2013

Predicting the spatial distribution of an invasive plant species and modeling tolerance to herbivory using Lythrum salicaria L. as a model system

Shyam Thomas
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Ecology and Evolutionary Biology Commons

Recommended Citation
Thomas, Shyam, "Predicting the spatial distribution of an invasive plant species and modeling tolerance to herbivory using Lythrum salicaria L. as a model system" (2013). Graduate Theses and Dissertations. 13563. https://lib.dr.iastate.edu/etd/13563

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Predicting the spatial distribution of an invasive plant species and modeling tolerance to herbivory using *Lythrum salicaria* L. as a model system

by

Shyam Mathew Thomas

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Ecology and Evolutionary Biology

Program of Study Committee:
Kirk A. Moloney, Major Professor
Luke C. Skinner
Karen C. Abbott
Lisa Schulte Moore
W. Stanley Harpole
John Nason

Iowa State University
Ames, Iowa
2013

Copyright © Shyam Mathew Thomas, 2013. All rights reserved.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER 1 GENERAL INTRODUCTION &amp; THESIS ORGANIZATION</td>
<td>1</td>
</tr>
<tr>
<td>Invasive Species Distribution</td>
<td>1</td>
</tr>
<tr>
<td>Herbivory Tolerance in Perennials</td>
<td>5</td>
</tr>
<tr>
<td>References</td>
<td>6</td>
</tr>
<tr>
<td>CHAPTER 2 HIERARCHICAL FACTORS IMPACTING THE DISTRIBUTION OF AN INVASIVE SPECIES: LANDSCAPE CONTEXT &amp; PROPAGULE PRESSURE</td>
<td>9</td>
</tr>
<tr>
<td>Abstract</td>
<td>9</td>
</tr>
<tr>
<td>Introduction</td>
<td>10</td>
</tr>
<tr>
<td>Methods</td>
<td>14</td>
</tr>
<tr>
<td>Results</td>
<td>20</td>
</tr>
<tr>
<td>Discussion</td>
<td>23</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td>30</td>
</tr>
<tr>
<td>Tables</td>
<td>36</td>
</tr>
<tr>
<td>Figures</td>
<td>38</td>
</tr>
<tr>
<td>Appendix</td>
<td>42</td>
</tr>
<tr>
<td>CHAPTER 3 INVASIVE SPECIES DISTRIBUTION MODEL: COMBINING THE EFFECTS OF SURROUNDING LAND-USE &amp; PROPAGULE PRESSURE</td>
<td>44</td>
</tr>
<tr>
<td>Abstract</td>
<td>44</td>
</tr>
<tr>
<td>Introduction</td>
<td>45</td>
</tr>
<tr>
<td>Methods</td>
<td>49</td>
</tr>
<tr>
<td>Results</td>
<td>59</td>
</tr>
<tr>
<td>Discussion</td>
<td>62</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>69</td>
</tr>
<tr>
<td>References</td>
<td>69</td>
</tr>
<tr>
<td>Tables</td>
<td>76</td>
</tr>
<tr>
<td>Figures</td>
<td>78</td>
</tr>
<tr>
<td>Appendix</td>
<td>84</td>
</tr>
<tr>
<td>CHAPTER 4 EXPLORING HERBIVORY TOLERANCE IN PERENNIALS WITH STORED BELOWGROUND BIOMASS</td>
<td>88</td>
</tr>
</tbody>
</table>
Abstract .................................................................................................................. 88
Introduction ........................................................................................................... 89
Methods .................................................................................................................. 92
Results & Discussion ............................................................................................... 99
Acknowledgements ................................................................................................. 107
References ............................................................................................................... 107
Tables ..................................................................................................................... 113
Figures ................................................................................................................... 114

CHAPTER 5 GENERAL CONCLUSIONS ................................................................. 120
ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Kirk Moloney, who has constantly been supportive and encouraged me to develop my own research ideas irrespective of the outcome. I have benefitted enormously from his laid-back yet constructive approach towards guiding my research and ensuring that I get a good job done at the end. I am also extremely thankful to Dr. Karen Abbot for helping me immensely in developing research techniques in an area that was entirely new to me at the beginning of my PhD. Her encouragement and enthusiasm towards my project was both instructive and vital in ensuring that I get a good job done. In short, as someone with almost no background or research experience in ecological modeling prior to my PhD I am happy to say that I have come a long way, thanks to all the patience and encouragement from Kirk and Karen.

I am also thankful to my committee members – Lisa, Stan, John and Luke for suggesting various ways in which I could improve my research in terms of both quality and usefulness. I am particularly grateful to Luke and Adam (from MN-DNR) for helping me immensely in the field validation work and giving me a very valuable real-world perspective on the problems encountered when it comes to managing invasive plants like loosestrife.

Given the long stretch of a PhD, I am thankful to a lot of friends and colleagues with whom I have shared some of the very best times in graduate school, and who have helped me in various possible ways through the last 6 years. Friends and colleagues within the department that I would like to thank – Brad, Jim, Tim, Erika, Andres, Renae,
Lauren, Paul, Forest, LeAnn & Ryan for helping and assisting with any pressing issues that would come up and more importantly, for some wonderful shared times that brought me closer to the wonderful and diverse American culture. As an international student, I am also thankful to friends from my home country: Anand, Asha, Najeeb, Sujith, Adarsh, Kanan, Dinesh, Merwyn & Abhijit for making it a lot easier to adapt to the demands of living away from home and for some memorable moments that shall stay with me always. A very special thank you to Omana aunty & Jose uncle for the most memorable times shared around fantastic home-cooked food and lively conversations, and most importantly for creating a family like atmosphere for students like me away from home.

The long stretch of PhD can be demanding both intellectually and emotionally. And, around four years back I was fortunate to find someone who is now a very special friend and partner in almost every aspect of my life. I know that a mere thanks won’t suffice for the many ways in which you have helped me and encouraged me to keep going during the process of my PhD. But yet, I would like say a big thank you to Nu for just being there for me as a friend, critic, emotional support and much more. In fact, this dissertation work is yours as much as it is mine! And, finally I must add that a lot of whatever I have achieved to date would never have been possible if it wasn’t for the constant support and encouragement from my family. In short, if it wasn’t for my parents willingness and encouragement, especially with respect to pursuing a career of my own choice, I really doubt I would have come this far. So, thanks a lot Amma, Achachan and Elsa for everything and for all the sacrifices you have made for me and just being there for me whenever I needed you!
ABSTRACT

Understanding the ecological factors behind the landscape-level distribution of invasive species is a rapidly growing area of research with strong applied implications. In a major part of my thesis, which comprises chapters 2 & 3, the focus is on spatial pattern analyses and predictive modeling of an invasive wetland plant: Purple Loosestrife (*Lythrum salicaria* L.). More specifically, the first part of my thesis (i.e. chapter 2) considers a novel hierarchical approach, wherein the spatial distribution of loosestrife in a human-modified landscape was found to be the consequence of three key hierarchical factors: wetland habitat availability, disturbance prone surrounding land-use conditions around the wetland habitat, and propagule pressure. In chapter 3, the spatial factors and ecological processes characterized in chapter 2 were put-together and several logistic and autologistic regression models were developed to predict locations of loosestrife occurrences. Incorporating propagule pressure as an autocovariate was found to be crucial in making accurate predictions of loosestrife invasion risk. However, in the absence of propagule pressure, the surrounding land-use model highlighted the role of anthropogenic edges in defining the invasibility of wetland habitats. From an applied perspective, the model based risk maps assist conservationists and land managers in predicting and checking the spatial spread of invasive loosestrife.

In the fourth and last research chapter of my thesis, a mathematical model is developed to explore herbivore tolerance in perennials with long-term belowground storage. The inspiration behind this model is loosestrife, an invasive perennial, and its biocontrol insect herbivores. More specifically, a discrete time model was built to explore the role of belowground allocation of biomass in a perennial plant with distinct growing
season and under regular seasonal defoliation by herbivores. The model addresses the role of two co-occurring traits like utilization of stored reserves for early-season growth and post-herbivory regrowth and consequent tolerance potential. The model results highlighted that belowground biomass allocation plays a critical role as it allows the plant to persist despite severe periodic defoliation by herbivores. The model findings also indicated that when highly efficient early-season use of stored reserves is coupled with high belowground biomass allocation potential the plant biomass and herbivore population can show sustained cycles. From the perspective of invasive perennials, the model suggests that brief periods of intense seasonal herbivory is incapable of extirpating the invasive plant population as long as the latter can efficiently allocate biomass belowground.

Overall, my dissertation work focuses on ecology of invasive Purple Loosestrife and involves a series of studies that start with understanding the landscape-level distribution pattern of loosestrife, characterizing the ecological processes that result in the distribution pattern, and finally making predictions by developing predictive models that incorporate the identified ecological factors and processes. The final chapter is essentially a consumer-resource model that involves loosestrife and its biocontrol agent, and highlights the utility of developing an exploratory mathematical model in understanding the role of belowground biomass allocation as an herbivory tolerance strategy.
CHAPTER1. GENERAL INTRODUCTION & THESIS ORGANIZATION

Invasive Species Distribution: Patterns, Processes & Models

The distribution of species in space and time has always fascinated naturalists and ecologists. In this respect, one of the fundamental challenges and endeavors of ecology since its earliest days is to understand the distribution patterns of the myriad life-forms that inhabit our world. And, since the earliest published works in species distribution patterns by Grinnell (1904) the field has grown by leaps and bounds (Elith et al. 2009). Not surprisingly, the science of ecology is also often defined as the study of distribution and abundance of species (Krebs 1972; Townsend et al. 2008). In more recent times, the significance associated with understanding species distribution patterns has increased dramatically, given the various conservation threats faced by plants and animals at both global and local scales. From the applied conservation perspective, defining the spatial distribution of a given species is the key starting point to more advanced predictive model building that eventually contributes to preservation of the species. Thus, it is not surprising that, over the last couple of decades, developing predictive species distribution models has taken a center-stage in the field of conservation biology and applied ecology (Guissan and Thuiller 2005; Elith et al. 2009). At the heart of the science of species distribution analyses and modeling lies the fundamental conceptual basis that species distribution and occurrence is correlated with geographic and/or environmental gradients.

Studies of species distribution analyses and modeling have taken a wide variety of approaches. However, with the recent advances in computational technology and availability of remotely-sensed satellite imagery, the focus has shifted from local-scale
field studies of species distribution to more broad-scale spatial pattern analyses and modeling. At a more conceptual level, species distribution research has moved from studies that focused essentially at one spatial scale to a more thorough multi-level approach that combines ecological processes and factors operating at multiple spatial scales. This hierarchical, multi-level3 approach has more robust and accurate ecological underpinnings, since the large-scale distribution of any given species is the outcome of multiple factors that operate at different spatial scales (Latimer et al. 2009). Similarly, at the landscape-level, species distribution is often the outcome of nested ecological processes like availability of suitable habitat, composition in the neighborhood (i.e. landscape context), especially with respect to human affected land-use changes and species specific ecological processes, like dispersal or movement between populations and home range or territorial behavior.

Availability of remotely-sensed satellite images of landscape composition, in terms of land-use and land cover information, is one of the driving forces behind the recent surge in species distribution modeling. In this respect, standardized raster datasets like National Land Use Land Cover Database (NLCD) is a rich source of environmental information that provides a solid base for developing large-scale species distribution models. Moreover, the 15 NLCD categories of land-use/land-cover types provides sufficient detail at a resolution of 30 m, with respect to environmental conditions that represent natural, semi-natural and human modified components. Unlike climate- and bioclimate-based distribution models that define physiological limits of a species’ geographic distribution, land cover information adds substantial variation, especially in cases where species occurrence is correlated with human land use practices (Thuiller et
The application of land-use land cover information as an environmental variable in species distribution modeling is all the more pertinent when one considers the human dimension involved. In other words, land cover changes through time shows an increasingly similar pattern at a global scale, wherein the percentage of wild and natural land cover has decreased rapidly at the expense of human affected land cover changes in the form of range lands, cultivated agricultural lands and densely populated areas (Ellis and Ramankutty 2008). Taken together, it is quite evident that land cover information, as in the NLCD at 30 m resolution, captures environmental variability at the regional scales that can be particularly useful in developing regional level species distribution models.

Invasive species distribution modeling is a more recent, specialized, sub-discipline within species distribution modeling that focuses on risk analyses and developing risk maps for non-native invasive species (Václavík et al. 2009; Latimer et al. 2009). Unlike native species, the distribution and abundance of alien invasive species is often positively correlated with human activities and land-use practices (Lilly and Vellend 2009). In the case of invasive plants, it is well acknowledged that disturbances are often a pre-requisite for their successful establishment (Hobbs and Huenneke 1995). In this respect, distribution modeling of invasive plant species often needs to take into consideration the role of human-induced disturbances in the vicinity and along edges of habitat (Vilà and Ibáñez 2011). In other words, it is speculated that surrounding land-use conditions are likely to differ significantly between locations that are invaded and uninvaded by alien plants.

A recurrent problem in developing predictive distribution models for invasive species is the assumption of ‘pseudo-equilibrium’ with the environment (Guisan and
Thuller 2005). And, in the case of invasive species this is a particularly serious issue as invasive species are rarely in equilibrium with the environment in the invaded region (Peterson 2003). Modelers and ecologists working with invasive species have taken a cautious approach when dealing with the assumption of ‘pseudo-equilibrium’ by selecting for regions that have dense distributions of the modeled species. Moreover, to account for the inherent spatial dependency in the distribution of invasive species, like exotic plants, one needs to account for dispersal or propagule-pressure. A well recognized approach to account for spatial dependency involves incorporating an auto-covariate term to simulate propagule-pressure as a distance-constrained model (Miller and Franklin 2007). Incorporating propagule-pressure into predictive invasive species distribution models based on surrounding land-use will further strengthen the predictions by selecting only for those suitable locations that are within dispersal distance. Taken together, robust invasion risk maps can be developed based on a combination of surrounding land-use information and propagule-pressure. Such risk maps, once developed can assist land managers and conservation biologists in checking the spread of the invasive species.

The first two chapters of my thesis share a common goal: identifying the factors that influence the distribution of an invasive wetland plant, purple loosestrife (*Lythrum salicaria* L.) and then putting together the identified nested ecological factors into a predictive model. In the first chapter, I examine the hypothesis that the spatial distribution of invasive loosestrife is the outcome of multi-level hierarchical processes, which includes availability of wetland habitats, surrounding land-use conditions, and propagule-pressure. Using a combination of survey recorded loosestrife distributions for the state of Minnesota and land-use land cover information based on NLCD 2001
database, I analyze the distribution pattern of loosestrife with ArcGIS tools and python scripts. In the second chapter, I combine the key ecological processes that were identified in my first chapter to develop a spatially-explicit auto-logistic model within a Bayesian framework to predict locations under high invasion risk. In short, the overarching goal in this chapter is to develop a predictive model for loosestrife that takes into account spatial variability associated with surrounding land-use conditions and propagule-pressure.

**Herbivory Tolerance in Perennials: Processes and Models**

The second part of my thesis focuses on herbivory tolerance in perennial plants with long-term belowground storage. Plant defense against herbivory is an active and contentious area of empirical and theoretical research in ecology. Tolerance to herbivory essentially involves reducing the impact of herbivore induced damage on plant fitness, such that the plant can regrow and/or reproduce despite the damage (Strauss and Agrawal 1999). Resistance on the other hand implies plants are capable of reducing herbivore induced plant damage by producing chemical compounds and substances that deter or kill insect herbivores (Meijden 2000). However, much of research on plant defense to-date has focused on resistance as a defense strategy as opposed to tolerance and on the evolutionary tradeoffs between tolerance and resistance traits (Strauss and Agrawal 1999; Tiffin 2000). In the particular case of perennials, like purple loosestrife, with long-term belowground storage and a distinct growing season, studies have shown that the ability to allocate biomass to roots in response to aboveground defoliation by specialist insect herbivores allows plant to persist, despite severe periodic defoliation (Meijden et al. 2000). However, biomass allocation to belowground storage as a tolerance mechanism is
poorly understood in the context of two other simultaneously occurring factors: early-
season utilization of stored reserves for aboveground growth and post-herbivory regrowth
capacity. Previous models that have focused on the role of belowground storage as a
defense strategy against defoliating herbivores have shown that belowground allocation is
beneficial, when damage is periodic and not too frequent within the growing season
(Iwasa and Kubo 1997; Jong and Meijden 2000). In this respect, the fourth chapter of my
thesis focuses on developing an exploratory mathematical model for perennials with
long-term belowground storage that are periodically defoliated by a specialist insect
herbivore. However, unlike previous models, my model is based on perennials with a
distinct growing season, as in the case of purple loosestrife, and includes seasonal
herbivory, as showcased by the biocontrol insect herbivores of loosestrife. Therefore,
there are strong applied implications, and in this respect the study also explores the
implications of the model findings with respect to the biocontrol management of invasive
perennial plants.

References
Elith, J. & Leathwick J.R. (2009) Species distribution models: ecological explanation and
prediction across space and time. Annual Review of Ecology Evolution & Systematics, 40, 677-697
the world. Frontiers in Ecology and the Environment, 6, 11-18
Grinnell J. (1904) The origin and distribution of the chestnut-backed chickadee. Auk 21,
364–65


CHAPTER 2. HIERARCHICAL FACTORS IMPACTING THE DISTRIBUTION OF AN INVASIVE SPECIES: LANDSCAPE CONTEXT AND PROPAGULE PRESSURE

Modified from a paper published in Landscape Ecology

Shyam M Thomas\textsuperscript{1,2} and Kirk A Moloney\textsuperscript{1}

\textsuperscript{1} Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA- 50011

\textsuperscript{2}Email: smthomas@iastate.edu (corresponding author)

Abstract

Distribution of invasive species is the outcome of several processes that interact at different hierarchical levels. A hierarchical approach is taken here to analyze the landscape level distribution pattern of Purple loosestrife (\textit{Lythrum salicaria}), an aggressive wetland invader. Using land use/land cover (LULC) data and loosestrife presence records we were able to identify and characterize the key processes that resulted in the observed large-scale distribution. Herbaceous wetlands, edges of open water sites, and developed open spaces were identified as loosestrife’s preferred LULC types. Analysis of spatial neighborhoods of these key land cover types revealed that disturbance modified open water edges and herbaceous wetlands were more likely to be invaded by loosestrife. Moreover, developed open spaces appear to hold loosestrife only if there is water rich conditions in the immediate neighborhood. Neighborhood analyses also showed that wetlands and open water edges embedded within a neighborhood matrix of grassland and agricultural environments is less likely to contain loosestrife. Finally, there is strong evidence of propagule pressure. Open water edges and wetlands invaded by
loosestrife had on an average more loosestrife as neighbors than uninvaded lake edges and wetlands. Taken together, it is apparent that loosestrife’s landscape level distribution is the outcome of three nested hierarchical factors: habitat preference, the spatial neighborhood and propagule pressure. The patterns characterized suggests that occurrence of an invasive species is not merely contingent on availability of suitable habitat but is also influenced by human actions within its proximity, and is further constrained by dispersal limitation.

