-reduction. The interventions consist of hull-washing stations at boat launches which I incorporate into a gravity modelling framework. By focusing on the economic behaviours of the human vectors of aquatic invasive species, I analyse the projected effects of place-based interventions and suggest that their regional efficacy can be offset by behavioural feedbacks.

CHAPTER 5

Modelling responses to management intervention for controlling the spread of freshwater invasives

Corey Chivers[1] , Brian Leung[2] [3]

A version of this chapter will be submitted to *Biological Invasions*.

---

[1] Department of Biology, McGill University.Stewart Biology Building 1205 ave Docteur Penfield Montreal, Quebec. 514-398-1833. corey.chivers@mail.mcgill.ca.

[2] Department of Biology, McGill University

[3] School of Environment, McGill University. brian.leung2@mcgill.ca

5.1 Abstract

Human-mediated dispersal is the main mechanism of spread of non-indigenous species. As such, numerous models have been developed to describe movement of human vectors and predict the future spread of invaders. To be maximally useful for guiding management, such predictions need to be integrated with management models of how different policies change human behaviour, how behaviours interact with and may vary across the spatial landscape, and how these interactions moderate the invasion progress. We use the dispersal of fresh water organisms by recreational boaters as our study system and mandatory location-specific boat cleaning as our management strategy. The cost of such strategies at specific boat launches imposes an additional factor which may influence a boaters' decision to use a given lake. We conducted a survey of recreational boaters (n=580 respondents, t=2354 boating trips) in Ontario, Canada in order to assess how boater behaviour might change under location-specific management strategies. By building on the concept of gravity models, we demonstrate how to integrate boaters' behavioural feedbacks to management actions and analyse the projected efficacy of various scenarios at both the local and landscape levels. In our analysis of 10 aquatic invasive species in Ontario, we found that the addition of as few as 16 mandatory cleaning stations could protect nearly 86% of currently uninvaded lake area over a ten year time horizon.

5.2 Introduction

Preventing the introduction and establishment of exotic and potentially invasive species into novel ecosystems is a central challenge for resource managers (Leung et al., 2012). Once a species has been detected in a new area, efforts often switch toward controlling secondary spread (Vander Zanden and Olden, 2008; Epanchin-Niell and Hastings, 2010; Parry et al., 2013). This is done to limit the new range in order to minimize the impacts exerted on the surrounding recipient ecosystems, or if detected early to contain the spatial extent while an eradication effort is attempted (Edwards and Leung, 2009). For freshwater invasive species, secondary spread occurs across a landscape of discrete patches (lakes), connected by vectors of overland dispersal. In the North American glacial lakes regions, this overland dispersal has been found to be primarily mediated by human vectors - mostly trailered boats on which species 'hitchhike' from one body of water to another (Johnson and Carlton, 1996; Johnston et al., 2001). By quantifying the movement patterns of recreational boaters, previous studies have used gravity models to successfully predict spatial trajectories of the spread of aquatic invasives (Bossenbroek et al., 2001; Leung et al., 2004, 2006; Gertzen, 2010).

While the utility of Gravity Models (GMs) for the prediction of spread of invasive species has been well demonstrated, their use as a decision support tool for managing secondary spread has not heretofore been investigated. Previous work has suggested that Random Utility Models (RUMs), an alternative approach to modelling boater behaviour, are required in order to predict the consequences of management (MacPherson et al., 2006; Timar and Phaneuf, 2009). These models impute the price of visiting a site as the travel cost needed to gain access, and the utility derived from visitation as some function of characteristics of the site (eg. water quality, lake area, sport fish presences, etc.) (Timar and Phaneuf, 2009). Predictions from this approach suggest that the imposition of any additional cost at a given site will reduce demand, and hence reduce the number of trips taken. The endogenous treatment of the behaviour of individual boaters in RUMs allows for counterfactual analysis of the effectiveness of management actions. However, Chivers and Leung (2012) demonstrated that RUMs and GMs can be reduced to alternative functional forms of boaters' trip-taking distributions, and that in a study of recreational boaters in Ontario, a GM provided a better fit to observed trip-taking behaviour. Given the explanatory power of the GM for predicting the spread of aquatic invasive species, it is a natural starting position for exploring how boaters' behaviours change in response to management and how these changes are likely to affect the spread of the myriad species which they may be inadvertently transporting.

We conducted a survey of recreational boaters, and used it to assess how trip decisions may be influenced by local, as well as landscape level management strategies which employ location-specific mandatory boat hull cleaning. We formalize the behavioural responses within a gravity model of boater movement and assess the impact of these responses on the efficacy of such policies to limit the spread of aquatic invasive species to inland lakes. Our approach provides a generalizable decision support tool for managers of lake resources who are aiming to mitigate the spread of invasive species. The decision support tool developed here has several desirable features: 1) It captures both the reduction in overall trip-taking, as well as redistributive trip-taking behaviours due to management. By separating the response to management into these two components, we are able to account for avoidance behaviours which, while protective for a particular target lake, may lead to an increased risk of invasion to nearby or similar lakes. 2) It is not data intensive. We show that the behavioural responses to management can be estimated using a simple survey design employing randomized counterfactual scenarios and stated preferences techniques (Heckman and Vytlacil, 2007; Dillman, 2000), and could be updated with observational data post policy implementation (Adamowicz et al., 1994). 3) It can be used to evaluate the cost effectiveness of any number of policy scenarios, allowing the resource manager to select a management strategy which meets their objectives. 4) It can be extended for application beyond the realm of freshwater invasives to any species with a primarily human-mediated vector of dispersal across a landscape of discrete patches (for example; boring forest pests spreading across a landscape of forest patches via the movement of infested firewood (Prasad et al., 2010).)

In this study, we first outline how behavioural responses to management can be incorporated into a gravity model formulation of human-mediated dispersal. We describe a Bayesian approach requiring information about trip taking behaviour under *status quo* and management conditions at only two lakes for any given survey respondent. We estimate model parameters using data gathered from a survey of recreational boaters in Ontario in 2011. We then apply the management model to investigate the projected efficacy of various intervention scenarios on the future spread of 10 aquatic invasive species in Ontario.

5.3 Methods

5.3.1 Boater survey

During January and February of 2011, survey invitations were sent out by mail to 5000 Ontario Outdoor Cards holders (recreational licenses) to complete an online survey of their boating behaviours during the 2010 boating season. The participants were selected using a stratified random sampling scheme where approximately 100 invitations to participate were mailed to each of 47 geographic regions, as defined by the first two letters of their postal code.

We developed an online survey tool which allows for the rapid collection of spatial information from respondents. Using an interactive map interface, respondents were asked to identify which lakes they visited as well as how many times they launched their boat into each lake during the 2010 boating season (see Figure 5–1). By collecting these responses using an interactive map, we were able to avoid the tedious and error-prone disambiguation process of reported lakes based on lake names and nearest towns alone. Of the 405,997 water bodies in Ontario, only 14,767 (3.6%) have officially recognized names, and 62, 59, and 44 of those are named Long Lake, Mud Lake, and Otter Lake, respectively. Additionally, several lakes are known to local residents by various conflicting names which do not have any 1:1 match in the Ontario GIS database.

Once boaters identified all lakes which they had visited during 2010 and had answered a few additional questions about their boating behaviours they were presented with a counterfactual management scenario. Counterfactual analysis is used to compare what actually happened with what would have happened in the presence (or absence) of intervention (Phaneuf, 2013). In a hypothetical scenario, a mandatory boat cleaning station is implemented at the launch site of one of the respondent's two most frequently visited lakes. The scenario included a description of the process involved including an estimated time to complete the cleaning process of 15 minutes. In addition to the general policy information which was the same for each respondent except for the location of implementation, a randomized estimate of the cost (0-16$) that would be incurred out of pocket at this wash station was also presented to the respondent (see Appendix C.1.1 for the complete text of the cleaning policy description). After reading this description, the respondent was reminded how many times they reported launching their boat at the policy lake and was then asked how many times they would have launched, were the scenario described to have been implemented at that location. Next, they were asked how many times they would have visited the non-policy lake under the policy scenario. This process of eliciting stated preferences for modelling the choice behaviour of individuals has been employed in the econometrics literature (Haider, 2002). The responses to these questions were then used to fit the management modified gravity model (see section 5.3.2). Additionally, we asked several other questions pertaining to boating behaviour as well as questions about whether they currently engage in various preventative measures relating to the transport of aquatic invasive species.

Using the survey design method of Dillman (2000) we field tested our survey tool using a pilot run on location at several boat launch locations in the District Municipality of Muskoka in central Ontario. By directly observing respondents interact with the questionnaire and identifying common problems and any misinterpretations of the instructions, we were able to streamline the user experience and clarify the written guide before sending out invitations to our main sample.

5.3.2 Gravity model

To estimate boater traffic among Ontario lakes, a production-constrained gravity model was used (Gertzen and Leung, 2011; Chivers and Leung, 2012). In this formulation, boaters are assumed to travel from their home location to a lake, and return to their home location before making another trip. While there may be instances where boaters travel directly from one lake to another (competitive anglers, for instance), the production-constrained formulation has been found to be the most predictive and least data-intensive version of the model for explaining the human-mediated spread of aquatic invasives (Muirhead and MacIsaac, 2011). Following the disaggregated notation of Chivers and Leung (2012), the site selection probability distribution for an individual boater n can be written as:

$$P(T_{nj}) \quad = \quad A_n W_j^e D_{nj}^{-d}, n = 1, \ldots, N, j = 1, \ldots, J. \tag{5.1}$$

Where W_j is the attractiveness of lake j (expressed as lake surface area in hectares), and D_{nj} is the distance between lake j and the home location of individual boater n (where they keep their boat when not in use). The parameters e and d describe the shape of the relationship between the 'pull' of lake i with lake size and distance from source j. A_n is the total 'pull' of all lakes:

$$A_n = 1/\sum_{k=1}^{J} W_k^e D_{nk}^{-d}. \tag{5.2}$$

This term imposes the constraint that $\sum_{j=1}^{J} T_{nj} = 1$, making equation 5.1 a proper probability distribution representing the probability that boater n will choose lake j on a given outing. The formulation given by equations 5.1 and 5.2 is similar to the formulation used by Leung et al. (2004), with the addition of a non-linearity on the W_j term. Further details of the base gravity model can be found in Chivers and Leung (2012).

5.3.3 Boater response to management

In order to limit the risk of introductions by trailered boats, some jurisdictions have begun to trial mandatory cleaning stations located at selectively chosen boat launches (Rothlisberger et al., 2010). The cleaning stations aim to reduce the total propagule pressure of non-indigenous species entering and leaving the selected lake(s). As with any policy intervention, however, there are costs involved. Whether or not the operational costs are borne directly by the users in the form of a per-use fee, there is a time cost involved in carrying out the cleaning which may affect the behaviour of individual boaters. We consider the two most likely *a priori* aspects of this behavioural change. First, faced with the cost of compliance, boaters may choose to reduce the number of times they visit the policy lake, substituting some other non-boating activity instead. In this case, the policy has successfully limited the propagules entering and leaving the lake, although this also has the negative effect of discouraging recreational boating overall. Secondly, boaters may choose to visit the policy lake less often but instead visit alternative, otherwise similar, lakes. In this situation, propagule pressure will be reduced at the policy lake, but the redistribution of boater traffic has the effect of increasing the flow of propagules between similar lakes in the vicinity. Both of these behavioural responses are consistent with that which would be predicted using traditional utility-based econometric models of recreational demand (Timar and Phaneuf, 2009; Fenichel et al., 2012).

To demonstrate how we estimate the strength of these behavioural responses to a management action, let us first define the two quantities of interest. First we wish to quantify the proportional reduction in attractiveness θ_m that occurs as a result of the implementation of a management intervention at lake m. The effect of this quantity will be to make the policy lake appear smaller or to have less 'pull' relative to the alternative lakes, to use the gravity analogy. Secondly, we wish to quantify the proportional reduction to the overall number of trips taken ϕ_m by individual boaters as a result of a management intervention at lake m. These two quantities enter into our original gravity equations as modifiers on the relative attractiveness of the policy lake m and non-redistributed reduction of the trip-taking probability to lake m, respectively.

$$P^*(T_{nj}) \quad = \quad A_n \theta_j W_j^e D_{nj}^{-d} \phi_j, n = 1, \ldots, N, j = 1, \ldots, J. \tag{5.3}$$

$$A_n \quad = \quad 1 / \sum_{k=1}^{J} \theta_k W_k^e D_{nk}^{-d}. \tag{5.4}$$

At all non-policy lakes $w = w_1, \ldots, w_W$ the vectors θ_w and ϕ_w are equal to 1. At all policy lakes $m = m_1, \ldots, m_M$ where the management intervention has been implemented, θ_m and ϕ_m are $\in [0,1]$. Notice that θ appears in both the numerator and the denominator, while ϕ only appears in the numerator. We can see from this that $\sum_{j=1}^{J} P^*(T_{nj}) = 1$ in the strictly redistributive case where $\theta_m \leq 1, \phi_m = 1$. We term this effect 'redistribution' because the net result is that there are no fewer trips taken by boaters under management than under the *status quo* condition, rather only that the trips have been (at least partially) redistributed from policy lakes to non-policy lakes. However, in the case of some reduction in overall trip-taking where $\phi_m \leq 1$, then $\sum_{j=1}^{J} P^*(T_{nj}) \leq 1$. In this case there is some probability $1 - \sum_{j=1}^{J} P^*(T_{nj})$ that an individual boater will choose to abstain from taking a boating trip entirely. We call this reduction effect 'loss' as it represents a net loss of boating activity resulting from management intervention. A conceptual diagram of how each θ_m and ϕ_m impact trip taking distributions to result in various combinations of loss and redistribution is given in Figure 5–2.

