This is to certify that the
dissertation entitled

DEVELOPING DEMOGRAPHIC MODELS TO INFORM SELECTION OF *ALLIARIA PETIOLATA* (GARLIC MUSTARD) BIOLOGICAL CONTROL AGENTS

presented by

Jeffrey A. Evans

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Entomology and Ecology, Evolutionary Biology, and Behavior

Major Professor's Signature

8/7/2009

Date

DEVELOPING DEMOGRAPHIC MODELS TO INFORM SELECTION OF *ALLIARIA PETIOLATA* (GARLIC MUSTARD) BIOLOGICAL CONTROL AGENTS

By

Jeffrey A. Evans

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Entomology
and
Ecology, Evolutionary Biology and Behavior

2009

ABSTRACT

DEVELOPING DEMOGRAPHIC MODELS TO INFORM SELECTION OF *ALLIARIA PETIOLATA* (GARLIC MUSTARD) BIOLOGICAL CONTROL AGENTS

By

Jeffrey Adam Evans

Biological control is often considered a safe and effective method for controlling invasive plant species. While methods are available for predicting biocontrol agent host specificity, biocontrol practitioners currently lack effective tools for predicting agent efficacy. Demographic models which account for spatial and temporal variation in population dynamics promise to improve the predictability of weed biological control programs, while lowering the risks they pose to non-target species. *Alliaria petiolata* (garlic mustard) is an obligate biennial forb that is invasive in North American forests. I analyzed sources of demographic variation in twelve unmanaged *A. petiolata* populations in Michigan and Illinois, USA, and over three plant generations. These data were used to parameterize matrix population models of *A. petiolata* population dynamics, analyze *A. petiolata* responses to simulated management, and inform the selection of effective biological control agents for potential release in North America. Hierarchical, generalized linear mixed models (GLMMs) were used to analyze the spatial and temporal structure of variability in each demographic transition. The degree of variation observed in *A. petiolata* demographic rates was greater than expected based on previous studies of this species. This variation was highly structured in space and time and exhibited negative density dependence and positive response to precipitation across most of the life cycle. Estimates of the population growth rate (λ) ranged from 0.48 to 5.88 across all sites and years. Within sites λ was temporally variable, ranging from 0.80 to 5.88 within one site.

A megamatrix model was used to summarize variation in growth within sites. Site growth rates (λ_M) ranged from 0.83 to 3.54 (mean = 1.90). Sensitivity and elasticity analyses of matrix population models indicated the importance of the seed bank to *A. petiolata*'s success. Sensitivity and elasticity rankings varied with λ, indicating that the transitions with the largest impacts on population growth differ for growing and declining populations, and within populations during good and bad years, rendering management options a moving target. Rosette survival (summer and winter) consistently emerged as the transition with the greatest effects on λ in populations with positive growth, as did germination of new seeds and transitions affecting fecundity. This result is consistent with past predictions that rosettes should be targeted by management. The model raises the caveat that rosette survival is only a an effective target when growth is positive; its proportional effect on λ decreases as λ decreases. These models predict a lower probability of suppressing *A. petiolata* with biocontrol than past studies. The simulations predict that *Ceutorhynchus scrobicollis* could control up to 5 of 12 populations if introduced alone. Introducing a second species could extend control 9 populations, although the probability of success is < 0.1 at 4 of these 9. Better data on the distribution of agent impacts are necessary to refine predictions. Variance in survival was negatively density dependent, even when mean survival was not. modeling residual variance in each vital rate as a function of density, demographic variance and stochasticity themselves become density dependent functions. This has potentially important consequences for populations of management concern, as small populations may become more susceptible to local extinctions. The predictive power of future weed management models may be improved by incorporating density dependent demographic stochasticity in their designs.

To my mother and father

and to Courtney

ACKNOWLEDGEMENTS

I owe my gratitude to many great individuals and groups. I want to thank my two co-advisors who have guided me through the fray: thanks to Doug Landis, who brought me to Michigan, encouraged me, and kept me on track for the last six years. Thanks equally to Doug Schemske, who has taught me to always think big. Their passion for the wild things in this world is an inspiration. I look forward to working with each of them for a long time to come. I would not have made it to this point without the guidance and friendship of my collaborator Adam Davis.

Thanks especially to my family. To my partner and wife, Courtney, whose belief in me and in us has been my only true need. To my parents who remember how hard first grade was for me but encouraged me at every step. And to my sister, Jocelyn, who just called me "nature boy".

This work could not have been done without the help of many people: Thank you Anna Fiedler, Grace Chen, Megan Woltz, Rob Ahern, and my other friends in the Landis and Schemske labs for your help in the field and for listening to my crazy ideas. Huge thanks are due to Ashok Ragavendran, my friend and late night numbers guru who has probably sweat as much as I have over some of these analyses. Thanks to S. Raghu for the hard earned data presented here and for more than one critical read of these chapters. Thanks to Shaun Langley for creating the map in Figure 2.1. Thank you to Tara Lehman, Molly Murphy, Rob Johnson, Marc Wegener, Susan Post, Jordan Shelley, Erin Haramoto, Brian Schutte, and the many others who assisted with field and lab work. The

Nature Conservancy, The Edward Lowe Foundation, Michigan Department of Natural Resources, Kent County Parks, Ottawa County Parks, the Shiawassee YMCA Camp, and the Gasinski and Bartelt families all provided access to field sites. Thank you to Mark McPeek, Rebecca Irwin, and the Dartmouth College Biology Department for providing a temporary home away from home for me during the last year of my graduate program. This work was supported by the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service (CSREES), grant # 2005-35320-15312, USDA Specific Cooperative Agreement 58-3611-6-109, and by the Michigan State University Plant Sciences Fellowship, MSU CNS Dissertation Completion Fellowship, and EPA Science to Achieve Results (STAR) Graduate Fellowship FP-91650101.

PREFACE

Invasive species constitute a significant threat to native biodiversity globally, and managing them effectively is an increasing challenge. In some cases invasions are so devastating that the only way to confront them is by literally clearing the earth and starting over. During my first year of graduate school I visited Everglades National Park and saw just that. Brazilian pepper tree (*Schinus terebinthifolius*) had invaded vast tracts of fallow farmland that had been incorporated into the park in the 1970s. As I drove into the area, Brazilian pepper was practically the only plant visible – an impenetrable thicket of waxy leaves dotted with bright red berries. The soil that supported them, I was told, had been created to promote agriculture. Using heavy equipment, the limestone bedrock had been macerated to a consistency fine enough that crops could be planted in it. Because the soil had been so radically altered, when the fields were abandoned, they were rapidly invaded by Brazilian pepper, which was able to out compete the marl prairie wetland community native to the area. Three decades later, huge earth moving machines were being used to scrape away the topsoil that had been disturbed during a century of agricultural land use, revealing the limestone bedrock beneath. Where this had been done rich, native vegetation returned to its natural place of prominence within months, and the Everglades were once again a river of grass.

Sometimes the best way to control invaders and protect native biodiversity is to scour or scorch the earth. But this is rarely the case. Because plant invaders are more often diffuse targets of management, standing intermixed with the very species we wish

to preserve, a cautious, more delicate approach to controlling them is necessary. In some cases, when an invasion is small, hand pulling or spot herbicide treatments can effectively control a target weed, and quarantines can be used to stop its spread. But when the target is widespread and well established a more autonomous approach is necessary.

Enter biological control. Biological control, or biocontrol, is conceptually simple if not elegant: Everyone has enemies. Use this to your advantage by letting your enemy's enemy do the dirty work. More technically, biological control is the practice of using the natural enemies of a targeted pest species to suppress the pest population. In the case of weedy invasive plants, herbivorous insects and occasionally plant pathogens are most frequently used as biocontrol agents. Although past biocontrol programs have been criticized for imprudently releasing agents that have harmed non-target - and even threatened – species, the culture of biological control practitioners has changed. Great lengths are now taken to reduce the risk of introducing biocontrol agents that feed on non-target plant species. Extensive pre-release tests of agent feeding preferences and host specificity typically take a decade or more to complete, and these still do not guarantee that permits will be issued to make releases.

But what of the plant? Until recently, the target plant has been something of an afterthought in the formulation of new biocontrol programs. Biocontrol systems have historically been 'designed' (and I use the term rather loosely) by seeking to introduce as many host specific biocontrol agent species as can be found until the target is suppressed. This approach has resulted in a enormous investments of time and money in the development, testing, and of release of biocontrol agents that ultimately have not

controlled their targets, and that in some cases had wildly unexpected indirect effects on other species and food webs. There must be a better way...

A more directed approach to finding 'the right agent', the one that ultimately does the heavy lifting and suppresses the target, was proposed by Peter McEvoy and Eric Coombs. The idea, again, is audaciously simple: Study the plant, determine its vulnerabilities, and target them explicitly. Drawing from a growing number of studies on invasive weed demography and conservation biology, they proposed that population models could be developed for a target invasive species that would highlight the stage or stages in its life history with the greatest effects on population growth. This stage is the plant's 'Achilles' heel', and is where management should be directed. Knowing this information up front, finding 'the right agent' would begin to look less like a lottery and much more like a directed search. Demographic models, layered with the already mandatory process of testing agent host specificity, could lead biocontrol researchers more efficiently towards effective biocontrol solutions for invasive weeds. Perhaps more importantly, by releasing fewer species of biocontrol agents, the risks to non-target species posed by biological control programs could simultaneously be reduced.

Adam Davis was the first to apply McEvoy and Coombs' proposal. Davis and his coauthors used available data to build a mathematical model of the invasive weed *Alliaria petiolata* (garlic mustard) and characterize its vulnerabilities. This model was then used to make recommendations about how garlic mustard should be managed. My dissertation work builds from Davis' model with new studies and analyses of this same species, garlic mustard. It examines the structure and scale of spatiotemporal variability

in demographic rates and how this variability affects projections of population viability and the effectiveness of proposed management strategies.

One of the major themes in my dissertation research is variability. A number of researchers have studied garlic mustard, and a number of these have developed models of its population dynamics. Looking across these studies and at my own research, what is really amazing is how incredibly different this one species can be. In some places 10% of the seedlings survive. In others, 90% survive. In some places the population grows explosively, while in others it barely persists from year to year. Within populations there is also incredible variation in survival and growth from year to year as populations cycle through phases of boom and bust dynamics. In one of my garlic mustard study populations a year of tremendous growth was followed by two years of population decline. All of this suggests that to truly understand this plant, we can't think of its population biology as a static entity. It does not have one growth rate, nor does it have one density dependent function. It is unlikely, too, that it has one optimal management target. This picture only begins to come into focus after watching a dozen populations for several years.

In the following chapters I will present my dissertation research on garlic mustard. The central questions I address The first chapter contains background information on garlic mustard biology, biological control, and some of the goals of my research. Chapter two describes the statistical analysis of garlic mustard survival and reproductive rates from the twelve study sites. In this chapter, I examine the spatial and temporal scales of variability in demographic rates as well as their relationships to

population density and abiotic factors. Chapter three is a detailed critique of a recently published modeling study of garlic mustard population dynamics by Pardini et al. which also makes management recommendations. This is a necessary detour, as there were a number of errors in this model's construction and parameterization that led to issuing incorrect management recommendations which some land managers have already begun implementing. The fourth chapter presents a linear model of garlic mustard population dynamics which incorporates annual variation in demographic rates within each site. The model is used to estimate the probability of controlling garlic mustard by simulating the effects of either single or multiple biocontrol agent species across a range of agent efficacies. This is the main result of the study. Finally, chapter five contains a reanalysis of the survival and fecundity data focused on density dependence of demographic stochasticity. This analysis suggests an important relationship between population density and the stochasticity of survival probability that should be explored further in future statistical and population models.

I didn't think I wanted to be a modeler when I started graduate school. Turns out... I do. Who knew?

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Brian (lecturing to large crowd): You're ALL individuals!

The Crowd: Yes! We're all individuals!

Brian: You're all different!

The Crowd: Yes, we ARE all different!

Man in crowd: I'm not...

Life of Brian

Monty Python, 1979

CHAPTER 1: DEVELOPING DEMOGRAPHIC MODELS TO INFORM SELECTION OF *ALLIARIA PETIOLATA* (GARLIC MUSTARD) BIOLOGICAL CONTROL AGENTS

INTRODUCTION

Biological invasions are a major contemporary management problem. *Alliaria petiolata* (garlic mustard (M. Bieb.) Cavara and Grande) (Brassicaceae) is one of the most damaging invasive weeds in North American forests. A native of Eurasia, its substantial negative impacts on litter and nutrient cycling and its allelopathic and anti-mycorrhizal effects can disrupt native forbs and tree regeneration (Blossey 1999, Meekins and McCarthy 1999, Prati and Bossdorf 2004, Stinson et al. 2006). Conventional strategies have failed to yield effective long term control of any but the smallest *A. petiolata* infestations (Nuzzo 1991, 1994, 1996, Nuzzo et al. 1996), and a search for biological control agents was initiated in 1998 (Blossey et al. 2001b). Biological control is often considered an environmentally safe alternative to conventional management. However, a growing awareness of the potential risks to non-target species posed by some weed biocontrol agents (Louda et al. 1997, Callaway et al. 1999) has prompted calls for increased rigor in new biocontrol programs (Simberloff and Stiling 1996, McEvoy and Coombs 2000, Louda et al. 2003a) with emphasis placed on *a priori* selection of host specific agents that have strong, negative impacts on their intended target plants (McEvoy and Coombs 2000, Pearson and Callaway 2003, Pearson and Callaway 2004, Thomas et al. 2004). Matrix population models of target plants and demographic analysis can improve both the safety and efficacy of weed biocontrol (Shea and Kelly 1998, McEvoy and Coombs 1999, Rees and Hill 2001). Using demographic models to identify plant life-stages that are most likely to affect a population's growth rate if damaged, agent selection and testing can be restricted to species which affect those optimal target life-stages (McEvoy and Coombs 1999). Additionally, models can suggest

whether single or multiple agents will be necessary to achieve suppression across the target's geographic range (Parker 2000). A preliminary model of *A. petiolata* constructed from published data (Davis et al. 2006) suggests that overwintering rosettes should be targeted initially. However, a more robust, spatially explicit data set is required to address questions about the range of conditions under which single or multiple agent biological control are projected to be successful. Developing safe, effective, and economical weed control strategies will require a combination of new and established empirical and theoretical ecologically based approaches. This has been the focus of my doctoral studies.