Keywords: Invasion, neighborhood, loosestrife, spatial pattern, surrounding land use, wetlands

**Introduction**

Understanding and predicting the distribution of invasive species is central to controlling their spread and mitigating the impact of biological invasions. Inevitably, recent studies have concentrated on predicting the potential distribution of invasive species (Rouget et al. 2004, Evangelista et al. 2008, Ibáñez et al. 2009). Unlike earlier studies that assumed invasion to be a relatively homogenous process, it is now understood that invasion is a heterogeneous process, with the spatial distribution of invasive species resulting from several interacting factors (Pearson and Dawson 2003, Rouget and Richardson 2003, Melbourne et al. 2007). To date, invasion ecology has taken a three-pronged approach to understand these interacting factors and to untangle the invasion Gordian knot. These factors include (1) species specific traits that make a species a successful invader (Kolar and Lodge 2005), (2) environmental conditions that make a region or habitat more invasible (Hobbs and Huenneke 1992, Pauchard and
Alback 2004), and (3) propagule availability, i.e., propagule pressure (Lockwood et al. 2005, Chytry et al. 2008). The complex nature of these interactions often occurs at different scales making invasion and the eventual distribution of an invasive species challenging to predict. Given this inherent complexity, identifying the key processes that determine an invasive species’ distribution in a coherent hierarchical manner is necessitated (Pearson and Dawson 2003, Guisan and Thuiller 2005).

The hierarchical approach towards analyzing species distribution patterns is logical since distribution patterns emerge from a series of nested processes and their interactions (Pearson and Dawson 2003, Milbau et al. 2009). Hierarchical theories are well entrenched in ecology, particularly in the field of landscape ecology, where complex landscape-level processes can be decomposed to simpler lower level processes (Kotliar and Wiens 1990, Wu and Loucks 1995). In other words, a hierarchical approach is an efficient way to break down complexity, and bring about order and clarity (Wu and David 2002). Therefore, identifying nested processes is the first critical step towards understanding higher level species distribution patterns. Once identified, the nested processes can be put together to produce a hierarchical model that can predict with greater accuracy species occurrence in a heterogeneous environment. Several recent studies have taken a hierarchical modeling approach to determine the probability of invasion or species occurrence given a set of nested hierarchical processes (Loarie et al. 2008, Ibáñez et al. 2009, Latimer et al. 2009). In a recent paper, Milbau et al. (2009) developed a hierarchical framework for biological invasions in order to organize the variable findings of several invasibility studies and experiments. In their hierarchical framework of biological invasions different factors affect the probability of invasion
success at different scales; climate being the dominant factor at the continental scale, while topography, land use and land cover matter most at the regional scale, and finally at small scales disturbances, soil conditions and biotic interactions play a deterministic role (Milbau et al. 2009).

In this study, we take a similar hierarchical approach to identify the processes that determine the large scale spatial distribution of invasive purple loosestrife (*Lythrum salicaria* L.), an aggressive invader of wetlands. Loosestrife is classified as a noxious weed that can potentially alter ecosystem functioning and impoverish native biodiversity of wetlands (Blossey et al. 2001). Recent studies have also shown that loosestrife is capable of plastic adaptability, which allows it to succeed in a wide range of wetland habitats (Chun et al. 2007, Moloney et al. 2010). Therefore, analyzing and understanding the regional distribution of loosestrife is a formidable challenge. Previous studies on loosestrife’s distribution pattern have failed to clearly characterize the ecological processes that are likely to be involved at different levels (Welk 2003, Anderson et al. 2006). Anderson et al. (2006) used the rule-based model, GARP, with coarse scale remotely sensed vegetation data to make broad scale predictions of regions within Kansas that are under high risk of loosestrife invasion. In an earlier study by Welk (2003), continental scale predictions were made using the GARP and DOMAIN models, and even at the large scales prediction accuracy was found to strongly depend on the quantity of data used to train the model. Unlike the previous studies that make predictive models of loosestrife distribution, the goal of the work we present here is to identify and characterize the ecological processes that result in the observed landscape level distribution pattern.
We consider a three-level hierarchical approach beginning with the identification of loosestrife’s preferred land cover types, followed by neighborhood analyses to discern if loosestrife occurrence is further influenced by surrounding land use. We define neighborhoods as circles of predetermined radiiuses around cells that have been identified as loosestrife’s preferred land-cover type. The significance of landscape context or surrounding land use is well acknowledged for mobile animals (Steffan-Dewenter et al. 2002, Umetsu et al. 2008), but relatively understudied for sessile plants (Jules and Sahani 2003, Murphy and Lovett-Doust 2004). It is however, assumed that colonization and establishment of a suitable habitat patch by a plant is often influenced by the neighborhood landscape composition. For invasive plants, neighborhoods that are more disturbed or prone to disturbances are known to be conducive to establishment (Hobbs and Huenneke 1992, Ibáñez 2009, Vilà and Ibáñez 2011). Thus, identifying characteristic neighborhood patterns that influence the establishment success of loosestrife can be used to make better predictions of the suitability of habitat patches for loosestrife. Finally, the role of propagule pressure is evaluated by determining if loosestrife invaded habitats have more loosestrife as neighbors than similar uninvaded habitats. Propagule pressure is viewed as a relatively independent chance-event factor, nevertheless a driving force in determining the heterogeneous nature of invasive species distributions (Lockwood et al. 2005). In essence, we hypothesize that loosestrife’s distribution is the outcome of three key nested factors and their interactions, viz.: availability of loosestrife’s favored land cover (habitat) types, its spatial neighborhood and propagule pressure.
Methods

Focal Species and Study Area

Minnesota has a long history of loosestrife invasion; the earliest known record in the year 1907 is from the city of Duluth (Blossey et al. 2001). Loosestrife, an aggressive perennial herb is native to Eurasia and was introduced in North America in the early 19th century (Thompson et al. 1987). Loosestrife is known for its high fecundity, and can potentially produce as many as a million seeds per individual plant (Thompson et al. 1987, Gaudet and Keddy 1995). The seeds are tiny, and are easily dispersed long distances via streams and creeks (Thompson et al. 1987). The predominant loosestrife habitat is nutrient rich wetlands with sandy soils; however recent studies have shown loosestrife is capable of invading relatively mesic wetland habitats (Thompson et al. 1987, Moloney et al. 2009). Not surprisingly by 1990 more than a thousand wetland and lake shores in Minnesota were invaded by loosestrife and this number doubled by the end of 1999 (Blossey et al. 2001). Minnesota’s landscape is particularly vulnerable to invasion by loosestrife, making it an interesting system to study landscape level distribution patterns of an invasive plant.

We selected the four adjacent counties of Washington, Hennepin, Anoka and Ramsey to study the landscape-level distribution pattern of loosestrife in Minnesota (Figure 1). These counties were selected because of their higher loosestrife density. This is evident from the fact that the cumulative loosestrife distribution comprises 1605 presence records for the entire state of MN, out of which 574 occur within the four selected counties. The selection of counties with high densities of loosestrife implies that loosestrife is likely to be in equilibrium with its environment.
GIS Data: Loosestrife Distribution and Land Use Land Cover

The loosestrife distribution data used in this study were obtained from the survey records of MN-DNR’s Invasive Species Program. The bulk (~90%) of loosestrife occurrence information comes from surveys conducted between 1981 and 1999, while the earlier records dating from 1938 are derived from herbarium. All locations were georeferenced using the Universal Transverse Mercator (UTM) system. It may be noted that loosestrife occurrences represent variable population sizes ranging from large dense populations to scattered individual presences. Nevertheless, in the effort to characterize spatial patterns it was assumed that each loosestrife occurrence, irrespective of the population size, is an unbiased indicator of environmental conditions suitable for the establishment of loosestrife and contributes equally to propagule pressure. Given that loosestrife can potentially colonize a broad range of disturbed wetland habitats such an oversimplified assumption is reasonable. Despite the uncertainties associated with the data, we refrained from trimming the data in any manner since each loosestrife occurrence is an independent record that provides valuable information on the spatial processes that result in the eventual spatial distribution.

Landscape environmental variables were procured through the National Land Cover Database 2001 (NLCD 2001) downloaded from MN-DNR’s data deli website (www.deli.dnr.state.mn.us) in raster format at 30m resolution. The NLCD is standardized raster data, where land use and land cover (LULC) is classified into 15 different categories (see Table 1). Urban conditions, which entail four “developed” categories, were common within the chosen study region, with the twin cities of Minneapolis and St. Paul being the approximate centers of the chosen study region.
In order to speed up computational procedures used in our analysis, each 30 m by 30 m raster cell was converted to a coarser scale of 60 m by 60 m by merging 4 adjacent cells into a single new cell. The newly created cell was assigned the median category, since it preserves the categorical nature of the LULC variables, moreover the alternative choices of “maximum” and “minimum” statistic types often results in a minority (less common) cell category being assigned to the newly created cell.

Spatial Pattern Analyses and Statistical Methods

Identifying land use land cover types

In order to identify the key land use land cover (LULC) types, we hypothesized that the proportion of observed loosestrife that fall within raster cells of various LULC types would be significantly different from the proportions that would fall within those same types when the loosestrife distribution is simulated as a random process. In other words, the LULC types that contain a significantly higher proportion of observed loosestrife compared to a random distribution will be identified as the key LULC types.

Prior to testing the above hypothesis, we narrowed our sampling to four circular areas of radius of 10 km that have high densities of loosestrife occurrence (Figure 1). This was done because there is considerable heterogeneity in the observed loosestrife distribution within the study region of four counties. While it is evident from the map (Figure 1) that the four circular areas have a high degree of overlap, the rationale behind the selection was to ensure that the data came from areas where all the suitable habitat sites were likely to be invaded by loosestrife. The large area sampled also ensured that the considerable spatial heterogeneity of the LULC distribution was captured. In the four sampled areas, urban land cover types were predominant in the two central areas.
sampled, while deciduous forests and pastures were abundant in the other two samples. Nevertheless, the environmental data contains a strong urban component, which implies the characterized patterns may not hold for a region with few urban developments. Despite these constraints the selected region is the best choice for several factors; high loosestrife density (~ one loosestrife observation for every 3 square km), a thoroughly surveyed region given proximity to the DNR office in St. Paul, and finally the selected region also had some of the oldest records of loosestrife occurrence in the entire state. Taken together, the strategy behind the sample selection was to strike a balance between the species-environment equilibrium assumption and environmental heterogeneity.

The random distribution of loosestrife was simulated by using the GIS tool, ‘Create Random Points’ by specifying the sampled circular areas of 10 km radius within the LULC layer as the extent and the number of random points generated was set to be equivalent to the observed number of loosestrife occurrences within that area. Random point generation was repeated 100 times. For each randomization, the LULC variables associated with the randomly distributed loosestrife occurrences were extracted. The proportions of random loosestrife incidences associated with each LULC type were then calculated. Also, prior to randomization we modified contiguous raster cells representing ‘open water’ LULC types like lakes and rivers by retaining only edge cells. Contiguity was defined by the 8-cell neighborhood rule. The edge cells were defined by a negative buffer of 60m around contiguous patches of open water LULC, and erasing all the cells that fall beyond this buffer length.

Identification of the preferred habitat for loosestrife was accomplished by determining the LULC types that had an overrepresentation of loosestrife occurrences
relative to a random distribution pattern. The proportion of observed occurrences of loosestrife for a specific LULC was calculated as the number of occurrences observed in that type divided by the total number of occurrences within each circular study area. The same metric was calculated for each of the 100 randomized distributions.

All randomizations were scripted and executed using Python 2.5 within the ArcGIS 9.3 framework (see Appendix A). From the iterated distribution of random loosestrife, estimates of the average of random loosestrife occurrences were made. In addition, 95% confidence intervals around the average proportion for the random patterns were determined from the standard error of the proportions for each LULC type. Finally, Pearson’s Chi-square analysis was used to confirm if there was a significant difference between the average proportion of randomly distributed loosestrife and the proportion of observed loosestrife distribution associated with each land use land cover type. For each LULC type within each sampled circle, the number of loosestrife that fell and did not fall in a given LULC type under the categories of observed and randomly simulated distributions were entered as columns and rows in the Chi-square matrix. All the Chi-square analyses were executed on R version 2.11.0 using the Stats package.

Exploring neighborhood landscape composition

After identifying the key LULC types (viz.; open water, developed open space or herbaceous wetland), we evaluated the location of each of these key LULC types with respect to landscape position to determine if the composition of the surrounding landscape (i.e., neighborhood) influenced loosestrife presence and absence. The neighborhoods were defined in two ways: as circles around a key LULC cell that contained loosestrife, and as circles around a key LULC type that did not contain any
loosestrife. In order to account for variation in spatial locations of invaded sites, the entire study region was included, i.e. all the four counties of Washington, Hennepin, Anoka and Ramsey (Figure 1). However, due to the large spatial extent of the selected study region, the number of potential but uninvaded open water (opw), developed open spaces (dop) and herbaceous wetlands (hwt) raster cells far exceeded the invaded cells. So, for each key LULC category we explored the spatial neighborhoods of 500 randomly selected “uninvaded” raster cells. The radii of the circular neighborhoods around the invaded and randomly chosen uninvaded cells were assigned in an increasing series of 100 m, 200 m, 400 m, 800 m and 1600 m each. Eventually, the proportions of each LULC type surrounding each focal cell for all combinations of neighborhood radii were calculated by executing a sequence of ARC GIS tools (Appendix B). The proportions obtained for each key LULC focal cell were then compared across the increasing neighborhood scales to assess if surrounding land use pattern differed between loosestrife invaded locations and loosestrife “uninvaded” locations. And, in order to gain a clearer picture of how strong the differences between the proportions of surrounding land use were, the overall proportion of the selected surrounding LULC type in the entire study region was calculated and used as a baseline for comparison. Hence, if the proportion of a particular LULC type in a neighborhood of given radius was above the baseline when the central, focal cell contained loosestrife, it clearly suggests that there is a unique land use pattern associated with loosestrife invaded sites and this pattern is strongly manifested in the selected neighborhood scale.
Evaluation of loosestrife occurrence as neighbors (proxy propagule pressure estimate)

We hypothesized that loosestrife containing patches have more loosestrife as neighbors, and hence receive higher propagule pressure from the surrounding landscape. In order to evaluate this hypothesis, raster cells that were classified as herbaceous wetlands were grouped into contiguous patches of wetlands, and those classified as open water edges were grouped together into contiguous open water edge patches using the 8-neighbor cells rule. Neighborhoods were then defined as buffers, and analyzed individually for each contiguous habitat patch type (viz; wetlands and open water edges) under conditions of loosestrife presence or absence. The buffers were defined from the patch edge at scales of 0.5 km, 1 km, 2 km, 4 km, 6 km, and 12 km. Finally, for each of the buffer scale categories the mean number of loosestrife within the neighborhood was calculated separately for contiguous wetland patches and open water patches. The 95% confidence intervals were defined around the mean loosestrife neighbors to exclude values that fall beyond the upper and lower 2.5% extremities. We limited this analysis to herbaceous wetlands and open water edges because they were capable of independently containing loosestrife, unlike developed open spaces, which contained loosestrife only if they fall within 200 meters of open water or herbaceous wetland sites.

Results

Developed open spaces, herbaceous wetlands and open water edges are the key LULC types

Out of the 15 land use land cover types, open water (opw), developed open spaces (dop) and herbaceous wetland (hwt) showed significantly higher proportions of
loosestrife occurrence than would occur if loosestrife occurrences were randomly distributed (Table 2). This was the most consistent pattern through all the four circular areas sampled (Table 2, Figure 2). As expected, the proportion of randomly simulated loosestrife occurrences tracked the availability of land use land cover types, but observed loosestrife proportions were significantly higher for certain LULC types like; opw, dop, and hwt. Deciduous forest (dcf) land cover occasionally had a higher proportion of loosestrife, but chi-square test results for the dcf category showed no significant difference between the observed and randomly simulated proportions. Moreover, a separate analysis showed that most (~95%) of loosestrife occurrences in deciduous forest cells were situated within 100 m of open water, developed open spaces or herbaceous wetland sites, suggesting occurrences in dcf could be a case of misclassification.

Role of surrounding land use

The analyses of spatial neighborhoods show that, irrespective of the key LULC type that holds loosestrife, loosestrife locations have a characteristic neighborhood composition (Figure 3). Analysis of open water (opw) edges as the focal LULC type showed that loosestrife containing opw edge cells had a higher proportion of developed open space, developed low intensity and developed medium intensity neighborhood LULC types than randomly selected opw cells not containing loosestrife (Figure 3). Similarly, herbaceous wetland (hwet) cells that contained loosestrife had a higher proportion of developed open space, developed low intensity and developed medium intensity neighborhood LULC types compared to any randomly chosen hwet cell not containing loosestrife (Figure 3). Moreover, the proportions of these characteristic LULC types were often higher within the immediate neighborhood (100 m – 400 m) of
loosestrife locations than the overall baseline proportion for the entire study region; this is particularly true for open water cells. In the case of hwet cells, only developed open space had higher proportion within immediate neighborhood compared to region’s baseline proportion. It is also interesting to note that absence of loosestrife in opw and hwet sites is often associated with certain neighborhood compositions. Loosestrife invaded opw and hwet sites were found to have on average lower proportions of grassland, pasture and deciduous forest as neighborhood as compared to randomly selected sites of the same kind. For other LULC types within the defined neighborhood there were no clearly discernible differences in the patterns (hence, not shown in Figure 3).

Analysis of developed open space (dop) as a focal LULC type showed that loosestrife containing developed open space cells has a characteristic neighborhood pattern, and it is strongly restricted to a neighborhood of 100m – 200m. These neighborhoods were characterized by a higher proportion of open water, herbaceous wetlands and also woody wetlands. In other words, developed open sites containing loosestrife have a neighborhood pattern that strongly suggests the presence of water in proximity.

*Evaluation of propagule pressure*

For both contiguous patches of wetland and open water edges the average number of loosestrife occurrence as neighbors were higher when patches contained loosestrife compared to those that did not contain loosestrife (Figure 4). The difference was more apparent at larger distance scales (> 2 km), and the average number of loosestrife
occurrences in the neighborhood increased more rapidly for invaded wetlands and lake edges.