In order to estimate θ_m and ϕ_m from our survey data, we reformulate the trip outcomes under each the control (reported actual trips taken) and the counterfactual policy condition (trips that would have been taken given the hypothetical scenario) as random samples from binomial distributions. Within each, the binomial success probability parameter is described by the probabilities given by equations 5.1 and 5.3. We can then construct the likelihood function by substituting the number of trips taken to each the policy and non-policy lakes before and after implementation.

$$L(\theta_m, \phi_m \mid D) = \prod_{n=1}^{N} \underbrace{\binom{S_{na}}{S'_{na}} \delta_{na}^{S'_{na}} (1 - \delta_{na})^{S_{na} - S'_{na}}}_{\text{Policy lake}} \underbrace{\binom{S'_{nb}}{S_{nb}} \delta_{nb}^{S_{nb}} (1 - \delta_{nb})^{S'_{nb} - S_{nb}}}_{\text{Non-policy lake}} \tag{5.5}$$

Where $\delta_{na} = \frac{P^*(T_{na})}{P(T_{na})}$ and $\delta_{nb} = \frac{P(T_{nb})}{P^*(T_{nb})}$, which are the proportional changes in the probability of visiting each lake under *status quo* (equation 5.1) management (equation 5.3). S_{na} and S'_{na} are the number of trips taken by boater n to the policy lake under *status quo* and management, respectively. S_{nb} and S'_{nb} are the number of trips taken by boater b to the non-policy under the same two conditions.

To understand what is going on here, imagine that a boater made 10 trips to lake a in 2010, and 5 trips to lake b. After considering a mandatory cleaning policy implemented at lake a, the boater decides that they would have visited only 6 times, and redistributed the remaining 9 trips to lake b. The likelihood function can be broken down into two binomial likelihoods. The first half of equation 5.5 is the likelihood of having observed 6 (S'_{na}) out of 10 (S_{na}) trips at the policy lake if the probability has been modified by the fraction δ_{na}. The second half can be thought of as an inverse process at the non-policy lake, where we calculate the likelihood of observing 5 (S_{nb}) out of 9 (S'_{nb}) trips taken between the *status quo* and policy conditions, under the modification in probability predicted by the gravity model (δ_{nb}). We validated the theoretical ability to recapture θ_m and ϕ_m with this formulation using simulated survey response data (see Appendix C.1.2).

Incorporating cost to boaters

In addition to the general behavioural effect, and in keeping with the predictions of econometric models of recreational demand, we expect the behavioural responses to be a function of the direct cost to the boater imposed by the policy. Recall that each respondent was presented with a randomized cost estimate between 0-16$. To factor this in, we model θ_m and ϕ_m in equations 5.3 and 5.4 as a function of the cost:

$$\begin{aligned} \theta_m &= f(cost) \\ &= exp(-\alpha_\theta(cost + \gamma_\theta)) \end{aligned} \tag{5.6}$$

$$\begin{aligned} \phi_m &= g(cost) \\ &= exp(-\alpha_\phi(cost + \gamma_\phi)) \end{aligned} \tag{5.7}$$

Where α_θ, γ_θ, α_ϕ, and γ_ϕ are shape parameters estimated using MCMC by substituting equations 5.6 and 5.7 into equation 5.5 (Calder et al., 2003). The functional forms for $f(cost)$ and $g(cost)$ were chosen as they capture several properties that we would logically expect. First, they range between 0 and 1, matching our definitions of ϕ_m and θ_m. Second, they are flexible enough to capture an intercept greater than zero (non-zero θ_m, ϕ_m via translation through γ), as there may be a behavioural response even when the monetary cost to the boater is zero, since there is an unavoidable time cost involved. Finally, they are able to capture how the strength of the response changes (accelerating or decelerating) as monetary costs increase through the α parameters. In addition to these functional forms, we also fit a linear model to each, as well as a null model in which θ_m and ϕ_m are both independent of cost and performed model selection using Bayesian Information Criterion (Burnham and Anderson, 2002).

5.3.4 Application to 10 invasives

To assess both the local and landscape-level efficacy of boat washing stations, we apply the model to mitigating the future spread 4 plant and 6 invertebrate aquatic invasive species (Table 5–1). Data on the time and location of confirmed presences of each species was collected from the Early Detection & Distribution Mapping System (EDDMapS, 2013). The current spatial extent, environmental tolerences, and spread dynamics were estimated for each species using the presence-only modelling approach described in Chapter 4. From this model, we project the future spread using stochastic simulations, resulting in estimated probabilities of invasion across space and time.

The modified gravity model defines the new network of dispersal corridors which emerges when cleaning stations are implemented at various lakes (termed 'policy locations'). The modified dispersal networks have two main properties distinguishing them from the dispersal network under *status quo* conditions. First, absent any behavioural response resulting from additional costs borne to boaters, the net propagule outflow from a policy location will of course be reduced as a direct consequence of the cleaning itself. The use of high-pressure sprayers for cleaning has been estimated to effectively remove around 90% of small bodied organisms and other macrophytes from boat hulls and trailers (Rothlisberger et al., 2010). Completely emptying any bilge water and bait buckets, and inspecting angling equipment for gear fouling species like waterfleas can be expected to further reduce the diversity and abundance of outbound propagules. We capture this direct effect of cleaning stations by reducing the net predicted propagule outflow from policy locations by 90%, beginning during the first year that the policy is implemented. Second, the redistributed and lost trips are incorporated into the predicted dispersal network by modifying the gravity equations as described in section 5.3.3. These changes also take effect during the first year that the policy is implemented.

5.3.5 Scenario analysis

To investigate the effects of management strategies which would employ the use of mandatory boat wash stations at selected lakes, we analysed 17 different policy scenarios. The baseline scenario represents the projected spread of all 10 species under *status quo* conditions and is equivalent to that predicted in Chapter 4. For our active policy scenarios, we simulate offensive strategies, where we focus efforts on those sites which are already most likely to be invaded. This is in contrast to a defensive strategy, which prioritizes uninvaded sites aiming to sanitize boats on their way into pristine, uninvaded locations. An offensive strategy is conceptually similar to a quarantine effort, however it differs in that we acknowledge that given limited resources, there will be some uncontained locations which are invaded and therefore potentially emanating propagules to uninvaded sites. Even in this situation, however, provided that fewer than half of the locations are invaded, offensive strategies have been shown to be most effective at reducing overall rates of spread (Drury and Rothlisberger, 2008).

Locations at which to implement the policy were selected by ordering locations based on the total amount of outbound propagule pressure under the *status quo* condition. That is, for each location, we took the total potential outbound propagule pressure (outbound traffic) and multiplied by the probability that the site is currently invaded. We then summed this quantity across each of the 10 species and selected the top M locations at which to implement the cleaning policy. We investigated the consequences of implementing the policy at the top 1,2,4,8,16,32,64, and 128 locations. For each number of policy locations, we ran two scenarios: one in which no cost is incurred directly by boaters and one at the full 16$ per-use fee.

To account for stochastic uncertainty associated with future spread, for each scenario we simulated 5000 stochastic realizations of spread over the ten year time horizon. Uncertainty in the parameters of both the underlying spread and establishment models for each species as well as the parameters of the behavioural model were incorporated by conducting the repeated forward simulations each time using random draws from the posterior predictive distributions of each (Calder et al., 2003). For each ensemble of simulations, the mean number of lakes invaded, total area invaded, as well as the risk of invasion at each location at every time step was recorded. All of our policy scenarios begin implementation in 2013 and are projected forward 10 years to 2023.

5.4 Results

5.4.1 Survey results

Of the 580 respondents (11.6% response rate), 146 reported visiting multiple lakes. The number of unique lakes visited by individual respondents was 3.78 ± 1.2

(mean ± SD), with a maximum number of unique lakes reported visited by a single respondent of 7. As our model treats the behavioural responses to management as homogeneous, population-level effect, we tested for correlations between respondents' behavioural changes under the management condition and several boater specific and spatial factors. No significant relationships (Pearson's r, $p > 0.05$) were found that would indicate that any of boater location, lake location, log(lake surface area), distance to lake, or type of boat outing (angling, water skiing/wake boarding/tubing, sight seeing) were correlated to the number of trips diverted (both redistribution and loss). While we expect heterogeneity to be non-zero, we did not find evidence of any significant boater specific or spatially driven differences between individuals in terms of their responce to management.

5.4.2 Behavioural response to management

The estimated proportion both of trips which boaters would redistribute to other lakes, as well as trips that would be entirely foregone under management intervention both increased with the magnitude of cost incurred by the boater. Of the three models tested, our full model (equations 5.6 and 5.7) was strongly selected, with BIC improvement of 39.1 over the linear model and 411.3 over the null model (Table 5–2). Even at no direct monetary cost, 2.8% [1.2-4.9%] (posterior mean [95% BCI]) of trips to policy lakes would be lost and 5.5% [2.0-6.6%] of remaining trips replaced by trips to alternative locations. Both curves show an accelerating relationship between proportion of trips redistributed/lost and cost. At the maximum cost of 16$, 64% [56-73%] of the trips were lost and of the remaining trips 62% [53-70%] were redistributed (Figure 5–3). Overall, the behavioural model captured 87% of the variance in the total change between the number of trips taken under *status quo* and the reported number of trips that would be taken under the counterfactual management scenarios (Figure 5–4).

5.4.3 Scenario analysis

The expected number of mitigated invasions over the ten year time horizon increased with the number of policy locations as expected. In all cases, when cleaning stations were provided at no cost to boaters the result was more mitigated invasions compared to requiring boater payments (Figure 5–5). However, as the number of policy lakes increased to the maximum of 128 locations, the additional benefit from providing the cleaning stations at no cost was diminished, and in some cases eliminated. This occurred as the level of mitigation reached effective quarantine (ie. zero new invasions), and hence there was no additional benefit to be gained by waiving the fee to minimize traffic redistribution. Over the time horizon that we analysed (2013-2023), the largest numbers of expected mitigated invasions were 23 for Spiny waterflea (2.9% of sites), and 21 (2.7% of sites) for Purple loosestrife, which resulted when the policy was implemented at the maximum number (128) of locations. For most species, at least 8 policy locations were required in order to mitigate an appreciable number of new invasions, especially for the boater pays scenarios (Figure 5–5).

The expected total surface area of the lakes at which invasions were mitigated reveals a different picture than when just considering the number of mitigated invasions. The expected benefits when looked at this way are now largest for Quagga mussel and Flowering rush with expected mitigations of $11,700$ and $5,000km^2$ of lake area which would have become invaded under the *status quo* condition, respectively (Figure 5–6). The difference between the boater pays and no fee conditions is also much less pronounced in terms of mitigated area and shows little to no effect at all for several species. Further, the expected mitigation reaches very nearly full quarantine level for all species using as few as 16 policy locations.

5.5 Discussion

Predicting how control measures will impact the future spread of invasive species is an important component of the management decision making process (Simberloff et al., 2005; Stohlgren and Schnase, 2006; Epanchin-Niell and Hastings, 2010). Here we have built upon the concept of gravity models (Bossenbroek et al., 2001; MacIsaac et al., 2004; Leung et al., 2006), and extended the formulation to incorporate boaters' behavioural responses in order to model how location-based policies may change the configuration of human-mediated dispersal networks, and how those changes may affect the efficacy of such measures for controlling spread. We have demonstrated how the main behavioural responses can be captured in two simple coefficients in the gravity model and we have derived an observation model for estimating them using a survey of recreational boaters. Our results suggest that behavioural responses which redistribute boating trips from a costly policy site to alternative locations can reduce the efficacy of an intervention when measured in terms of the number of mitigated invasions. These effects however are most prominent when the intervention is implemented at only a few locations, and have a much smaller impact on the overall mitigated area of invasion than on the number of mitigations.

The notion of boaters 'redistributing' or reducing their trips in the face of costly policies as described in this paper is similar to the behavioural response that would be predicted by a Random Utility Model of recreational demand, as the additional costs involved make boaters more likely to choose alternative destinations where they would derive the same, or more, net utility (MacPherson et al., 2006; Timar and Phaneuf, 2009). Indeed, that boaters would change their behaviour in the face of additional costs is not surprising, however the perceived benefits of compliance may dampen the extent of these behavioural changes. For instance, boaters who derive and appreciate the benefits of infestation-free waters may be willing accept additional costs in order to contribute to the control of invasive species. Our findings indicate that while some boaters are willing to pay up to a 16$ launch fee (those respondents who did not reduce or redistribute any trips in the face of a full cost cleaning policy), others will adapt their behaviour to avoid incurring even the minimum of 15 minutes of additional time cost. It is the combination of all of these different behaviours that will determine the reduction in propagule pressure. In a study of boaters in five US states, individual boaters were willing to pay an average of $1.90 in additional fees for boater registrations to fund AIS prevention efforts (Jensen, 2010), suggesting that there are non-zero perceived benefits associated with invasion mitigation policies.

While we are proximally interested in the effect that management intervention will have on propagule pressure, we are ultimately interested in how this change will modify the risk of establishment (Wonham et al., 2013). This study provides new tools for addressing two of the three components of optimally controlling established invasive species which were identified in a review by Epanchin-Niell and Hastings

(2010). Specifically, we capture the invasion dynamics, as well as the projected dynamics under various control efforts and their relative costs. In order to determine the best course of action, however, the third component – a monetary measure of invasion damages – is needed for each of the species analysed here. This study therefore represents a part of a more integrated assessment to fully inform policy decisions (Leung et al., 2012; Ibáñez et al., 2014). Without an accounting of the relative potential impacts, both ecological and economic, of the 10 species in our study, we designed our policy scenarios to preferentially target locations at which the greatest overall reduction in outbound propagule pressure was equally weighted across all species could be acheived. This strategy, however, could be refined to one in which the selection of which sites to target is done by weighting the species-specific outflows at each site by the estimated impact of each invader, multiplied by the expected number of lakes facing future invasion risk. This way, an intervention strategy would get the most benefit by preferentially reducing the spread of the most potentially impactful species. Importantly, a complete evaluation of any particular management plan will require that costs be measured in the same currency as damages, and hence benefits of mitigation, such that a cost-benefit ratio can be determined (Epanchin-Niell and Hastings, 2010). Here we have assumed that the fixed costs of implementing a cleaning policy at each additional site is both quantifiable and constant. In reality, however, some locations will require more resources than others as the number of access points is variable between sites and therefore the number of cleaning stations required in order to achieve the desired effect is also likely to differ.