BIOLOGICAL CONTROL

> "Biological control is the use of parasitoids, predators, pathogens, antagonists, or competitor populations to suppress a pest population, making it less abundant and thus less damaging than it would otherwise be." (Van Driesche and Bellows 1996)

The use of natural enemies to control targeted pest species dates back millennia, but the use of scientifically rigorous testing to select biological control agents has a relatively shorter history (Van Driesche and Bellows 1996). Early biological control efforts were trial and error based and principally designed by intuition. However, as awareness and concern over impacts to non-target organisms grew during the 20th century, agent selection became a more focused process and began to address some of these concerns.

Host Range Testing

Weed biological control in the United States has been regulated since 1957 by the Technical Advisory Group for Biological Control Agents of Weeds (TAG), a multi-agency federal panel headed by USDA-APHIS whose purpose is to advise on the practice of weed biological control "based on consideration of potential non-target impacts and conflicts of interest" (USDA-APHIS 2006). The goal of agent testing is to predict which non-target plant species an herbivore is likely to attack if released in a new environment and is described by Van Driesche and Bellows (1996). The host specificity of candidate biocontrol agents is evaluated using a centrifugal phylogenetic approach (Wapshere 1974, 1989). A list of test plants is assembled which includes close relatives of the target, species likely to co-occur with it, and species of economic importance. There are three stages of testing which are used to discern the agent's physiological host range (i.e. what it is capable of feeding and developing on in a contrived, no-choice situation) and its ecological host range (i.e. what it is likely to feed and develop on under natural field conditions) (Louda et al. 2005a): (1) in larval feeding trials, larvae of candidate species are offered test plants tests under no choice conditions in confinement. Rejection by the larva indicates that it is outside its host range. If it feeds on the plant, (2) a no-choice adult oviposition trial is conducted in confinement. If eggs are not laid on the plant, it is deemed not a host. If eggs are laid, (3) a multiple choice oviposition test is conducted in field cages or in the open field. If eggs are not laid on the plant, it is not a potential host. If they are laid under these more natural conditions, the non-target plant is within the agent's host range, and the agent may be rejected if there is reason to protect the non-target plant from harm (Wapshere 1989 in Van Driesche and Bellows 1996).

Benefits of Biological Weed Control

Biological control of weeds has several potential advantages over conventional control methods. Arguments in support of biocontrol include its (1) potential effectiveness, (2) low resource input requirements, (3) self-perpetuation, (4) low output of pollutants, (5) reduced non-target impacts, and (6) overall compatibility with alternate management strategies (McEvoy and Coombs 2000). Cost to benefit ratios for successful biological control programs can be lower than 1:145 by some estimates (Hoddle 2004a) and as low as 1:12,698 by others (Huffaker et al. (1976) in Gutierrez et al. 1999) with gains increasing over time. The successful biological control of St. Johnswort (*Hypericum perforatum* L.) (Louda et al. 1997, Whitten and Hoy 1999) and purple loosestrife (*Lythrum salicaria* L.) (Blossey et al. 2001a, Landis et al. 2003) in North America and prickly pear (*Opuntia spp.*) in Australia and the Caribbean (Bellows 1999) are frequently cited as evidence of weed biological control's enormous potential.

Pitfalls of Biological Weed Control: Non-Target Impacts

Non-target impacts of weed biological control programs have been documented in many systems (e.g. Howarth 1991, Louda et al. 1997, Stiling and Simberloff 2000, Louda and O'Brien 2002, Louda et al. 2003a, Pearson and Callaway 2003, Louda and Stiling 2004), and have inspired discussion about the role biological control should play in weed management (e.g. Simberloff and Stiling 1996, Pearson and Callaway 2003, Hoddle 2004b, Louda and Stiling 2004, Pearson and Callaway 2004, Thomas et al. 2004). The debate focuses on tradeoffs between potential gains from weed biological control

programs and losses incurred through direct and indirect non-target impacts and our ability (or inability) to accurately forecast the outcomes of these biological control programs. Decision makers evaluating whether to release new agents must strike a balance between these. Ultimately, we are forced to weigh the unknown consequences of introducing an organism into a novel environment against the consequences of either using conventional control methods or of doing nothing at all.

Direct Non-Target Effects

Most documented cases of direct non-target impacts of weed biological controls relate either to older weed-control efforts or reflect times when either host-specificity was not considered important or when certain non-target plants were within the host range of the proposed agent but were not considered valuable for conservation (e.g. Louda et al. 2005a). Such is the well-documented case of the weevil *Rhinocyllus conicus* (Coleoptera:Curculionidae) which was released to control invasive European thistles (*Carduus* spp.), despite knowledge of its feeding and development on North American *Cirsium* thistles from host specificity testing prior to release (Louda et al. 2003b, Rose et al. 2005). At the time of its release in North America, it was believed that *R. conicus*'s preference for *Carduus* spp. in host specificity trials would limit its impacts on native thistles (Louda et al. 1997). Additionally, most thistles, native or otherwise, were generally considered rangeland weeds without conservation value. Non-target feeding was reported almost immediately after the initial releases in 1969, and in 1993 the rare Platte thistle (*Cirsium canescens*) was identified as a preferred host of *R. conicus* (Louda et al. 1997). Further interstate re-distribution of *R. conicus* was prohibited effective in

2000 (Louda et al. 2003b). Preferential feeding by *R. conicus* on *C. canescens* has resulted in a decrease in *C. canescens* fecundity and population growth rates and is predicted to lead to global extinction of *C. canescens* (Rose et al. 2005). Demographic modeling indicates that the federally threatened *C. pitcheri*, a sister species of *C. canescens* will rapidly be driven extinct if *R. conicus* spreads into its habitat (Louda et al. 2005a).

The value of each weed biocontrol program varies regionally. *Cactoblastis cactorum*, the same biocontrol agent touted as being enormously successful in the control of *Opuntia* spp. in Australia and the Caribbean and the "poster child of biological control" dispersed naturally from the Caribbean into Florida in the United States in 1989 (Stiling 2002) where it now threatens to drive the endangered native Florida semaphore cactus *O. corallicola* extinct (Louda and Stiling 2004). Although it had been considered for intentional introduction into the United States on several occasions, introduction was rejected out of concern for native *Opuntia* and commercial prickly pear production in Mexico (Louda et al. 2003b). Because its introduction into North America was unintentional, some (Hoddle 2004b) suggest that the threat it poses in North America should not be considered a non-target effect of biological control. However, its deliberate introduction into Caribbean islands set up the possibility for natural spread to the mainland United States. This indicates shortsighted planning, as the potential for redistribution could have been anticipated and considered before introductions were made (Louda et al. 2003b).

Indirect Non-Target Effects

Indirect impacts on non-target species are more difficult to anticipate and quantify than direct non-target impacts, and there are relatively fewer examples in the literature. Conventional biological control theory, based on simple predator-prey interaction models, postulates that biological control agents impose negative impacts on their target hosts and thereby confer positive indirect effects on desirable native species. In turn, reductions in the target host plant negatively impact the control agent and impose regulation through negative feedbacks. This system is expected to work when control agents are both highly host-specific and exert strong negative pressures on target populations (Pearson and Callaway 2003).

The consequences of direct non-target feeding by biocontrol agents can be severe (Louda et al. 1997, Stiling 2002), but these risks can be quantified *a priori* (Pemberton 2000). Thus, agent selection processes have increasingly emphasized host specificity (McEvoy 1996). Host specific agents are presumed to have neutral effects on non-target organisms. This assumption is used to justify the release of multiple host-specific agents per target plant when the agents are believed to be host specific (Pearson and Callaway 2003). In this approach, agents are chosen without regard for their potential effectiveness, and a "lottery" of chance is established in the new environment to determine which of the agents will succeed in suppressing the target (McEvoy and Coombs 2000). As a result of this approach, insect species introduced as biological control agents are now more numerous than their invasive weed target species (McEvoy and Coombs 1999, McEvoy and Coombs 2000). Despite this, the large majority of weed biological control agents released do not control their target hosts (McEvoy and Coombs 1999, Denoth et al.

2002). A growing body of evidence suggests that indirect non-target effects can be mediated by host specific biocontrol agents though "ecological replacement, compensatory responses, and food-web subsidies" (Pearson and Callaway 2003). These undesirable indirect effects are much more likely to occur when the introduced control agents do become established in the introduced range but exert weak pressure on their targets and fail to control them (Cory and Myers 2000, Pearson and Callaway 2003).

Pearson and Callaway (2003) describe the case of two gall flies, *Urophora* spp., introduced in the 1970s to control the non-native, invasive knapweeds *Centaurea maculosa* and *C. diffusa*. The flies became established in North America and have remained highly host specific. However, they have not controled *Centaurea* spp. populations and subsequently have become highly abundant. Deer mice (*Peromyscus maniculatus*) are generalist predators and have increased their overwintering survival though feeding on the introduced gall fly larvae. *Peromyscus maniculatus*, in turn, has become more abundant in this system and negatively impacts native plants and insects though feeding, other small mammals through competition, and may increase predator abundances. Importantly, as a vector of Hanta virus, *P. maniculatus* may negatively impact human populations (Pearson and Callaway 2003).

IMPROVING WEED BIOLOGICAL CONTROL

The key to minimizing the risks of both direct and indirect non-target effects is to select only highly host specific, highly effective agents and release the minimum number of agent species necessary to suppress the target (Pearson and Callaway 2003). Techniques for testing host specificity are relatively well defined, but quantitative

approaches for predicting biocontrol agent efficacy have not yet been used in selection of agents prior to release. A flush of interest in developing new methods promises to improve the success rate of biological control and decrease the risk and frequency of non-target impacts (McEvoy and Coombs 1999, Briese 2006, Davis et al. 2006, Raghu and van Klinken 2006, van Klinken and Raghu 2006).

The Use of Transition Matrix Models

Matrix population models can be used to interpret a weed species' population dynamics and to guide the selection of new biological control agents (McEvoy and Coombs 1999, Raghu and van Klinken 2006). Caswell (2001) describes their properties, construction, and interpretation. In past studies, matrix models have been used to interpret the mechanisms by which biocontrol has succeeded (McEvoy and Coombs 1999), failed (Shea and Kelly 1998, Parker 2000), or had variable outcomes in different locations (Shea et al. 2005). Recently, attention has turned toward applying them to the development of new biocontrol programs. While the details of model construction will be presented in Chapter 4, I will briefly describe some of the ways the model can be used.

A matrix model is an algebraic expression whose components describe the mean probability of each individual in a population surviving from one life stage to the next and its reproductive output at each stage in discrete time. These probabilities correspond with transitions in the organism's life cycle. The transitions are arranged into a matrix, abbreviated **A** ('the **A** matrix') from which a number of useful population statistics can be calculated. The most pertinent statistics to population management are the population growth rate (λ, lambda), and the sensitivities and elasticities of lambda to the matrix

elements. The population growth rate, λ, is related to the intrinsic rate of increase r from continuous models as $\lambda = e^r$ or $r = \ln \lambda$ (Caswell 2001). When λ is equal to one, each individual exactly replaces itself during its lifetime. Values of λ greater or less than one indicate expanding or declining populations, respectively. The goal of biological control is to affect the survival probabilities of a weed such that λ is driven below one. By simulating the effects of biological control agents or other management on survival or reproductive rates, the model can be used to assess the efficacy that management of a particular stage in the plant's life cycle would need to achieve to reduce λ below 1.

The sensitivity of λ to the each demographic rate is the local slope of λ as evaluated at a particular value of the demographic rate. It indicates how much λ will change in absolute terms if the transition is perturbed. Elasticities are calculated from the sensitivities, but are scaled by the proportional contribution of each demographic transition to λ. The elasticity of each demographic rate is the proportional change in λ expected from a proportional perturbation of the demographic rate. It indicates how much λ will change in relative terms if the transition is perturbed. Transitions with large elasticities and large sensitivities are predicted to have large impacts on population growth and theoretically represent the optimal targets for management.

Elasticity analyses have been used previously to interpret the relative success of established biological control programs. McEvoy and Coombs (1999) used a *post hoc* elasticity analysis of a successful biocontrol program for tansy ragwort *Senecio jacobaea* L. (Asteraceae) in Oregon. They attributed successful control of the target to just one of the two agents released, deeming the other redundant and an unnecessary risk. Parker (2000) similarly modeled patterns of demographic variation in invasive Scotch Broom

Cytisus scoparius along North America's west coast. She showed significant spatial heterogeneity in *C. scoparius* demographic parameters, but elasticity analysis identified no obviously vulnerable life stage. Her model predicted that nearly 100% and 70% of seeds would have to be destroyed in prairie and urban populations, respectively, to reduce λ to less than one, which the introduced biocontrol agents were incapable of achieving. Shea and Kelly (1998) and Shea et al. (2005) used demographic models to interpret differential outcomes of biological control agents released against *Carduus nutans* in New Zealand and Australia. They found major life history differences between the two *C. nutans* populations that drove differences in their elasticity structures. Dynamics of New Zealand populations were driven by early life stages, whereas rosette longevity was more important in Australian populations of *C. nutans*. Of three agents considered in their analysis, none were able to control *C. nutans* alone in New Zealand, while in Australia two were predicted to reduce λ below one.