**Discussion**

The landscape-level patterns suggest loosestrife invaded locations are not randomly distributed across the landscape. Instead there is a hierarchical structure, starting with a strong propensity for loosestrife occurrence in three key LULC types *viz.*; herbaceous wetlands, open water edges, and developed open spaces. However, the analyses of neighborhoods of these LULC types show that loosestrife is more likely to occur among wetlands, open water edges and developed open spaces with a certain characteristic neighborhood landscape composition. In other words, loosestrife’s habitat is not merely a function of a few favored LULC types but also depends on the composition of surrounding land use and land cover. Finally, it is evident that, besides the right suite of environmental conditions, loosestrife occurrence is also determined by proximity to neighboring loosestrife populations. This proximity to neighboring loosestrife locations suggests that propagule pressure plays a critical role in the colonization and spread of invasive loosestrife. Taken together, the hierarchical approach allows us to narrow down and identify the processes that determine loosestrife’s distribution, and the probability of loosestrife occurring in any given location can be formulated more precisely. Our findings, thus argue in favor of recently postulated hierarchical approaches towards understanding and predicting species distributions (Pearson and Dawson 2003, Pearson et al. 2004, Guissan and Thuiller 2005, Milbau et al. 2009).
Loosestrife prefer open water edges and herbaceous wetlands affected by light to moderate degree of disturbances

The higher occurrence of loosestrife in herbaceous wetlands and open water edges is not surprising given that loosestrife is an aggressive and opportunistic invader of wetlands and riparian habitats (Thompson 1987). It is also known that loosestrife is more likely to invade wetlands and riparian habitats that are more prone to disturbances, and tend to have more open canopy (Gaudet and Keddy 1995, Blossey et al. 2001). Studies have also shown that loosestrife is commonly encountered along roadside ditches (Wilcox 1989). In summary, loosestrife like many other invasive plants are more likely to occur within disturbance modified habitat conditions (Hobbs and Huenekke 1992). Our study also show that disturbances play a crucial role in determining if a wetland or open water edge is likely to be invaded by loosestrife. By explicitly defining neighborhoods at multiple scales, we were able to highlight the importance of surrounding land use for loosestrife invasion. It is striking to note that of all the available LULC types, the ones that are strongly associated with loosestrife invaded locations have light to moderately intense human induced disturbance components (developed open spaces, developed low intensity, developed medium intensity) and the LULC’s associated with “uninvaded” loosestrife locations have relatively higher proportions of pastures, grasslands or deciduous forests within the defined neighborhoods. In the case of wetlands and open water edges that are invaded, the patterns also suggest that low intensity disturbances, as found in developed open spaces, mattered the most, followed by developed low intensity and developed medium intensity habitats. This gradient in neighborhood composition suggests that there is a range of disturbances that enhances loosestrife invasion, beyond
which disturbances of high intensity are likely to impede invasion. Few studies have shown through explicit multi-scale neighborhood analysis the clear-cut difference in the associated surrounding landscape composition for invaded and uninvaded habitat locations. While studies have shown that landscape context or matrix can have a strong influence on invasion (Pauchard and Alback 2004), few appear to have explicitly analyzed the same. In a more recent study by Ibáñez et al. (2009) landscape neighborhoods defined at two different scales were found to significantly modify probability of occurrence of several selected invasive plants, moreover characteristic landscape neighborhood composition patterns were found for each invasive plant. In another comparable study, extinction of several grassland species in remnant grassland patches located within a rural-urban spatial gradient were found to be influenced by habitat quality, and the latter is largely dictated by the composition of the landscape matrix (Williams et al. 2006).

The predominance of disturbance prone LULC types (i.e. developed open spaces, developed low intensity, and developed medium intensity) in the surrounding landscape seem to contribute positively to loosestrife invasion by modifying the habitat quality of herbaceous wetlands and open water edges. In a recent paper, Vilà and Ibáñez (2011) highlighted the importance of considering surrounding landscape in determining the probability of invasion of a local patch. Surrounding landscape can influence the local patch conditions by altering light availability, creating disturbed habitat edges, and more indirectly by affecting propagule availability (Vilà and Ibáñez 2011). Moreover, it is interesting to note that certain neighborhood compositions appear to buffer herbaceous wetlands and open water sites embedded within them from loosestrife invasion. The
predominance of deciduous forests, pastures or grasslands within the neighborhood apparently results in minimal loosestrife invasion of open water and wetland sites within it. The closed canopy conditions of deciduous forests during the growing season are likely to impede the colonization and establishment by loosestrife. However, it is less clear how grasslands and pastures behave as buffers against invasion. A possible explanation behind the low loosestrife incidence in wetlands and open water edges with higher amounts of pasture in the surrounding landscape perhaps lies in the fact that most pastures are actively managed, or in private landholdings, which limits accessibility during surveys. In a study by Lovett-Doust et al. (2003), land ownership was found to play a vital role in determining the diversity and distribution of rare biota. Studies have also found strong effects of urbanization in the surrounding landscape on plant invasion (Lindenmayer and McCarthy 2001, Borgmann and Rodewald 2005), but to date the effects of other land cover types like agriculture and pastures is poorly understood (Vilà and Ibáñez 2011). This lack of consensus on the role of “non-urban” LULC types in the surrounding landscape is partly due to the fact that local variability in plant invasion is also attributable to landscape configuration (Vilà and Ibáñez 2011). However, similar to our findings, Lindenmayer and McCarthy (2001) have found that occurrence of exotic pine and blueberry in an Australian landscape varied strongly with respect to landscape context, and was found to be lowest in sites embedded within continuous areas of native forest. Though our findings do not provide any clear mechanistic understanding, it is evident that surrounding landscape composition is important in determining the vulnerability of wetlands and open water edges to loosestrife invasion.
Developed open space: a secondary anthropogenic habitat for loosestrife

Developed open spaces contain significant amounts of loosestrife making it also a key land cover type for loosestrife invasion. Developed open spaces are regions with a relatively higher proportion of green vegetation cover within the urban-suburban gradient; which includes recreational parks, golf courses, and private gardens or backyards. Given the use of loosestrife as an ornamental plant, it is not surprising that developed open spaces as a category have a high prevalence of loosestrife. Urban ecological studies have repeatedly shown that urban environments are characterized by the presence of economically valued exotic species (Niemela 1999). But for other LULC types within the range of urban-suburban conditions, increase in intensity of human activities and proportional decrease in green vegetation cover seems to be disadvantageous to loosestrife. A recent study by Gulezian and Nyberg (2010) has shown a similar pattern wherein abundance of invasive and exotic plants decreased with increasing intensity of human activities. Even more interesting is the finding that loosestrife containing developed open spaces always had moist, water abundant conditions in their immediate vicinity. The observed association of the surrounding landscape pattern suggests that developed open spaces are likely to be poor quality habitats that are capable of holding loosestrife only when there is a wetland or lake in close proximity. It is apparent that proximity to water not only makes the developed open space more suitable for loosestrife, but such sites are also likely to receive more loosestrife propagules.
Loosestrife invasion of a suitable habitat is furthered by propagule pressure

Besides the influential role played by preferential land cover and landscape context there exists a strong spatial dependence in the loosestrife distribution pattern. A key mechanism that results in spatial dependence is dispersal (Miller et al. 2007). In the case of species distribution studies the significance of dispersal is well demonstrated (Craft et al. 2002, Miller et al. 2007). The presence of dispersal constraints is evident in our study since loosestrife locations are on average closer to other loosestrife invaded locations. Our findings supports the recent study by Yakimowski et al. (2005), where dispersal limitation was found to be the immediate factor responsible for the patchy distribution of loosestrife invaded wetlands before micro-site conditions tend to further limit establishment. A similar pattern was detected for zebra mussel invaded lakes where proximity to invaded lakes was one of the key factors that determined the spread of zebra mussels (Craft et al. 2002). Similarly, studies focusing on landscape models and patterns have established the role played by propagule pressure in invasion by exotic plants (Rouget and Richardson 2003, Deckers et al. 2005). However it may be noted that, given the phenotypic plasticity attributed to loosestrife in reproductive and vegetative characteristics it is likely that loosestrife’s distribution and abundance is not merely the outcome of dispersal limitation, but is also determined by finer environmental factors like soil moisture and nutrients (Mal and Lovett-Doust 2005, Chun et al. 2007, Moloney et al. 2009).

Conclusion

Using a simple systematic approach, our study characterizes the landscape level distribution pattern for purple loosestrife. It is apparent that there is a latent hierarchy in
the processes that leads to loosestrife’s broad scale distribution. While our study does not attempt to provide a mechanistic understanding of how these processes affect the eventual distribution, the tacit assumption of such a hierarchy is useful in understanding the distribution pattern of loosestrife. Moreover, both of the previous studies on loosestrife (Welk 2003, Anderson et al. 2006) used low resolution environmental data, only to make broad scale predictions with little information on the ecological processes involved. Comparatively, we were able to better explain and identify the likely processes behind loosestrife’s distribution.

Discerning patterns in a clear manner is the first and perhaps most critical step towards understanding any ecological process (Levin 1992). There has been a recent renaissance in the development of invasive species distribution models (see reviews by Peterson 2003, Guisan and Thuiller 2005); however there also exists a high degree of variability among these models in their predictive ability (Murphy and Lovett-Doust 2007, Ward 2007, Evangelista et al. 2008). While robust predictive species distribution models (SDM’s) are extremely useful, the patterns (and underlying processes) upon which the SDM’s are built are equally insightful. Studies have mostly focused on identifying the landscape-level factors responsible for the distribution pattern of invasive plants (Lindenmayer and McCarthy 2001, Bradley and Mustard 2006). Our results show that detection of the pattern in a hierarchical manner allows us to posit loosestrife invasion in a well-defined ecological framework. Taken together, it is evident that while certain LULC types like wetlands, open water edges and developed open spaces are most likely to be invaded, their vulnerability is spatially variable and depends on the
composition of surrounding land use, and the final probability of a suitable site being invaded by loosestrife lies in its proximity to neighboring loosestrife locations.

Acknowledgements

This project is indebted to Luke Skinner and the MN-DNR’s Invasive Species program for providing us with the loosestrife distribution data. We are grateful to MN-DNR for the financial support provided in the form of summer stipend to the first author. The first author benefitted a lot from DNR summer student interns who provided helpful on-field location specific insights, and is thankful for the same. Finally, we thank two anonymous reviewers for providing valuable comments on improving the manuscript.

References


Table 1. Number of raster cells and their equivalent proportions for each of the fifteen LULC types within the selected study region (see Fig. 1 inset).

<table>
<thead>
<tr>
<th>LULC type (abbreviation)</th>
<th>Number of raster cells</th>
<th>Proportion of raster cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open water (opw)</td>
<td>90445</td>
<td>0.0765</td>
</tr>
<tr>
<td>Developed open space (dop)</td>
<td>125709</td>
<td>0.106</td>
</tr>
<tr>
<td>Developed low intensity (dlo)</td>
<td>211878</td>
<td>0.179</td>
</tr>
<tr>
<td>Developed medium intensity (dmd)</td>
<td>90653</td>
<td>0.0766</td>
</tr>
<tr>
<td>Developed high intensity (dhi)</td>
<td>40722</td>
<td>0.0344</td>
</tr>
<tr>
<td>Barren (bar)</td>
<td>2379</td>
<td>0.002</td>
</tr>
<tr>
<td>Deciduous forest (dcf)</td>
<td>184207</td>
<td>0.155</td>
</tr>
<tr>
<td>Evergreen forest (evf)</td>
<td>26022</td>
<td>0.022</td>
</tr>
<tr>
<td>Mixed forest (mxf)</td>
<td>2159</td>
<td>0.0018</td>
</tr>
<tr>
<td>Shrub / scrub (shr)</td>
<td>14066</td>
<td>0.0118</td>
</tr>
<tr>
<td>Grassland (grs)</td>
<td>31864</td>
<td>0.0269</td>
</tr>
<tr>
<td>Pasture / hay (pas)</td>
<td>162904</td>
<td>0.1377</td>
</tr>
<tr>
<td>Cultivated crops (cult)</td>
<td>124953</td>
<td>0.1056</td>
</tr>
<tr>
<td>Woody wetlands (wwt)</td>
<td>2709</td>
<td>0.0022</td>
</tr>
<tr>
<td>Herbaceous wetlands (hwt)</td>
<td>71616</td>
<td>0.0605</td>
</tr>
</tbody>
</table>
Table 2. Identification of key LULC types that are preferred by loosestrife. Chi-square analysis of all the LULC variables identified three key LULC types; Open Water, Developed Open Space, and Herbaceous Wetland as loosestrife’s favored habitats. The identified key LULC types showed the most consistent significant difference between proportions of observed loosestrife and proportions of random loosestrife occurrence among all the four sampled circular regions shown in the map (Fig. 1). The significance was established at P ≤ 0.05 (indicated in bold).

<table>
<thead>
<tr>
<th>Circle</th>
<th>open water (opw)</th>
<th>developed open space (dop)</th>
<th>herbaceous wetland (hwt)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2 = 8.90$, df = 1, P = 0.002</td>
<td>$\chi^2 = 4.43$, df = 1, P = 0.035</td>
<td>$\chi^2 = 2.49$, df = 1, P = 0.114</td>
</tr>
<tr>
<td>Circle 1</td>
<td>$\chi^2 = 1.70$, df = 1, P = 0.198</td>
<td>$\chi^2 = 0.510$, df = 1, P = 0.47</td>
<td>$\chi^2 = 8.29$, df = 1, P = 0.003</td>
</tr>
<tr>
<td>Circle 2</td>
<td>$\chi^2 = 2.21$, df = 1, P = 0.136</td>
<td>$\chi^2 = 5.72$, df = 1, P = 0.016</td>
<td>$\chi^2 = 8.76$, df = 1, P = 0.003</td>
</tr>
<tr>
<td>Circle 3</td>
<td>$\chi^2 = 10.8$, df = 1, P = 0.001</td>
<td>$\chi^2 = 2.95$, df = 1, P = 0.085</td>
<td>$\chi^2 = 11.8$, df = 1, P = 0.0005</td>
</tr>
</tbody>
</table>
Figure 1. Map of Minnesota showing loosestrife distribution (solid black dots). The pull-out box shows loosestrife distribution for the selected four counties (thick black outlines) that comprise the overall study region. The circles within these counties represent the sampled areas with high loosestrife densities, and are indexed by numbers 1 to 4. Location of Minneapolis is indicated by the star.
Figure 2. Relative distributions of observed loosestrife occurrence and randomly simulated loosestrife occurrence for the four sampled circular regions. Each graph shows the proportions of average random and observed loosestrife across different land use land cover types for each of the four sampled circular regions. The dark solid line represents the average loosestrife percentages resulting from 100 random iterations, and the broken lines that lie very closely around the solid dark line are the upper and lower 95% confidence levels. The grey solid line represents observed loosestrife percentages.

Figure 3. Arranged column-wise each plot shows the proportion of a specific LULC type around a key focal LULC type (i.e. open water, herbaceous wetlands and developed open space) as a function of neighborhood radius. The solid black lines represent the proportions around key focal LULC type selected randomly, while the solid grey lines represent the proportions around key focal LULC type that contain loosestrife. The broken lines around the solid lines are 95% confidence intervals. Proportion of LULC type available within the entire selected study area is shown by the horizontal dashed line.
Figure 4. Average number of loosestrife as neighbors around; (a) wetland habitats and (b) open water edges invaded by loosestrife (grey lines) and without loosestrife (black lines) as a function of increasing neighborhood distance (see methods for details on how then number of neighbors was determined). The broken lines are 95% confidence interval.
Appendix A

Python codes to generate random points and find the LULC type it falls on. Finally, iterating the same process over 100 times and calculating the frequency.

```python
import arcgisscripting
gp = arcgisscripting.create(9.3)
gp.CheckOutExtension("Spatial")
u'CheckedOut'
gp.OverWriteOutput = 1
for r in range (1,101):
    gp.CreateRandomPoints("F:/Loosestrife Proj/Module1","rp" + str(r) + ".shp","F:/Loosestrife Proj/Module1/855Sample.gdb/LsfBuffer","","129")
    gp.ExtractValuesToPoints_sa("F:/Loosestrife Proj/Module1/randompoint" + str(r) + ".shp","F:/Loosestrife Proj/Module1/855Sample.gdb/LandCover855","F:/Loosestrife Proj/Module1/extractedrandompoint" + str(r) + ".shp","NONE","VALUE_ONLY")
    gp.frequency_analysis("F:/Loosestrife Proj/Module1/exrp" + str(r) + ".shp","F:/Loosestrife Proj/Module1/855Sample.gdb/prsence" + str(r),"RASTERVALU")
```

Appendix B

Python codes to find the count of various LULC types within defined neighborhood scales of randomly simulated and observed loosestrife distribution.

```python
import sys, string, os, arcgisscripting
gp = arcgisscripting.create(9.3)
gp.Workspace = "F:\Loosestrife Proj\Module2\Results\Random1600855.gdb"
gp.MakeFeatureLayer_management("F:\Loosestrife Proj\Module2\LsfRnd_Select.shp","rp855layer")
<geoprocessing server result object object at 0x00A1A308>
gp.CheckOutExtension("Spatial")
u'CheckedOut'
gp.OverWriteOutput = 1
rows = gp.SearchCursor("rp855layer")
row = rows.Next()
```
while row:
    rowID = row.GetValue("FID")
    string = "FID" + "=" + str(rowID)
    gp.SelectLayerByAttribute_management("rp855layer","NEW_SELECTION",string)

    gp.Buffer_analysis("rp855layer","rpb855layer" + str(rowID),1600,","",","","")
    gp.ExtractByMask_sa("F:\Loosestrife Proj\Module1\855Sample.gdb\extlnd855","rpb855layer" + str(rowID),"Extr_r855layer" + str(rowID))

    gp.RasterToPoint_conversion("Extr_r855layer" + str(rowID),"r855lyr_16point" + str(rowID))

    row = rows.Next()
CHAPTER 3. INVASIVE SPECIES DISTRIBUTION MODEL: COMBINING THE EFFECTS OF SURROUNDING LAND-USE & PROPAGULE-PRESSURE

A paper to be submitted to *Journal of Applied Ecology*

Shyam M Thomas\(^1,2\) and Kirk A Moloney\(^1\)

\(^1\) *Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA- 50011*

\(^2\) Email: *smthomas@iastate.edu* (corresponding author)

**Abstract**

1. Distribution and spread of invasive plants across a landscape is largely governed by disturbance invoking anthropogenic land-use practices and propagule pressure. However, the spatial variability in disturbances and propagule pressure associated with the occurrence of invasive plant is poorly understood. In this study, we develop predictive models for an invasive wetland plant – Purple Loosestrife, by considering explicitly the spatial variability associated with surrounding land-use conditions and propagule pressure.

2. Using survey data of loosestrife distribution and remotely-sensed land use land cover information for the state of Minnesota, we first predicted loosestrife occurrences in two types of wetland habitat namely; herbaceous wetlands and open-water edges with a series of logistic regression models that incorporated surrounding land-use at three different neighborhood scales. The best-fitting surrounding land-use model was then combined with three different distance constraint models that simulated propagule pressure.
3. Loosestrife occurrence as a function of surrounding land-use showed best fit at the neighborhood radius of 400 m. Predictions made from the surrounding land-use model at 400 m scale were fairly accurate and invasion of wetland locations were correlated with the proportion of disturbance prone anthropogenic land-use conditions in the neighborhood. Inclusion of an autocovariate simulating propagule pressure improved model fit and performance significantly.