We also assume that the cleaning policies are in place and running over the entire ten year time horizon which we considered. For cleaning stations requiring ongoing operational costs, a dynamic bioeconomic model may provide deeper insights into where and for how long cleaning stations ought to be operational to achieve the optimal results (Leung et al., 2002). The spatial-dynamic aspect spreading invasions, however, may require simplifying assumptions in order to achieve analytical (non-simulation based) representations amenable to solving the optimal control problem and is the subject of ongoing research (Epanchin-Niell and Wilen, 2012). Another aspect of such an analysis could include the future benefits which would accrue as a result of mitigating the spread of currently unidentified invaders. Obviously, the expected benefits in terms of mitigated local invasions derived in this paper account only for those ten species which we analysed. Methods for estimating the rate at which novel, high-impact species will be introduced to a given region exist (Aukema et al., 2011) and could be used to estimate additional future benefits of a given control policy.

Interventions like those which we have analysed here take a landscape level view of the management of aquatic invasive species, and target specifically the inter-patch dispersal and colonization phases of secondary spread. If management jurisdictions are fragmented into multiple independent zones, individual managers may be less likely to implement these forms of control, as the likelihood of invasion from uncontrolled neighbouring jurisdictions may be high and may instead choose to focus resources on strategies aiming to control the proliferation of locally established populations (Shea et al., 2010; Epanchin-Niell et al., 2010; Ibáñez et al., 2014). Successful landscape-level efforts may require coordination among several agencies and jurisdictions, which is not always easy and for which there is no explicit legislative frameworks to facilitate this in the Ontario context (Smith et al., 2013). Additionally, while we have presented here a model for incorporating behavioural responses to management aimed at controlling the spread of aquatic invasive species, such management actions may represent just one aspect of a broader, multi-pronged, strategy for controlling secondary spread. When policies are implemented in concert with, for instance educational campaigns (Padilla and Williams, 2004) or auditing of fish stocking procedures to identify and remove non-target species (Davies et al., 2013), resulting interactions may need to be accounted for when assessing expected benefits. For example, consistent with a previous study of boater attitudes and behaviours regarding aquatic invasive species (Rothlisberger et al., 2010), a large majority of our survey respondents reported that they did not always take steps to clean their boats or check for fouling plants and animals between launches (86% ± 1% (mean ± SE)). Education campaigns aiming to improve these figures may have overall consequences which could lower the marginal benefits of location-specific policies.

In conclusion, we have developed here an extension of the widely used gravity modelling framework to incorporate behavioural responses to management interventions and have demonstrated the utility of this approach for estimating the efficacy of location-based control efforts. This work provides managers with a decision support tool which can be integrated into cost-benefit analyses and help to inform policy planning for the control of invasive species.

References

Adamowicz, W., J. Louviere, and M. Williams, 1994. Combining revealed and stated preference methods for valuing environmental amenities. *Journal of environmental economics and management* **26**:271–292.

Aukema, J. E., B. Leung, K. Kovacs, C. Chivers, K. O. Britton, J. Englin, S. J. Frankel, R. G. Haight, T. P. Holmes, A. M. Liebhold, D. G. McCullough, and B. Von Holle, 2011. Economic Impacts of Non-Native Forest Insects in the Continental United States. *PLoS ONE* **6**:e24587.

Bossenbroek, J. M., C. E. Kraft, and J. C. Nekola, 2001. Prediction of long-distance dispersal using gravity models: zebra mussel invasion of inland lakes. *Ecological Applications* **11**:1778–1788.

Burnham, K. P. and D. R. Anderson, 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer.

Calder, C., M. Lavine, P. Muller, and J. Clark, 2003. Incorporating multiple sources of stochasticity into dynamic population models. *Ecology* **84**:1395–1402.

Chivers, C. and B. Leung, 2012. Predicting invasions: alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects. *Journal of Applied Ecology* **49**:1113–1123.

Davies, G. D., R. E. Gozlan, and J. Robert Britton, 2013. Can accidental introductions of non-native species be prevented by fish stocking audits? *Aquatic Conservation: Marine and Freshwater Ecosystems* **23**:366–373.

Dillman, D. A., 2000. Mail and Internet surveys: the tailored design method. Psychology/Sociology/Marketing. J. Wiley, second edition.

Drury, K. L. S. and J. D. Rothlisberger, 2008. Offense and defense in landscape-level invasion control. *Oikos* **117**:182–190.

EDDMapS, 2013. Early Detection & Distribution Mapping System for Ontario. The University of Georgia - Center for Invasive Species and Ecosystem Health.

Edwards, P. K. and B. Leung, 2009. Re-evaluating eradication of nuisance species: invasion of the tunicate, Ciona intestinalis. *Frontiers in Ecology and the Environment* **7**:326–332.

Epanchin-Niell, R. S. and A. Hastings, 2010. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. *Ecology letters* **13**:528–41.

Epanchin-Niell, R. S., M. B. Hufford, C. E. Aslan, J. P. Sexton, J. D. Port, and T. M. Waring, 2010. Controlling invasive species in complex social landscapes. *Frontiers in Ecology and the Environment* **8**:210–216.

Epanchin-Niell, R. S. and J. E. Wilen, 2012. Optimal spatial control of biological invasions. *Journal of Environmental Economics and Management* **63**:260–270.

Fenichel, E. P., J. K. Abbott, and B. Huang, 2012. Modelling angler behaviour as a part of the management system: synthesizing a multi-disciplinary literature. *Fish and Fisheries* .

Gertzen, E. L., 2010. Assessing the relationship between propagule pressure and probability of establishment for the aquatic invader Bythotrephes longimanus using two complementary approaches Department of Biology. Masters, McGill University.

Gertzen, E. L. and B. Leung, 2011. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

Haider, W., 2002. Stated Preference & Choice Models - A Versatile Alternative to Traditional Recreation Research. In A. Arnberger, C. Brandenburg, and A. Muhar, editors, *Monitoring and Management of Visitor Flows in Recreational and Protected Areas Conference Proceedings*, pages 115–121.

Heckman, J. J. and E. J. Vytlacil, 2007. Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics*, volume 6, Part B of *Handbook of Econometrics*, chapter 70, pages 4779–4874. Elsevier.

Ibáñez, I., J. Diez, L. L. P. Miller, J. D. Olden, C. J. B. Sorte, D. M. Blumenthal, B. A. Bradley, C. D'Antonio, J. S. Dukes, R. Early, E. D. Grosholz, and J. J. Lawler, 2014. Integrated Assessment of Biological Invasions. *Ecological Applications* **24**:25–37.

Jensen, D. A., 2010. Assessing the effectiveness of aquatic invasive species outreach influencing boater behaviour in five states. Masters, University of Minnesota Duluth.

Johnson, L. E. and J. T. Carlton, 1996. Post-establishment spread in large-scale invasions: dispersal mechanisms of the zebra mussel Dreissena polymorpha. *Ecology* pages 1686–1690.

Johnston, L. E., A. Ricciardi, and J. T. Carlton, 2001. Overland dispersal of aquatic invasive species: a risk assessment of transient recreational boating. *Ecological*

Applications **11**:1789–1799.

Leung, B., J. M. Bossenbroek, and D. M. Lodge, 2006. Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biological Invasions* **8**:241–254.

Leung, B., J. M. Drake, and D. M. Lodge, 2004. Predicting invasions: Propagule presure and the gravity of allee effects. *Ecology* **85**:1651–1660.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti, 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings. Biological sciences / The Royal Society* **269**:2407–13.

Leung, B., N. Roura-Pascual, S. Bacher, J. Heikkilä, L. Brotons, M. a. Burgman, K. Dehnen-Schmutz, F. Essl, P. E. Hulme, D. M. Richardson, D. Sol, M. Vilà, and M. Rejmanek, 2012. TEASIng apart alien species risk assessments: a framework for best practices. *Ecology letters* **15**:1475–93.

MacIsaac, H. J., J. V. M. Borbely, J. R. Muirhead, and P. a. Graniero, 2004. Backcasting and Forecasting Biological Invasions of Inland Lakes. *Ecological Applications* **14**:773–783.

MacPherson, A. J., R. Moore, and B. Provencher, 2006. A dynamic principal-agent model of human-mediated aquatic species invasions. *Agricultural and Resource Economics Review* **35**:144–154.

Muirhead, J. R. and H. J. MacIsaac, 2011. Evaluation of stochastic gravity model selection for use in estimating non-indigenous species dispersal and establishment. *Biological Invasions* **13**:2445–2458.

Padilla, D. K. and S. L. Williams, 2004. Beyond Ballast Water: Aquarium and Ornamental Trades as Sources of Invasive Species in Aquatic Ecosystems. *Frontiers in Ecology and the Environment* **2**:131–138.

Parry, H., R. Sadler, and D. Kriticos, 2013. Practical guidelines for modelling post-entry spread in invasion ecology. *NeoBiota* **18**:41–66.

Phaneuf, D. J., 2013. Heterogeneity in Environmental Demand. *Annual Review of Resource Economics* **5**:227–244.

Prasad, A. M., L. R. Iverson, M. P. Peters, J. M. Bossenbroek, S. N. Matthews, T. Davis Sydnor, and M. W. Schwartz, 2010. Modeling the invasive emerald ash borer risk of spread using a spatially explicit cellular model. *Landscape ecology* **25**:353–369.

Rothlisberger, J. D., W. L. Chadderton, J. McNulty, and D. M. Lodge, 2010. Aquatic Invasive Species Transport via Trailered Boats: What is Being Moved, Who is Moving it, and What Can Be Done. *Fisheries* **35**:121–132.

Shea, K., E. Jongejans, O. Skarpaas, D. Kelly, and A. W. Sheppard, 2010. Optimal management strategies to control local population growth or population spread may not be the same. *Ecological applications : a publication of the Ecological Society of America* **20**:1148–61.

Simberloff, D., I. M. Parker, and P. N. Windle, 2005. Introduced Species Policy, Management, and Future Research Needs. *Frontiers in Ecology and the Environment* **3**:12–20.

Smith, A. L., D. R. Bazely, and N. Yan, 2013. Are legislative frameworks in Canada and Ontario up to the task of addressing invasive alien species? *Biological Invasions* .

Stohlgren, T. J. and J. L. Schnase, 2006. Risk analysis for biological hazards: what we need to know about invasive species. *Risk analysis : an official publication of the Society for Risk Analysis* **26**:163–73.

Timar, L. and D. J. Phaneuf, 2009. Modeling the human-induced spread of an aquatic invasive: The case of the zebra mussel. *Ecological Economics* **68**:3060–3071.

Vander Zanden, M. J. and J. D. Olden, 2008. A management framework for preventing the secondary spread of aquatic invasive species. *of Fisheries and Aquatic Science* **65**:1512–1522.

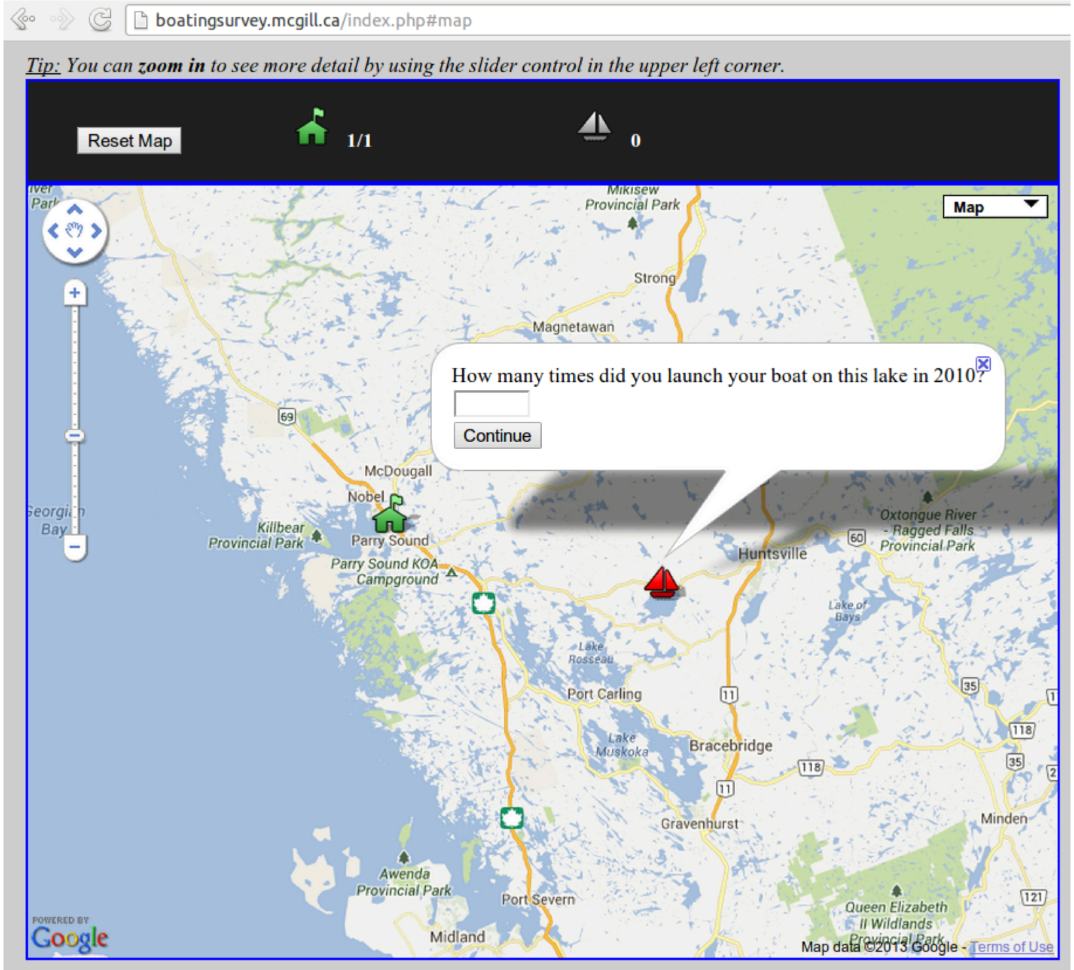
Wonham, M., J. Byers, E. D. Grosholz, and B. Leung, 2013. Modeling the relationship between propagule pressure and invasion risk to inform policy and management. *Ecological Applications* .