Most recently, Davis et al. (2006) used a matrix population model to make *a priori* predictions about which stages in the *A. petiolata* life cycle would be most susceptible to biological control agents and predicted the levels of mortality that must be induced by single or multiple hypothetical control agents at each life stage to reduce λ below one. Their model, parameterized with data from multiple published sources, makes specific recommendations about which of the potential biocontrol agents being considered for *A. petiolata* are most likely to be effective against North American populations. Davis et al. (2006) found that the greatest elasticity corresponded to the rosette to flowering transition followed by transitions affecting seed production. Incorporating the range of impacts the potential agent species have on *A. petiolata* in

laboratory studies, they concluded that multiple agent introductions would be necessary to extend control of *A. petiolata* to the greatest number of populations. The weevil *Ceutorhynchus scrobicollis*, which disrupts the rosette to flowering transition by mining in the overwintering rosettes and reduces fecundity by feeding on foliage, was recommended as a priority agent to be released first. It is predicted to control all but the most vigorous populations of *A. petiolata*. For some populations, another supporting agent that affects fecundity was recommended to be released with *C. scrobicollis* when one becomes available. Post-release monitoring will allow biocontrol managers to interpret which sites and conditions necessitate multiple agents. Taking this "plant first" approach allows biocontrol practitioners to understand the target's weaknesses and make informed agent release decisions. By eliminating ineffective agents from consideration, the risks of causing non-target impacts can be substantially reduced (McEvoy and Coombs 1999). In combination with rigorous host-specificity testing, this technique has the potential to transform the development of new biological control programs from the traditional "lottery" approach of agent selection (McEvoy and Coombs 2000) into a directed search for host-specific agents that affect particular life history stages or transitions.

Demographic Variability

One strategy used to select effective biocontrol agents is to choose and introduce agents which have been successful elsewhere (Harris 1991). In the case of *Cactoblastis cactorum*, redistribution from Australia to the Caribbean resulted in control of the target in the new location, although the non-target impacts of this choice were costly (Stiling

2002). However, replicating a strategy that had been successful against *Carduus nutans* in Australia failed in New Zealand (Shea et al. 2005). The relatively few demographic studies of plants across spatial or temporal gradients have found significant variability in demographic rates and statistics (e.g. Bierzychudek 1982, Bierzychudek 1999, Parker 2000). Horvitz and Schemske (1995) studied a tropical rainforest herb for five years in four locations (n = 16 **A** matrices) and found variability in λ, in the correlation structure of demographic parameters, and in the sensitivity and elasticity structure of the populations both in space and over time. The variation they observed in parameter sensitivity and demography was not always correlated with variation of other elements in the environment that the plant could profit from (e.g. presence/abundance of pollinators during peak flowering). The uncoupled relationship between plant and environmental or exogenous biotic factors means that variation in a species' demography does not guarantee changes in fitness. Rather, demographic variability presents fitness opportunities or hazards only in concert with conditions that permit realization of a change in fitness. This suggests that if the target species' demography and sensitivity structure vary spatiotemporally, what constitutes an optimal biocontrol strategy will be conditional on the form and scale of demographic variability and the degree of parallelism between plant and agent performance and demography across the range of conditions. Shea et al's (2005) findings reflect the importance of such variability for biological control.

Horvitz and Schemske's (1995) analysis took a factorial approach to characterizing and interpreting the significance of space (plot) and time (year) in their model. Another method developed by Horvitz and Schemske (1986), used by Pascarella

and Horvitz (1998), considers the environment itself as a demographic entity with properties that change over time, in which the study species has demographic rates that vary as a function of the environmental state. They studied a tropical understory shrub across a gradient of forest-canopy openness created by a hurricane and the response of the shrub to the closure of the canopy over time. The progressive closure of the canopy was characterized by one set of transition matrices, and the plant's performance under each canopy condition was characterized by another. Each patch could transition between any of seven canopy states, and each plant could transition between any of eight developmental stages. They nested each of the 8x8 plant demographic matrixes within each element of the 7x7 forest canopy matrices to create a 56x56 megamatrix which encapsulated both the environmental dynamics of the system and the organismal dynamics within each of the possible environmental states. Similar to Horvitz and Schemske (1995), Pascarella and Horvitz (1998) observed differences between the elasticity structures of their matrices when considered separately versus when considered together in their megamatrix. Individually, the matrixes suggested a stable population dominated by large individuals in closed canopy conditions, whereas the megamatrix indicated a rapidly growing population whose spread was dependent on the existence of open patches cleared by the hurricane.

Davis et al.'s (2006) model of *A. petiolata* biological control was a critical first step towards the incorporation of predictive ecological models into the decision making process of invasive species managers. Their approach made use of available data on a well studied target species and allowed rapid formulation of management recommendations. However, the data used to parameterize their model were collected

from disjunct natural and laboratory populations from Ontario to Kentucky over more than two decades. These were then pooled to generate ranges across which the demographic parameters were varied in the simulation analysis. This approach does not account for the structured correlations between parameters that are shown to have been important in other study systems. In cases where time is limited, generation of a demographic model from multiple, unrelated existing data sources to guide selection of effective agents is much preferred to the "lottery" approach of blindly releasing all available agents (McEvoy and Coombs 2000). However, when it is possible to do so, using data collected from distinct populations across the target plant's spatial range and over multiple generations or years will allow much more robust management conclusions to be reached.

STUDY SPECIES

Distribution

Alliaria petiolata is a frequent component of temperate forest understory and edge communities. It is native to Eurasia where it occurs from England east to Czechoslovakia and from Sweden and Germany south to Italy (Nuzzo 1993b, 2000). It has been redistributed into Central Asia, New Zealand (Bangerter 1985) and much of North America (Nuzzo 2000, Welk et al. 2002). *Alliaria petiolata* was first collected in North America on Long Island, New York, in 1868 (Nuzzo 1993a), where it was likely introduced by immigrants from the old world. In North America, *A. petiolata* is most abundant in New England and the Midwest with populations now present in at least 36 U.S. States and 4 Canadian Provinces (Nuzzo 2000, USDA-NRCS 2007) from the east

coast to Alaska (The Nature Conservancy 2002, Ellen Anderson, USDA Forest Service personal communication November, 2005). Climate based models of its potential North American distribution project further range expansion in the future (Welk et al. 2002, Peterson et al. 2003).

Life Cycle

Evans (2006) reviewed *A. petiolata*'s biology and life history. *Alliaria petiolata* is a disturbance adapted species which profits from anthropogenic and natural disturbances (Pyle 1995) and can tolerate harsh growing conditions such as lead contaminated soils (Pichtel et al. 2000). Optimal photosynthetic rates are achieved under light conditions typical of forest edges, although it can grow under conditions ranging from closed-canopy forest shade to full sunlight (Dhillion and Anderson 1999, Meekins and McCarthy 2000, Myers et al. 2005). In forest interiors *A. petiolata* often colonizes light gaps where trees have fallen or been removed (Luken et al. 1997).

North American *A. petiolata* populations have an obligate biennial life cycle (Cavers et al. 1979) which can be decomposed into three basic developmental stages: seeds, first year plants, and second year plants (henceforth "adults"). First year plants can be further separated into seedlings and rosettes. Seedlings emerge from early spring through early summer. Rosettes are distinguished from seedlings at some point during the early summer as those first year plants that survived the germination period early in the growing season and no longer bare cotyledons. Seeds and rosettes are the only stages present in autumn and winter.

Seeds of *A. petiolata* require cold stratification to germinate. Dormancy for the majority of seeds varies from one to two winters (Cavers et al. 1979, Roberts and Boddrell 1983), although dormant seeds remain viable for nine years or more (Nuzzo and Blossey, unpublished data, ongoing experiment). Germination begins and peaks in early spring (Baskin and Baskin 1992) under high light, low competition conditions prior to leaf-out of canopy trees and prior to germination of most native ground layer species (Myers and Anderson 2003). Seedlings form a low, tight canopy over the forest floor with population densities as high as 20,000/m^2 (Trimbur 1973). Seedlings which survive the summer overwinter as basal rosettes, bolt the following spring, and flower in early summer (Cavers et al. 1979). High seedling mortality results in only 5-9% of seedlings surviving to form rosettes and only 2-4% of rosettes survivors reaching reproductive age (Cavers et al. 1979). Nuzzo (1993c) estimated 21.4% winter rosette survival resulting in mean spring rosette densities of 39.9 (range = 4-102 rosettes/m^2) with 9% of the variance in overwintering surival attributable to fall rosette density. Mature adult plants reach heights up to 1.8 m (Evans 2006).

Flowers are primarily self pollinating but are visited by generalist pollinators including Diptera: Syrphidae, and Chironomidae or Ceratapogonidae (described only as "midges" by Cavers et al. 1979) and Hymenoptera: Apidae, Andrenidae, Halictidae (Cavers et al. 1979, Anderson et al. 1996, Cruden et al. 1996). *Alliaria petiolata* reproduces exclusively by seed (Cavers et al. 1979). Seed production is variable among individuals (Susko and Lovett-Doust 2000) and populations, with per capita fecundity ranging up to 7900 seeds (Nuzzo 1993b). In dense stands, seed production per square meter can exceed 100,000 (Cavers et al. 1979).

Several factors affect seed production in *A. petiolata*. Experiments by Susko and Lovett-Doust (1999) showed that removal of 50-100% of cauline leaves (leaves along the stem) from adult plants reduced seed production by 25-46%, and removal of 50-75% of the root mass of adult plants decreased seed production by 8-13% and reduced the proportion of seeds maturing by ≈ 4%. Position of flowers within inflorescences and plant size also affect seed production (Susko and Lovett-Doust 2000). In North America plants senesce following seed production, although European *A. petiolata* can perenniate by production of adventitious buds (Cavers et al. 1979). Dispersal of seeds is limited, with the majority falling near the parent plant. Long distance dispersal is facilitated by humans, deer, and mice that transport seeds in muddy footwear and hooves, in fur or clothing, and automobile tires (Nuzzo 1993a). Seeds have limited ability to float but can disperse along riparian corridors (Cavers et al. 1979, Nuzzo 1993a).

In newly established populations, first and second year plants are typically not intermixed within localized patches, creating an effective alternation of generations. Over time, the seed bank moderates this effect and first and second year plants are found in unevenly mixed patches. Seedlings that germinate under cover of second year plants have very high mortality which keeps the generations locally segregated in many areas (Winterer et al. 2005). In areas where it is invasive, *A. petiolata* spreads in a moving front as satellite populations ahead of the core establish and fill out, with a positive net rate of spread in Illinois averaging 5.4 m/y (Nuzzo 1999). A study of seven invaded forests in southern Michigan documented spatial expansion of *A. petiolata* in 100% of sites over a four year period (Evans 2006).

Invasiveness of *Alliaria petiolata*

Release from pests or other natural enemies has been proposed as contributing to the success of some invasive species (Williamson 1996). Damage from herbivore natural enemies and plant pathogens are frequently found on North American *A. petiolata*. However, they do not have significant impacts on its survival or reproduction, suggesting that natural enemy release may play a role in *A. petiolata*'s invasiveness (Evans 2006).

Allelopathy has also recently been identified as contributing to increased invasive ability in several plant species (Bais et al. 2003, Call and Nilsen 2003, Grant et al. 2003, Weston and Duke 2003, Wolfe and Klironomos 2005). Vaughn and Berhow (1999) extracted several phytotoxic substances from *A. petiolata* tissues that negatively impacted the growth of forbs and grasses. They proposed that these compounds or their derivatives might additionally inhibit the growth of arbuscular mycorrhizal fungi (AMF). Later laboratory and field studies showed that AMF growth, abundance, and associations with vascular plants were reduced or eliminated in seeds or soils treated with *A. petiolata* extracts or in which *A. petiolata* had previously grown (Roberts and Anderson 2001, Stinson et al. 2006). Prati and Bossdorf (2004) demonstrated direct allelopathic inhibition of germination of a native North American forb grown in soils in which *A. petiolata* had been grown, while a congeneric forb native to Europe responded positively to the soil treatment. Disruption of AMF in natural communities could have significant repercussions for regeneration of trees dependent on AMF associations (Stinson et al. 2006).

CONTROL STRATEGIES

Conventional Control

Many conventional methods have been explored to control *Alliaria petiolata*. Use of prescribed fire (Nuzzo 1991, Nuzzo et al. 1996, Schwartz and Heim 1996, Luken and Shea 2000), herbicide applications (Cavers et al. 1979 and references therein, Nuzzo 1991, 1994, 1996, 2000 and references therein, Carlson and Gorchov 2004), flooding (Nuzzo 1999, Evans 2006), and mechanical removal (Nuzzo 1991) have resulted in unsatisfactory control.

Biological Control

In 1998 a search for appropriate biological control organisms for *A. petiolata* was launched in Delemont, Switzerland through the cooperative efforts of CABI Bioscience in Switzerland, Cornell University, and the University of Minnesota (Hinz and Gerber 1998). Blossey et al. (2001b) and Hinz and Gerber (2005) have summarized the search for biocontrol agents for *A. petiolata*. From an initial survey of the literature which identified 69 species of phytophagous insects and 17 fungi in Europe that are associated with *A. petiolata*, four weevils belonging to the subfamily Ceutorhynchinae (Coleoptera: Curculionidae) in the genus *Ceutorhynchus* are currently considered candidate agents.

Larval *Ceutorhynchus alliariae* Bristout and *C. robertii* Gyllenhal mine in stems and leaf petioles of *A. petiolata* from March to May and pupate in the soil. Adults emerge later the same summer and feed on leaves of *A. petiolata*. Adults overwinter in the litter and soil, emerge early in the spring, and oviposit in *A. petiolata* stems and leaf petioles. In field surveys in Europe, *C. alliariae* and *C. robertii* were found either separately or

together in 81-100% of *A. petiolata* plants dissected (Gerber et al. 2002, Gerber and Hinz 2005, Gerber et al. 2008a). Feeding damage from these species results in reduced fecundity of adult *A. petiolata* plants.

Ceutorhynchus constrictus (Marsham) has the narrowest host range of the candidate biocontrol agents tested to date (Hinz and Gerber 2005). Larvae feed on seeds from May to July and then leave the host plant to pupate in the soil. Adults emerge the following April to feed on leaves and mate. Each larva consumes and destroys up to three *A. petiolata* seeds during its development.