4. *Synthesis and applications.* Spatially explicit incorporation of surrounding land-use yields an ecologically realistic spatial projection of invasion risk wherein disturbance prone habitat edges tend to be more invasible. Combining this prediction with location specific estimates of propagule pressure further reduces uncertainty by spatially constraining areas of high invasion risk. Our approach is applicable to other invasive plants since it is based on two key drivers of plant invasion: disturbance and propagule-pressure.

Keywords: autologistic model, GIS, omission and commission errors, principal component regression, loosestrife, risk map, validation

**Introduction**

Biological invasion is one of the most pervasive and pernicious effects of human mediated environmental changes. Over time, invasions can homogenize species diversity, and drastically alter the structure and function of ecosystems. Such ‘emerging ecosystems’ or ‘novel biomes’ are rapidly arising with the increase in the human driven collapse of biogeographical boundaries (Hobbs et al. 2006; Lindenmayer et al. 2008). Moreover, these human modifications of the landscape are producing a progressively
convergent pattern globally – increase in urban areas and expansion of agricultural lands at the expense of natural land cover (Ellis & Ramankutty 2008). In short, land-use is an indicator of human population distribution. And, invasive species are typically associated with anthropogenic land-use (Vitousek et al. 1996). Therefore, increasingly similar global land use patterns imply vulnerability to invasion by a common pool of opportunistic species. Given this present-day scenario, it is imperative to understand the impact of human land use on biological invasions, and apply the lessons learnt to the management of invasive species.

Predicting regions of high invasion risk is a critical first step for the management and control of alien plant species. If left unmanaged invasive exotics can cause severe ecological as well as economic loss (Hobbs 1995; Pimentel et al. 2005). Not surprisingly, there has been a recent surge in the development of tools and modeling approaches geared towards effective prediction of regions under high invasion risk (Guisan & Thuiller 2005; Elith 2006). Climatic suitability models are among the most common approaches employed in predicting areas under high risk of invasion (Kriticos et al. 2003; Elith et al. 2006). However, climatic / bioclimatic models are coarse-scale models that are best suited to delineate an invasive species’ potential distributional range (Pearson & Dawson 2003; Hampe 2004). To elicit a more accurate picture of the spatial variability of invasion risk, fine scale environmental variables like land use & land cover (LULC) are required, since they are ideal for analyzing distribution patterns at both the landscape (10 – 200 km) and local scale (1 – 10 km) (Pearson & Dawson 2003; Milbau et al. 2009). Conservation plans are often drawn at the landscape level, while ecological processes at the local scale play a key role in defining the variability in landscape level invasion risk.
Efficacy of conservation plans requires accurate identification of areas under high invasion risk across a landscape; hence the choice of LULC as an environmental variable is warranted.

Hobbs (2000) described land-use as a measure of ‘the extent of modification and predominance of different types of disturbances in any given area’. Despite this clear link between land-use and disturbance, few studies to date have explored the role of land-use and land cover more explicitly (Vilà et al. 2007). Surrounding land-use is known to modify local habitat conditions making a location more vulnerable to invasion, particularly when there is a strong anthropogenic component (Vilà et al. 2011). And, for invasive plants surrounding land-use in the form of anthropogenic disturbances affects a focal habitat or a site’s invasibility by either altering the quality of the habitat or by increasing the number of propagules arriving at the site. More recently, efforts have been made to differentiate between the invasibility of a site and the actual amount of invasion occurring at a site (Chytrý et al. 2008; Eschtruth & Battles 2011). In these recent works, invasibility is defined as a site (location) or habitat specific feature, independent of the impact of propagule pressure. Thus, combining location specific measures of invasibility with a consideration of spatially variable propagule pressure is a sound approach to measure the risk of invasion by an invasive species across a spatially heterogeneous landscape.

Propagule pressure is often approximated indirectly using models that measure fine scale spatial dependence among the response variables. Models that incorporate fine scale spatial autocorrelation are auto-regressive models, where spatial lag of a response variable is measured as a function of the distance to its nearest neighbor or based on a
pre-defined neighborhood (Miller et al. 2006). When the response variable is presence / absence of a species at a given location it is referred to as an autologistic model (Augustin et al. 1996). In autologistic models, an autocovariate term is used to characterize the influence of presences within a defined distance or neighborhood on the probability of occurrence of a species.

In this paper, autocovariate is defined by three different distance constraint models to simulate propagule pressure. This includes two kinds of nearest-neighbor models where the strength of propagule pressure is calculated first as an inverse distance function, and then as a negative exponential distance function to the nearest source population. In the final distance-constraint model, the autocovariate is estimated as a function of multiple sources within a predefined neighborhood and their cumulative effect on propagule pressure. The use of three different distance-constraint models with a slight increment in complexity at each level provides a clearer understanding of the variability in propagule pressure quantification and its consequence in the eventual prediction of loosestrife’s probability of occurrence.

In short, we predict the locations under high risk of invasion by Purple loosestrife (*Lythrum salicaria* L.), an aggressive invader of wetland habitats, by first developing logistic regression models that considers the effect of surrounding land use, and then fine-tune the models by developing autologistic models that combines propagule pressure with surrounding land-use as predictor variables. Thus, we combine spatial variability of processes that operate at two different levels of ecological organization. By explicitly accounting for the spatial variability, accurate risk maps of loosestrife invasion can be
developed, which is of indispensable value in managing loosestrife invasion at the landscape scale.

**Methods**

*Focal Species and Study Area*

Purple loosestrife (*Lythrum salicaria* L.), an erect perennial herb of Eurasian origin, is among the worst invasive alien plants in North America (Blossey et al. 2001, but see Lavoie 2008). First recorded in the early 1800’s, loosestrife soon became a semi-naturalized species with established populations in wetlands of northeast United States (Blossey et al. 2001). By mid 1900, it had spread through most of the northeastern United States, often forming dense monocultures, with very scant native vegetation amidst (Blossey et al. 2001, Lavoie 2008). An aggressive wetland invader, loosestrife has since spread rapidly to colonize wetlands through most of the continental United States and adjacent Canadian provinces.

Loosestrife is capable of invading a broad range of wetland habitats; however disturbance mediated openings is a key requirement (Rachich & Reader 1999). Not surprisingly, loosestrife is considered an indicator of disturbed wetlands (Lavoie 2008). Once established, loosestrife can produce large amounts of seeds, most of which are dispersed by water, and a majority end up as a seed bank in the soil (Thompson et al. 1987). Loosestrife has a prominent rootstock that overwinters and re-grows rapidly in spring. Overall, eliminating a dense stand of loosestrife from a wetland is a challenging task.
Given loosestrife’s affinity for wetlands it is not surprising that Minnesota’s wetland rich landscape is heavily invaded. From the earliest known records of loosestrife in 1938, the loosestrife invasion rate showed a logistic growth pattern, and by 1999 nearly 2000 wetlands were invaded (Blossey et al. 2001). Most loosestrife invasion is within a predominantly urban environment in and around the twin cities of Minneapolis and St. Paul, which suggests loosestrife’s establishment and spread is driven by human affected disturbances and changes.

*Loosestrife Distribution Data and Environmental Variables*

We obtained loosestrife occurrence data for the entire state of Minnesota from the MN-DNR’s Invasive Species Program. All of the 1604 loosestrife presence points were geo-referenced using the Universal Transverse Mercator system to an accuracy of the nearest 20 meters. However, we narrowed the study area to four conterminous counties (Washington, Hennepin, Anoka and Ramsey) with the twin cities of Minneapolis and St. Paul as the approximate center. The selected study area is predominantly urban with a high prevalence of loosestrife invasion – 574 recorded occurrences. The selected area also has some of the oldest records of loosestrife occurrence. Thus, the chosen study area is likely to have a loosestrife distribution approaching equilibrium.

Environmental variable used in all the models was the National Land Use Land Cover Database (NLCD 2001) for the state of Minnesota, which was obtained from MN-DNR’s data deli (www.deli.dnr.mn.us) in raster format at a resolution of 30 meters. The land use land cover (LULC) data comprised of fifteen standardized categories of land-use and land cover types. Given the predominantly urban nature of the selected study area, the most common land use category was the “developed” category, which includes
developed open space, developed low intensity, developed medium intensity and
developed high intensity LULC types. Other common land use categories include
deciduous forests, cultivated crops and pastures. Prior to analysis of loosestrife
distribution the resolution of the LULC raster layer was reduced to 60 meters per pixel.
The coarser resolution was used for analysis as it speeds up iterative computational
procedures executed at the pixel level.

Model Structure and Details

We developed the predictive habitat risk models for loosestrife based on the
spatial distribution of two key habitat types - herbaceous wetlands and open water edges
(Figure 1). The choice of these two land cover types as loosestrife’s preferred habitats was
based on the results of prior exploratory analyses (Thomas & Moloney 2013), which
ascertained that these two land cover types have a significantly higher proportion of
loosestrife occurrence compared to other LULC types. Out of the total 574 loosestrife
occurrences in the study area, herbaceous wetlands and open water edges contained
14.3% and 12.3% of the occurrences, respectively, despite the low availability of these
two LULC types within the study area. And, from hereon we use the term ‘wetland’
while referring to both herbaceous wetlands and open water edges.

For each wetland type (i.e. herbaceous wetlands and open water edges), the
probability of loosestrife occurrence within wetland cells was predicted first as a function
of the proportion of surrounding LULC types at three different neighborhood radii (400
m, 800 m, and 1600 m). The minimum radius was set to 400 m because it was the
minimum scale at which the spatial neighborhoods contained all the 14 different
surrounding LULC types. The different radii at which the surrounding LULC types were
accessed allows comparing landscape context at different scales, and determines if there is any scale specific effect of landscape context. Once identified, the surrounding LULC information from the best-fitting neighborhood scale was combined with an autocovariate to evoke the additional effect of propagule pressure.

Survey recorded occurrences of loosestrife within each of the wetland types were used as the loosestrife presence data. Since, the data available on loosestrife was presence-only data, loosestrife absence data were simulated by randomly generating points that fell within raster cells of each wetland category excluding those cells that already had a recorded loosestrife presence. Given that the absences are essentially randomly generated points, they should be considered as ‘pseudo-absences’. More so, all the absences were located within land cover types (i.e. herbaceous wetlands and open water edges) preferred by loosestrife, which makes the model predictions more stringent and closer to the actual observed distribution of loosestrife. While studies have suggested that pseudo-absences be derived from environmentally unsuitable regions, the latter approach is better suited if the goal is to predict species’ potential distributions or ranges (Václavík & Meentemeyer 2009). Another consideration regarding our approach is that loosestrife presences were from variable population sizes, ranging from individual loosestrife plants to large contiguous populations. Models trained with such inherent population size variability among locations are likely to be less precise in predicting invasion hot-spots. However, given loosestrife’s inherent ability to colonize a variety of disturbed wetland habitats, such variability is useful in making robust predictions of loosestrife’s habitat suitability.
The response variables used in developing our models were presences and pseudo-absences of loosestrife in wetland cells of either the herbaceous wetland category or the open water edge category. Hence, we begin with logistic models for each wetland category where surrounding land-use information is the independent variable. Given the high degree of collinearity among the fourteen types of LULC, we employed a principal component analysis method to reduce the dimensionality of the environmental variables to a set of uncorrelated principal component (PC) axes. Loosestrife occurrences were then regressed against all the 14 PC scores to identify the significant PC predictors at the best-fitting neighborhood scale. Each model is essentially a logistic regression model based on a specific neighborhood scale, where principal component scores \((x_1, x_2, ..., x_{14})\) were ‘modified’ environmental covariates with intercept \((\beta_0)\) and associated regression coefficients \((\beta_1, \beta_2, ..., \beta_{14})\), and \(\epsilon\) is the error term (Table 1). The model then takes the form:

\[
\text{logit} \ (p_i) = \log \left( \frac{p_i}{1-p_i} \right) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_{14} x_{14i} + \epsilon_i
\]

where \(p_i\) is the probability of loosestrife occurrence within a given wetland cell. Once the best-fitting neighborhood scale was identified, we systematically dropped non-significant PC predictors till the best-fitting combination of PC predictors was identified for the best-fitting neighborhood scale.

To evaluate the additional effects of spatial dependence due to propagule pressure, regression models were developed by combining surrounding land use with a consideration of propagule pressure as an autocovariate. A proxy measure of propagule pressure was estimated from the spatial distribution of loosestrife-invaded locations in the selected study region. Inclusion of a proxy propagule pressure measure as an
autocovariate makes the model an auto-logistic regression model (Augustin et al. 1996; Wu & Huffer 2001), which is similar to the logistic regression model, but includes the autocovariate term $z_i$ and its associated coefficient, $\rho$:

$$\text{logit} (p_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots + \beta_n x_{ni} + \rho z_i + \epsilon_i$$

Three different autocovariate measures were formulated, representing three unique distance constraint models: (i) the nearest neighbor linear model (NN), (ii) the nearest neighbor exponential decay model (NED), and (iii) the cumulative distance model (CD). Each distance constraint model autocovariate term was evaluated in a model containing all of the 14 PC predictors (Table 1). Combining the autocovariate with all 14 PC predictors allows for selecting the significant PC predictors after accounting for the effect of propagule pressure.

For nearest neighbor models (NN), propagule pressure in any given wetland cell (or location) was defined as an inverse function of the distance to the nearest loosestrife invaded cell.

$$Z_{NN,i} = 1/d_{ij}$$

where $Z_{NN,i}$ represents the effective propagule pressure at location $i$. This was calculated as the inverse of the Euclidian distance $d_{ij}$ between source location $j$ and target location $i$. An inverse distance function implies a decrease in propagule pressure with increasing distance to the nearest invaded location.

A modification of the NN model is the nearest neighbor exponential decay model (NED) where the effective propagule pressure at a target location decreases exponentially with distance from source location. Mathematically, the nearest neighbor exponential model is:
\[ Z_{\text{NED},i} = \exp(-b*d_{ij}) \]  
(2)

where \( d_{ij} \) is the same as in the nearest neighbor model. The parameter \( b \) is the rate of decay of the effective propagule pressure with distance. High values of \( b \) indicate strong dispersal limitation, as propagule pressure falls off rapidly with increasing distance, whereas low values indicate the opposite. Unlike in NN model, the effective propagule pressure in NED model decays slowly with increase in distance from a source location (Appendix A).

In the cumulative distance model (CD) multiple propagule source locations within a specified neighborhood radius can cumulatively affect propagule pressure at a target location. Hence, for a pre-determined neighborhood containing \( n \) loosestrife invaded source locations, the CD model can be formulated as:

\[ Z_{\text{CD},i} = \sum_{j=1}^{n} \exp (-b*d_{ij}) \]  
(3)

\( Z_{\text{CD},i} \) represents the effective propagule pressure at location \( i \), and \( d_{ij} \) is the distance between target location \( i \) and loosestrife containing propagule source location \( j \). We selected the smallest neighborhood radius that resulted in a significant positive effect of the autocovariate by running the autologistic model with training data for a series of neighborhoods with increasing radiiuses. As in the NED model, parameter \( b \) is the rate of decay of propagule pressure with increasing distance. For both models, the best-fit for parameter \( b \) was estimated by maximizing the log-likelihood estimates of the distance constraint models by using training data comprised of randomly selected loosestrife presences. Finally, to evaluate the role of propagule pressure in determining the probability of loosestrife occurrence, we also predicted loosestrife occurrence as a function of the best-fitting distance constraint model alone. Comparing the performance
of the distance constraint model with the combined autologistic model elucidates the distinct role of propagule pressure and environmental factors.

Model Training and Initiation

Prior to running the above defined logistic and autologistic models to make predictions of loosestrife occurrences within each wetland type, the models were trained using a set of loosestrife presences and pseudo absences (training data). This training data comprised of randomly selected 33 out of 69 locations of known loosestrife occurrences in herbaceous wetlands, along with 400 randomly selected ‘uninvaded’ loosestrife pseudo-absence locations to form the training data for herbaceous wetlands. Similarly, 39 out of 77 randomly selected open water edge locations of known loosestrife occurrences along with 400 randomly selected ‘uninvaded’ loosestrife; pseudo-absence locations formed the training data for open water edges. Both the logistic and auto-logistic models for each wetland type were trained using the same set of presence-absence data.

Given the potential biases and gaps in survey data, a Bayesian approach was taken in model fitting as it allows for flexibility in parameter estimation by defining a prior distribution, and allows for direct comparison of multiple models. Regression parameters for all models in our study were given a vague prior distribution: \( \beta_n \sim \text{Normal}(0, 1000) \). The large standard deviation (1000) around the mean (0) implies parameter estimates are essentially uninfluenced by the prior. In short, a vague or uninformative prior is the Bayesian equivalent of an ‘objective’ approach wherein parameter estimation is driven by the data (Latimer et al. 2006). The Deviance Information Criterion (DIC) is essentially the Bayesian equivalent of AIC, which allows for direct comparisons among models. Like AIC, DIC is a measure of relative goodness of fit, where lower values of DIC imply
a better fit. The DIC values were used to find the best-fitting model among the LULC models specified at different neighborhood scales and among LULC models with different combinations of PC predictors for the identified best-fitting neighborhood scale. Finally, DIC values were used to compare results among the best-fitting LULC model, the propagule pressure alone models and the combined autologistic model that integrated the LULC model with each of the three propagule pressure models. Bayesian models employ the MCMC (Markov Chain Monte Carlo) algorithm for simulating the posterior distribution of regression parameters, which yields the mean parameter estimate along with 97.5th and 2.5th credible intervals. The regression parameters thus estimated, were derived from 30000 sampling iterations of the MCMC algorithm with a “burn-in” of 5000 iterations. All the above-mentioned simulations were executed using WinBUGS version 1.4 (Spiegelhalter et al. 2003).

**Invasion Risk Maps and Model Validation**

Two different invasion risk maps were developed, one as a function of surrounding land use (the logistic model) and another as a function of both surrounding land-use and propagule pressure (the auto-logistic model). For the surrounding land use risk map the probability raster for each wetland type was built based on the significant PC predictors of loosestrife occurrence and their loadings. For the surrounding land-use plus propagule pressure risk map, the probability raster also incorporated the additional effect of propagule pressure.