Common Name	Scientific Name	Year first detected
Banded Mysterysnail	*Viviparus georgianus*	1998
Chinese Mysterysnail	*Cipangopaludina chinensis*	1996
Eurasian Water-milfoil	*Myriophyllum spicatum L.*	1992
European Common Reed	*Phragmites australis*	1948
Flowering Rush	*Butomus umbellatus L.*	1949
Purple Loosestrife	*Lythrum salicaria L.*	1991
Quagga Mussel	*Dreissena bugensis*	1992
Rusty Crayfish	*Orconectes rusticus*	1964
Spiny Waterflea	*Bythotrephes longimanus*	1982
Zebra Mussel	*Dreissena polymorpha*	1988

Table 5–1: **10 Aquatic invasive species in Ontario.** Data from (EDDMapS, 2013).

Model	k	ΔBIC
$\theta_m = exp(-\alpha_\theta(cost + \gamma_\theta))$ $\phi_m = exp(-\alpha_\phi(cost + \gamma_\phi))$	4	0
$\theta_m = m_\theta(cost) + b_\theta$ $\phi_m = m_\phi(cost) + b_\phi$	4	39.1
$\theta_m = c_\theta$ $\phi_m = c_\phi$	2	411.3

Table 5–2: **Model comparison of relationship between cost and θ_m and ϕ_m.** Three alternatives models of the relationship between the direct cost of a cleaning policy at lake m and the proportion of lost (ϕ_m), and redistributed (θ_m) trips to a lake m with a cleaning policy. k is the number of estimated parameters for each model.



boatingsurvey.mcgill.ca/index.php#map

*Tip: You can **zoom in** to see more detail by using the slider control in the upper left corner.*

How many times did you launch your boat on this lake in 2010?

Figure 5–1: **Screenshot of the interactive online survey tool.** Respondents identified each of the locations at which they launched their boat, as well as how often. Next they were asked how often they would have boated at each location under a hypothetical management scenario.

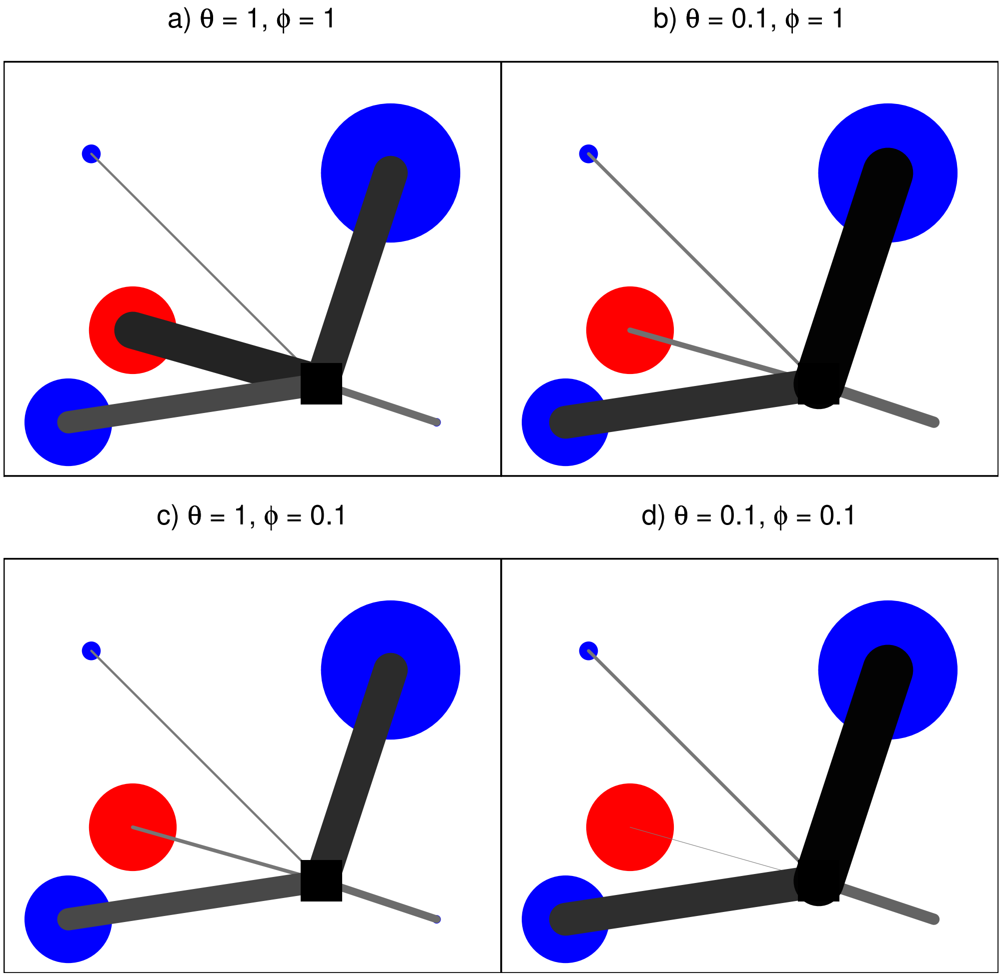


Figure 5–2: **Conceptual figure demonstrating the two aspects of behavioural responses to management.** Circles indicate lakes. The policy lake is shown in red. Squares are the home location of a hypothetical boater and the width of the line segments is proportional to the number of trips taken to each lake. (a) shows the null case where there is no effect of management on boaters' trip-taking distribution θ_m and ϕ_m are both equal to 1. (b) an entirely redistributive response where the boater takes the same number of trips that they otherwise would have, but distributes them to the non-policy lakes ($\theta_m < 1$). (c) shows an entirely reductive (loss) response where the trips not taken to the policy lake are not redistributed ($\phi_m < 1$). Finally (c) shows a combination of reduced and redistributive trip taking ($\theta_m < 1, \phi_m < 1$).

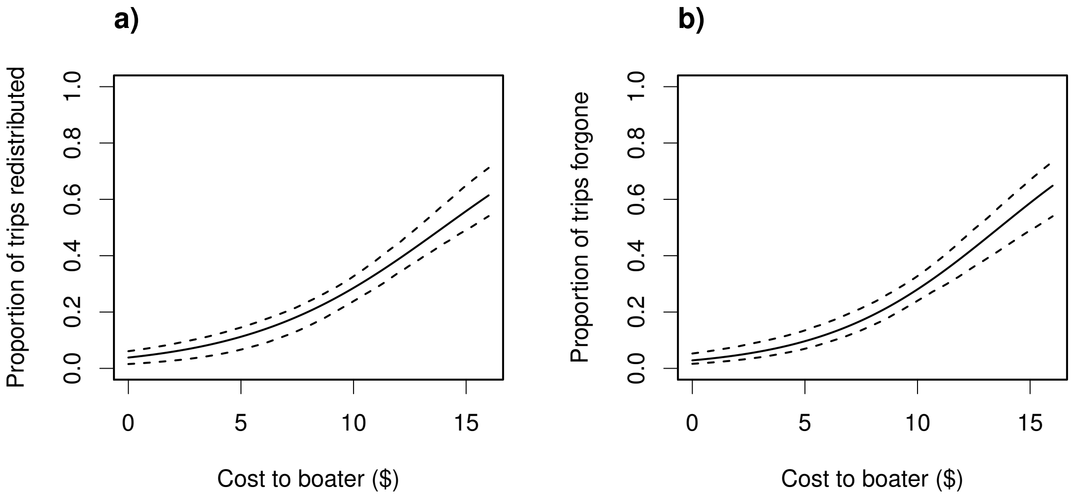


Figure 5–3: **Proportion of boater trips diverted as a function of the cost incurred by boaters.** (a) The proportion of the trips that would have been taken to a policy lake under the *status quo* which were diverted to alternative lakes $(1-\theta_m)$. (b) The proportion of trips that would have been taken to a policy lake under the *status quo* which were were forgone due to the policy $(1-\phi_m)$. Solid lines are posterior mean relationships and dashed lines are 95% credible intervals.

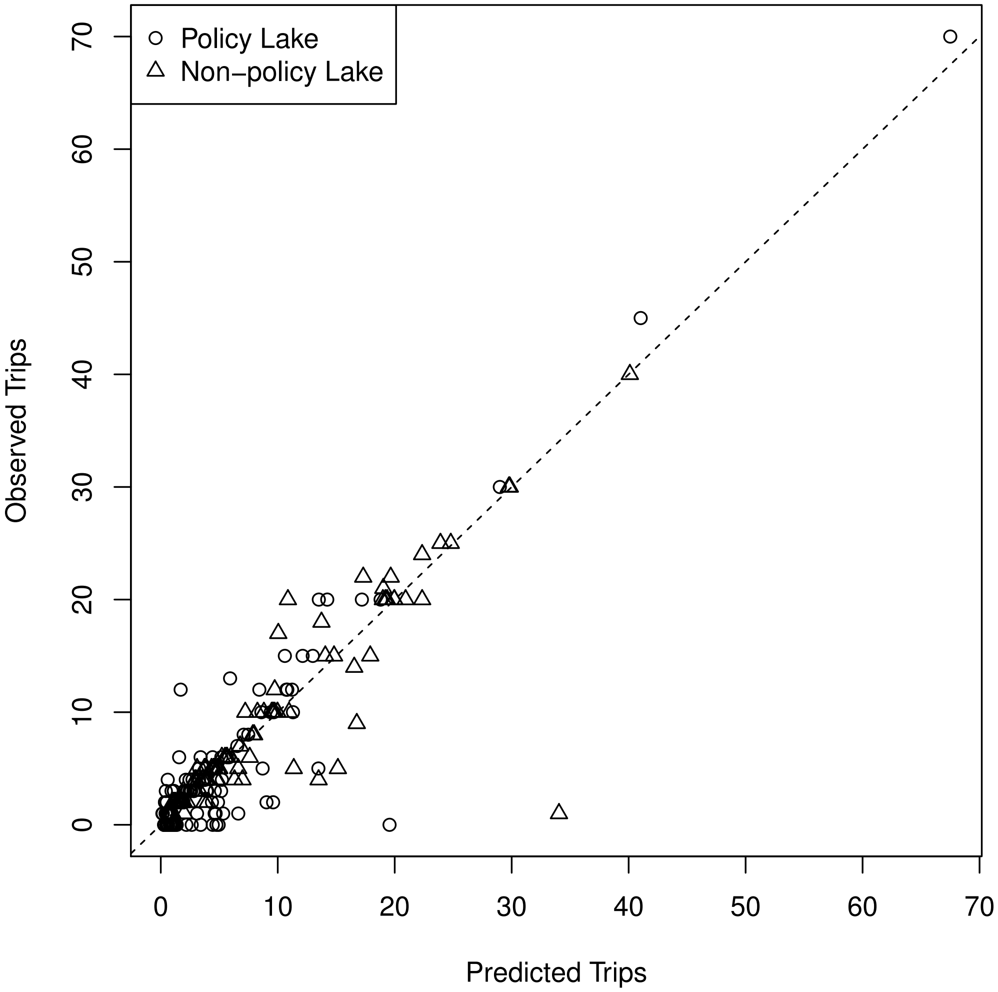


Figure 5–4: **Comparison of model predictions vs. observed number of trips taken under policy scenarios.** The number of trips that a boater reported they would take to both a lake with a mandatory cleaning and inspection station (open circles) and without (open triangles) under the counterfactual management scenario. The model predictions capture 87% of the variance in the reported behaviours. The dashed line represents the 1:1 equivalence line.