Ceutorhynchus scrobicollis Nerensheimer & Wagner larvae feed in leaf petioles, buds, and root crowns of overwintering *A. petiolata* rosettes. Larvae leave the plants to pupate in the soil by late April. Adults emerge from May to June and aestivate during summer. Females begin laying eggs in mid September and oviposit continually through winter into spring. Individual females can produce viable eggs for at least three consecutive years, although few survive that long and fecundity decreases with age. In field surveys in Berlin, Germany, *C. scrobicollis* was found attacking 4-100% of *A. petiolata* plants collected and dissected (Gerber et al. 2002, Gerber and Hinz 2005). Damage from *C. scrobicollis* leads to direct mortality in overwintering rosettes as well as reduced fecundity in adult plants.

A petition to release *C. scrobicollis* in North America was declined by the TAG in 2008 citing the need for host specificity testing on more plant species and testing of more western North American species.

SUMMARY

Alliaria petiolata is an invasive weed that has the potential to radically alter native North America forest plant communities. Concerns about the safety and efficacy of biological control programs require new predictive tools for ensuring that only the most effective biocontrol agents are released in future programs. The *A. petiolata* biocontrol program presents an excellent opportunity to develop and test such tools.

A preliminary plant-based model has been developed to evaluate which stages of the *A. petiolata* life-cycle represent optimal targets for management efforts (Davis et al. 2006). However, is model did not account for spatiotemporal variation and covariation in demographic parameters seen in other studies of this species. Such variation can significantly affected population trajectories and dynamics (e.g. Horvitz and Schemske 1995, Pascarella and Horvitz 1998). Capturing variability in *A. petiolata* demographic rates across its range and over multiple years could be critical to projecting the frequency of conditions where single or multiple agent biocontrol are projected to succeed. The linear models used in initial *A. petiolata* analyses (Davis et al. 2006) may not capture the true dynamics of natural populations. Observations of early and late spring populations of *A. petiolata* suggest that survival is density dependent.

Building from Davis et al.'s (2006) models, we need to ask targeted questions about the dynamic interactions between *A. petiolata* and populations of the insect herbivores that have been proposed as potential control agents. Understanding how populations of the weed and control agent behave in the presence of one another and ultimately in a spatial environment will be an important step towards projecting the long-term outlook for *A. petiolata* biocontrol.

CHAPTER 2: THE SCALE AND SOURCES OF DEMOGRAPHIC VARIATION IN *ALLIARIA PETIOLATA*

ABSTRACT

Variability in demographic rates among natural plant populations can have large impacts on population structure and cause populations to exhibit subtle or radical differences in behavior over time. As population modeling studies are increasingly called upon to guide policy and management decisions, it is important that they accurately represent the dynamics of their study systems. Quantifying the sources of variability across the life history of an organism is the first step in this process. I studied the demography of 12 natural populations of the invasive forb *Alliaria petiolata* (garlic mustard) over three generations of its life cycle. Generalized linear models (GLMs) and hierarchical, generalized linear mixed models (GLMMs) were used to analyze the spatial and temporal scales of structured variability in each lower level demographic transition and the intrinsic, edaphic, and climate-driven mechanisms which underlie them. I developed statistical approaches to deal with common discontinuities, such as zero- and one-inflation issues, that made conventional analyses problematic. Population density and climate variables played important roles both in directly affecting morality and in predicting the fates of individuals later in their life histories. Variation in germination and survival rates was significantly structured across sites and years, but the distribution of fecundity varied only across sites. Persistence of dormant seeds in the soil seed bank was high, with a mean viability of 91.9% after one year. Germination, seedling, and summer rosette survival were influenced by negative density dependent feedbacks. Overall winter rosette survival was not density dependent, although the probability of extremely high (1) and low (0) winter survival was negatively density dependent. Fecundity was negatively dependent on final flowering plant density during dry years, but positively dependent

during wetter years. The frequency and amount of precipitation early in the life history were associated with increased survival for the remainder of the life cycle, but increasing summer precipitation and temperature negatively affected survival. The degree of variation observed in *A. petiolata* demographic rates encompasses the results of almost all previous studies of this species. This variation is highly structured in space and time in response to biotic and abiotic conditions. Previous studies that have explored subsets of *A. petiolata*'s demographic parameter space may therefore be limited in the scope and applicability of their predictions.

INTRODUCTION

Spatiotemporal variation in demographic rates of plant populations arises through a combination of stochastic effects, environmental drivers, and intrinsic factors. The magnitude and form of this variation has important consequences for both local and regional population and metapopulation dynamics, particularly with respect to extinction probabilities. Incorporation of environmental and stochastic demographic variability and population regulation can greatly improve the accuracy and predictive power of population models which would otherwise behave asymptotically (e.g. exponentially). As modeling studies are used more frequently to guide population management, it is important that they accurately represent the behaviors of the populations they are meant to simulate (McEvoy and Coombs 1999, Buckley et al. 2003a, b, Briese 2006, Davis et al. 2006). A single plant species can vary in its life history and demography across its range to the point where entirely different approaches could be necessary to successfully manage it in different locations (Parker 2000, Shea et al. 2005). Accounting for variation among populations becomes especially important for regionally to continentally applied management such as weed biological control, which utilizes autonomously redistributing herbivorous or pathogenic agents. Complex indirect non-target impacts can arise when biocontrol agents spread to areas where they are not effective at controlling the target species (Pearson and Callaway 2003, Pearson and Callaway 2005). Thus, quantifying demographic variation among populations and incorporating it explicitly into new management models presents an opportunity to improve the safety and efficacy of management tools like biological control.

Alliaria petiolata (garlic mustard, Brassicaceae [M. Bieb] Cavara and Grande) is a broadly distributed, invasive weed in North America that is the target of active management using various techniques. Numerous studies have quantified aspects of its biology and ecology, but a comprehensive analysis of variation in its population structure and dynamics across the landscape is incomplete. A promising initial evaluation of the known variability in *A. petiolata* demography assembled from published sources indicated a gradient of susceptibility to management at specific life history transitions (Davis et al. 2006). In their study, Davis et al. evaluated the responses of simulated *A. petiolata* populations to management of varying intensity that targeted single or multiple demographic transitions simultaneously. Building from this, the next step towards predicting the probability of successful *A. petiolata* management is to determine the distribution of the species' demographic rates among natural populations.

Individual *A. petiolata* survival and reproductive rates reported in the literature each vary broadly when compared across multiples studies (Davis et al. 2006, Table 1, and Pardini et al. 2009, Appendix A). For example, Cavers et al. (1979) measured the annual survival probability of dormant seeds as 0.99 from 1975-1978 in southern Ontario, Canada. . Thirty years later, Meekins and McCarthy (2002) reported values between 0.30-0.32 from 1996-1998 for the same parameter measured in Athens, Ohio. Variation in seed survival of this magnitude could mean the difference between a successful management program versus a failed one because of the resulting differences in longevity of the seed bank. This is evident in the changing elasticity structures among populations with different seed survival rates (see Davis et al. 2006, Figure 4A for an illustration).

Results of previous *A. petiolata* studies have suggested two principal guiding questions from which the experimental and analytical designs were crafted. First, I asked whether the variation in each demographic transition is hierarchically structured in space and/or time. In the example above, one would want to know whether the differences in seed survival resulted from site effects, year effects, both, or if the variation was random. Also, because *A. petiolata* growth and photosynthetic rates respond to light intensity (Dhillion and Anderson 1999, Meekins and McCarthy 2000, Myers et al. 2005), I also asked whether there were differences within sites between forest edge and forest interior habitats, as such differences could explain possible variation in population dynamics within each site. The answers to these questions will be used to determine the hierarchical levels at which each demographic rate should be pooled for use in future simulations of *A. petiolata* population dynamics.

Second, I asked whether the variability in survival and reproductive rates could be explained by local variation in biotic and abiotic conditions like population density, soil or climate properties. This second question seeks a mechanistic explanation of demographic variablity. From this perspective, I would expect any site, year or other categorical differences to be the net result of differences in the conditions that are biologically meaningful to plant growth and survival. Addressing this question will allow parameterization of survival and reproductive functions which can be combined to build a more generalized, dynamic model of *A. petiolata* population growth.

In this study I quantified recruitment, survival, and reproduction of over 60,000 *A. petiolata* individuals and seeds at 12 study sites in Michigan and Illinois from 2005 through 2008. I conducted two sets of analyses using generalized linear mixed models

(GLMMs) and generalized linear models (GLMs) to address each of the questions posed above. First I analyzed the spatial and temporal scales at which demographic rates varied among populations and years. These analyses generated estimates of demographic parameter distributions for use in predictive population models. Second, I parameterized survival and reproduction models of each life history stage as a function of population density and abiotic variables. These functions will be used in future discrete-time dynamic models of *A. petiolata* populations. Within each set of analyses I used multi-model inference to rank and select the best supported model from a set of a priori candidate models.

METHODS

Study Species

Alliaria petiolata is an understory forb native to western Eurasia. It has been documented in North America since the 1860s (Nuzzo 1993b) and now occurs in at least 36 U.S. states and six Canadian provinces (USDA-NRCS 2007). *A. petiolata* is a shade and cold tolerant herb (Anderson et al. 1996, Dhillion and Anderson 1999, Meekins and McCarthy 2000, Myers and Anderson 2003, Myers et al. 2005). Dormant seeds can remain viable in the soil seed bank for at least ten years (Nuzzo and Blossey, unpublished data). Density dependence has been observed for survival and fecundity (Meekins and McCarthy 2000, Winterer et al. 2005, Pardini et al. 2008, Pardini et al. 2009), but not in germination or recruitment rates. Like other Brassicaceae, *A. petiolata* has complex allelochemistry that negatively affects competitors through disruption of soil microbial and fungal communities (Vaughn and Berhow 1999, Roberts and Anderson 2001, Prati

and Bossdorf 2004, Stinson et al. 2006, Stinson et al. 2007, Callaway et al. 2008, Wolfe et al. 2008). Few pathogens or herbivores cause significant damage to *A. petiolata* in its invasive range (Renwick 2002, Evans 2006, Evans and Landis 2007, Keeler and Chew 2008), although Yates and Murphy (2008) recently found three naturalized herbivore species that successfully develop on *A. petiolata* in southwestern Ontario, Canada. They proposed that these species may merit further investigation as possible biological control agents for *A. petiolata*.

The *A. petiolata* life cycle can be broken into four distinct life stages: seeds, seedlings, rosettes, and flowering plants (Figure 2.2). Second year flowering plants are frequently referred to as "adults" in the *A. petiolata* literature. Seeds of *A. petiolata* germinate in early spring after a period of cold stratification (Baskin and Baskin 1992, Raghu and Post 2008) and produce dense carpets of seedlings. In the North Central region of the USA, surviving seedlings mature into low rosettes of petiolate leaves. Rosettes grow through the summer and fall, holding most of their leaves through the winter. Surviving rosettes bolt in late April or May the following year in southern Michigan (mean height = 71 cm, data from Evans and Landis 2007) and flower from May through June. Flowers are predominantly self pollinated (Durka et al. 2005). Seeds develop in slender fruits (siliques) along the upper stem and are shed from August through September after all flowering plants have senesced. Individual siliques contain 14.3 to 20.7 seeds on average (Evans and Landis 2007). Mean per capita fecundity estimated in eight Michigan forests in a previous study ranged from 0 to 446 (mean = 207) (Davis et al. 2006). Dispersal of seeds is limited. Most fall beneath the parent plant, although deer are believed to move seeds that get caught in their fur or hooves, and mice

are thought to frequently cache them. Past studies have demonstrated large variation in all *A. petiolata* vital rates .

Study Sites and Data Collection

Alliaria petiolata demographic rates were measured in 12 forested sites in Michigan (7) and Illinois (5) from 2005 through 2008 (Figure 2.1, Table 2.1). Sites were selected based on the presence of established, unmanaged *A. petiolata* populations and accessibility. Seven demographic rates were quantified in the field. Four replicate measures of each demographic rate of were taken at each site for each of three cohorts: 2005-2007, 2006-2008, and 2007-2009. Four groups of sampling quadrats were established within each site: two near the forest edge and two in the interior. This allowed comparisons between plants growing in different light environments within each site. Sampling areas were spaced 20 to 150 m apart as space permitted within each site. Three types of permanent quadrats were established within each sampling area to estimate rates of 1) seed germination and seed survival, 2) seedling survival, and 3) rosette survival and fecundity. A timeline illustrating how data were collected is shown in Table 2.2.

Germination and Seed Survival

I estimated germination rates in 20 x 20 x 2.5 cm deep wire screen trays buried at the soil surface. Each June (2005, 2006, 2007), topsoil was collected from 3-5 cm depth in an area free of *A. petiolata* near each sampling area and distributed to each of four screen trays. Two trays were then randomly selected to receive an addition of locally collected *A. petiolata* seeds at rates estimated from mature plants in the rosette survival

quadrats. The other two trays served as controls for background seed contamination in the soil. Adult plants were cleared within 2 meters of the seed trays to prevent the introduction of unrecorded seeds. The following year seedlings were counted destructively as they emerged from February until germination was complete. Germination rates of newly shed seeds (g_1, Figure 2.2) were calculated as the total number of seedlings in the seed addition trays minus the number of seedlings in the control trays divided by the number of seeds added. One pulse tray and one control tray were collected in June and the remaining seeds were elutriated, counted, and stained with 2,3,5-tetrazolium chloride (AOSA 2000) to determine viability (seed bank survival rate s_s). The second pair of trays was left out for a second winter. Germination rates of older seeds from the seed bank were measured the following spring, two years after being sown (g_2, Figure 2.2), using the same technique.

Seedling Survival

Estimates of seedling survival to the rosette stage (s_r) were made each spring (2006, 2007, 2008) in 25 x 25 (Michigan sites) or 20 x 20 cm (Illinois sites) quadrats. Seedling locations were marked on transparent plastic sheets laid over the quadrats every one to three weeks. Seedling survival rate was estimated as the ratio of seedlings surviving to the rosette stage in June to the peak number of seedlings observed on any date in the spring. Second year rosettes were also marked, although they did not occur in all seedling quadrats.