Models were evaluated by testing the predictions made for each wetland category using a validation dataset that excluded the training data. More specifically, the validation dataset included loosestrife occurrences within wetland cells that were completely
independent of the presences used in the training data and also included loosestrife occurrences that fell outside but within 100 meters of the nearest wetland cell. The latter choice (i.e. extra-wetland loosestrife occurrences) is based on the assumption that occurrence within immediate vicinity of a wetland site (i.e. raster cell) is indicative of loosestrife invasion of that wetland site. Using ArcGIS, the nearest wetland cell within a radius of 100 meters was located and identified for each of the extra-wetland occurrences. Both true occurrence locations and extra-wetland occurrence locations were combined, which resulted in a total of 142 and 152 validation points for herbaceous wetland habitat and open water edge habitat, respectively (Figure 1). We evaluated each model using the receiver operating characteristics (ROC) method, which quantifies the trade-off between true-positive rate and false-positive rate of model predictions (Fielding & Bell 1997). In a ROC plot, the trade-off between true-positive rate and false-positive rate typically results in a curve called the ROC curve. The area under the ROC curve (AUC) is a measure of the fit of the model with respect to the validation data used, the higher the AUC measure the better the model fit. A null model yields an AUC value of 0.5. Thus, higher the AUC value relative to the null model value - 0.5, the better the performance of the model. In order to calculate the AUC for the logistic and auto-logistic regression models associated with each wetland type, the predicted probabilities of loosestrife validation points associated with each wetland type were compared with an equal number of randomly selected locations from the entire region of the selected wetland type. This is equivalent to finding how well the logistic model and the auto-logistic model are able to discern loosestrife’s suitable habitat relative to the entire sampled area represented by the wetland type.
Minimal predicted area (MPA) is another measure of accuracy, where the goal is to find the model that minimizes the predicted surface for a predefined probability threshold (Engler et al. 2004). The predefined threshold is typically set at a value that contains at least 90% of the loosestrife validation points. In other words, a good habitat-suitability model should be able to discern loosestrife’s habitat by minimizing the amount of area predicted as suitable. In our application of MPA, the reduction in number of cells was calculated as an indirect measure of area. Thus, the lower the MPA measure relative to the total available area within each wetland category for a model, the better the model performance, and the more accurate the area predicted as suitable is. We predicted MPA for both the logistic and auto-logistic models associated with each wetland type at the 90% threshold level.

Results

Among logistic regression models assessed at different neighborhood scales of surrounding land use, the lowest DIC values for both open water edges and herbaceous wetlands was attained at 400 m. Therefore, we selected 400 m as the best-fitting neighborhood scale. However, the DIC values at larger neighborhood scales were higher by a very narrow margin suggesting there is no significant difference in the model fit among the three neighborhood scales (Table 1). The best-fitting significant PC-predictors selected lowered the DIC values substantially at the best-fitting neighborhood scale of 400 m (Table 1). The propagule pressure alone models yielded a poor fit relative to best-fit LULC model as evident from the lower DIC values for both wetland types (Table 1). This suggests that propagule pressure alone is not a good predictor of loosestrife
occurrence. The autologistic models generally proved to be a better fit with lower DIC values than the logistic model for both wetland types (Table 1, Figure 2b & c). However, the improved fit of the autologistic models varied according to the distance constraint model used. For both wetland habitats, model fit improved significantly when the distance constraint was incorporated as an NED model or a CD model.

In the absence of an auto-covariate, the best-fitting significant PC predictors at the 400 m neighborhood scale for open water edges were PC 1 and PC 3, and for herbaceous wetlands PC 1 and PC 4 were identified as the best-fitting, significant predictors (Figure 2a). It may be noted that PC 2 was marginally significant for herbaceous wetlands; however its inclusion resulted in poor model fit, and hence was excluded (not shown in Table 1). The loadings of the significant principal components suggest there is a surrounding land use pattern that is associated with the probability of loosestrife occurrences (Appendix B). This relationship between the significant PC predictors and loosestrife occurrence becomes clearer when PC scores of the significant PC predictors are correlated against the original LULC variables (Figure 3). Overall, the probability of loosestrife occurrence is positively correlated with increasing proportion of ‘developed’, anthropogenic conditions in the neighborhood and low when surrounded by high proportion of ‘agricultural’ land-use. However, not all ‘developed’ land-use conditions have equal effect on loosestrife occurrence. Notably, loosestrife occurrence is relatively lower when the surrounding ‘developed’ land-use categories are representative of a high intensity of anthropogenic developments (or disturbances), and this is evident in the correlations associated with the significant predictors PC 3 and PC 4 of the open water edges and herbaceous wetlands, respectively. Also of interest is the inconsistent effect of
deciduous forests with respect to loosestrife occurrence in open water edge habitats,
wherein deciduous forests within the neighborhood had both positive and negative impact
on loosestrife occurrence (Figure 3).

In the presence of an autocovariate, the autologistic model selected a different set
of PC predictors. For herbaceous wetlands, the NED based autologistic model selected
PC 2 and PC 6 as the best-fitting predictors, while the CD based autologistic model
selected PCs 1, 2, 4 and 6 as the best-fitting predictors. For open water edges, both the
NED and CD based model selected only PC 1 as the best-fitting significant predictor.

The final raster probability maps show the risk of loosestrife invasion in open
water edges and herbaceous wetlands as: (i) a function of the best-fitting significant
predictors derived from surrounding land use information (Figure 4), and (ii) as a
combination of surrounding land use and best-fitting distance constraint model, i.e. the
NED model (Figure 5). It may be noted that for open water edges the CD and NED
models had similar DIC values suggesting that the former is a best-fit model too.
Validation results corroborated the better fit of the auto-logistic model, when propagule-
pressure is incorporated as a nearest distance function (Table 2). Both NED and CD
based autologistic models registered high AUC and low MPA values indicating that
inclusion of propagule pressure as an autocovariate improves the model’s performance
substantially. However, the CD based autologistic models performed bit poorly compared
to the NED based autologistic models (Table 2). In other words, including multiple
source locations, as in the CD model, did improve model performance but not beyond
that of the NED based autologistic models.
Discussion

The models developed in this paper highlight the importance of surrounding land use and proxy measures of propagule pressure in predicting loosestrife invaded sites located in two kinds of wetland habitat. Models incorporating only surrounding land use performed poorly in predicting loosestrife occurrence relative to models that included propagule pressure as an autocovariate. However, the surrounding land use based models highlighted that loosestrife invasion within a wetland habitat type is affected by the composition of surrounding land-use, particularly when it involves disturbance prone human-induced edges. The addition of propagule pressure as an autocovariate was critical in making more accurate prediction of loosestrife occurrence. Taken together, it is evident that loosestrife invasion is a spatially constrained phenomenon, wherein the combined effects of surrounding land-use conditions and propagule pressure determine the invasion risk.

Surrounding Land-Use as a Predictor of Loosestrife Invasion

The best-fitting logistic regression models based on surrounding land-use indicate that surrounding land-use as an explanatory variable is able to discern, to a reasonable extent, loosestrife invaded sites from ‘uninvaded’ sites. The lack of scale-specific differences in the measure of model fit between loosestrife invaded locations and surrounding land-use is perhaps due to the effect of surrounding land-use conditions averaging out at 400 m. The significant PC predictors selected by the best-fitting logistic regression models also indicate that there is a signature land-use pattern around loosestrife-invaded sites, which differs from the land-use pattern surrounding ‘uninvaded’ sites. This is evident from the PC loadings associated with the significant PC predictors.
The correlation between PC 1 and the original environmental variables suggests that for both wetland habitat types disturbance-prone, ‘developed’ land-use categories in the neighborhood are positively correlated with loosestrife invaded sites. The influence of anthropogenic land-use in determining exotic species distribution is not surprising since human induced habitat modifications is among the key driving factors (Maheu-Giroux & de Blois 2007, Vilà & Ibanez 2011). In the case of open-water edge habitats, PC 1 loadings and correlations also suggest that pastures and croplands are correlated with ‘uninvaded’ sites. However, for herbaceous wetlands, pasture and crop LULC types showed a weak positive correlation with uninvaded sites. Instead, uninvaded herbaceous wetland sites were more strongly correlated with deciduous forests and other herbaceous wetlands in the neighborhood. In other words, it appears that predominance of certain landscape elements around wetland habitat locations significantly lowers the chance of loosestrife invasion. Knowledge of such ‘buffering’ landscape elements indicates that invasibility of wetland habitats in a landscape is highly variable and depends on landscape context. The surrounding land-use model also suggests that wetland edges defined by developed land-use categories are the most suitable invasion sites. This qualitative difference in wetland edge attributes and their influence on invasibility upholds a key spatial property of plant invasion, wherein habitat edges are recognized as the initiation points of invasion (Pauchard & Alback 2006). Similarly, Cadenasso & Pickett (2001) found that the structure of edge vegetation determines the flux of exotic seeds that invade forest interiors.

Inclusion of the second significant PC predictor (PC 3 for open water edges, and PC 4 for herbaceous wetlands) adds information on the role of surrounding land-use land
cover conditions. For both wetland types, correlations and loadings associated with the second PC predictor suggest that intensely developed land-use categories are negatively correlated with loosestrife-invaded sites. There are also differences in the surrounding land-use patterns around herbaceous wetland invaded sites and open-water edge sites. In the case of herbaceous wetlands, PC 4 indicates that presence of open-water edges in the neighborhood are positively correlated with loosestrife invaded sites, suggesting open-water edge habitats like rivers, lakes, and creeks are perhaps ‘carriers’ of loosestrife propagules. The conflicting pattern associated with deciduous forest is speculated to be the outcome of the variation invoked by spatial configuration with respect to both the wetland habitat types within the landscape, and perhaps also due to the sharp difference in shape of the wetland habitat types. More specifically, the linear shape of open-water edges represented by lake shores and river banks has a wider spatial spread and tends to capture a higher degree of landscape heterogeneity as compared to herbaceous wetlands, which are generally more contiguous and spatially consolidated.

**Propagule Pressure as a Measure of Spatial Contagion**

Inclusion of a distance constrained autocovariate term improved model-fit for the NED and CD based autologistic models. However, there were differences between the two habitat types with respect to improvement in model-fit by NED and CD based autologistic models. The PC predictors selected by the autologistic models informs about the probability of loosestrife occurrence explained by covariates after accounting for propagule pressure. In this respect, it is interesting to note the differences between herbaceous wetlands and open water edge habitats. The selection of PC 2 and PC 6 instead of PC 1 and PC 4 (the best-fitting predictors in the absence of an autocovariate)
by the NED based autologistic model for herbaceous wetlands suggests that the presence of both herbaceous and woody wetlands in the neighborhood positively influenced loosestrife occurrence (see Appendix B, Table 1 for PC loadings). It is of interest that PC 1 and PC 4 were no longer significant predictors for the NED based autologistic model, which suggests that the spatial dependency invoked by the autocovariate was captured in the LULC alone logistic model by the presence of developed open spaces and open water conditions within the immediate neighborhood of herbaceous wetlands. However, the retention of PC1 and PC 4 predictors along with PC2 and PC 6 by the CD based autologistic model suggests that propagule pressure quantified as a cumulative distance function has a weak influence on the determination of loosestrife occurrence within herbaceous wetland cells and is unable to compensate for the spatial variability captured by PC 1 and PC 4 predictors. This weak influence of the CD based autologistic model is likely to be the outcome of the relatively large spatial scale (i.e. 1000 m) at which spatial dependency is realized by the autocovariate. Similarly, the elimination of PC 3 as a significant predictor in the autologistic models for open water edges suggests that the spatial contagion captured by the autocovariate is generated by the addition of PC 3 as a significant predictor in the LULC based logistic model.

The apparent change in the selected predictors as a consequence of including an autocovariate is not unusual and has been reported in previous studies involving autocovariate models (Dorman 2007; Betts et al. 2009). It is understood that the inclusion of a spatial component in the form of an autocovariate captures fine scale effects, as opposed to the PC predictors which operate at relatively broader spatial scales Hawkins et al. 2007).
The better performance of the NED model relative to the CD model suggests that, for loosestrife, propagule pressure from the nearest source location alone is critical in determining its invasion probability. It is however quite apparent that unlike the NED model, the CD based auto-logistic model highlights areas that have a high density of loosestrife occurrences, thus pointing to invasion hot-spots that need immediate management action (see Appendix C for CD model based map). The failure of the NN model, the simplest of all distance constraint models, to improve model-fit suggests that the spatial dependency resulting from propagule pressure is not a simple function of distance to nearest source location. In NN models the magnitude of propagule pressure decreases rapidly with distance, which is unlikely to hold for invasive species like loosestrife as occasional long-distance dispersal events occur along with the frequent short-range dispersal events.

**Combined Autologistic Model: Strengths & Limitations**

Results of the combined autologistic model are comparable to other invasive species distribution models that have similarly highlighted the strong role played by dispersal in predicting locations under future invasion risk (Rouget & Richardson 2003; Williams et al. 2008). In the absence of dispersal constraint, risk maps based on surrounding land-use conditions alone suggests that there are several potential locations that are highly vulnerable to invasion. However, inclusion of a dispersal constraint further limits the number of the potentially invasible locations predicted. Allouche et al. (2008) used different distance constraint models, and found they yielded patchy distributions of habitat suitability with better predictive accuracy. Similarly, it is evident from the risk maps developed for loosestrife that inclusion of dispersal constraints strongly affects the
estimates of invasion risk and yields a patchy distribution of suitable locations.
Incorporating propagule pressure as an autocovariate reduces omission errors (i.e. an individual occurs where the model predicts it should not occur). Reduction of omission error in our models is particularly evident for NED based autologistic models, since high AUC values essentially indicate low omission errors. However, omission errors can also result from sensitivity of the surrounding land-use variable to both landscape composition and configuration within a given area. In this respect, the poor predictions of loosestrife occurrences in the northern part of the study area are possibly the outcome of model’s poor sensitivity to regions with high amounts of deciduous forest relative to anthropogenic land-use.

Invasive species are unlikely to be in equilibrium with their environment, which makes it challenging to predict their occurrence with low commission errors (i.e. absence of species where they should occur). In our models, commission errors are apparent when one compares MPA measures of the surrounding land-use model with the NED based autologistic model for both wetland types. Inclusion of propagule pressure lowered the MPA value suggesting that only locations falling sufficiently close to a loosestrife invaded location would result in a higher suitability, while locations distant from invaded loosestrife location resulted in lower predicted suitability compared to the suitability predicted by the surrounding land-use alone model. Commission errors can also be attributed to interference by ecologically similar species. In our study area, *Typha angustifolia* and *T. latifolia*, are known to co-occur with loosestrife, and wetlands are often dominated by either *Typha* sp. or loosestrife (Blossey et al. 2001). The possibility of site pre-emption by *Typha* sp. was confirmed by a quick field survey of wetlands
around the twin cities of Minneapolis / St. Paul. In many wetlands sites dominated by hybrid *Typha*, loosestrife was either totally absent or relegated to a few individuals along wetland edges abutting a road.

From a management perspective, invasive species distribution modeling aims is to improve accuracy by minimizing omission errors (Jimenez-Valverde et al. 2011). In this respect, inclusion of an autocovariate mimicking propagule pressure significantly improved the predictive accuracy of the model. It is, however, difficult to pin-point the ‘cause’ behind the improved performance of autologistic model as both environmental conditions, such as disturbances in the immediate neighborhood, and biotic elements, such as propagule pressure, are equally plausible contributing factors. Despite this lack of clarity in the workings of an autocovariate, its incorporation as an endogenous source of spatial autocorrelation is beneficial, especially when the goal is to improve predictive accuracy in species distribution models (Betts et al. 2009).

**Conclusions**

A variety of methods and approaches are now available to develop predictive species distribution models (McNally 2000, Elith & Leathwick 2009). However, there is an increasing emphasis on building distribution models with ecologically relevant variables in a meaningful manner so that the predictions made are insightful. In this respect, our modeling approach makes accurate predictions without losing ecological meaning and relevance. In a recent paper, Hulme (2010) proposed the need for a ‘scenario planning’ approach as a way forward to overcome uncertainties and complexities involved in weed risk assessment. Similarly, we believe that for loosestrife risk assessment, the focus needs to shift to manageable regional scenarios, wherein
spatial uncertainties inherent in ecological processes across multiple scales are explicitly considered.

Acknowledgements
We are thankful to Minnesota-DNR’s Invasive Species Program for providing survey based loosestrife distribution data and for financially supporting the first author in the form of a summer stipend. The first author is grateful to Luke Skinner and Adam Doll of Minnesota-DNR for providing valuable on-field information regarding loosestrife invasion and spread. We are also thankful to members of plant lab group at Iowa State University for suggesting ways to improve the model presentation and write-up. The first author is also grateful to the EEB program at Iowa State, which provided research support through several semesters of support as a Research Assistant.

References


Tables

Table 1

Logistic and autologistic regression models predicting the probability of loosestrife occurrence within a raster cell of either herbaceous wetland category or open water edge category. The logistic regression models are classified according to the neighborhood scale at which the surrounding land-use was accessed, and the distance-constrained models used to simulate propagule pressure. In addition, for the best-fitting neighborhood scale (i.e. 400 m), the best-fitting combination of significant PC predictors was identified. The logistic regression model at the best-fitting neighborhood scale is then coupled individually with the three uniquely defined distance based auto-covariates to yield the autologistic regression models. For both logistic and autologistic models, the numerical values for the subscripts of the independent variables represent the respective PC axis values. Lower DIC values imply better model fit. Model parameters and variables are as defined in the Methods section.
Table 1 continued

<table>
<thead>
<tr>
<th>Herbaceous Wetlands</th>
<th>Open Water Edges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model Type</strong></td>
<td><strong>Model Structure</strong></td>
</tr>
<tr>
<td><strong>Logistic regression models</strong></td>
<td></td>
</tr>
<tr>
<td>Surrounding LULC alone (all 14 PC predictors)</td>
<td>$\logit(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_{14} x_{i14} + \epsilon_i$</td>
</tr>
<tr>
<td>Surrounding LULC alone (with best-fitting significant PC predictors)</td>
<td>$\logit(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_4 x_{i4} + \epsilon_i$</td>
</tr>
<tr>
<td>Propagule pressure alone (NN model)</td>
<td>$\logit(p_i) = \beta_0 + \rho_{zNN,i} + \epsilon_i$</td>
</tr>
<tr>
<td>Propagule pressure alone (NED model)</td>
<td>$\logit(p_i) = \beta_0 + \rho_{zNED,i} + \epsilon_i$</td>
</tr>
<tr>
<td>Propagule pressure alone (CD model)</td>
<td>$\logit(p_i) = \beta_0 + \rho_{zCD,i} + \epsilon_i$</td>
</tr>
</tbody>
</table>

**Autologistic regression models**

| Surrounding LULC (significant PC predictors & Propagule pressure nearest neighbor - NN model) | $\logit(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_4 x_{i4} + \rho_{zNN,i}$ | 400 | 220.6 | $\logit(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_3 x_{i3} + \rho_{zNN,i} + \epsilon_i$ | 400 | 250.8 |
Surrounding LULC (significant PC predictors) & Propagule pressure (negative exponential distance - NED model)

\[
\text{logit } (p_i) = \beta_0 + \beta_2 x_{2i} + \beta_6 x_{6i} + \rho z_{\text{NED},i} \quad 400 \quad 188.3
\]

Surrounding LULC (significant PC predictors) & Propagule pressure (cumulative distance - CD model)

\[
\text{logit } (p_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \rho z_{\text{CD},i} \quad 400 \quad 240.2
\]

\[
\text{logit } (p_i) = \beta_0 + \beta_1 x_{1i} + \rho z_{\text{CD},i} \quad 203.6
\]

Table 2
Validation results indicating performance of models using the AUC (area under characteristic curve) and MPA (minimal predicted area) method. Higher AUC values and lower MPA values imply better model performance. MPA values estimated at 90% threshold are relative to the total number of cells contained by each wetland category viz.; 42,022 for open water edges and 71,716 for herbaceous wetlands.