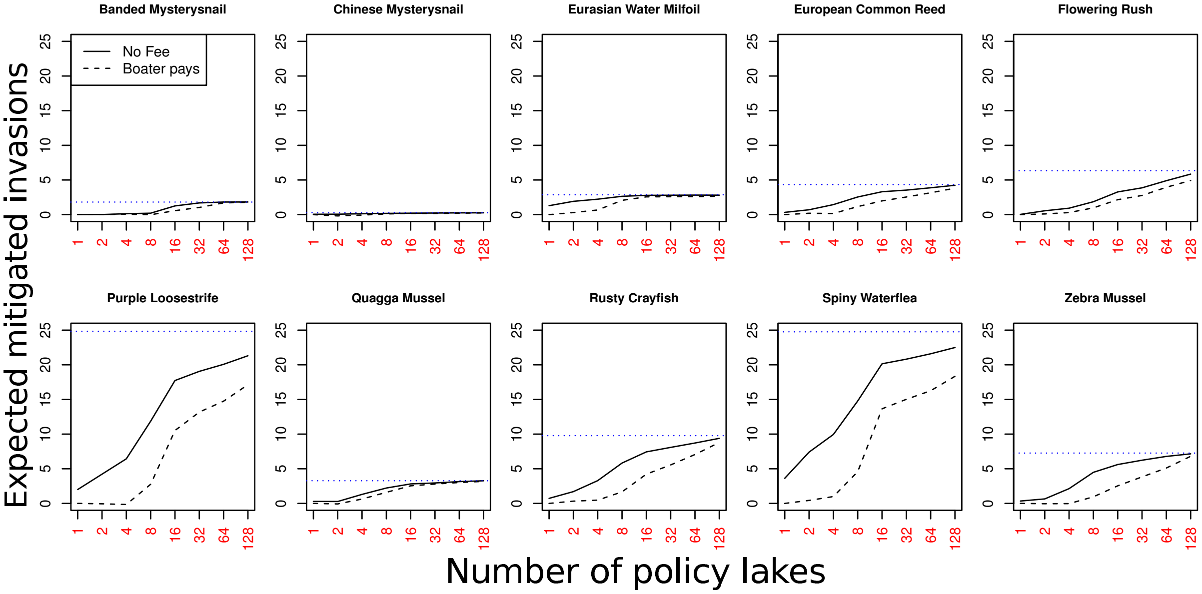


Figure 5–5: **Expected number of mitigated invasions over a 10 year time horizon (2013-2023).** Projected effect of implementing mandatory boat cleaning stations at from 1 to 128 locations. When the policy imposes a direct monetary cost to boaters (dashed lines) fewer invasions are mitigated due behavioural feedbacks. The blue dotted lines indicate the maximum expected number of invasions over the time horizon given the *status quo* condition. This indicates full quarantine, as this is the level at which no more invasions could have been mitigated.

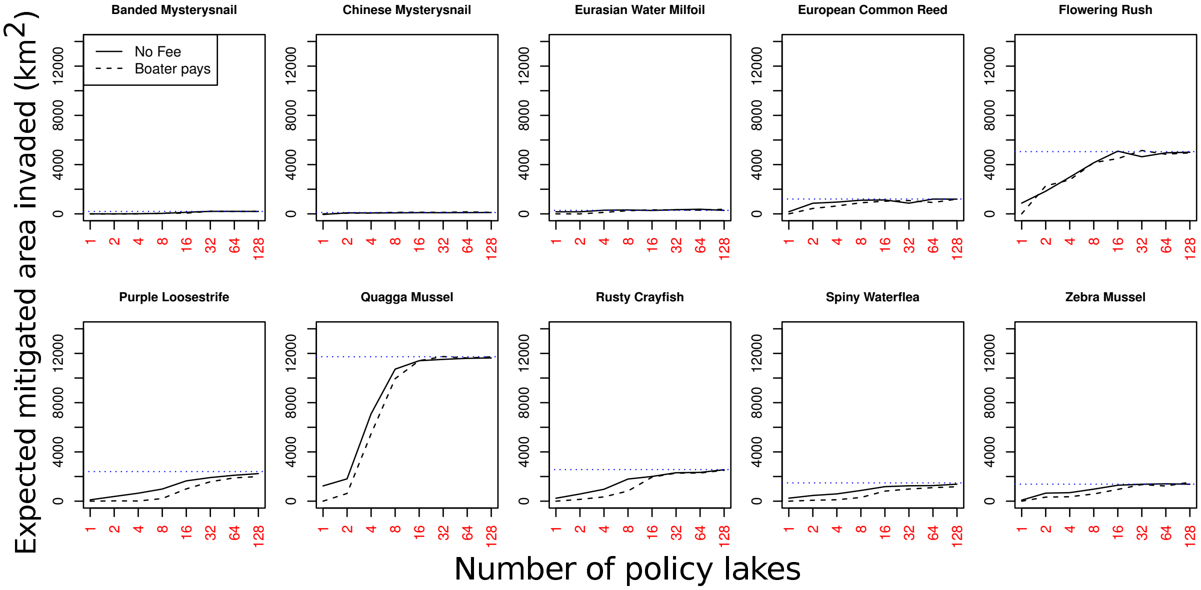


Figure 5–6: **Expected total lake surface area kept clear of invasion by each species over a 10 year time horizon (2013-2023).** Projected effect of implementing mandatory boat cleaning stations at from 1 to 128 locations. Redistributive effects of monetary cost to boaters (dashed lines) are much less pronounced in terms of area. The blue dotted lines indicate the expected maximum lake surface area projected to become invaded over the time horizon given the *status quo* condition. This indicates full quarantine, as this is the level at which no more invasions could have been mitigated.

CHAPTER 6
General Conclusion

6.1 Conclusion

The relatively recent emergence of invasive species as a major driver of global change (Ricciardi, 2007) has meant that anticipating where, when, and how species will spread in the future has become an emerging imperative for ecologists and policy-makers alike (Clark *et al.*, 2001). With their potentially large economic (Pimentel *et al.*, 2005) and ecological impacts (Dextrase & Mandrak, 2006), the ability to forecast and subsequently control invasive species could yield enormous benefits.

A common theme in this dissertation has been that the explicit acknowledgement and accounting of prediction uncertainty is vital to how we understand, predict, and manage invasive species. While it is never possible to account for every aspect of uncertainty pertaining to our models and predictions of ecological systems (or indeed any system), I have made attempts here to address those which are most likely to have an impact on the decisions that would be guided by them. There will always be trade-offs involved when attempting to quantify uncertainty in any prediction endeavour, and the secondary spread of invasive species is no different. For instance, addressing model uncertainty by comparing alternative model structures and their consequences for prediction as I did in Chapter 2 can become increasingly difficult as the computational complexity of the model increases. As models are necessarily abstractions of the real world, we can never expect to have the 'correct' model of a complex ecological system (Runge & Johnson, 2002). Even still, when alternative functional forms have drastic consequences for forecasting and managing populations, a 'better' model can go a long way toward improving predictions (Chivers & Leung, 2012). Similarly, the Bayesian models in chapters 4 and 5 incorporate and propagate parameter and state uncertainty by repeatedly simulating the underlying dynamics and by employing computationally costly Markov Chain Monte Carlo methods. These approaches are becoming more and more feasible as the cost of computational power falls, but there are still limits on the size and complexity of the models which are possible.

With any model (or ensemble of models), in order to provide credible decision support, their predictions need to be evaluated with their associated uncertainties accounted for. To that end, the new approach developed in Chapter 3 provides the inferential mechanisms for determining the degree to which predicted probabilities match up with actual outcomes, while incorporating both stochastic and epistemic prediction uncertainties. The application of this approach to a published model of the invasive zooplankton *Bythoterphes longimanus* in Ontario demonstrated how probabilistic species spread forecasts can be assessed as new data are collected.

Possibly the largest factor contributing to uncertainty is the limited data available with which to build forecasting models. While many current models of spread require data describing both where species are present and where they are absent (Ibáñez *et al.*, 2009; Gertzen & Leung, 2011; Catterall *et al.*, 2012), the much more common situation is that we only have records of species presences, and these records exist only for a limited subset of locations. Rarely the result of systematic sampling, presence-only records are often the result of either opportunistic sampling or citizen-reported sightings. This form of data, while ubiquitous, make efforts aiming to assess the risks posed to sites across space and time difficult. Without data on species absences, currently available techniques do not purport to estimate species prevalence or site-specific probabilities of presence. By viewing the data generating process as a combination of dispersal, establishment, and detection, the model developed in Chapter 4 provides a framework for projecting spread while incorporating the uncertainties inherent to presence-only data.

Despite uncertainties, decisions need to be made pertaining to what (if any) actions are to be taken to mitigate the impacts of invasions. Building off of the gravity model evaluated in Chapter 2 and the presence-only prediction framework developed in Chapter 4, in Chapter 5 I put these components together to evaluate the efficacy of place-based policies to slow the spread of currently established aquatic invasives. By focusing on the economic behaviours of the human vectors of aquatic invasive species, this model suggested that behavioural feedbacks can diminish the overall benefits of policy measures, but that it is dependent on how the impact of new invasions are measured.

Together, this thesis has developed and tested new techniques and modelling approaches for incorporating and quantifying uncertainty when predicting and managing the spread of invasive species.

References

Catterall, S., Cook, A.R., Marion, G., Butler, A. & Hulme, P.E. (2012). Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography*, 35, 001–011.

Chivers, C. & Leung, B. (2012). Predicting invasions: alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects. *Journal of Applied Ecology*, 49, 1113–1123.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.a., Lodge, D.M., Pascual, M., Pielke, R., Pizer, W., Pringle, C., Reid, W.V., Rose, K.a., Sala, O., Schlesinger, W.H., Wall, D.H. & Wear, D. (2001). Ecological forecasts: an emerging imperative. *Science*, 293, 657–60.

Dextrase, A.J. & Mandrak, N.E. (2006). Impacts of Alien Invasive Species on Freshwater Fauna at Risk in Canada. *Biological Invasions*, 8, 13–24.

Gertzen, E.L. & Leung, B. (2011). Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions*, 13, 2433–2444.

Ibáñez, I., Silander, J.a., Wilson, A.M., LaFleur, N., Tanaka, N. & Tsuyama, I. (2009). Multivariate forecasts of potential distributions of invasive plant species. *Ecological applications : a publication of the Ecological Society of America*, 19, 359–75.

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.

Ricciardi, A. (2007). Are modern biological invasions an unprecedented form of global change? *Conservation Biology*, 21, 329–36.

Runge, M. & Johnson, F. (2002). The importance of functional form in optimal control solutions of problems in population dynamics. *Ecology*, 83, 1357–1371.

Appendices

APPENDIX A
Appendix

A.1 Supplemental Material Chapter 2

A.1.1 Theoretical validation of model selection procedure

To assess the ability of our procedure to discern between the two human behavioural models, we simulated trip outcome data of the same form as that which we collected using our survey. That is, we created a simulated environment, in which individual boaters behave according to one or the other behavioural model. We then fit each model and conducted model selection using the approach outlined in the main paper.

To ensure that our simulations are applicable to the collected data set, we used the same number of boaters, making trip outcomes in a comparable choice space ($n = 145$ boaters choosing among 781 lakes). Each boater was simulated to take j trips according to:

$$j \sim Poission(\lambda) \tag{A.1}$$

Where the rate parameter is fitted from the data ($\lambda = 16$). Further, we incorporated spatial auto-correlation into the environment by modelling the lake size distribution as a function of its spatial location. Specifically, lake size was distributed as:

$$S_l \sim N(\mu = lat_l lon_l, \sigma^2 = lat_l lon_l) \tag{A.2}$$

The general effect of which is to cluster larger lakes together toward one corner of the simulated environment, and smaller lakes together toward the other (see Fig. A–1).

We then simulated trips taken by each boater according to either the gravity model, or the random utility model. We used uniformly distributed random parameters in the same range as our fitted parameters to the data, and retained the trip outcomes to use for model fitting and selection. We repeated this procedure, re-simulating using different parameter values 500 times for each generating model. For all simulations, our model selection procedure was able to correctly identify the generating model. We calculated the ΔAIC (AIC of non-generating model minus the AIC of generating model) for each instance of our simulation. Negative values would indicate the erroneous selection of non-generating model. For both generating models, this value was always positive, with minimums of 100.95 and 201.80 for the cases of a generating gravity model and generating random utility model respectively. These results suggest ample model discriminatory power within a spatially structured environment given our sample size.

We also wished to validate the theoretic properties of our model fitting procedure to estimate model parameters. For this we simply plot the fitted parameter values against the true parameter values. From this, we can visually inspect for any bias in the predictor. If there is no bias, the predicted/observed point cloud should fall symmetrically around the 1:1 line. Fig. A–2 and A–3 shows the performance of the parameter estimators for the gravity and random utility models, respectively. Our results show no signs of bias across each parameter or each model.

In this validation section, we have shown via simulation that given one of either the random utility, or gravity as the generating model of human behaviour, both our model fitting (parameter estimation) and model selection procedures work as expected. We have shown this to be the case even when spatial auto-correlation is present in the system.

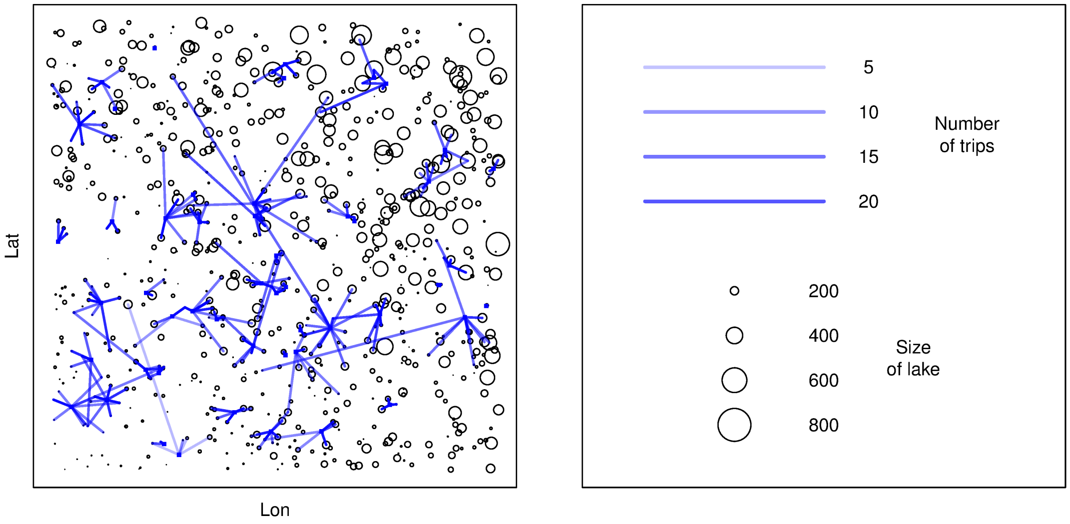


Figure A–1: **Simulated trip outcomes in a landscape of lakes with induced spatial auto-correlation.** Size of circle is proportional to the size of the simulated lake.