Rosette Survival and Fecundity:

Whereas Davis et al. (2006) estimated survival of rosettes to the flowering stage over a full year as s_{rf}, I split the measurements of the transition into two periods to provide greater temporal resolution (Figure 2.2). Summer (s_{sum}) and winter rosette survival (s_{win}) were measured in 50 x 50 (Michigan sites) and 40 x 40 cm (Illinois sites) quadrats. New rosettes were marked in June (2005, 2006, 2007) on transparent plastic sheets. Survivors were located during a fall survey in October or November and again the following June. Summer and winter survival rates were estimated as the ratio of fall to summer rosettes, and mature flowering plants to fall rosettes, respectively. Fecundity (f) was estimated for each surviving mature plant by counting the number of siliques and scaling by the number of seeds per fruit. This relationship was determined from a set of destructively harvested test plants using a breakpoint linear regression to account for nonlinearity in the ratio of seeds per fruit between large and small plants (Appendix 2.B.iv). Seed addition rates for the germination trays during the following year were estimated from these plants as the density of seed rain (seeds m^{-2}) produced by all surviving plants.

Germination trays vandalized or disturbed by animals during the experiment (21%) were not used in analyses of g_1, g_2, or s_s. All plant densities were converted to a common scale of plants m^{-2} for use as covariates in analyses of density dependence.

Climate Data

Temperature and precipitation data were assembled from the National Climate Data Center (NCDC) database (www.ncdc.noaa.gov) using the nearest available weather station for each site. Missing observations were filled in using data from the next nearest weather station. Observations from some stations that were difficult to reformat and assimilate into the database as provided by NCDC were downloaded from a commercial internet website (www.wunderground.com) which makes the same data available in a different format. Daily precipitation (mm), minimum and maximum temperatures (°C) values were compiled for each site from January, 2004 through October, 2008. These were then distilled into sets of monthly and seasonal summary statistics to be tested as predictors of *A. petiolata* vital rates.

Decisions about how to summarize the climate data were based on *A. petiolata*'s reported cold tolerance and early season growth (Anderson et al. 1996), and sensitivity to dry summer conditions (Byers and Quinn 1998) and then refined through exploratory graphical analysis. Abbreviations for climate statistics are italicized. Temperature data were processed in five ways. 1) Because *A. petiolata* is capable of photosynthesizing during winter when temperatures are above freezing, growing degree days were calculated using a base temperature of 0°C (*gdays*). I hypothesized that *A. petiolata* vital rates could be driven by threshold-based climate conditions. I calculated the number of days per month with minimum temperatures: 2) below 0°C (cold days, *cold*), 3) below -18°C (very cold days, *vcold*), and 4) above 0°C (warm days, *warm*), and 5) with maximum temperatures above 34°C (hot days, *hot*). Threshold temperatures used for

very cold and hot days were based on graphical analysis. Mean, minimum, and maximum temperatures had weak or no associations with demographic rates.

Precipitation was compiled into three monthly statistics: 1) total monthly precipitation (mm, *prcp*), 2) the number of days with > 0 mm precipitation (rainy days, *rain*), and 3) the number intervals between precipitation events during which the soil could dry out (drying events, *dry*). The drying statistic complements the rainy days statistic as a way of quantifying the frequency of precipitation events. For example, the rainy days statistic might indicate that it rained 14 days in a given month but doesn't convey any information about whether it rained for two weeks straight and then was dry for two weeks, or if it rained every other day. Knowing additionally whether there was one drying event versus 14, we can distinguish between these two patterns of precipitation.

Seasonal climate statistics were calculated by splitting the "garlic mustard year" into four stages based on *A. petiolata* phenology and summing monthly values within each stage for each year: spring (*Spring*, January – March of seedling year, used with g_1, s_r), summer (*Sum*, May – September of seedling year, used with s_{sum}), winter (*Win*, October of seedling year– March of flowering year, used with s_{win}), and flowering period (*fec*, February – May of flowering year, used with *f*). The spring and summer stages were split further for some analyses. January and February data were grouped, as this seemed likely to be an important pre-germination time interval, and August and September drying events, when drought stress seemed most likely to affect summer survival. Finally, precipitation data from March were used in analyses of seedling survival, and May precipitation and rainy days were used in summer survival analyses.

Soil Data

Selected soil chemical and physical properties relevant to plant growth, including gravimetric water content, particle size distribution, water holding capacity, organic carbon, pH, P, K, Mg, Ca, and inorganic N were measured at the quadrat level for all locations. Ten soil cores 1.9 cm in diameter were taken to a depth of 10 cm from the perimeter of each adult census quadrat at the time of census in June 2007 and bulked to form a composite sample. Gravimetric soil moisture was immediately determined on a 25 g subsample by oven drying at 65 °C to constant weight (Klute 1986). The remainder of the composite sample was sent to A&L Great Lakes Laboratories in Fort Wayne, Indiana, where all other soil analyses were performed. Particle size distribution was measured using the hydrometer method (Gee and Bauder 1986). Soil water holding capacity was measured at a matric potential of -1/3 bar (-33 kPa) using the pressure membrane method (Klute 1986). This matric potential was chosen for measuring soil moisture retention as it represents field capacity, the point at which all water drainage due to gravity has occurred, after a soil has been fully saturated (Brady and Weil 1996). Organic carbon was measured by loss upon combustion (Nelson and Sommers 1994) and soil pH was measured in aqueous solution with a hydrogen selective electrode (Thomas 1996). Available P, K, Mg and Ca were measured using ICP (inductively coupled plasma spectroscopy) methods (Soltanpour et al. 1996). Finally, inorganic N ions, including both NO_3-N and NH_4-N, were measured using automated colorimetric procedures (Mulvaney 1996). The site mean value and standard error of each soil variable are included in Appendix 2.A.

Statistical Analysis

The data analysis addressed two principal goals. First, I wanted to determine the spatiotemporal scales at which most variation occurred for each *A. petiolata* vital rate to facilitate pooling of variables at the appropriate scales. For example, if g_1 varies significantly at the site level but not across years, estimates of germination probability can be drawn from a common distribution within each site. These parameter estimates will be used to assemble a set of **A** matrices for use in future projection models, loglinear analyses, and Life Table Response Experiments (LTREs) to analyze how variation in survival rates contributes to differences in population growth and possible responses to management. Second, I wanted to explore mechanistic sources of variation in vital rates, and thus constructed a separate set of environmental models which included climatic, edaphic, and population density covariates

Hierarchical Mixed Model Analyses

I used fixed effects GLMs and GLMMs, which incorporate both fixed and random effects, to evaluate how each lower level demographic rate varied across the spatial and temporal structure of the sampling design. These are generalized forms of familiar linear and linear mixed models which can accommodate data with non-normal error distributions (Bolker et al. 2009). This is done via a link function that transforms the predictor to meet the distributional assumptions made about the data. This allows a linear model to then be fit. The link function used depends on the underlying distribution of the data. Models of g_1, g_2, s_r, s_{sum}, and s_{win} were fit using a binomial error distribution and

logit link. The natural logarithm of fruit production + 1, $\ln(f+1)$, was modeled with a normal distribution and identity link using the mean siliques/plant as the response variable weighted by the number of plants in the quadrat. Although I initially modeled the raw fecundity observations at the plant level, residual diagnostics were unacceptable. Distributions used to model s_s are described below. All analyses were conducted in the GLIMMIX procedure in SAS version 9.2 (SAS Institute 2008) except where noted.

For each vital rate, a set of candidate models was developed that embodied alternative hypotheses about which spatiotemporal levels of the data structure were important. I then used a multi-model inferential approach based on maximum likelihood to evaluate which models were best supported by the data (Burnham and Anderson 2002). Models of s_{win}, f, g_2, and s_s were evaluated and ranked using AICc, an information criterion corrected for small sample sizes (Anderson 2008). Models of g_1, s_r, and s_{sum} were overdispersed and were evaluated with Quasi-AICc (QAICc), which includes an additional correction for overdispersion (Bolker et al. 2009). AICc and QAICc will both be referred to as "AIC" in the main text for readability, but will be distinguished in the tables and appendices where important. The model with the lowest AIC score (AIC_{min}) within a set of competing models is considered the best model, given the data. More details about the model fitting and ranking process are provided in Appendix 2.B.i. An explanation of how spuriously supported models with "pretending variables" (Anderson 2008) were identified is described in Appendix 2.B.ii.

For g_1, g_2, s_r, s_{sum}, s_{win}, and f, the full GLMM structure included fixed terms for Site, Year, Site*Year interaction, and Location (edge vs. interior). Site was included as a

fixed effect in this analysis because we are ultimately interested in making site-specific predictions about *A. petiolata* population growth and its potential response to management. Quadrats were modeled as random effects to account for correlations between observations made within individual quadrats. Although year would ideally have been modeled as random, there were not enough levels of the year variable (n=3) to test the assumption that the levels of each random effect are normally distributed. Reduced models included all factorial combinations of the fixed effects in the full model plus an intercept-only model for a maximum of ten models per demographic rate.

The seed bank (s_S) viability data had an inflated frequency of ones, i.e. 100% viability. These properties made it difficult to fit a GLM or GLMM using a standard error distribution. Instead, I used a two stage conditional modeling approach similar to that described by Cunningham and Linenmayer (2005). The data were analyzed by first modeling probability π of observing 100% seed viability in a sample with a binary GLM. Second, I modeled the non-zero observations with a beta error distribution (see Appendix 2.B.iii for details). As Cunningham and Lindenmayer (2005) point out, this two step method has an advantage over using zero-inflated or mixture distributions because the component analyses are orthogonal and can be driven by independent processes. Because a number of observations were discarded (described above) the 113 "good" observations of s_S were unbalanced with respect to sites and quadrats across years. Although there was at least one empirical observation of s_S per site per year, random quadrat effects and edge/interior location effects were excluded to avoid overfitting.

Environmental Models

The hierarchical analyses are useful for analyzing the structure of variability in vital rates across sites and years. However, they are not mechanistic and therefore are limited in how broadly their results can be generalized beyond the study system. I constructed a second set of GLMs and GLMMs to explore how *A. petiolata* demographic rates responded to differences in soil properties, climate variation, and population density. Following the example of Buckley et al. (2003a), I used GLMMs to evaluate the response of each lower level demographic transition in the *A. petiolata* life cycle to a set of extrinsic and intrinsic variables. For each transition, a set of candidate models with different combinations of fixed and random effects was defined *a priori*. Because of the large number of possible explanatory variables and interactions, preliminary graphical analysis of the data was necessary to reduce the number of variables included during the model fitting process. Models were fit to the data in GLIMMIX and ranked using AICc or QAICc when overdispersion was present. Extrinsic factors included as fixed effects in models for each transition were combinations of climate statistics for time intervals relevant to the transition and soil variables in each quadrat. If there were no strong *a priori* hypotheses about specific soil variables, versions of full models with different combinations of soil variables were fit and ranked using AIC. Soil variables from the best supported models were included in the development of subsequent reduced models. In all cases, the only soil variables retained in the final models were those which had *a priori* support. Intrinsic effects included the natural logarithm of *A. petiolata* population density [ln(plants m^{-2})] at the start of the transition as well as observed *A. petiolata* densities and survival rates from previous transitions. For example, both October rosette densities and

June adult plant densities were included in analyses of f. Quadrats were modeled as random effects. Fecundity models were fit using a Poisson error distribution and survival models were fit with binomial errors except as noted below.

The environmental models of g_1, s_r and s_{sum} were overdispersed and were ranked using QAICc (Appendix 2.B.i). The significance of random quadrat effects was tested for each model with a likelihood ratio test. The residual deviance explained by the full GLMM was compared to the residual deviance of a reduced GLM with no random terms. The difference in deviance explained was compared to a mixture of χ^2 distributions with 1 degree of freedom using the covtest / glm option in GLIMMIX to test the null hypothesis that the random effects did not improve the amount of deviance explained by the model.

Winter survival had an extreme-value inflated distribution with a high proportion of both zeros and ones (n = 37). This was accommodated by splitting the winter survival analysis into a two step conditional process like the seed survival analysis. In the first step, I estimated the probability of an observation being extreme versus not extreme as a binary response to a set of predictor variables (model $s_{win}1$). Second, if the observation was extreme, I modeled the probability of it having 0 versus 100% survival as a binary response (model $s_{win}2a$). If the observation was not extreme (i.e. $0 < s_{win} < 1$), it was modeled with a binomial GLMM (model $s_{win}2b$). Random quadrat effects did not improve the fit of the binary models and were dropped from $s_{win}1$ or $s_{win}2a$. Note that model $s_{win}1$ predicts the probability of an observation being extreme or not. This determines whether to proceed to $s_{win}2a$ or $s_{win}2b$. Model $s_{win}2a$ predicts the probability

of an extreme observation having a survival probability of either 0 or 1, conditional on knowing that the observation is extreme. Model s_{win}2b predicts quadrat mean survival probabilities between 0 and 1, conditional on knowing that the observation is not extreme.

G_2 was not measured during the third and final season (spring 2009) but was estimated by modeling the relationship between g_1 and g_2,and the seed bank size during the first two years of the study and projecting forward (Appendix 2.B.v). Because the model used pooled data, there are only site level estimates of g_2 during the third year. Hierarchical models of g_2 did not include the estimated rates from the third year. No environmental analyses were run on g_2.

RESULTS

Alliaria petiolata demographic rates were highly variable across the twelve sites and over the three years of the study. For survival there were clear site by year interaction effects, as the indicted by the changes in ranking of the 12 sites over the study period (Figure 2.3). These interaction plots are also revealing of the relative contributions of site, year, and site by year interactions in structuring the vital rates spatiotemporally. For example, variation in germination of new seeds (g_1) appears to have been dominated by year effects, while the site by year interaction had much stronger effects on seedling survival. These data agree with my general observation that summer survival and g_1 were highest during the first year of the study and then decreased during subsequent years,

while patterns of variation in seedling survival over time was much less consistently across sites.