<table>
<thead>
<tr>
<th>Habitat type</th>
<th>Model type</th>
<th>AUC</th>
<th>MPA (at 90% threshold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPW (open water edges)</td>
<td>Logistic</td>
<td>0.77</td>
<td>22,841</td>
</tr>
<tr>
<td></td>
<td>Autologistic – NED</td>
<td>0.97</td>
<td>884</td>
</tr>
<tr>
<td></td>
<td>Autologistic – CD</td>
<td>0.93</td>
<td>6,915</td>
</tr>
<tr>
<td>HWT (herbaceous wetlands)</td>
<td>Logistic</td>
<td>0.76</td>
<td>42,359</td>
</tr>
<tr>
<td></td>
<td>Autologistic – NED</td>
<td>0.98</td>
<td>2,679</td>
</tr>
<tr>
<td></td>
<td>Autologistic – CD</td>
<td>0.95</td>
<td>10,401</td>
</tr>
</tbody>
</table>
Figure 1. Map of the four conterminous counties selected as the study area showing the two key wetland habitat types – open water edges (OPW) and herbaceous wetlands (HWT) with loosestrife invaded locations used for model validation. The solid triangles...
are loosestrife validation points for open water edge habitats, and solid circles (●) are loosestrife validation points for herbaceous wetland habitats.

Figure 2. Posterior means of regression coefficients with 95% credible intervals predicting loosestrife occurrence within wetland habitats estimated from (a) logistic models based on surrounding land-use conditions alone as represented by the first six leading PC axes, (b) NED-based autologistic model and (c) CD-based autologistic model based on the combined effects of autocovariate simulating propagule pressure and surrounding land use conditions. The solid black bars represent herbaceous wetland habitat, while the dotted bars represent open water edge habitat. Parameter values that do not overlap the vertical zero line are considered to be significant predictors. (Note: the choice of showing only first six PC axes was because their eigen-values were greater than or equal to 1 and also for visual clarity).
Figure 3. Correlation between principal component loadings and the original environmental (LULC) variables for the best-fitting significant principal component predictors (PC1 and PC4 for herbaceous wetlands, and PC1 and PC3 for open water edges) evaluated at the neighborhood scale of 400 m. The sign associated with the regression coefficients were accounted for so that LULC types that positively influence loosestrife occurrence show positive loadings while LULC types with negative loading values have the opposite effect. Significance of correlation was accessed at the following thresholds: *p<0.10, **p<0.05, & ***p<0.001
opw = open water, dop = developed open spaces, dlo = developed low intensity, dmd =
developed medium density, dhi = developed high intensity, barr = barren, dcf = deciduous
forest, evf = evergreen forest, mxf = mixed forest, shr = shrub / scrub, grs = grassland, pas =
pasture, crp = cultivated crops, wwt = woody wetlands, & hwt = herbaceous wetlands
Figure 4. Map of the four conterminous counties selected as the study area showing loosestrife invasion risk for herbaceous wetlands and open water edges as predicted by the logistic regression model using surrounding land-use information. The empty space represents the matrix around the wetland habitats, which comprise of the remaining 12 land-use land cover categories.
Figure 5. Map of the four conterminous counties selected as the study area showing loosestrife invasion risk for both herbaceous wetlands and open water edges as predicted by the auto-logistic regression model with propagule pressure estimated using the negative exponential distance (NED) model. The empty white space represents the matrix around the wetland habitats, which comprise of the remaining 12 land use land cover categories.
Appendix A

Comparison of the fitted decline in propagule pressure estimates between nearest neighbor model (NN) and nearest negative exponential decay model (NED) as a function of increasing distance for herbaceous wetlands. For the NED model, the decline in propagule pressure was estimated by finding the best-fitting rate of decay parameter (i.e. $b = 0.0019$). See ‘Methods’ for details on model structure and parameter estimation. (Note: For open water edges the difference in propagule pressure patterns is overall similar.)
## Appendix B: PC loadings at 400 m radius for herbaceous wetlands and open water edges

Table 1

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
<th>PC8</th>
<th>PC9</th>
<th>PC10</th>
<th>PC11</th>
<th>PC12</th>
<th>PC13</th>
<th>PC14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opw</td>
<td>-0.05</td>
<td>0.11</td>
<td>-0.06</td>
<td>0.70</td>
<td>0.36</td>
<td>0.17</td>
<td>-0.39</td>
<td>0.16</td>
<td>0.25</td>
<td>-0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.29</td>
</tr>
<tr>
<td>Dop</td>
<td>-0.41</td>
<td>0.02</td>
<td>0.10</td>
<td>0.26</td>
<td>-0.13</td>
<td>-0.11</td>
<td>0.11</td>
<td>-0.25</td>
<td>-0.45</td>
<td>0.14</td>
<td>0.22</td>
<td>-0.56</td>
<td>0.14</td>
<td>-0.22</td>
</tr>
<tr>
<td>Dlo</td>
<td>-0.50</td>
<td>0.02</td>
<td>0.07</td>
<td>0.05</td>
<td>-0.09</td>
<td>0.03</td>
<td>0.11</td>
<td>-0.17</td>
<td>-0.18</td>
<td>0.01</td>
<td>-0.18</td>
<td>0.56</td>
<td>0.50</td>
<td>-0.25</td>
</tr>
<tr>
<td>Dmd</td>
<td>-0.49</td>
<td>0.15</td>
<td>-0.04</td>
<td>-0.26</td>
<td>0.04</td>
<td>-0.05</td>
<td>-0.12</td>
<td>0.10</td>
<td>0.14</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.28</td>
<td>0.72</td>
<td>-0.13</td>
</tr>
<tr>
<td>Dhi</td>
<td>-0.37</td>
<td>0.16</td>
<td>-0.03</td>
<td>-0.43</td>
<td>0.11</td>
<td>-0.08</td>
<td>-0.23</td>
<td>0.28</td>
<td>0.32</td>
<td>-0.10</td>
<td>0.06</td>
<td>-0.45</td>
<td>0.44</td>
<td>-0.07</td>
</tr>
<tr>
<td>Dcf</td>
<td>0.25</td>
<td>0.04</td>
<td>0.60</td>
<td>-0.19</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.15</td>
<td>0.11</td>
<td>0.01</td>
<td>0.55</td>
<td>-0.15</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.43</td>
</tr>
<tr>
<td>Evf</td>
<td>-0.06</td>
<td>-0.11</td>
<td>0.38</td>
<td>0.22</td>
<td>-0.51</td>
<td>-0.05</td>
<td>0.29</td>
<td>0.03</td>
<td>0.56</td>
<td>-0.25</td>
<td>0.22</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.09</td>
</tr>
<tr>
<td>Mxf</td>
<td>-0.06</td>
<td>-0.24</td>
<td>0.18</td>
<td>0.10</td>
<td>0.34</td>
<td>-0.29</td>
<td>0.41</td>
<td>0.67</td>
<td>-0.24</td>
<td>-0.12</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Shr</td>
<td>0.04</td>
<td>-0.29</td>
<td>0.13</td>
<td>-0.13</td>
<td>0.51</td>
<td>-0.37</td>
<td>0.01</td>
<td>-0.49</td>
<td>0.19</td>
<td>-0.01</td>
<td>0.43</td>
<td>0.12</td>
<td>0.05</td>
<td>-0.06</td>
</tr>
<tr>
<td>Grs</td>
<td>0.14</td>
<td>-0.19</td>
<td>-0.04</td>
<td>0.01</td>
<td>-0.36</td>
<td>-0.51</td>
<td>0.12</td>
<td>-0.26</td>
<td>-0.29</td>
<td>0.08</td>
<td>0.12</td>
<td>0.02</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>Pas</td>
<td>-0.02</td>
<td>-0.59</td>
<td>-0.15</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.10</td>
<td>0.04</td>
<td>-0.16</td>
<td>0.13</td>
<td>-0.21</td>
<td>-0.58</td>
<td>0.11</td>
<td>-0.36</td>
<td></td>
</tr>
<tr>
<td>Crp</td>
<td>0.02</td>
<td>-0.42</td>
<td>-0.29</td>
<td>-0.17</td>
<td>-0.15</td>
<td>0.47</td>
<td>-0.01</td>
<td>0.22</td>
<td>-0.08</td>
<td>0.18</td>
<td>0.55</td>
<td>0.09</td>
<td>0.04</td>
<td>-0.25</td>
</tr>
<tr>
<td>Wwt</td>
<td>0.07</td>
<td>0.12</td>
<td>0.46</td>
<td>-0.17</td>
<td>0.17</td>
<td>0.46</td>
<td>-0.12</td>
<td>-0.09</td>
<td>-0.28</td>
<td>-0.63</td>
<td>0.10</td>
<td>-0.01</td>
<td>0.05</td>
<td>-0.06</td>
</tr>
<tr>
<td>Hwt</td>
<td>0.32</td>
<td>0.46</td>
<td>-0.34</td>
<td>-0.10</td>
<td>0.01</td>
<td>-0.16</td>
<td>0.34</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.20</td>
<td>0.07</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.61</td>
</tr>
<tr>
<td>Proportion of Variance</td>
<td>0.19</td>
<td>0.13</td>
<td>0.1</td>
<td>0.08</td>
<td>0.078</td>
<td>0.075</td>
<td>0.066</td>
<td>0.062</td>
<td>0.055</td>
<td>0.053</td>
<td>0.045</td>
<td>0.037</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Eigen-value</td>
<td>1.63</td>
<td>1.35</td>
<td>1.21</td>
<td>1.1</td>
<td>1.05</td>
<td>1.03</td>
<td>0.96</td>
<td>0.93</td>
<td>0.88</td>
<td>0.86</td>
<td>0.8</td>
<td>0.72</td>
<td>0.53</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>PC loadings for Open Water Edges</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
<th>PC8</th>
<th>PC9</th>
<th>PC10</th>
<th>PC11</th>
<th>PC12</th>
<th>PC13</th>
<th>PC14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opw</td>
<td>-0.01</td>
<td>0.63</td>
<td>0.15</td>
<td>-0.02</td>
<td>0.33</td>
<td>-0.10</td>
<td>0.30</td>
<td>-0.12</td>
<td>0.04</td>
<td>-0.03</td>
<td>-0.11</td>
<td>-0.01</td>
<td>0.05</td>
<td>0.59</td>
</tr>
<tr>
<td>Dop</td>
<td>-0.31</td>
<td>-0.19</td>
<td>-0.46</td>
<td>-0.24</td>
<td>-0.04</td>
<td>-0.15</td>
<td>-0.15</td>
<td>0.07</td>
<td>-0.15</td>
<td>-0.24</td>
<td>-0.59</td>
<td>0.14</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Dlo</td>
<td>-0.42</td>
<td>-0.25</td>
<td>-0.26</td>
<td>-0.18</td>
<td>0.10</td>
<td>0.03</td>
<td>-0.04</td>
<td>-0.06</td>
<td>0.03</td>
<td>0.19</td>
<td>0.12</td>
<td>0.58</td>
<td>-0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>Dmd</td>
<td>-0.42</td>
<td>-0.23</td>
<td>0.29</td>
<td>0.32</td>
<td>-0.07</td>
<td>-0.01</td>
<td>-0.03</td>
<td>0.06</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.16</td>
<td>0.71</td>
<td>0.18</td>
</tr>
<tr>
<td>Dhi</td>
<td>-0.32</td>
<td>-0.15</td>
<td>0.44</td>
<td>0.43</td>
<td>-0.10</td>
<td>-0.07</td>
<td>0.02</td>
<td>0.08</td>
<td>0.03</td>
<td>-0.11</td>
<td>-0.06</td>
<td>-0.32</td>
<td>-0.58</td>
<td>0.14</td>
</tr>
<tr>
<td>Dcf</td>
<td>0.31</td>
<td>0.03</td>
<td>-0.34</td>
<td>0.35</td>
<td>-0.10</td>
<td>-0.14</td>
<td>-0.27</td>
<td>0.46</td>
<td>-0.24</td>
<td>-0.28</td>
<td>0.25</td>
<td>0.06</td>
<td>-0.01</td>
<td>0.38</td>
</tr>
<tr>
<td>Evf</td>
<td>0.18</td>
<td>-0.20</td>
<td>-0.25</td>
<td>0.29</td>
<td>0.18</td>
<td>0.42</td>
<td>0.30</td>
<td>0.33</td>
<td>0.40</td>
<td>0.36</td>
<td>-0.25</td>
<td>-0.12</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>Mxf</td>
<td>0.10</td>
<td>-0.16</td>
<td>-0.24</td>
<td>0.38</td>
<td>-0.03</td>
<td>0.24</td>
<td>0.36</td>
<td>-0.57</td>
<td>-0.47</td>
<td>-0.16</td>
<td>-0.07</td>
<td>0.05</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Shr</td>
<td>0.27</td>
<td>-0.08</td>
<td>0.00</td>
<td>0.27</td>
<td>0.06</td>
<td>-0.54</td>
<td>-0.30</td>
<td>-0.30</td>
<td>0.00</td>
<td>0.58</td>
<td>-0.20</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Grs</td>
<td>0.32</td>
<td>-0.29</td>
<td>0.13</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.16</td>
<td>0.00</td>
<td>-0.16</td>
<td>0.46</td>
<td>-0.56</td>
<td>-0.34</td>
<td>0.30</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Pas</td>
<td>0.30</td>
<td>-0.37</td>
<td>0.23</td>
<td>-0.19</td>
<td>0.18</td>
<td>0.06</td>
<td>0.08</td>
<td>-0.21</td>
<td>0.11</td>
<td>0.08</td>
<td>0.65</td>
<td>-0.26</td>
<td>0.04</td>
<td>0.28</td>
</tr>
<tr>
<td>Crp</td>
<td>0.18</td>
<td>-0.27</td>
<td>0.33</td>
<td>-0.34</td>
<td>0.14</td>
<td>0.19</td>
<td>-0.12</td>
<td>0.23</td>
<td>-0.54</td>
<td>0.08</td>
<td>-0.44</td>
<td>0.00</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Wwt</td>
<td>0.02</td>
<td>0.22</td>
<td>0.06</td>
<td>0.13</td>
<td>0.03</td>
<td>0.59</td>
<td>-0.68</td>
<td>-0.30</td>
<td>0.15</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>Hwt</td>
<td>0.14</td>
<td>0.09</td>
<td>0.07</td>
<td>-0.16</td>
<td>-0.88</td>
<td>0.10</td>
<td>0.15</td>
<td>-0.06</td>
<td>0.08</td>
<td>0.20</td>
<td>-0.05</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.30</td>
</tr>
<tr>
<td>Proportion of Variance</td>
<td>0.2</td>
<td>0.13</td>
<td>0.1</td>
<td>0.09</td>
<td>0.08</td>
<td>0.077</td>
<td>0.068</td>
<td>0.063</td>
<td>0.058</td>
<td>0.046</td>
<td>0.041</td>
<td>0.036</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Eigen-value</td>
<td>1.66</td>
<td>1.36</td>
<td>1.18</td>
<td>1.12</td>
<td>1.03</td>
<td>1</td>
<td>0.98</td>
<td>0.93</td>
<td>0.81</td>
<td>0.76</td>
<td>0.71</td>
<td>0.53</td>
<td>0.0001</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Map of the four conterminous counties selected as the study area showing loosestrife invasion risk for both herbaceous wetlands and open water edges as predicted by the auto-logistic regression model with propagule pressure estimated using the cumulative distance (CD) model. The empty white space represents the matrix around the wetland habitats, which comprise of the remaining 12 land use land cover categories.
CHAPTER 4. EXPLORING HERBIVORY TOLERANCE IN PERENNIALS WITH STORED BELOWGROUND BIOMASS

A paper to be submitted to Functional Ecology

Shyam M Thomas\textsuperscript{1,2} Karen C Abbott\textsuperscript{1} and Kirk A Moloney\textsuperscript{1}

\textsuperscript{1} Department of Ecology, Evolution, and Organismal Biology, Iowa State University,
Ames, IA- 50011

\textsuperscript{2}Email: smthomas@iastate.edu (corresponding author)

Abstract

1. Plant defense against herbivory is a well-studied topic of research with many unanswered questions. Between the two recognized categories of plant defense strategies, resistance and tolerance, tolerance is relatively poorly studied. Perennials with long-term, belowground stored biomass are examples of plants that can tolerate high amounts of aboveground defoliation or biomass loss without major fitness consequences.

2. We developed a mathematical model for a perennial plant with seasonal growth and its specialist insect herbivore using three state variables: aboveground biomass, belowground biomass and herbivore population density. The model provides a clear understanding of how three closely associated traits—(i) belowground biomass allocation to roots, (ii) spring utilization of stored reserves, and (iii) post-herbivory regrowth capacity—modulate the persistence of the perennial plant population under different intensities of herbivory. We particularly focus on the role played by belowground allocation to storage as a tolerance mechanism against aboveground herbivory.
3. Model results suggest that allocation to belowground storage is indeed a critical mechanism that can determine tolerance to herbivory in perennials. Low belowground biomass allocation resulted in the extinction of both the herbivore and plant population, while stable coexistence of plants at low manageable biomass, along with its specialist insect, required a moderate amount of post-herbivory belowground allocation to roots. High values of belowground allocation and stored reserve utilization resulted in sustained cycles of the herbivore and plant populations. Also, herbivore traits that limit plant damage, such as stronger intraspecific competition for food or increased herbivore interference, further increases belowground storage.

4. In the context of invasive perennials, model findings suggest that biocontrol agents (i.e. specialist insect herbivores) causing periodic short-term defoliation are incapable of extirpating plants that can efficiently allocate biomass belowground. In doing this, our model combines, for the first time, three key traits attributable to perennials with long-term belowground storage (i.e., belowground biomass allocation, spring utilization of stored reserves and post-herbivory regrowth capacity) into a consideration of the effects of herbivory and offers a fresh perspective on how variation in these three key traits can influence the success of biocontrol management.

Keywords: biomass allocation, biocontrol, insect herbivores, food limitation, herbivore interference, plant damage, regrowth

**Introduction**

Plants have evolved several different strategies to avoid and recover from herbivore-induced damage. These plant defense strategies are classified into two key
categories: tolerance, wherein the impact of herbivory on plant fitness is reduced, and resistance, wherein the absolute amount of damage caused by an herbivore is minimized (Crawley 1989; Nunez-Farfan et al. 2007). Understanding the roles of tolerance and resistance is especially significant in the context of invasive plants when insect herbivory is used for biocontrol. Biocontrol programs are often considered to be the most ecologically sound and effective approach to check the spread of invasive species (Thomas & Willis 1998, Delfosse 2005). Exotic plants generally have low resistance to insect herbivory; however this vulnerability to herbivory is often counterbalanced by high tolerance potential (Bossdorf et al. 2004; Ashton & Lerdau 2008). Empirical and theoretical studies on plant tolerance are fewer when compared to studies focusing on resistance, thus making it a poorly understood ecological phenomenon (Tiffin 2000).