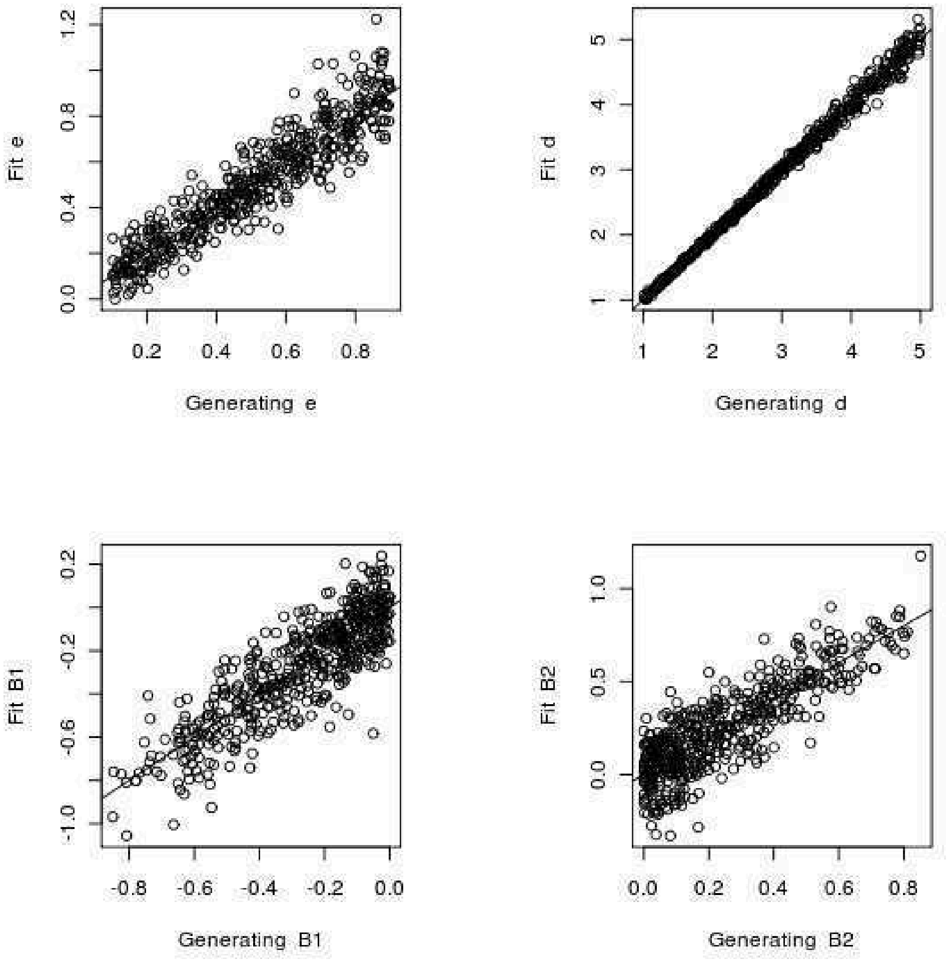


Figure A–2: **Parameter estimation performance for RUM and GM.** Generating vs maximum likelihood estimates for the four parameters (panels) of the gravity model. The 1:1 line is also plotted for comparison.

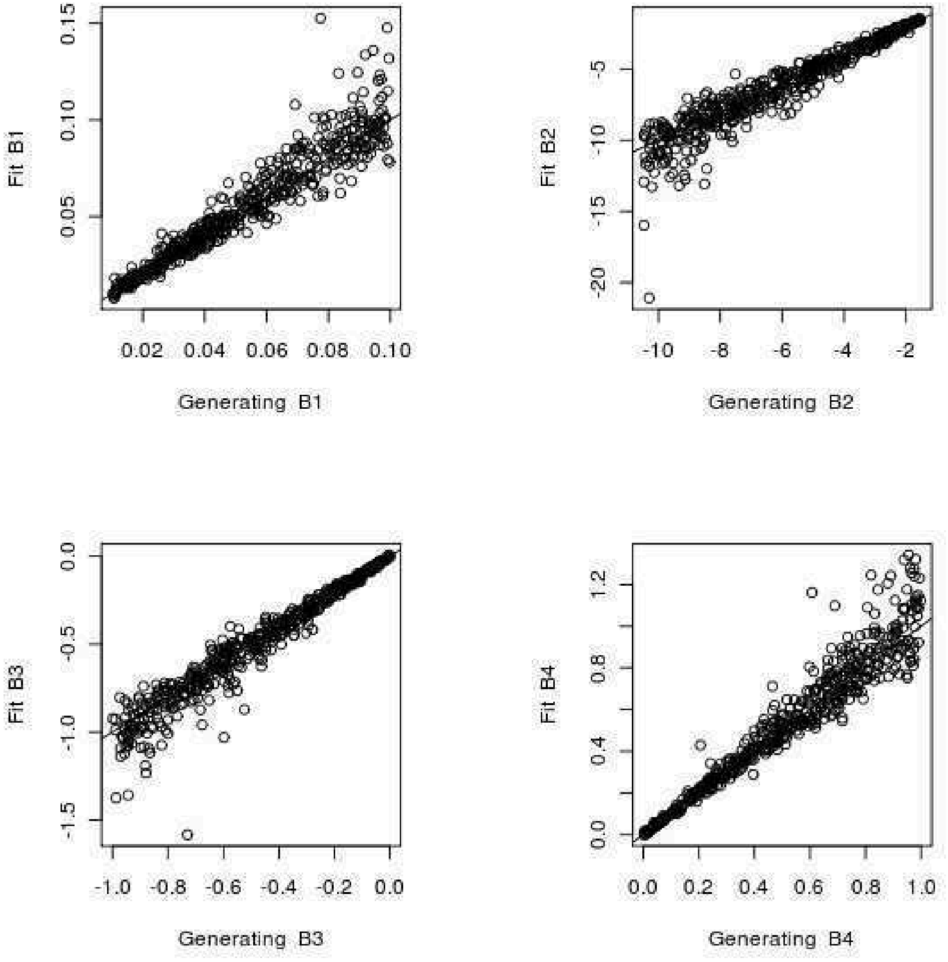


Figure A–3: **Parameter estimation performance for Random Utility Model.** Generating vs maximum likelihood estimates for the four parameters (panels) of the Random Utility Model. The 1:1 line is also plotted for comparison.

APPENDIX B
Appendix

B.1 Supplementary Material for Chapter 3

B.1.1 VMAPP Methods

Derivation of the δ estimator

For the binary variables S and R, let $P(S = 1) = \hat{p}$ and $P(R = 1) = p$, and by extension $P(S = 0) = 1 - \hat{p}$ and $P(R = 0) = 1 - p$. S and R are the simulated outcomes from model predictions and the real outcomes from a validation data set, respectively. From this, the probability that $S \neq R$ can be written as

$$P(S \neq R) = p(1 - \hat{p}) + \hat{p}(1 - p) \tag{B.1}$$

and the probability that $S > R$ given that $S \neq R$ can be written as

$$P(S > R \,|\, S \neq R) = \frac{\hat{p}(1 - p)}{p(1 - \hat{p}) + \hat{p}(1 - p)}. \tag{B.2}$$

Which is the fraction of discrepancies ($S \neq R$) in which $S = 1, R = 0$. Let us call $P(S > R \,|\, S \neq R) = A$ and $P(S \neq R) = B$, and show that

$$\delta = \hat{p} - p = 2(A - 1/2)B. \tag{B.3}$$

By substituting (B.1) and (B.2) into (B.3):

$$\hat{p} - p = 2\left(\frac{\hat{p}(1 - p)}{p(1 - \hat{p}) + \hat{p}(1 - p)} - 1/2\right)\left(p(1 - \hat{p}) + \hat{p}(1 - p)\right). \tag{B.4}$$

Which can be solved to either the identity $\hat{p} = \hat{p}$,

$$\begin{aligned}\hat{p} &= 2\left(\frac{\hat{p}(1-p)}{p(1-\hat{p})+\hat{p}(1-p)} - 1/2\right)(p(1-\hat{p})+\hat{p}(1-p)) + p \\ &= 2\hat{p} - 2p\hat{p} - p + p\hat{p} - \hat{p} + p\hat{p} + p \\ &= \hat{p},\end{aligned}$$

or $p = p$,

$$\begin{aligned}p &= -2\left(\frac{\hat{p}(1-p)}{p(1-\hat{p})+\hat{p}(1-p)} - 1/2\right)(p(1-\hat{p})+\hat{p}(1-p)) + \hat{p} \\ &= -2\hat{p} + 2p\hat{p} + p - p\hat{p} + \hat{p} - p\hat{p} + p \\ &= p.\end{aligned}$$

Hence if we can estimate the quantities (B.1) and (B.2), we can estimate δ using (B.3). Further, we can estimate (B.1) and (B.2) each as a function of some aspect of the predictive model itself.

Functional forms of f_1 and f_2

The exact functional forms of both $P(\boldsymbol{S} \neq \boldsymbol{R})$, estimated by f_1, and $P(\boldsymbol{S} > \boldsymbol{R} \,|\, \boldsymbol{S} \neq \boldsymbol{R})$, estimated by f_2 will not typically be known in advance for a given model. However, we can define these functions such that they capture the essential features that are likely to be encountered. The forms used and their rational are outlined in turn here.

The f_1 functional form

This function estimates the relationship between $\hat{\boldsymbol{p}}$ and $P(\boldsymbol{R} \neq \boldsymbol{S})$. In our analysis, we estimate $\boldsymbol{\delta}$ as a function of $\hat{\boldsymbol{p}}$. We therefore further require the function $f_1(\hat{\boldsymbol{p}})$ to exist over the domain $\in [0,1]$ (or at least $\in [\hat{p}_{min}, \hat{p}_{max}]$), and since it is a probability, to have range $\in [0,1]$. Under the null (where $\hat{\boldsymbol{p}} = \boldsymbol{p}$), we can see that $P(\boldsymbol{R} \neq \boldsymbol{S}) = 2\hat{\boldsymbol{p}}(1-\hat{\boldsymbol{p}})$. With the additional observation that $P(\boldsymbol{R} \neq \boldsymbol{S})$ must tend toward zero as $\hat{\boldsymbol{p}} \to 0$ or $\hat{\boldsymbol{p}} \to 1$, we seek a hump-shaped function over the 0-1 range.

$$f_1(\hat{\boldsymbol{p}}) = (1 - e^{-a_1 \hat{\boldsymbol{p}}^{b_1}})(1 - e^{-a_2 (1-\hat{\boldsymbol{p}})^{b_2}}) - \gamma \tag{B.5}$$

This formulation (B.5) meets our criteria and is flexible enough to capture various deviations from the null case as illustrated in Fig. SB–1.

The f_2 functional form

This function estimates the relationship between $\hat{\boldsymbol{p}}$ and $P(\boldsymbol{S} > \boldsymbol{R} \,|\, \boldsymbol{R} \neq \boldsymbol{S})$. Since like f_1, f_2 is estimating a probability, we wish to construct a function capable of expressing the range $\in [0,1]$. Also like $f_1(\hat{\boldsymbol{p}})$, the range and domain of $f_2(\hat{\boldsymbol{p}})$ must be $\in [0,1]$. We aim to capture the misspecification pathologies described in the main text (specifically general over or under bias, directional changes in bias across the range of $\hat{\boldsymbol{p}}$, as well as the null case where $P(\boldsymbol{S} > \boldsymbol{R} \,|\, \boldsymbol{S} \neq \boldsymbol{R}) = 0.5$, which occurs when $\boldsymbol{p} = \hat{\boldsymbol{p}}$). We propose:

$$f_2(\hat{\boldsymbol{p}}) = (e^{-a_1 \hat{\boldsymbol{p}}^{b_1}} - c_1)(e^{-a_2 (1-\hat{\boldsymbol{p}})^{b_2}} - c_2) \tag{B.6}$$

With free parameters $\theta_{f2} = \{a_1, a_2, b_1, b_2, c_1, c_2\}$, which is flexible enough to meet the above mentioned criteria. For instance, the parametrization $a_1 = 0, a_2 = 0, b_1 = 0, b_2 = 0, c_1 = 1 - \sqrt{0.5}, c_2 = 1 - \sqrt{0.5}$ captures the null expectation of $f_2(\hat{\boldsymbol{p}}) = 0.5$ for all $\hat{\boldsymbol{p}} = \boldsymbol{p}$. Examples of alternative parametrizations capturing the desired cases are shown in Fig. B–2.

While the functional forms of $f_1(\hat{\boldsymbol{p}})$ and $f_2(\hat{\boldsymbol{p}})$ described here are flexible enough to capture the expected behaviours of the true functions for the model miscalibration pathologies we have described, it is possible that other functional forms will be more appropriate in other situations. For this reason, we have designed the rVMAPP R package to allow user specified functional forms (details in the package documentation online at `https://github.com/cjbayesian/rvmapp`).