The distributions of most transitions were heavily right skewed (Figure 2.4), with high frequencies of zero % survival of summer (25 of 143 observations, mean s_{sum} = 0.22 ± 0.03) and winter (30 of 121, mean s_{win} = 0.33 ± 0.03) rosettes. This contrasts with the more uniform distribution of rosette survival (mean s_r = 0.49 ± 0.04), with only four observations of zero survival out of 144. Seed survival was high in contrast (100% 70 of 113 times, mean s_s = 0.92 ± 0.02). Mean g_1 (0.31 ± 0.03) and g_2 (0.11 ± 0.03) were lower than comparable mean rates compiled from the literature by Pardini et al. (2009), and are closer to the means of rates assembled by Davis et al. (2006). Few other *A. petiolata* studies have reported measures of s_{sum} and s_{win}. These are more commonly combined multiplicatively into s_{rf}, the probability of surviving from the rosette to the flowering stage over a full year. The mean estimated s_{rf} (0.1048) calculated from s_{sum} and s_{win} was also considerably lower than others' estimates (mean of rates compiled in Davis et al. 2006 = 0.548).

Natural peak seedling densities in the study system ranged from 75-6025 m^{-2} (median = 877.5), and June rosette densities in the system ranged from 31.25-3344 m^{-2} (median = 232). The high densities I observed occurred during the first year of the study, reaching maxima (median) of 5696 (1096), 6025 (890), and 2650 (704) plants m^{-2} for seedlings and 3344 (368), 1092 (210), and 976 (212.25) m^{-2} for rosettes in 2005-2007, respectively.

Hierarchical Mixed Model Analyses

Demographic rates of *A. petiolata* were substantially differentiated among sites and years. Likelihood based evaluations most strongly supported models of g_1, g_2, s_r, s_{sum}, and s_{win}, with terms for site, year, and site*year interactions (Table 2.3). This indicates that sites differed in their responses to year effects for these parameters. For each of these models except s_{win}, the second best supported model also included a term for forest edge versus interior location. Each had a ΔAIC of approximately 3, but the extra variable didn't explain any additional variance in the response, evidenced by the similarity of the ln(l) values for these models. The additional location term in these models appears to be a "pretending variable" (Appendix 2.B.ii) and these models were discarded from consideration. After dropping these models from consideration, the remaining models in each of these sets had no support from the data.

The best supported model of fecundity included a significant fixed effect for site only ($F_{11,45}$ 5.00, $P < 0.0001$). Estimated mean values for f are therefore statistically the same for all years within each site. Models with terms for location also appeared to be well supported as described above and were excluded from model weighting.

Seed viability in the seed bank was evaluated in two stages because of the high proportion of ones in the data. In the first step, the data best supported the binomial GLM with a fixed year effect only. The probability of a population having 100% viability was lower in 2006 than in either 2005 or 2007 (ls-means difference $t_{110} = 3.17$, $P = 0.0020$, and $t_{110} = -3.32$, $P = 0.0012$, respectively), while the probabilities were approximately

equal in 2005 and 2007. Among the remaining observations with less than 100% seed survival, variability in seed survival was randomly distributed among sites and years with a mean viability of 0.7847 and beta distributed errors. Note that the second best model for $s_S < 1$ also appears to be a "pretending variable" model (Table 2.3). Combining the probabilities from the two seed survival models yields predictions of $s_S = 0.9462$, 0.8668, and 0.9508 for all sites for the three consecutive years. Plots of the model conditional predictions (i.e. including random effects) are shown in Figure 2.5.

Environmental Models:

Density dependent feedbacks and responses to climate and intrinsic factors from previous life stages were common features in models of most demographic rates. In some cases an earlier population density, climate or survival metric was predictive of a later survival or reproductive outcome. This suggests that individual success, however measured, is developed though a cumulative process that conditions individual plants over their life time. For most *A. petiolata* demographic rates a single model clearly received the greatest support from the data, indicated by the Akaike weights and ΔAIC scores. The exception to this was $s_{win}1$, for which the best model received 34% of the Akaike weights. Because the next best model was only half as well supported, I chose the top ranked model. The parameter estimates and their standard errors for the best supported model of each demographic rate are presented in Table 2.4-Table 2.8. Also shown are the type-3 *F*-tests of the null hypotheses that each parameter value is not different from zero and the χ^2 likelihood ratio tests of the random effects. In models that

included and retained random quadrat effects, the random effect is assumed to be normally distributed with mean 0 and variance as estimated. Random quadrat effects greatly improved the fit of models of g_1, s_r, s_{sum}, $s_{win}2b$, and f but were dropped from the two binary models $s_{win}1$ and $s_{win}2a$. The goodness of fit of each model to the data is illustrated in Figure 2.6. This shows the observed and predicted values from each model generated from the best linear unbiased predictors (BLUPs) which are based on the parameter estimates for both the fixed and random effects. Interpreting how the independent variables in each model affect the response can be challenging when there are significant interaction terms and because the regression coefficients are in a transformed data scale. Plots of the observed and predicted demographic rates versus the independent variables help to illustrate these relationships (Appendix 2.C).

Germination

Germination probability of newly shed seed (g_1) was most strongly influenced by climate conditions in January and February, just prior to germination (Table 2.4). The best model of g_1 was supported by 85.3% of the Akaike weights and was separated from the next best model by ΔAIC of 3.7. Germination was reduced by the frequency of days with measurable precipitation (*rain*) in January and February and increased by the number of warm January and February days. There was an interaction between warm days and soil water holding capacity that weakened the relationship between warm days and *W.H.C.* as W.H.C. increased. Because of its correlation with g_1, g_2 was not modeled.

Seedling Survival

Seedling survival was regulated by a combination of negative density dependent mortality plus the additive and interactive effects of several climatic and edaphic factors (Table 2.5). The best supported model of s_r received 95.8% of the Akaike weights, while the next best model had a ΔAIC of 7.8 (Table 2.5). Seedling survival is negatively dependent on March precipitation ($prcp_{mar}$, mm), the natural logarithm of soil water holding capacity (*whc3bar*), and the frequency of warm days in January and February ($warm_{jf}$, t_{min}>0°C) but is positively dependent on total soil inorganic N. The frequencies of hot days in summer (hot_{sum}, tmax>34°C) and rainy days in autumn ($rain_{fall}$) during the year prior to germination affect s_r through an interaction illustrated in Figure 2.7. Fall rain has a negative effect on s_r following a summer with a high number of hot days, but has a positive effect after a summer with few hot days. Because survival rates are bounded on the closed interval [0, 1], the relative magnitude of any single effect is dependent on the combined magnitudes of all the other factors in the model, as is seen in the effect of seedling density (Figure 2.7). Although only seedling plots in Michigan contained both seedlings and second year rosettes, including these in the models with additive or non-additive terms did not explain any additional variance in seedling survival.

Summer Survival

The best model of summer rosette survival (s_{sum}) was supported by 99.3% of the Akaike weights, and was separated from the next best model by a ΔAIC of 10.0. After

accounting for the variability in survival due to other extrinsic factors in the model, the strength of density dependent mortality during the summer (slope = -0.2951, Table 2.6) was approximately equal to the rate in spring (slope = -0.2320, Table 2.5). During summer, this translates to a decrease in survival probability from 1.0 to 0.427 ($=e^{-0.2951}$) for each increase of 1 unit of the log of June rosette density, holding all other variables constant at zero. The range of June density was ~ 4.6 from the most to the least dense quadrat, in log transformed units. Density alone would be expected to reduce s_{sum} from 1.0 to 0.257 ($=e^{-.2951*4.6}$), again holding other variables constant at zero. In nature the variables do co-vary and are non-zero, so the strength of density dependence will again be scaled by the contributions of other variables. Because of a positive correlation between the quantity ($prcp_{may}$) and frequency ($rain_{may}$) of rain events (Pearson's r = 0.7699), increases in either variable positively affect s_{sum}. There were also negative correlations between summer hot days hot_{sum} and the number of drying events in August and September dry_{as}. This indicates s_{sum} was lower during cooler summers that associated with more frequent late summer rain events. Although such strong correlations among variables can indicate multicollinearity, models including the interaction between these two variables were more strongly supported than models containing polynomial functions of either individual variable and other simpler models. Summer heat and May rain were positively correlated (Pearson's r = 0.5133), so it is possible that the positive effect of hot summer days is a time-lagged outcome of earlier conditioning from ample spring rains. Summer survival responded positively to increasing percent sand in soils,

which was generally higher in Michigan sites, but negatively to increasing soil N through its interactions with spring rain ($prcp_{May}$).

Winter Survival

The three component sub-models of the conditional winter rosette survival analysis predicted that: (1) the probability of having an extreme survival rate, either 0 or 100%, was negatively density dependent and also decreased with the frequency of summer rainy days (Table 2.7, Model $r_{win}1$). (2) Among quadrats that did have extreme values, the probability of having 100% survival (versus 0%) was greatest at very low densities and was negatively dependent on summer growing degree days (Table 2.7, Model $r_{win}2a$). (3) Among quadrats that did not have extreme values ($n = 84$), there was a positive correlation between summer and winter survival probabilities (Table 2.7, Model $r_{win}2b$). Survival in these plots was also negatively dependent on the frequency of summer rain, soil pH, and the frequency of very cold winter days, but was not density dependent.

Fecundity

Mean quadrat per capita fecundity, modeled as the mean number of fruits (siliques) per plant was predicted by two opposing density dependent functions, each conditional on climate interactions (Table 2.8). First, adult plants were negatively affected by the rosette density they experienced as rosettes during June of their first year. The strength of this effect on fecundity was greatest during cool summers (low hot_{sum})

but was tempered during hot summers (high hot_{sum}). Second, both the frequency of rain events during a plant's second spring until flowering ($rain_{fec}$) and the log of the final density of adult plants were positively correlated with per capita fecundity.

DISCUSSION

Comparisons to Previous Studies

The magnitude of variation in *A. petiolata* demographic rates from this study was much greater than expected. At the quadrat level, the range of values of each transition observed in the study system in four years was inclusive of almost all *A. petiolata* demographic rates published previously over a thirty year period that I am aware of (Figure 2.4). The frequency distributions of most vital rates in my data parallel the distributions of these published rates, evidenced by the correspondence between the histograms and the clustering of published rates in Figure 2.4. The means of demographic rates in this study were not significantly different from those of the compiled published rates based on overlapping 95% confidence intervals using site by year means. The exceptions to this were geometric mean per capita fecundity (f) [mine: 52.643 (41.51, 66.75); published: 204.194 (93.17, 447.50)] and rosette to flowering survival (s_{rf}) [mine: 0.09 (0.054, 0.126); published: 0.548 (0.443, 0.653)]. The means of each of these rates were lower in my system. Reduced fecundity may have been related to the same processes that lead to the low rosette survivorship I recorded. This is suggested by the mutual dependency of summer survival and fecundity on both June rosette density and hot summer days.

Past studies have also found evidence of stage-specific density dependent mortality and reproduction (Winterer et al. 2005, Pardini et al. 2008, Pardini et al. 2009) or a lack thereof (Rebek and O'Neil 2006) as well as density dependent population growth (Meekins and McCarthy 2000, but see Appendices D and E for a discussion of their calculations, 2002). The range of first year rosette densities observed in June and in October in this study was greater than those evaluated by most previous investigators (Table 2.9). Meekins and McCarthy's (2002) first year rosette densities fell above the center of my data range (June from their Figure 3), while Pardini et al.'s (2009) early May rosette densities fell mostly below it (from their Figure 2A). Rebek and O'Neil's (2006) fall October densities and Pardini et al.'s (2009) August rosette densities each encompassed the center of my fall density range. While the Rebek and O'Neil measurements are comparable to my fall rosette density, Pardini et al.'s August sample is phenologically earlier than mine. What they measured as s_1 from early April-mid May gives a much more compressed seedling survival period than mine. Similarly, my measure of summer rosette survival covered June-late October, while their s_2 covered early May-August. Finally, my winter survival measure ran from October-June, while their s_3 ran from August-May and included most of the mortality attributed to summer losses. This may be why Pardini et al. found compelling evidence of strong "winter" density dependent mortality while I did not. Rebek and O'Neil's finding of no density dependent mortality from October-June concurs with my own.

Seedling densities in my study quadrats fell within the range of other published accounts (Table 2.9). Winterer et al. (2005, J. Winterer personal communication) recorded seedling densities within the range I observed. Seedling densities in quadrats

used to estimate seedling survival by Pardini et al. (2009) were below this range. Their empirical measure of s_r was based on the survival of 469 marked individuals spread across 40 1 m^2 quadrats for an average density of 11.725 seedlings m^{-2}. These estimates and my own are all well below Trimbur's (1973) scaled estimate of 20,000 seedlings m^{-2}. The maximum seedling density recorded in my study system was 142 100 cm^{-2} (=14,200 m^{-2}) at the Shiawassee site, but this was outside of the study quadrats (J. Evans, personal observation).

Second year plants have been shown to negatively impact seedling survival (Meekins and McCarthy 2002, Winterer et al. 2005). Although I saw this as well in seedling survival environmental models that included a term for second year plant density, these models were not strongly supported by the data. Other investigators have shown that *A. petiolata* has a plastic response to variation in light availability but reaches maximum photosynthetic rates under high light. (Dhillion and Anderson 1999, Meekins and McCarthy 2000, Myers et al. 2005) but I found no differences in any demographic rates between forest edge versus interior plots. It is possible that the light conditions at the edge plots were more shaded than those studied by others, or that other unmeasured differences between edge and interior locations offset the effects of light exposure.