Tiffin (2000) highlighted the need to move plant herbivory research from studies that focus largely on evolutionary trade-offs and selection pressure on tolerance and resistance traits to future studies that emphasize the actual mechanisms of tolerance and resistance. Identification of putative mechanisms of tolerance, or plant traits associated with tolerance, will bring substantial clarity to our understanding of plant-herbivore interactions, and thus facilitate the design of appropriate experiments to test and ascertain the role of specific tolerance traits. Controlling the spread and establishment of invasive plants through biocontrol programs requires a thorough understanding of tolerance mechanisms in plants, including plant responses to herbivory and the effects of these responses on herbivore dynamics.

In perennials, storage of resources belowground is considered to be an effective defense strategy against aboveground herbivory (Iwasa & Kubo 1997; Strauss & Agrawal
1999; de Jong & van der Meijden 2000; Tiffin 2000). Models and empirical studies have shown that the use of belowground stored reserves for compensatory post-damage aboveground regrowth is a key tolerance trait that allows perennial plants to persist despite defoliation (Iwasa & Kubo 1997; van der Meijden et al. 2000; de Jong & van der Meijden 2000; Bossdorf et al. 2004; Paula & Ojeda 2009). Empirical studies of invasive plants have pointed to herbivore-induced increases in allocation to belowground stored reserves as the explanation for their high tolerance (Schat & Blossey 2005; Wang et al. 2011; Maguire et al. 2011). In addition to the allocation of resources to belowground storage, two additional plant traits affect tolerance: spring depletion of stored reserves to produce new aboveground structures (shoots, leaves) and post-herbivory aboveground regrowth ability, again using stored reserves. Increased allocation to belowground storage is often accompanied by rapid regrowth after herbivory (van der Meijden et al. 2000; Maguire et al. 2011), resulting in high tolerance.

In this paper, we develop an exploratory model that simulates the effect of aboveground insect herbivory on a temperate perennial plant with long-term stored reserves belowground. The model is based on the biology and life-cycle of an invasive wetland perennial, purple loosestrife (*Lythrum salicaria* L.) and its biocontrol agent, a leaf-feeding insect herbivore (*Galerucella* sp.). Like most temperate perennials, loosestrife shows strong seasonality in its life cycle, wherein the active aboveground growth occurs in spring from stored reserves in the belowground biomass rich rootstock. Because it is the part of the plant that persists from year to year, the amount of stored reserves in the overwintering rootstock is critical in determining the long-term survival of the plant. Since successful biocontrol requires the persistence of herbivores at
sufficiently high densities, our model includes local herbivore population dynamics. This makes our work distinct from earlier models that have explored the role of belowground storage as a plant defense strategy, which considered aboveground biomass removal to be an unpredictable, exogenous event (Iwasa & Kubo 1997, de Jong & van der Meijden 2000). Finally, we discuss the significance of our model’s results and findings in the context of the biological control of invasive perennial plants with long-term belowground storage.

Methods

The Plant-Herbivore System

Purple loosestrife (Lythrum salicaria L.) is a well-studied emergent herbaceous perennial plant that invades disturbed wetlands. Among the several key adaptive features that make loosestrife a hard-to-control weed is its ability to store substantial amounts of biomass belowground. It is well recognized that once established, loosestrife can persist for several years with annual sprouting of shoots every spring (Schatt & Blossey 2005). The introduced biocontrol beetles (Galerucella pucilla and G. calmariensis) emerge as adults soon after spring sprouting of aboveground leaves and actively feed on the meristematic tissue followed by oviposition, deposition of eggs and a larval stage. The larvae are the principal consumers of aboveground biomass and feed primarily on young leaves (Malecki et al. 1993; Blossey 1995). Larvae eventually leave the plant to pupate in litter below the plant, where they also overwinter (Malecki et al. 1993). Thus, during all stages in the life cycle, the insect herbivores are dependent on loosestrife. Towards the end of the growing season, in late fall, loosestrife dies back leaving behind a large
rootstock that overwinters and initiates the next cycle of aboveground growth in spring. It is important to note that from a temporal perspective, the active feeding by biocontrol beetles occurs for a brief period in spring after spring aboveground growth (Katovich et al. 2008). Defoliation by beetle larvae and adults ends by late spring, which provides loosestrife an ample amount of time before fall for regrowth and thus recovery of aboveground biomass lost to herbivory (Malecki et al. 1993). This ability to recover lost aboveground biomass may serve as a critical factor in modulating the eventual end-of-season biomass allocation to belowground storage, and thus the long-term persistence of loosestrife in the presence of these herbivores.

Short-term studies have shown that artificial removal of aboveground vegetation or herbivory by insects induces rapid regrowth with an increase in the number of stems and shoots, thus changing the overall plant architecture (Steffensen et al. 2001; Shatt & Blossey 2005). Very few studies have focused on biomass allocation by loosestrife under insect herbivory, especially over multiple seasons of growth (Katovich et al. 1998; Nötzold et al. 1998; Katovich et al. 1999). These studies have suggested that aboveground insect herbivory has a weak immediate effect on belowground biomass and it may take more than two to three years of consistent heavy defoliation to substantially deplete belowground biomass and cause plant mortality (Katovich et al. 1999). As previously stated, biomass allocation to belowground storage can interact with other key traits found in temperate perennials, like spring belowground resource utilization and post-herbivory regrowth capacity, and result in a range of potential outcomes. To gain a clearer understanding of these outcomes, we developed a simple mathematical model based on loosestrife and its specialist insect herbivores used in biocontrol programs. We
are particularly interested in understanding how belowground allocation to roots
determines the persistence of a perennial plant with a distinct growing season under
regular, seasonal herbivory.

The Model

We developed a population-level model with three state variables to explore the
dynamics of herbivory-induced regrowth and its effect on belowground stored biomass.
We model total plant biomass using two variables: aboveground biomass \( V_t \) and
belowground stored biomass \( S_t \). Herbivore population density \( H_t \) was the third state
variable. In our model, aboveground vegetation is entirely dependent on availability of
belowground storage. In other words, the plant dies when the belowground stored
biomass is exhausted.

At the beginning of each growing season, we assume that the plant initiates spring
growth by converting stored belowground biomass into aboveground biomass. We
assume that each unit of belowground biomass can yield at most \( y \) units of aboveground
biomass \( (0 < y \leq 1) \), but that the actual amount of aboveground biomass produced may be
less due to density dependence. After initial production, aboveground biomass grows
(using energy from photosynthesis) by a factor of \( r \) before the onset of herbivory. Using
the Beverton-Holt model for density dependence, aboveground biomass in year \( t \) before
herbivore feeding, \( V'_t \), is,

\[
V'_t = \frac{ryS_{t-1}}{1+bs_{t-1}}
\]

\((1)\)

where \( S_{t-1} \) is the amount of stored belowground biomass left at the end of the previous
growing season. The parameter \( b \) governs the strength of density dependence so that even
under exceedingly high amounts of belowground storage, the initial amount of
aboveground vegetation will saturate (at \( V' = ry/b \)) because we assume it is constrained by competition for other limiting resources (e.g. space, light). The amount of stored biomass in the same year is reduced due to the allocation to aboveground vegetation,

\[
S'_t = S_{t-1} - \frac{ys_{t-1}}{1+bs_{t-1}} \tag{2}
\]

where \( S'_t \) is the amount of stored biomass after spring growth of aboveground tissue.

Next, herbivores arrive and begin to feed. The amount of biomass removed by the herbivores is a saturating function of the ratio of herbivore population size to aboveground biomass (Abbott et al. 2008):

\[
d_t = \frac{h_t}{p + \frac{h_t}{V_t}} = \frac{H_t}{pV'_t + H_t} \tag{3}
\]

The constant \( p \) determines how steeply aboveground biomass loss increases with herbivore density, where low values of \( p \) imply that damage escalates quickly as herbivores increase from low density. In other words, the parameter \( p \) is a measure of mutual interference among herbivores, where high values of \( p \) mean high amount of interference among herbivores as they feed on plant tissue. Because \( d_t \) is the amount of herbivore damage in year \( t \), the amount of aboveground biomass remaining after herbivory, \( V''_t \), is

\[
V''_t = (1 - d_t)V'_t \tag{4}
\]

After herbivory, the plant has the potential to compensate for the aboveground biomass lost by rapid regrowth using energy from belowground stores. After a fraction \( d_t \) of aboveground biomass is lost, we assume a fraction \( cd_t \) of the remaining belowground stores can be converted to new aboveground biomass. The parameter \( c \) represents the
regrowth capacity of the plant. After post-harvest regrowth, the above and belowground biomasses in year $t$ are,

$$V_t = V''_t + cd_tS'_t$$

$$S''_t = S'_t - cd_tS'_t.$$  \hspace{1cm} (5)

Energy gained through photosynthesis by aboveground tissues may be used to replenish the belowground stores. We assume that a fraction of the aboveground biomass remaining after herbivory will eventually be converted into belowground biomass, where it will be stored over winter. This fraction can be up to $q$ ($0 \leq q < 1$), but will be lower due to density dependence if the belowground stores are already plentiful. We assume a linear decline in belowground allocation as $S''_t$ approaches its carrying capacity, $K_s$, and no allocation if $S''_t > K_s$. Given these considerations we have

$$S_t = S'_t + qV''_t \left(1 - \frac{S'_t}{K_s}\right)$$

where the last term is replaced by 0 if negative. Equation (7) gives the end-of-season belowground biomass that will be used to initiate spring growth (Equation (1)) the following season.

We model herbivore population dynamics by assuming that their within season survival, $h_t$, increases with per capita food availability (Abbott et al. 2008),

$$h_t = \frac{\frac{V_t}{H_t}}{k + \frac{V_t}{H_t}} = \frac{V_t}{kH_t + V_t}.$$  \hspace{1cm} (8)

Higher values of $k$ imply more intense food limitation and thus lower herbivore survival at a given level of food availability. Assuming that the surviving herbivore population grows at a net per capita rate $r_p$, herbivore dynamics are described by,

$$H_t = r_p h_{t-1} H_{t-1}.$$  \hspace{1cm} (9)
Model Simulations and Analyses

Model simulations tracking change in the amount of belowground biomass ($S_t$), aboveground biomass ($V_t$) and herbivore population size ($H_t$) were iterated for 200 years. Initial densities and parameter ranges considered during the simulations are given in Table 1. We used values that we believe are reasonable for loosestrife and similar systems, and we explore the sensitivity of our results to changes in the parameter values as described below. Nonetheless, since the values in Table 1 do not represent real-world estimates, note that we used our model output to gain a qualitative understanding of the role of tolerance traits in plant-herbivore dynamics; our results are not meant to represent quantitative predictions. Throughout our analysis, we focus on the effects of the three key tolerance parameters ($q$, $y$, and $c$). We identified and characterized the behavior of the plant and herbivore population as showing either cyclic dynamics or a stable equilibrium, depending on the parameter values. In order to ascertain when the dynamics were cyclic, we calculated the coefficient of variation of belowground biomass through the last 100 years for each parameter combination considered. We then identified those parameter combinations that yielded belowground biomass dynamics with a coefficient of variation greater than 0.1 as systems not at equilibrium and possibly exhibiting cycles. The presence of sustained herbivore cycles with no damping was confirmed by simulating the model for parameter combinations that produced a CV>0.1 and visually examining the dynamics. For parameters producing a stable equilibrium, we made a further distinction between dynamics that resulted in the persistence of the plant and herbivore populations and those that resulted in extinction. We also determined the effect
of each of the key parameters on the amount of aboveground and belowground biomass at the final time step for parameters resulting in a stable equilibrium.

To further understand the role of belowground allocation in determining persistence of plant populations, we study the model across different combinations of $q$, $y$ and $c$ that represent different plant growth and storage strategies. For example, slow growing plants with high belowground storage capacity would be modeled by parameter values representing low rates of conversion of root biomass to spring regrowth (i.e., low values for $y$) and high rates of root storage from aboveground biomass after herbivory (i.e., high values of $q$). Such plants may be viewed as poor competitors but may also be capable of long-term persistence under high defoliation. Similarly, in the presence of herbivores, plants with low $c$ (low regrowth capacity) and high $q$ represent plants that respond to herbivory by less aboveground regrowth to compensate for the lost biomass, but have the potential to efficiently convert the remaining undamaged aboveground biomass to belowground storage. For the different tolerance strategies produced by varying values of $q$, $y$ and $c$, we systematically varied the herbivore traits to mimic different types of herbivores by increasing either the effect of food limitation ($k$) or the effect of herbivore interference ($p$) in small increments of 10% across a range of values from 1.0 through 6.0, while keeping all other parameters fixed. We then determined the impact of different types of herbivores on the amount of above- and belowground biomass in the final time step.

In a final analysis, the effectiveness of the insect biocontrol agents in controlling loosestrife was considered by comparing aboveground or belowground biomass in the
absence of insect herbivores to biomass in the presence of herbivores. The measure of biocontrol effectiveness for aboveground biomass, $E_V$, was calculated as

$$E_V = \frac{V_{(-H)}}{V_{200}},$$  \hspace{1cm} (10)$$

where $V_{200}$ represents the amount of aboveground biomass at the 200th (final) time-step in the presence of biocontrol herbivores and $V_{(-H)}$ is the final aboveground biomass under identical conditions except in the absence of herbivores. Along the same lines, the measure of biocontrol effectiveness for belowground biomass was calculated as

$$E_S = \frac{S_{(-H)}}{S_{200}},$$  \hspace{1cm} (11)$$

where $S_{(-H)}$ and $S_{200}$ represent the amount of belowground biomass at the final time-step in the absence and presence, respectively, of herbivores. The goal of a biocontrol program is to reduce the plant’s biomass relative to its carrying capacity in the absence of herbivores. Therefore, parameter combinations with larger $E_V$ and $E_S$ result in more effective biocontrol.

**Results & Discussion**

From our initial conditions, model simulations generally showed a characteristic pattern wherein stored resources decreased rapidly with an initial increase in aboveground biomass and, after a certain amount of stored reserve depletion, both aboveground biomass and belowground biomass eventually reached a stable equilibrium (Figure 1a & b). In most cases, the herbivore population showed a rapid increase in the population with the initial short-lived increase in aboveground plant biomass, which was then followed by a crash in the herbivore population that eventually stabilized at a much lower population size, as both stored resource and aboveground biomass of the plant are
depleted (Figure 1a). In some of the cases where we observed CV of belowground biomass > 0.1, the populations exhibited long-period damped oscillations with a slow convergence to equilibrium (Figure 1b). For a small subset of parameter combinations with CV > 0.1 for belowground biomass, sustained cycling of both aboveground and belowground biomass components, along with the herbivore population, was produced (Figure 1c).

Sustained cycling was limited to combinations of high end-of-season allocation to storage ($q$) and spring storage utilization ($y$), and then only when spring plant growth rate ($r$) was also high and aboveground plant density dependence ($b$) was low (Figure 2). It is of interest that sustained cycles are associated with large values of $q$ and $y$, suggesting that plants escape herbivory by having traits that allow for efficient addition of biomass to below- and aboveground components. In contrast, the parameter representing regrowth capacity after herbivory ($c$) does not appear to influence the qualitative dynamics of the model, as it has no effect on whether there is cycling or a stable equilibrium. Sustained oscillations in our model always seem to appear as short-period $2^n$-point cycles (e.g. Figure 1c), which are typical of the discrete-logistic model, and are thus likely the result of density dependence in the belowground storage compartment in our model (Equation (7)), rather than a result of interactions with the herbivore. Oscillatory dynamics are well recognized in discrete-time single-species models like the discrete logistic and are due to the potential for strong overcompensation at high population growth rates (May 1976). The sustained cycles we observed were generally characterized by very high levels of aboveground biomass (as in Buckley et al. 2005), due to high accumulation of aboveground biomass in spring after initial sprouting from the rootstock (high $r$) coupled
with high allocation to storage (high $q$) and high subsequent conversion to spring aboveground biomass at sprouting (high $y$).

For lower values of $r$ and higher values of $b$, our model has a stable equilibrium (Figure 2), sometimes approached via long-period, damping oscillations (as in Figure 1b). Long-period cycles are characteristic of consumer-resource interactions (Murdoch et al. 2002) and thus likely result from interactions with the herbivore. In our study, unlike many consumer-resource models, these long-period oscillations converged on a stable end-point for all parameter combinations considered, perhaps due to the stabilizing influence of the Beverton-Holt density dependence we imposed on aboveground biomass (equation (1)) (Botsford 2012).

Extirpation of the plant population and subsequent extinction of the herbivore was observed when the plant was poor at allocating biomass to belowground stores (low $q$) or at using stored energy for spring growth (low $y$; Figure 2). Both the herbivore and the plant population persist as long as the parameter values are sufficient to sustain the plant population in the absence of herbivory. Post-herbivory regrowth capacity ($c$) generally did not affect extinction. From the perspective of biocontrol-based management of invasive plants, it is noteworthy that herbivory was never successful at driving a viable plant population to extinction. As we will discuss below, however, herbivory was quite effective in some cases at reducing plant biomass.

Among the three key parameters, large $q$ (allocation to belowground biomass) had a strong positive impact on final belowground biomass and a somewhat weaker positive impact on aboveground biomass (Figure 3a). An increase in $y$ (spring allocation of belowground stores to aboveground biomass) also increased final above- and
belowground biomass, but in contrast to $q$, the effect on aboveground biomass was the stronger of the two effects (Figure 3b). Unlike the other key parameters, an increase in $c$ (post-herbivory regrowth potential) had a strong negative impact on both above- and belowground biomass (Figure 3c). Moreover, higher values of $c$ showed a stronger negative effect on belowground biomass relative to aboveground biomass, suggesting that high regrowth capacity is a strain on belowground resources, when coupled with poor belowground storage allocation.

Increasing the degree of herbivore food-limitation ($k$) or mutual interference ($p$) resulted in an interesting overall pattern wherein both above- and belowground biomass increased identically, while the herbivore population showed sharply contrasting response patterns (Figure 4). More specifically, the direct effect of increasing $p$ or $k$ was a decrease or increase, respectively, in the herbivore population, and this invariably resulted in an increase in the amount of above- and belowground biomass. It is interesting that the ensuing increase in above- and belowground biomass with increase in $p$ or $k$ is identical, although it is unclear why this pattern arises.