Algorithm for estimating f_1, f_2, and δ

1. For j in $j = 1, \ldots, J$, DO:
2. Simulate a predicted outcome for each point in the validation set ($S_{ij} \sim Bernoulli(\hat{p}_i),\ i = 1, \ldots, n$). If the model includes parameter uncertainty (bootstrapped, or Bayesian posteriors of $\hat{\boldsymbol{p}}$), simulate outcomes from the j^{th} series of predictions (randomly selected) in $\hat{p}_{ij}$.
3. Compute the difference between the simulated and observed outcomes ($S_{ij} - R_i$).
4. Map $S_{ij} \neq R_i \rightarrow \Gamma_{ij} = 1$, and $S_{ij} = R_i \rightarrow \Gamma_{ij} = 0$. (Fig. 1, row 2).
5. Fit $f_{1j}(\hat{p})$ using MLE. $\mathcal{L}(\theta_{f_1} \mid \Gamma_{1j}, \ldots, \Gamma_{nj}, \hat{p}_{1j}, \ldots, \hat{p}_{nj}) = \prod_{i=1}^{n} \begin{cases} f_1(\hat{p}_{ij}), & \text{if } \Gamma_{ij} = 1 \\ 1 - f_1(\hat{p}_{ij}), & \text{if } \Gamma_{ij} = 0 \end{cases}$

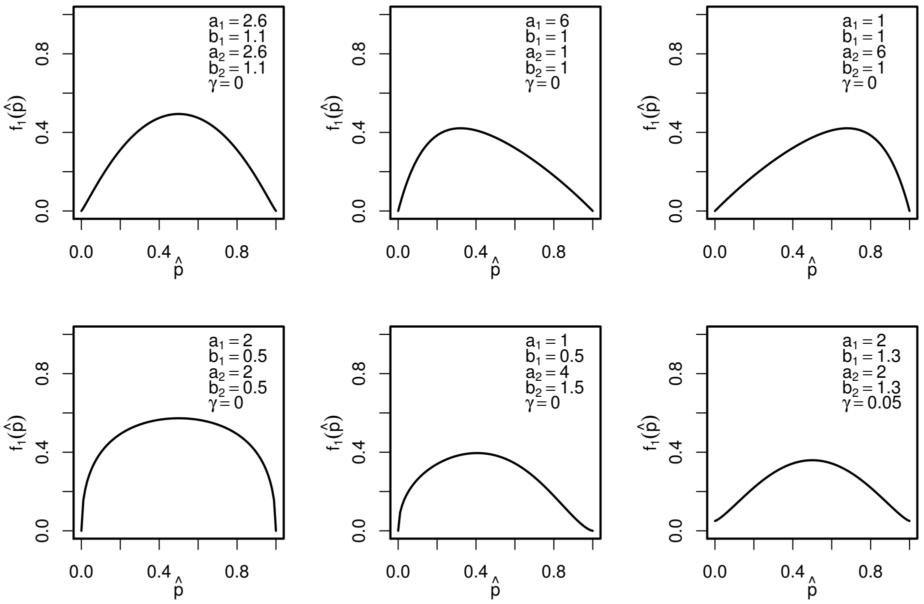


Figure B–1: **Example parametrizations of $f_1(\hat{\boldsymbol{p}})$ capturing each of the miscalibration pathologies.**

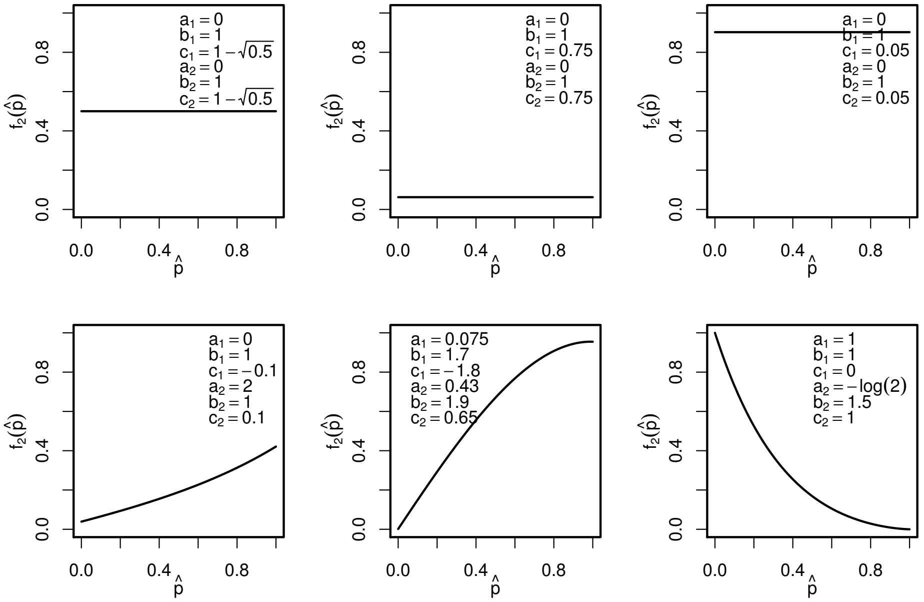


Figure B–2: **Example parametrizations of $f_2(\hat{\boldsymbol{p}})$ capturing each of the miscalibration pathologies.**

6. Map $S_{ij} > R_i \rightarrow \Delta_{ij} = 1$, and $S_{ij} < R_i \rightarrow \Delta_{ij} = 0$. Exclude any cases where $S_{ij} = R_i$ (Fig. 1, row 3).
7. Fit $f_{2j}(\hat{p})$ using MLE. $\mathcal{L}(\theta_{f_2} \mid \Delta_{1j}, \ldots, \Delta_{nj}, \hat{p}_{1j}, \ldots, \hat{p}_{nj}) = \prod_{i=1}^{n} \begin{cases} f_2(\hat{p}_{ij}), & \text{if } \Delta_{ij} = 1 \\ 1 - f_2(\hat{p}_{ij}), & \text{if } \Delta_{ij} = 0 \end{cases}$
8. Compute $\hat{\boldsymbol{\delta}}_j(\hat{\boldsymbol{p}})$ as $2(f_{2j}(\hat{\boldsymbol{p}}_j) - 0.5)f_{1j}(\hat{\boldsymbol{p}}_{\boldsymbol{j}})$
9. END FOR LOOP.

While here we estimate f_1, f_2 as functions of $\hat{\boldsymbol{p}}$, it is possible to generalize this algorithm to estimate $f_1, f_2, \boldsymbol{\delta}$ as functions of any quantity of interest (eg. model covariates, space, or some additional variable which was not originally included in the predictive model).

An example using the R package rvmapp

The R package for computing both the miscalibration goodness-of-fit statistics as well as estimating δ is available at `https://github.com/cjbayesian/rvmapp`. Installation instructions are provided in the README.md file.

To run a validation analyses using ‘rvmapp‘, two data objects are required. First, a vector containing the observed outcomes in the validation data. This can be either numeric, containing 1’s and 0’s or a logical vector. Second, either a vector of predicted probabilities generated by a model, or a matrix or data frame of probabilities. In the latter case, where uncertainty in the predictions is included, each row of the matrix represents a random draw from the predictive distribution.

The following example will walk through a simple simulated example to demonstrate how to use rvmapp and interpret the output. We’ll start by using a model which generates a single prediction vector, then show how to incorporate prediction uncertainty into the validation procedure using VMAPP.

Start by simulating some example data. Here we'll simulate data from a standard logistic model.

```
library(rvmapp)
set.seed(123)

## Simulate data on which to build a model ##
x <- runif(100,-2,2)        ## independent vars
p <- 1/(1+exp(-x))          ## simple logistic probabilities
outcomes<-bs(p)             ## Binary outcomes
```

Next, we'll fit a logistic regression (note that we have the 'correct' model).

```
model.fit <- glm(outcomes ~ x, family=binomial(logit))
```

Then we'll simulate an independent set of validation data from the same process which generated the original data.

```
## Simulate some validation data
x_val <- runif(100,-2,2)    ## New independent vars
p_val <- 1/(1+exp(-x_val))  ## Probabilities from the same model as before
outcomes_val <- bs(p_val)   ## Binary outcomes (validation data)

## Predict the probabilities associated with our validation data
p_hat <- predict(model.fit, newdata = data.frame(x=x_val),type="response")
```

Now that we have predictions from a fitted model and a set of validation data, we can pass these values to VMAPP.

```
## Run VMAPP ##
VMAPP <- vmapp(d=outcomes_val, pred=p_hat)
```

Resulting in an object `VMAPP` of class `vmapp` which has default print and plot methods:

```
> VMAPP
###########################################################
Validation Metric Applied to Probabilistic Predictions

100 validation data points used.
  Test for overall bias:
    Two-tailed test P-value: 0.738

  Test for direction change in bias:
    Two-tailed test P-value: 0.004  **
      ** Direction of deviation:  Slope of bias greater than 0.
###########################################################
plot(VMAPP)
```

From which we can see that we have found that our predictions are significantly miscalibrated. Specifically, there is a significant changing bias over the range of $\hat{p}$ from under to over-prediction. This is an example of a type-I error (since we know that our model is correct), which we have shown in the main paper will occur more often than theoretically expected if the parametric uncertainty of our predictions is not accounted for. Next we will use bootstrapping to account for this uncertainty and re-run the VMAPP analysis.

```
n_boot <- 1000 ## Number of bootstraps
p_hat_boot <- array(dim=c(n_boot,100)) ## An array to store predictions

for(i in 1:n_boot)
{
    ## Sample original data with replacement
    boot_index <- sample(1:100,replace=TRUE)
    x_boot <- x[boot_index]
    model.fit <- glm(outcomes[boot_index] ~ x_boot,
        family=binomial(logit))
    p_hat_boot[i,] <- predict(model.fit,
        newdata = data.frame(x_boot=x_val),
        type="response")
}

## Run VMAPP again, this time with parameter uncertainty included
VMAPP2 <- vmapp(d=outcomes_val, pred=p_hat_boot)
VMAPP2
plot(VMAPP2)
```

Note that the confidence intervals of $\hat{\delta}$ need not exclude zero anywhere in the range of $\hat{p}$ in order to find significant departure from the null goodness-of-fit hypothesis. This is because the inferential mechanism of VMAPP uses strictly the discrepancies $P(S > R \,|\, S \neq R)$ between simulated predicted outcomes and observed outcomes, whereas a quantitative estimate of δ requires inclusion of the probability of those discrepancies $P(S \neq R)$.

Now that we have conducted the VMAPP goodness-of-fit test and plotted the δ estimations, we can also use the predict method to calculate the estimated δ distribution for any given point in our validation set. Simply pass the `VMAPP2` object to the predict function along with the index of the validation data point of interest.

```
> predict(VMAPP2,1)
Using a 0.95 % CI
        mean    lowerCI    upperCI
1 -0.1191105 -0.2648331 0.05347423
```

Which gives the estimated deviation ($\hat{\delta}$) and confidence intervals for the model predictions of the first data point in the validation set. In this case, the expected difference between prediction and actual risk is -12% meaning that the predicted risk is expected to be is 12% lower than the true value (with a confidence range +17%/-15%). For more help with this, and other functions provided in the `rvmapp` package, use `?predict.vmapp, ?plot.vmapp` etc.

B.1.2 Description of HL and Cox's method

While many model validation studies in the ecological literature do not directly assess the calibration problem, there are two existing methods for doing so which

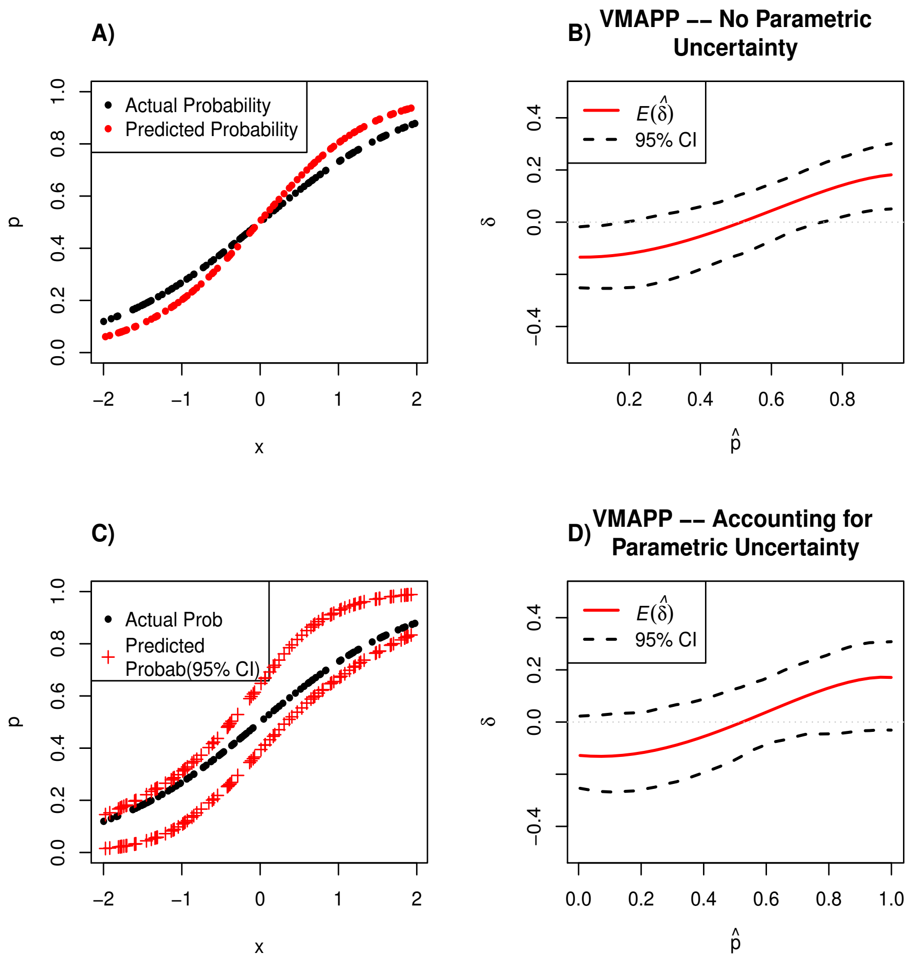


Figure B–3: **Running VMAPP with and without accounting for parametric uncertainty.** Panel A): Actual (black) and predicted (red) probabilities from a fitted logistic regression model. Panel B): δ estimation via VMAPP on 100 validation points from drawn from the actual probabilities. VMAPP detects significant non-zero δ, negative (indicating underestimation) for low predicted values and positive (overestimation) for higher predicted values. Panels C) and D) repeat the procedure with an accounting for parameter uncertainty using bootstrapping. VMAPP does not detect significant deviation from the null hypothesis of $p = \hat{p}$.

we will compare with VMAPP. Hosmer and Lemeshow (2000) describe a test which employs a strategy of binning validation observations and comparing observed and expected rates in each bin using a χ^2 goodness-of-fit test. The formulation is:

$$H = \sum_{g=1}^{n} \frac{(O_g - E_g)^2}{N_g \hat{p}_g (1 - \hat{p}_g)} \tag{B.7}$$

With O_g, E_g, N_g, are number of observed positive, expected positive, and total number of outcomes in each group g, respectively. The expected positive outcomes are calculated by summing the predicted probabilities $\hat{p}_g$ within each of n groups. The resulting statistic is then compared against a χ^2 distribution with $n - 2$ degrees of freedom.