The hierarchical analyses show that there is a significant amount of spatial and temporal structure underlying this variation (Figure 2.3). I found significant site, year, and site*year effects in all germination and survival transitions. Germination rates of new seeds were highest during the first and third years of the study, and were generally greatest in the southernmost and northernmost sites during all years. Seedling survival

exhibited the most pronounced spatial structuring, increasing with latitude and from west to east within each year. It is possible that this structure is partially driven by climate related differences in germination phenology, which was up to several weeks later in the more northern sites during some years. Experimental hierarchical models in which I first adjusted survival rates to correct for phenological differences among sites produced the same results (not shown), so I used the original data. The distributions of summer and winter rosette survival were much more temporally than spatially structured. During the first year, when summer survival was highest across most sites, it was greatest in the southernmost Illinois sites and decreased northward. There was much less landscape-scale spatial structure during the second and third summers, although most individual sites maintained the same rank in summer survival that they had the first year (Figure 2.3). In other words, the "best" sites for summer survival in a good year were still the best, even when conditions were less favorable across the region. Similar maintenance of site rank was seen across most sites in germination of new seeds and winter survival probabilities. Winter rosette survival showed an almost opposite temporal pattern to summer survival, increasing in the second year in most sites and then decreasing the next. Per capita fecundity was very consistent within sites over time. Only the Homer Lake site varied in fecundity over time, dropping during the second year and then increasing in the third. Finally, seed bank persistence was relatively unstructured. It varied by year only in the proportion of quadrats with 100% seed viability, which was not spatially driven. The differences in the spatiotemporal structures of these transitions may translate into differences in their responses to management of specific life history stages and how locally specialized management of different stages must be (see Chapter 4).

Conclusions

These results suggest several important conclusions. First, *A. petiolata* does not perform equally well at all times and locations. It has good years and bad years (or better years), and these vary by site (Figure 2.3). Additionally, the magnitude and direction of site and year effects differ across life history stages. Germination of new seeds, summer survival, and winter survival of most sites closely tracked each other over time, indicating the importance of regional influences on these demographic rates. In contrast, mean seedling survival did not show such strong inter-site trends across years. Instead, seedling survival variation was much more locally driven, as evidenced by the lack of parallel lines in the seedling plot in Figure 2.3.

From the perspective of a natural resource manager, this means that the target plant is not likely to perform in the same way at all places and times and, thus, should not be expected to respond to management identically at all locations. The consequences of making this assumption have been demonstrated previously in other systems where multiple populations of an invasive weed responded differently to a common management effort (Shea et al. 2005). In the case of weed management using classical biological control, implementation of ineffective strategies is of particular concern because herbivorous or pathogenic natural enemy species cannot be recalled once released and can be damaging to non-target species (Pearson and Callaway 2003). It will be important to determine whether differences among life history stages in the spatiotemporal scales of variation in vital rates will translate into differences in population dynamics at the same scales. If so, it could mean that life history stages like

seedling survival with stronger local structure would require more locally optimized management plans. Transitions such as germination that are largely shaped by regional effects may be better suited to implementation of a common management approach across sites. Future models of *A. petiolata* population dynamics will have to account for these spatial and temporal sources of variation in vital rates by making site-specific models that account for year to year differences in vital rates.

Second, the variation in vital rates is driven by a common set of underlying mechanisms. Multiple competing environmental models were evaluated to identify the mechanisms that are collectively quantified as categorical site and year effects in the hierarchical models. For each demographic transition, a single explanatory model emerged as the most probable. These models show that individual plants are affected by both their immediate and past environments at each step through their life history. For example, seedling survival is greater where there is higher soil N availability, but this comes at the cost of lower summer rosette survival. Increased rain in May helps condition rosettes to endure the hottest days of summer. However, plants that experienced high densities and hot summer days during their first year on average produce fewer seeds when they mature. Elements like soils and climate patterns and population densities which feed from one life stage into the next are consistent within locations over time and tie the individual life history stages together into coherent covariance signatures. Based on this, it may be worth exploring the data structure further using structural equation modeling or generalized linear latent and mixed models (GLLAMM) in the future (Skrondal and Rabe-Hesketh 2004) The models as currently parameterized will form the basis for exploring the consequences of site differences on projected population

performance and can be used to predict how populations will respond to perturbations. Further, because these models quantify the site in terms of continuous variables, it may be possible to predict the performance of populations at other locations for which the relevant data are available.

Third, demographic rates of natural *A. petiolata* populations are more variable than I expected. The existence of variability in *A. petiolata* vital rates documented here is not surprising. What is noteworthy, though, is that nearly the entire range of previously known vital rates occurs in just a small subset of *A. petiolata*'s North American distribution and over a relatively short timescale. Modeling studies based on a single site or a single year of data are very likely to misrepresent the longer term dynamics of the population. For example, mean seedling survival (s_r) at the Bob-Peoria site was measured as 0.843, 0.149, and 0.094 over three consecutive years. No single rate characterizes s_r at the site, and it is unlikely that three rates are truly sufficient. Additionally, s_r is linked with other vital rates in the *A. petiolata* life cycle through density dependent functions and dependence on common exogenous factors. Previous modeling studies of this species have explored how population dynamics change across a range of parameter values (Davis et al. 2006, Pardini et al. 2009) but have been unable to account for the covariances among parameters. These models likely capture a subset of the range of population dynamics exhibited by natural *A. petiolata* populations. However, by holding dynamic parameter values as fixed during simulations or breaking apart the correlations among vital rates, they may misunderestimate (*sensu* George W. Bush, November 6, 2000) the complexity or elasticity structure of natural population dynamics (J. Evans, unpublished).

An important next step to interpreting the behavior of *A. petiolata* populations will be to compare a stochastic population model based on a fixed set of matrices for each site and year with a dynamic model based on the intrinsic and environmental functions parameterized here. These can then be used to assess the potential interactions between the different *A. petiolata* populations and the suite of proposed management tactics.

Tables: Chapter 2

Table 2.1. Names and locations of study sites. Rainfall estimates represent mean values summed from climatic data used in analyses. See methods section for details.

Site (Abbreviation)		State	Latitude	Longitude	Site Description	Annual Rainfall (mm)
Bob-Peoria	(B)	IL	40.72205	-89.5055	Hardwood floodplain forest	818
Edward Lowe Foundation	(ELF)	MI	41.96451	-85.9962	Second growth northern hardwood forest along edge of old field.	978
Farmdale	(F)	IL	40.67668	-89.4878	Hardwood floodplain forest	818
Healy Road	(HR)	IL	42.10862	-88.2148	Mesic second growth hardwood forest	936
Holland State Park	(HSP)	MI	42.77765	-86.2025	Hardwood forest in dunes along Lake Michigan shore	774
Homer Lake	(HL)	IL	40.06325	-87.9787	Upland second growth hardwood forest	993
Illini Plantations	(IP)	IL	40.07925	-88.2107	Upland second growth coniferous forest	960
Ives Road	(IR)	MI	41.98147	-83.932	Hardwood forest bordered by restored tallgrass prairie, descending to hardwood floodplain forest	743
Johnson Park	(JP)	MI	42.92559	-85.7699	Hardwood oak forest in urban park	972
Rose Lake	(RL)	MI	42.81234	-84.4042	Second growth hardwood forest in abandoned crop field.	860
Russ Forest	(RF)	MI	42.01162	-85.9703	Old growth oak/maple forest	978
Shiawassee YMCA	(SH)	MI	42.88546	-84.0491	Black locust (*Robinia pseudoaccacia*) and degraded pine plantation	912

Table 2.2. Timeline of *Alliaria petiolata* sampling for each demographic parameter. S, F, W, and Sp indicate summer, fall, winter, and spring, respectively. Boxes enclose the time interval over which a parameter was measured. The unshaded box indicates the summer survival transition that was measured in the smaller seedling survival-sized quadrats in Michigan for the 2005-2006 cohort. Germination data are not presented in Chapter 2.

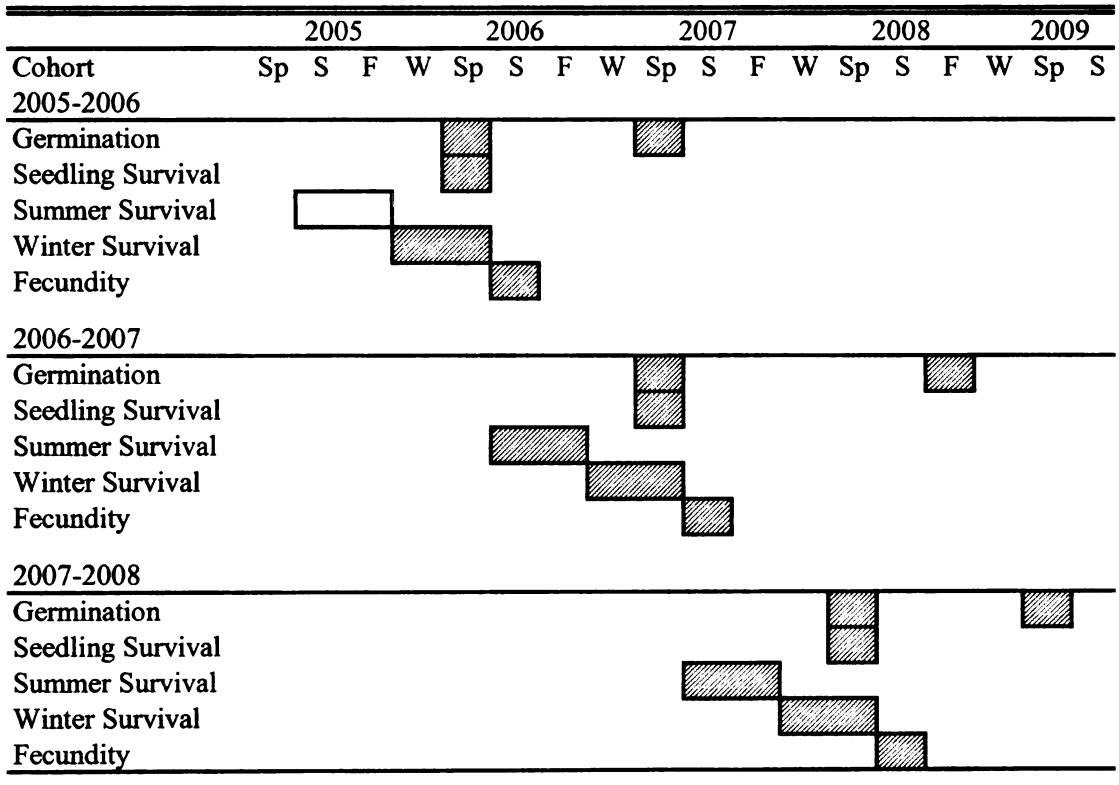


Table 2.3. Results of model selection for spatiotemporal differentiation of *A. petiolata* life history transitions. Details for each transition include the error distribution (dist) used in GLMs and GLLMs, the information criterion used for model weighting and selection (ic), and the dispersion parameter φ used to calculate QAICc where applicable. Variables: fixed effects (S=site, Y=year, L=location, I=intercept), Laplace approximated maximum likelihood (ln(l)), number of observations (n), levels of random quadrat effect (r), number of parameters including random effects (k), (Q)AIC: AICc or Quasi-AICc for overdispersed models, delta (Q)AIC (Δi) , and Akaike weight (w_i). Models with r=0 were GLMs with fixed effects only. A maximum of 5 models are shown per vital rate. All other models with $\Delta i < 11$ and $w_i > 0$ are shown.

Transition, Details	Fixed Effects	ln(l)	n	r	k	(Q)AICc	Δ_i	wi
germination 1 (g_1)	S, Y, S*Y	-828.11	137	48	37	352.6	0	0.872
dist=binomial	S, Y, S*Y, L	-828.08	137	48	38	356.4	3.8	0.128
ic=QAICc	Y	-1228.26	137	48	4	379.4	26.8	<0.001
φ=6.62	Y, L	-1228.24	137	48	5	381.5	28.9	<0.001
seedling survival (s_r)	S, Y, S*Y	-513.48	144	48	37	526.6	0	0.820
dist=binomial	S, Y, S*Y, L	-512.68	144	48	38	529.7	3.0	0.180
ic=QAICc	S, Y	-886.93	144	48	15	769.8	243.1	<0.001
φ=2.41	S, Y, L	-884.43	144	48	16	770.2	243.6	<0.001
summer srv. (s_{sum})	S, Y, S*Y	-511.10	143	48	37	387.9	0	0.843
dist=binomial	S, Y, S*Y, L	-510.47	143	48	38	391.3	3.4	0.157
ic=QAICc	Y	-846.68	143	48	4	484.0	96.0	<0.001
φ=3.56	Y, L	-846.39	143	48	5	485.9	98.0	<0.001
winter srv. (s_{win})	S, Y, S*Y	-240.22	121	48	36	584.1	0	0.541
dist=binomial	S, Y, S*Y, L	-238.30	121	48	37	584.5	0.3	0.456
ic=AICc	Y	-294.28	121	48	4	596.9	12.8	0.001
	Y, L	-293.31	121	48	5	597.1	13.0	0.001
Fecundity Pooled	S	-81.37	93	48	13	193.3	0	0.946
dist=lognormal	S Y	-79.99	93	48	16	199.1	5.8	0.052
ic=AICc	I Only	-100.77	93	48	2	205.7	12.3	0.002
	Y	-99.61	93	48	5	209.9	16.6	<0.001
germination 2 (g_2)	S, Y, S*Y	-253.91	82	45	25	581.0	0	0.852
dist=binomial	S, Y, S*Y, L	-253.51	82	45	26	584.5	3.5	0.148
ic=AICc	Y	-378.88	82	45	3	764.1	183.0	<0.001
	Y, L	-378.48	82	45	4	765.5	184.5	<0.001
seed srv. (s_s) Ones	Y	-66.97	113	0	3	140.2	0	0.994
dist=binomial	I Only	-75.07	113	0	1	152.2	12.0	0.002
ic=AICc	S, Y	-59.52	113	0	14	151.3	11.2	0.004
	S	-68.61	113	0	12	164.3	24.2	<0.001
seed srv. (s_s) < 1	I Only	25.05	43	0	2	-45.8	0	0.849
dist=beta	Y	25.66	43	0	4	-42.3	3.5	0.145
ic=AICc	S, Y	40.79	43	0	15	-33.8	12.0	0.002
	S	36.85	43	0	13	-35.1	10.7	0.004
	S, Y, S*Y	45.78	43	0	25	34.9	80.7	<0.001

Table 2.4. Best supported environmental model of germination of newly shed seeds after one winter (g_1). Parameter estimates from GLMM with binomial errors and logit link. Model ranking was evaluated with QAICc. Akaike weight = 0.853, ΔQAICc of next best model = 3.7. All climate data are from time periods which precede germination. Random quadrat effect was evaluated with a likelihood ratio test by comparing the change in the -2ln(l) from dropping quadrat from the model to a mixture of χ^2 distributions in GLIMMIX (SAS Institute 2008).