As with previous studies, our model highlights the importance of investing in belowground storage (high $q$) for the persistence of a perennial plant population, when loss of aboveground biomass due to herbivory or other forms of aboveground damage is highly likely (Iwasa & Kubo 1997; de Jong & van der Meijden 2000). However, it is also evident that, in addition to allocation to storage, other closely associated traits, like spring utilization of stored reserves and regrowth capacity, have an additional affect on the biomass stored belowground and thus the ability to persist in the presence of herbivores.
Interactive effects of allocation to storage \((q)\) and spring utilization of stored reserves \((y)\) on belowground and aboveground biomass

In the presence of herbivores, given a moderate post-herbivory regrowth capacity \((c = 0.5)\), both aboveground and belowground biomass showed a similar pattern in response to increase in parameters \(y\) and \(q\) (Figures 5a-b). Low \(q\) values were always detrimental to the plants suggesting again that belowground allocation is a critical trait that strongly influences a perennial’s long-term persistence. However, a very low \(y\) value was also found to be detrimental to the plant as it produces little aboveground biomass. When the effects of \(q\) and \(y\) are considered together, it is apparent that low values of \(q\) and \(y\) result in poor to almost no detectable aboveground and belowground biomass, while an increase in both \(q\) and \(y\) results in a sharp increase in above- and belowground biomass. Overall, it is clear that belowground allocation in the presence of herbivores, along with adequate spring conversion of stored reserves to aboveground biomass, plays a critical role in determining the plant’s ability to persist.

Strengthening herbivore food-limitation and the degree of interference among herbivores, \(i.e.,\) increasing the values of \(p\) and \(k\), resulted in higher amounts of belowground biomass for the combination of high \(q\) and low \(y\), while for the low \(q\) and high \(y\) combination aboveground biomass was higher (Figures 4a & b). This is not all that surprising given that an increase in competition for food \((p)\) or interference among herbivores \((k)\) results in an increase in the amount of aboveground biomass that eventually is converted into belowground biomass. Studies have reported that persistent herbivory leads to remarkable changes in plant architecture and biology, wherein plant
size aboveground is significantly reduced with increased secondary and tertiary branching, flowering and fruit or seed production is either delayed or reduced, and most plant biomass is stored belowground (Schat & Blossey 2005). Therefore, as our model suggests, it is possible that in the presence of herbivores and persistent herbivory, selection for perennials with high allocation to belowground biomass (i.e., \( q \)) is likely.

**Interactive effects of allocation to storage (q) and regrowth capacity (c) on belowground and aboveground biomass**

In the presence of herbivores, given a high amount of spring growth (\( y = 1 \)), the interactive effects of \( q \) and \( c \) on above- and belowground biomass are slightly more complex (Figure 5c & d) than was observed for \( q \) and \( y \). Both above- and belowground biomass showed a rapid increase in response to increases in \( q \). Aboveground biomass increases with \( c \) when \( q \) is high, but this pattern is reversed for lower values of \( q \). In contrast, belowground biomass always decreases with \( c \). The effect of increasing \( p \) or \( k \) for two different combinations of \( q \) and \( c \) parameter values revealed that the amount of above- and belowground biomass was relatively higher for the high \( q \) and low \( c \) combination (Figure 4c & d).

Empirical studies have shown that resprouting after defoliation is not only dependent on the amount of stored reserves available, but it also depends on the time taken for the depleted reserves to be replenished by allocating biomass back to the roots from the resprouts (Schultz et al. 2009; Paula & Ojeda 2009). Studies have also shown that the efficiency of regrowth-driven replenishment of belowground storage is highly variable across species and environmental conditions (Olano et al. 2006). In a similar vein, inclusion of regrowth capacity in our model clearly articulates the need for efficient
replenishment of stored reserves after post-herbivory regrowth (i.e., high $q$ values), if a plant is to survive under heavy herbivore loads. In a short-term study by van der Miedjen et al. (2000), regrowth after a single episode of defoliation did not incur any cost on belowground storage, but it was speculated that multiple episodes of herbivory might yield a different outcome. Overall, perennial plants with high regrowth ability are less likely to persist if reallocation to belowground biomass is not sufficient to make up for the regrowth induced loss in stored reserves.

**Implications for Biocontrol of Invasive Perennial Plants & Model Limitations**

Our model shows that the effectiveness of biocontrol agents in maintaining low plant biomass depends on three key traits ($q$, $y$ & $c$; Figure 6). When plants have lower biomass allocation to roots ($q$) and low utilization of stored reserves during spring growth ($y$), herbivores are most effective at controlling below- and aboveground biomass. On the other hand, high post-herbivory regrowth capacity ($c$) often allowed more effective control of below- and aboveground biomass. Our model thus sheds light on the role of variation in three key life-history traits and the implications this has for the outcome of biocontrol programs, thus allowing field biologists and managers to identify problematic species that can be effectively managed by biocontrol herbivores. Studies have similarly shown that from a management perspective, success of biocontrol will depend on the variability in life-history strategy within a population (Strauss and Agrawal 1999; Muller-Scharer 2004; Dino-Garcia et al. 2003).

Empirical studies substantiate our model’s key finding that belowground storage plays a critical role in determining a perennial plant’s ability to tolerate herbivory. Studies have reported that in the native range with high densities of specialist insect
herbivores, perennial plants are able to persist by rapidly allocating high amounts of biomass to belowground storage (Maguire et al. 2011). In the case of invasive loosestrife, a recent study found that, in more stressful mesic habitats within the invaded domain, loosestrife exhibits a shift in life-history strategy by allocating more biomass to roots with a concomitant decrease in plant height (Moloney et al. 2010). Empirical work on loosestrife has also shown that compensatory regrowth of aboveground tissues following herbivory occurs at the expense of belowground storage, resulting in substantial reduction in both above- and belowground biomass with repeated seasonal herbivory (Nötzold et al. 2005).

Competition for light is crucial in determining loosestrife’s persistence, and the odds of biocontrol successfully extirpating a loosestrife population are higher when surrounded by native wetland vegetation that competes for light (Blossey et. al. 2001). Intense competition for light can also result in higher biomass allocation to aboveground parts or a phenological shift towards earlier spring sprouting from belowground storage (Bastlová and Květ 2002). It is beyond the scope of our current model to incorporate the effects of light competition or phenological shifts on loosestrife-herbivore dynamics. Studies have also reported that intense seasonal herbivory over time can have significant negative effect on aspects of loosestrife biology not included in our model, such as inflorescence size and seed production (Schat & Blossey 2005). Therefore, herbivory tolerance through biomass allocation to belowground storage and the ensuing reduction in aboveground biomass can have substantial negative impact on reproductive success, which may further slow down the spread of invasive perennials like loosestrife. Despite these deficiencies, our model does integrate the findings of empirical studies that focus
on the role of different potential tolerance traits in the success of invasive perennial plants
despite aboveground herbivory. We are confident that our model offers a fresh
perspective for future research on tolerance as a response to herbivory in perennial plants.

Acknowledgements
The first author is thankful to plant lab group members for assistance in improving the
model, especially from an empirical perspective. We thank Stan Harpole, Lisa Schulte-
Moore, John Nason and Luke Skinner for helpful feedback at earlier stages of this
project.

References
and inducible resistance on herbivore population dynamics *Theoretical
Population Ecology, 73*, 63-78
Aston, I.W. & Lerdau, M.T. (2008) Tolerance to herbivory, not resistance, may explain
differential success of invasive, naturalized and native North American vines.
*Diversity and Distributions* **14**, 169-78
Batalová, D. & Květ, J. (2002) Differences in dry weight partitioning and flowering
phenology between native and non-native plants of purple loosestrife (*Lythrum
Distribution, adult phenology, and oviposition. *Oikos, 74*, 225-34


Delfosse, E.S. (2005) Risk and ethics in biocontrol. Biological Control, 35, 319-329


de Jong, T. & van der Meijden, E. V.D. (2000) On the correlation between allocation to defense and regrowth in plants. Oikos, 88, 503-08


Rosenzweig, M. L. (1971) Paradox of enrichment: destabilization of exploitation 
ecosystems in ecological time. *Science* **171**: 385-87

biomass allocation and plant architecture in Purple Loosestrife. *Environmental 
Entomology*, **34**, 906-914


feeding effects of Galerucella spp. on Purple Loosestrife. *Proceeding of 17th N.A. 
Prairie Conference*, 32-34

response of tolerance in lines selected for glucosinate content. *Evolution*, **52**, 703-
712


and Evolution*, **13**, 325-29

*Evolutionary Ecology*, **14**: 523-36


Symbols, their definitions, and the value or range of values for state variables and parameters used in the model. Abbreviations: BG = belowground, AG = aboveground.

<table>
<thead>
<tr>
<th>State variables / Parameters</th>
<th>Assigned value or Range of values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial values of state variables:</strong></td>
<td></td>
</tr>
<tr>
<td>$H_0$ (initial herbivore population density)</td>
<td>20</td>
</tr>
<tr>
<td>$K_s$ (maximum BG storage)</td>
<td>50 – 150</td>
</tr>
<tr>
<td>$S_0$ (initial belowground biomass)</td>
<td>100</td>
</tr>
<tr>
<td>$V_0$ (initial aboveground biomass)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Plant related parameters:</strong></td>
<td></td>
</tr>
<tr>
<td>$b$ (density-dependence in spring conversion of BG to AG biomass)</td>
<td>0.005 – 0.5</td>
</tr>
<tr>
<td>$c$ (capacity for AG regrowth from BG stores after herbivory)</td>
<td>0.1 – 1.0</td>
</tr>
<tr>
<td>$q$ (fraction of AG biomass converted to BG storage in fall)</td>
<td>0.1 – 1.0</td>
</tr>
<tr>
<td>$r$ (spring AG growth rate from photosynthesis)</td>
<td>2 – 7</td>
</tr>
<tr>
<td>$y$ (maximum fraction of BG biomass converted to AG biomass in spring)</td>
<td>0.1 – 1.0</td>
</tr>
<tr>
<td><strong>Herbivore related parameters:</strong></td>
<td></td>
</tr>
<tr>
<td>$k$ (strength of food limitation for herbivores)</td>
<td>1.0 – 6.0</td>
</tr>
<tr>
<td>$p$ (degree of herbivore interference)</td>
<td>1.0 – 6.0</td>
</tr>
<tr>
<td>$r_p$ (herbivore population growth rate)</td>
<td>1.0 – 3.5</td>
</tr>
</tbody>
</table>
Figure 1. Model simulations showing inter-annual dynamics of above- and belowground biomass along with herbivore population density for 200 years, highlighting examples of (a) rapid convergence to a stable equilibrium ($y = 0.6$, $c = 0.5$, $q = 0.8$, $k = p = 1$, $r = 6$, $r_p = 2.5$, $b = 0.05$, $K_s = 100$), (b) damped oscillations leading to a stable equilibrium ($y = 0.8$, $c = 0.9$, $q = 0.9$, $k = p = 1$, $r = 3$, $r_p = 2.25$, $b = 0.001$, $K_s = 100$), and (c) sustained oscillations of above- and belowground biomass and herbivore population density ($y = 0.8$, $c = 0.5$, $q = 0.8$, $k = 1$, $p = 4$, $r = 6$, $r_p = 2.5$, $b = 0.001$, $K_s = 100$). Initial conditions are as shown in Table 1.
Figure 2. Model behavior highlighting parameter combinations resulting in sustained cycles of the plant and herbivore populations (denoted by the label ‘cycles’, see methods for details on how cycling is defined and identified), persistence at a stable equilibrium (denoted by the label ‘stable coexistence’), or extinction of both species (denoted by the label ‘extinction’). Extinction occurs only at parameter values that are insufficient to
support a viable plant population, regardless of herbivory. The horizontal and vertical axes of each plot represent the range of values considered for each parameter in the model simulation as described in Table 1. The vertical axes of plots in the third row show two parameters because the patterns were identical for each of the parameters represented by the vertical axis. Parameters not shown on the axes are held constant at $r = 6$, $y = 0.9$, $p = 3$, $k = 3$, $c = 0.5$, $q = 0.9$, $r_p = 2.5$, $K_s = 50$, $b = 0.005$.

Figure 3. Aboveground (black solid diamonds) and belowground (grey solid squares) biomass at the final time step ($200^{th}$ year). Only parameter combinations resulting in a stable equilibrium are used here. Biomass is plotted against (a) fall allocation to belowground storage, $q$; (b) spring utilization of stored reserves, $y$; and (c) post-herbivory regrowth capacity, $c$. All other parameters were held constant at: $r = 6$, $p = 1$, $k = 1$, $c = 0.5$, $q = 0.5$, $y = 0.5$, $r_p = 2.0$, $K_s = 50$, $b = 0.05$. 
Figure 4. Effects of parameters $p$ and $k$ on above- and belowground biomass. Increasing $p$ and $k$ shows an identical effect on the amount of belowground biomass ($+$) and aboveground biomass ($\times$) for different parameter combinations of $q$ with $y$ (a & b) and $c$ (c & d). Unlike biomass, herbivore population size shows a distinct response pattern to an increase in parameter $p$ (●) and parameter $k$ (●) for different parameter combinations of $q$ with $y$ (a & b) and $c$ (c & d). All other parameters were held constant at: $y = 1$, $c = 0.5$, $k = 1$, $p = 1$, $r = 6$, $r_p = 2.2$, $K_s = 50$, $b = 0.05$. 
Figure 5. 3D plots showing the change in aboveground (left column) and belowground biomass (right column) in the presence of herbivores under different combinations of $q$ (allocation to belowground storage in fall), $y$ (conversion of belowground storage to spring aboveground biomass) and $c$ (regrowth capacity after herbivory). All other parameters that were held constant at: $y = 1$, $c = 0.5$, $k = 1$, $p = 1$, $r = 6$, $r_p = 2.25$, $K_s = 100$, $b = 0.05$. 
Figure 6. Log of effectiveness of biocontrol of aboveground ($E_V$, solid black diamonds) and belowground ($E_S$, solid gray squares) biomass plotted against increasing levels of parameters $q$, $y$, & $c$. We only show results for parameter values that allow the plant population to persist in the absence of herbivory. Effectiveness was calculated as in equations (10) & (11). The larger the effectiveness, the stronger the impact of the biocontrol program. All other parameters were held constant at: $r = 6$, $p = 1$, $k = 1$, $c = 0.5$, $q = 0.5$, $y = 0.5$, $r_p = 2.0$, $K_s = 50$, $b = 0.05$. 
CHAPTER 5. GENERAL CONCLUSIONS

The distribution of species in space is a fundamental piece of ecological information that forms the basis of conservation and management decisions. However, distilling useful ecological information from species distribution patterns can be methodologically complex and challenging. In this respect, it is mandatory that analyses of spatial distribution patterns have a clear approach that is based on sound ecological principles. One such approach that has been steadily gaining popularity among ecologists is the hierarchical approach, wherein nested biotic and abiotic processes interact to yield the emergent pattern. Much of my dissertation research focused on hierarchically characterizing the distribution of invasive purple loosestrife, and then eventually developing a predictive model based on the nested processes involved.

In the second chapter, I explored the distribution of invasive purple loosestrife in a human-modified landscape. Moreover, by taking a hierarchical approach the distribution of purple loosestrife across the largely anthropogenic landscape was found to be strongly tied to three key nested factors and ecological processes. Loosestrife occurrence was not merely contingent on the availability of wetlands, and showed additional spatial discretion by preferring locations that are disturbed or have high degree of human land-use modification in the immediate vicinity. Finally, loosestrife distribution showed strong spatial dependency wherein loosestrife occurrence is more likely if there is a loosestrife invaded location in its proximity. Presence of such spatial dependency is a clear indication of the role of propagule pressure in structuring the spatial distribution of loosestrife. Taken together, the chapter suggests that distribution of an invasive plant is a
complex spatial process that involves multiple interactive factors that operate at different scales to yield the distribution pattern. More importantly, unlike earlier studies that have typically attributed occurrence of species across a landscape to landscape context, the study posits loosestrife within an ecologically robust framework where multiple nested processes are involved in determining the distribution. Scientists seeking to analyze spatial distribution patterns will gain immensely by considering such well defined ecological frameworks and, more importantly, identifying the ecological processes involved in a clear manner will be very helpful in developing well-informed and ecologically meaningful predictive models.

The third chapter of my thesis extends on the finding of the second chapter by developing an autologistic regression model to predict the occurrence of loosestrife. The overarching goal in this chapter was to develop invasion risk maps for loosestrife based on surrounding land-use information and propagule-pressure. The developed risk maps and predictions made by the models were then validated by an independent dataset, and the accuracy of the various models developed was determined. In the absence of propagule pressure, the surrounding land-use model predicted loosestrife occurrences with a modest degree of accuracy. However, the surrounding land-use model also suggested a strong relationship between human disturbance-induced edges and loosestrife invasion. In short, the surrounding land-use model showed that plant invasion is essentially an edge driven phenomenon wherein human modifications of habitat edges are the key spatial drivers of invasion within the focal habitat. This unique potential of surrounding land-use as an environmental variable has not been reported previously. And, in this respect, my research points to the utility of using surrounding land-use as an
environmental variable, particularly in the study of invasive and non-native plant distributions over landscapes. Adding propagule pressure as an autocovariate refined our model and was instrumental in making highly accurate predictions. The improved performance of autologistic models with a distance-constrained autocovariate showed loosestrife invasion is an ongoing, spatially-constrained process, where only a fraction of all suitable sites within the proximity of a loosestrife invaded location are vulnerable to invasion.

Overall, the findings of my research on loosestrife distribution highlighted that, like in the case of other invasive and exotic plant species, loosestrife invasion is largely driven by human-induced disturbances in the landscape. However, unlike earlier studies the potential ecological processes and factors behind the invasion of loosestrife is more clearly articulated by the hierarchical analyses and predictive autologistic model. In short, both analyses and the models capture spatial variability associated with surrounding land-use conditions and propagule-pressure. From a conservation and management perspective, the study shows that the most efficient approach to check the spread of loosestrife is to keep a watch on locations that are disturbed and located close to an invaded location. More specifically, the invasion hotspots are likely to have both high density of human affected disturbance activities and loosestrife presences. Identifying such regional hotspots is extremely useful in prioritizing and planning the conservation and management actions over landscapes.

In the fourth chapter, a simple mathematical model simulating herbivory tolerance in perennials with long-term belowground storage was developed. My model revealed that in perennials with a distinct growing season and under seasonal herbivore attack,
allocation to belowground storage is beneficial. However, concurrent factors like early-season utilization of stored reserves for growth and regrowth after herbivory can interact with belowground allocation to yield considerable variability in the amount of biomass stored at the end of each growing season. One of the key findings from this exploratory study was that biomass allocation to storage needs to be efficient under high intensities of herbivory for the long-term persistence of the perennial plant. The findings from this study are particularly significant when considered from the perspective of invasive perennials like purple loosestrife and its biocontrol insect herbivores. In the case of invasive perennials, the ability to allocate biomass to storage is an effective tolerance strategy in response to intense periodic herbivory, which makes biocontrol management a poor approach to extirpate a noxious plant population if it is species capable of allocating to belowground storage efficiently.