A second approach to assessing predictive calibration was originally described by Cox (1958), and later its use was described in the context of species distribution modelling by Pearce and Ferrier (2000). In this approach, logistic regression is applied to the validation data as a function of the logit-transformed ($logit(x) = ln\left[\frac{x}{(1-x)}\right]$) predicted probabilities.

$$ln\left[\frac{P(y = 1 \,|\, x)}{P(y = 0 \,|\, x)}\right] = \beta_0 + \beta_1 ln\left[\frac{\hat{p}}{1 - \hat{p}}\right] \tag{B.8}$$

Cox (1958) shows that if $p = \hat{p} = P(y = 1 \,|\, x)$, then the expected value of the parameters β_0 and β_1 are 0 and 1, respectively. Pearce and Ferrier (2000) use likelihood ratios to test for significant departures from these expected values. Significant departures from these values represent miscalibration pathologies which match those which we have described in the main paper. Specifically, $\beta_0 \neq 0$ indicates general over, or under estimation, and $\beta_1 \neq 1$ indicates a change in the bias over the range of $\hat{p}$.

B.1.3 Bythotrephes Sampling Methods

Design of 311 lake survey in 2005/6

The overall purpose of the 2005/6 survey was to produce a large data set of presence/absence observations of *Bythotrephes* that could be used to model the risk of its spread and establishment. In 2005 and 2006 311 lakes or 19% of the lakes in the watershed were sampled. 92 of the 311 lakes were sampled in 2005, and the remainder in 2006. In 2005, survey lakes were selected from a watershed database compiled from Landsat (TM) images (Hélie et al., 1993). Lake locations and size were used as strata in an equally stratified, random sampling design (Hirzel and Guisan, 2002). Specifically, a 12-cell polygonal grid was mapped onto the watershed, with polygons approximating quaternary watershed boundaries. Within each cell, lakes were grouped into three size categories (1-10, 10.1-100 and >100 ha) and an equal number of lakes from each size category was randomly chosen for sampling, providing lakes with preset locations and a variation in size, and with a potentially wide but uncontrolled variation in chemistry (Neary, 1990), recreational development, and access to human traffic.

In 2006, 274 lakes (45 lakes were sampled in both 2005 and 2006) were selected, without replacement, using three criteria. The criteria were lake area, distance by connected water bodies from known-invaded lakes, and apparent propagule pressure - PP (Cairns et al., 2006). 101 lakes were selected for the hydrologic connectivity criterion. They were selected by hand, along chains of lakes both downstream and upstream of 16 lakes known to be invaded in 2005. 89 of the lakes were upstream of known invasions. Those lakes selected to reflect ranges in PP and area were randomly selected. To select lakes along the PP gradient we first dichotomously assigned all the lakes in the watershed into lakes accessible by road, i.e. within 100 m of an all season road, vs. those not accessible by road. 408 of the lakes were deemed accessible, and the remainder were not. An index of propagule pressure (Ipp) was calculated as the distance from the lake to the nearest known invaded lake plus the distance by road from the lake to the nearest town. Ipp was then split into 8 equal bins, and randomly chose 10 lakes from each bin giving us the first 80 lakes to be sampled in the PP category. 20 additional lakes were then selected from the inaccessible list and split into those that were within 500 m vs. further than 500m from an all season road. Finally, lakes were selected by size. There were 1083 lakes of 1-10 ha in size, 499 lakes with areas of 10.1 to 100 ha and 83 lakes larger than 100 had. 37 lakes were randomly selected from the two smaller size categories, an over-selection of 5 lakes per size category, to allow for sampling contingencies. The largest size category had only 30 lakes, as the remaining 53 had already been excluded as they were either known to be invaded as of 2005, or had been previously selected for sampling by the PP or hydrologic connectivity selection process. A full description of the lake selection process employing GIS layers is provided by Cairns et al. (2007).

Sampling Details

All lakes were sampled between mid-June and the end of August, when *Bythotrephes* is commonly abundant. Usually from a canoe, zooplankton samples were collected in each lake using a 63 um mesh conical tow net with a side length of 1.4 m, and mouth and cod end diameters of 30 cm and 5 cm, respectively. Such nets can detect the vast majority of small *Bythotrephes* populations on the Shield if the sampling protocol includes duplicate hauls at 3 offshore stations (Boudreau and Yan, 2004). To further increase detection probability, duplicate hauls were collected at 5 sites roughly equally distributed along a fetch of each lake, with the contents of the two hauls at each station being combined. In case persistent recent winds had removed *Bythotrephes* from offshore areas, a 6th horizontal haul was collected from a downwind location in each lake. All samples were condensed in the field and preserved with 5.5% buffered sugar-formalin.

The 2005/2006 311 lake survey doubled the number of known invasions in the watershed from 23 in 2004 to 46 in 2006. The largest parks in the watershed remained uninvaded in 2006. The invaded lakes were all accessible to the public and within 100m of an all-season road. They ranged in size from 23 to 12,000 ha. 40% of the lakes >100 ha in the watershed supported *Bythotrephes* populations by 2006. All the Bythoterphes data were recorded in an Access DB which was subsequently used by various CAISN investigators to model the risk of spread and establishment of the invader (Potapov et al., 2011; Gertzen and Leung, 2011; Wang and Jackson, 2011; Weyhenmeyer et al., 2010).

Quality of the presence/absence data from the 311 lake survey

A Taylor's Power Law plot of log variance vs. log abundance of *Bythotrephes* from the 5 vertical haul stations in the 23 invaded lakes was significant with a slope of 2.7 indicating that *Bythotrephes* spatial distribution was aggregated in the lakes.

Still *Bythotrephes* populations were generally large enough to be captured at multiple stations. On average, *Bythotrephes* were recovered at 2.8 of the 5 vertical hauls stations, and in only 3 cases was it detected at only one vertical haul station. The 6th, horizontal haul station was a worthwhile addition as this station produced 2 of the invasion records (Cairns et al., 2007). There was no relationship between sampling date and the number of stations at which *Bythotrephes* was detected, suggesting no seasonal detection bias.

Sampling in 2010

Cairns and Yan (2011) conducted an additional large survey of 135 lakes for *Bythotrephes* both in watershed 2EB and in a few neighboring tertiary watersheds in 2010 for two reasons. First, Gertzen and Leung (2011) concluded that the short term spread of *Bythotrephes* in the watershed was not yet over, and this prediction warranted testing. Secondly, the initial 300 lake survey had led to the production of 3 quite different approaches to modelling the invader's spread, i.e. models built on hydrological + human assisted propagule pressure (Gertzen and Leung, 2011), human propagule pressure + water quality (Potapov et al., 2011), and water quality, lake location and predation pressure (Wang and Jackson, 2011). The collaborators were keen to test the predictive abilities of these different models as the invader continued to spread in the watershed. Hence, in 2010, each modelling group was asked to provide Cairns and Yan with a list of 30-50 lakes within or neighbouring watershed 2EB, and 135 of these lakes were sampled in 2010. 25 of the sampled lakes were beyond the boundaries of watershed 2EB but still within the immediate area, i.e. 7 were in watersheds 2EA, 11 in 2EC, and 7 in 2HF (see (Cairns and Yan,

2011) for maps). Of the lakes sampled, 57 were selected by Lewis, 61 by Leung and 29 by Jackson. All field and lab methods for processing haul samples were identical to those of Cairns et al. (2007), with the additional collection of anions in the water quality samples, and phytoplankton samples.

The 2010, 135 lake survey documented 17 new invasions bringing the total number of invaded lakes in the region from 53 to 70, and bringing the number of known invaded lakes in watershed 2EB from 45 to 55. Of the 135 lakes, 67 were sampled in previous years, 3 in 2005 and 2010, 44 in 2006 and 2010 and 20 in all 3 years.

References

Boudreau, S. A. and N. D. Yan, 2004. Auditing the accuracy of a volunteer-based surveillance program for an aquatic invader Bythotrephes. *Environmental monitoring and assessment* **91**:17–26.

Cairns, A., M. Elliott, N. Yan, and E. Weisz, 2006. Operationalizing CAISN project 1.V technical report #1: Lake selection. Technical report, Ontario Ministry of the Environment, Dorset Environmental Science Centre.

Cairns, A., N. Yan, E. Weisz, J. Petruniak, and J. Hoare, 2007. The large, inland lake, Bythotrephes survey - Limnology, Database Design, and Presence of Bythotrephes in 311 south-central Ontario Lakes. Technical Report April, Ontario Ministry of the Environment, Dorset Environmental Science Centre.

Cairns, A. and N. D. Yan, 2011. New reports of Bythotrephes, summary of methodology and database design. Technical Report January, Canadian Aquatic Invading Species Network, Toronto, ON.

Cox, D., 1958. Two further applications of a model for binary regression. *Biometrika* **45**:562–565.

Gertzen, E. L. and B. Leung, 2011. Predicting the spread of invasive species in an uncertain world: accommodating multiple vectors and gaps in temporal and spatial data for Bythotrephes longimanus. *Biological Invasions* **13**:2433–2444.

Hélie, R. G., G. M. Wickware, and M. Sioh, 1993. Quantitative assessment of surface water at risk due to acidification in eastern Canada. Environment Canada, Ottawa, Ontario.

Hirzel, A. and A. Guisan, 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* **157**:331–341.

Hosmer, D. and S. Lemeshow, 2000. Assessing the Fit of the Model. In *Applied Logistic Regression*, chapter 5, pages 143–202. John Wiley & Sons, Inc., second edition.

Neary, B. P., 1990. The acidification of Ontario lakes: an assessment of their sensitivity and current status with respect to biological damage. Technical report, Ministry of Environment, Limnology Section, Dorset Research Centre.

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

Potapov, A., J. Muirhead, N. Yan, S. Lele, and M. Lewis, 2011. Models of lake invasibility by Bythotrephes longimanus, a non-indigenous zooplankton. *Biological Invasions* pages 2459–2476.

Wang, L. and D. a. Jackson, 2011. Modeling the establishment of invasive species: habitat and biotic interactions influencing the establishment of Bythotrephes longimanus. *Biological Invasions* **13**:2499–2512.

Weyhenmeyer, G., E. J. Weisz, and N. D. Yan, 2010. Relative value of limnological, geographic, and human use variables as predictors of the presence of Bythotrephes longimanus in Canadian Shield lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **67**:462–472.

APPENDIX C
Appendix

C.1 Supplemental Material for Chapter 5

C.1.1 Survey management scenario text

Upon removing your boat from the lake, you would be required to pass your boat and trailer through a hull, trailer and gear cleaning station.

At this station, you would be required to:

– Empty all bilges and live wells

– Ensure that all ropes, fishing lines, propellers, and trailer parts are clear of any plants or animals

Time required to complete the cleaning process will vary depending on the size of your vessel, but it is estimated to take approximately 15 minutes.

The cost of this mandatory procedure is XX $.

C.1.2 Theoretical validation of management GM

To ensure that we were able to recapture unbiased estimates of the behavioural response parameters θ_m and ϕ_m, we simulated boater trip taking behaviour under both *status quo* and management conditions. Retaining only that information which we get from our survey and using the same sample size and GM parameter values, we fit the behaviour model across a range of true generating values of θ_m and ϕ_m. Predicted-actual plots indicate close agreement between actual and predicted values (Figure C–1).

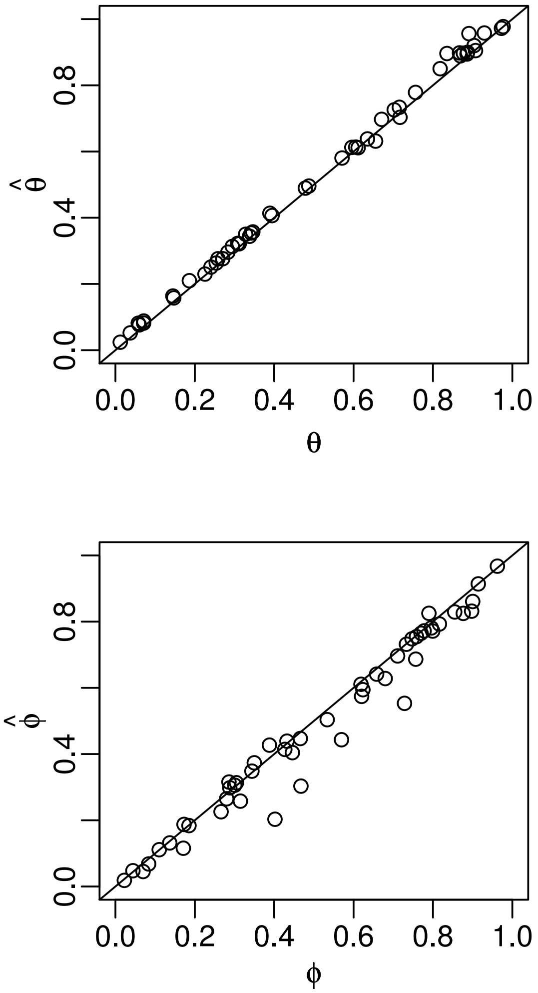


Figure C–1: **Predicted-Actual plots validating the management model.** A 1:1 relationship indicates perfect parameter estimation.