Parameter	Symbol	Test	Estimate	SE
Intercept	μ	NA	-2.7391	0.3961
WHC3bar	*wch*	($F_{1,85}$ = 20.74, $P < 0.0001$)	0.0624	0.0137
Warm Days (Jan., Feb)	$warm_{jf}$	($F_{1,85}$ = 647.31, $P < 0.0001$)	0.3534	0.0139
whc x $warm_{jf}$	*vcoldWhc*	($F_{1,85}$ = 406.09, $P < 0.0001$)	-0.00899	0.0004
Rainy Days (Jan., Feb.)	$rain_{jf}$	($F_{1,85}$ = 78.18, $P < 0.0001$)	-0.03662	0.0041
Quadrat (random effect)	e_q	(χ^2 =5442.66, df = 1, $P < 0.0001$)	0.9787	0.2116

Table 2.5. Seedling Environmental Model: Parameter estimates from GLMM of seedling survival from April until June of the first year with binomial errors and logit link. Model ranking was evaluated with QAICc. Akaike weight = 0.958, ΔQAICc of next best model = 7.8. Random quadrat effect was evaluated with a likelihood ratio test by comparing the change in the -2ln(l) from dropping quadrat from the model to a mixture of χ^2 distributions in GLIMMIX (SAS Institute 2008).

Parameter	Symbol	Test	Estimate	SE
Intercept	μ	NA	0.3937	0.8397
ln(Seedling Maximum Density)	$d_{sdl\,(y)}$	($F_{1,89}$ = 14.74, P = 0.0002)	-0.232	0.06043
Precipitation *(March,* mm)	p_{mar}	($F_{1,89}$ = 59.78, P < 0.0001)	-0.01301	0.00168
Warm Days (*Jan.-Feb.*)	$warm_{win}$	($F_{1,89}$ = 148.12, P < 0.0001)	-0.1238	0.01018
Hot Days (*Summer*)	hot_{sum}	($F_{1,89}$ = 241.14, P < 0.0001)	0.6021	0.03878
Rainy Days *(Oct-Dec)*	$rain_{fall}$	($F_{1,89}$ = 106.23, P < 0.0001)	0.09505	0.00922
hot_{sum} x $rain_{fall}$	$hsrf$	($F_{1,89}$ = 158.74, P < 0.0001)	-0.01661	0.00132
tN	tN_q	($F_{1,89}$ = 31.46, P < 0.0001)	0.07959	0.01419
ln[W.H.C. (1/3 Bar)]	$whc3_q$	($F_{1,89}$ = 21.31, P < 0.0001)	-0.8123	0.176
Quadrat (random effect)	e_q	($\chi 2$ = 404.74, df = 1, P < 0.0001)	0.5760	0.1409

Table 2.6. Summer Environmental Model (s_{sum}): Probability of a rosette surviving from June to October. Summer survival GLMM with binomial errors and logit link. Akaike weight = 0.993, ΔQAICc of next best model = 10.01. Random quadrat effect was evaluated with a likelihood ratio test by comparing the change in the -2ln(l) from dropping quadrat from the model to a mixture of χ^2 distributions in GLIMMIX (SAS Institute 2008).

Parameter	Symbol	Test	Estimate	SE
Intercept	μ	NA	-2.1092	1.6265
ln(June Rosette Density)	d_{June}	($F_{1,85}$ = 15.7, P = 0.0002)	-0.2951	0.07447
Rainy Days (*May*)	$rain_{May}$	($F_{1,85}$ = 21.61, P < 0.0001)	0.2825	0.06078
Precipitation (*May*, mm)	p_{may}	($F_{1,85}$ = 92.34, P < 0.0001)	0.09125	0.009496
$rain_{may}$ x p_{may}	$rmpm$	($F_{1,85}$ = 99.24, P < 0.0001)	-0.00741	0.000744
Soil % Sand	$sand_q$	($F_{1,85}$ = 24.94, P < 0.0001)	0.1100	0.02203
Rainy Days (*Spring*)	$rain_{spr}$	($F_{1,85}$ = 5.2, P = 0.0251)	0.0837	0.03672
Sand x Rain (*Spring*)	$sandxrain_{spr}$	($F_{1,85}$ = 14.6, P = 0.0003)	-0.00182	0.000475
Soil Total Inorganic N	tN	($F_{1,85}$ = 51.29, P < 0.0001)	-0.1814	0.02532
$rain_{may}$ x tN	rmN	($F_{1,85}$ = 53.16, P < 0.0001)	0.01072	0.00147
Hot Days (*Summer*)	hot_{sum}	($F_{1,85}$ = 49.52, P < 0.0001)	0.2580	0.03667
Drought Events (*Aug,Sept.*)*	dry_{sum}	($F_{1,85}$ = 199.02, P < 0.0001)	-0.5059	0.03586
Hot Days x Drought Events	hsd	($F_{1,85}$ = 31.42, P < 0.0001)	-0.02625	0.004683
Quadrat (random effect)	e_q	(χ2 =425.52, df = 1, P < 0.0001)	1.0298	0.2309

Table 2.7. Winter Environmental Model (s_{win}). Two-stage model of winter rosette survival (swin). Model $r_{win}1$ predicts the binary probability of observing an extreme value (either 0 or 100% survival) versus any other intermediate value and determines whether an observation proceeds to $s_{win}2a$ or $s_{win}2b$. Higher predicted probabilities from $r_{win}1$ are more likely to have extreme values and are passed on to $s_{win}2a$. Model $s_{win}2a$ predicts the binary probability of observing 0% survival versus 100% survival, conditional on knowledge that the outcome is one of these. Model $r_{win}2b$ is a binomial GLMM that predicts the survival probability of a quadrat that does not have an extreme value. For the first two binary models, the response coded as the "event" whose probability was modeled is indicated. Akaike weights = 0.343, 0.904, and 0.865 for the three models, respectively. ΔAICc of next best models = 1.4, 5.4, and 3.7, respectively.

Parameter	Symbol	Test	Estimate	SE
Model $r_{win}1$ event = extreme				
Intercept	μ	na	11.5502	3.0308
ln(Oct. Rosette Density)	d_{oct}	($F_{1,118} = 28.9$, $P < 0.0001$)	-2.5055	0.4661
Rainy Days (Summer)	$rain_{sum}$	($F_{1,118} = 3.95$, $P = 0.0492$)	-0.0775	0.0390
Model $r_{win}2a$ event = 100% surv.				
Intercept	μ	na	19.1297	9.2316
ln(October Rosette Density)	d_{oct}	($F_{1,33} = 4.99$, $P = 0.0324$)	-10.0220	4.4877
Growing Degree Days (*Spring*)	gd_{spring}	($F_{1,33} = 4.84$, $P = 0.0350$)	-0.07713	0.03507
d_{oct} x gd_{sdl}	$dr_{sum(i)}$	($F_{1,33} = 5.58$, $P = 0.0242$)	0.03738	0.01582
Model $r_{win}2b$				
Intercept	μ	na	5.0221	3.7513
Summer Rosette Survival	s_{sum}	($F_{1,35} = 17.19$, $P = 0.0002$)	1.3694	0.3428
Rainy Days (*Summer*)	$rain_{sum}$	($F_{1,35} = 9.00$, $P = 0.0050$)	-0.02786	0.06026
Soil pH	pH_q	($F_{1,35} = 9.28$, $P = 0.0044$)	-0.3706	0.1298
Cold Days *(Winter)*	$cold_{win}$	($F_{1,35} = 5.07$, $P = 0.0307$)	-0.02085	0.009443
Quadrat (random effect)	e_q	($\chi^2 = 150.4$, df = 1, $P < 0.0001$)	0.4366	0.1188

Table 2.8. Fecundity Environmental Model (*f*). GLMM with Poisson distributed errors, log link, and random quadrat effects. Model ranking was evaluated with AICc. Akaike weight = 0.999, ΔAICc of next best model = 16.3. Random quadrat effect was evaluated with a likelihood ratio test by comparing the change in the -2ln(*l*) from dropping quadrat from the model to a mixture of χ^2 distributions in GLIMMIX (SAS Institute 2008).

Parameter	Symbol	Test	Estimate	SE
Intercept	μ	NA	7.4679	1.1246
ln(June Rosette Density)	d_{June}	($F_{1,39}$ = 0.23, P = 0.6350)	-0.0497	0.1039
Hot Days (*Summer*)	hot_{sum}	($F_{1,39}$ = 25.56, P < 0.0001)	0.4391	0.08685
d_{June} x hot_{sum}	$d_{June}hot_{sum}$	($F_{1,39}$ = 20.61, P < 0.0001)	-0.0705	0.01553
Rainy Days (*Fecundity*)	$rain_{fec}$	($F_{1,39}$ = 21.72, P < 0.0001)	-0.1212	0.0260
ln(Flowing Adult Density)	d_f	($F_{1,39}$ = 19.08, P < 0.0001)	-1.6401	0.3754
d_f x $rain_{fec}$	$d_f rain_{fec}$	($F_{1,39}$ = 20.26, P < 0.0001)	0.0355	0.007875
Quadrat (random effect)	e_q	χ^2 = 181.18, df = 1, P < 0.0001	0.2701	0.07244

Table 2.9. Comparison of plant densities used in four studies of *A. petiolata*. Seedling, Summer, and Winter refer to survival measurements. Plant minimum and maximum densities have been converted to common units of plants m^{-2}. Geometric mean densities and 95% confidence intervals of the geometric mean are shown. The mean seedling density from Pardini et al. (2009) is the arithmetic mean and is based on their report of marking 469 seedlings across 40 1x1 m quadrats.

Transition	Time Interval	Min plants m^{-2}	Max plants m^{-2}	Mean (LCL UCL)	Scale of Measurement	Location	Ref.
Seedling	March-June	75	6025	806.9 (693.7-938.6)	0.04-0.0625 m^2	IL & MI	1
Seedling	April-May	-	-	11.725	1 m^2	MO	3
Seedling	March-October	207	2058	-	0.1 m^2	PA	2
Summer	June-October	31.25	3344	253.4 (218.0-294.5)	0.16 & 0.25 m^2	IL & MI	1
Summer	May-August	4	235	-	1 m^2	MO	3
Winter	October-June	4	540	47.1 (37.2-59.6)	0.16 & 0.25 m^2	IL & MI	1
Winter	August-May	1	278[a,b]	6.0 (4.1-8.7)	1 m^2	MO	3
Winter	October-June	4.47	108.2	-	1 m^2	OH	4
Fecundity	June	4	540	26.3 (21.1-32.9)	0.16 & 0.25 m^2	IL & MI	1
Fecundity	June-July	1	50[a]	4.4 (3.4 5.7)	1 m^2	MO	3

References cited are: 1) This study, 2) Winterer et al. (2005), 3) Pardini et al. (2009), and 4) Rebek and O'Neil (2006).

[a] Means and confidence intervals from Pardini et al (2009) were obtained by extracting x and y pixel coordinates from each data point in their printed figures 2B and 2C using photo editing software. Because they reported sampling winter survival in 34 quadrats but only 18 points are visible in their Figure 2B, I inferred that the remaining 16 quadrats had rosette densities of 1 plant m^{-2}.

[b] It appears that Pardini et al. (2009) mistakenly reported that 1795 rosettes were used in calculating winter survival. Data extracted from their figure 2B reveal that there were approximately 1340 rosettes used in this calculation. This corresponds closely with the 1346 rosettes they report marking for measurements of summer survival. Thus, I interpret that the number of plants indicated in the text for summer and winter rosette survival must be reversed in the text on page 390 of their paper. See Chapter 3.

Figures: Chapter 2

Figur

and Fi

Ives R

Forest

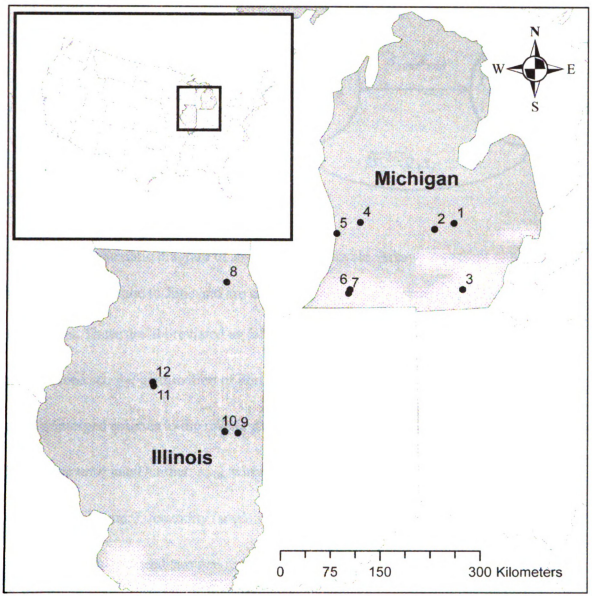


Figure 2.1. Locations of seven *A. petiolata* study sites established in Michigan in 2004 and five sites established in Illinois in 2005. Site Key: 1) Shiawassee, 2) Rose Lake, 3) Ives Road, 4) Johnson Park, 5) Holland State Park, 6) Edward Lowe Foundation, 7) Russ Forest, 8) Healy Road, 9) Homer Lake, 10) Illini Plantations, 11) Farmdale, 12) Peoria.

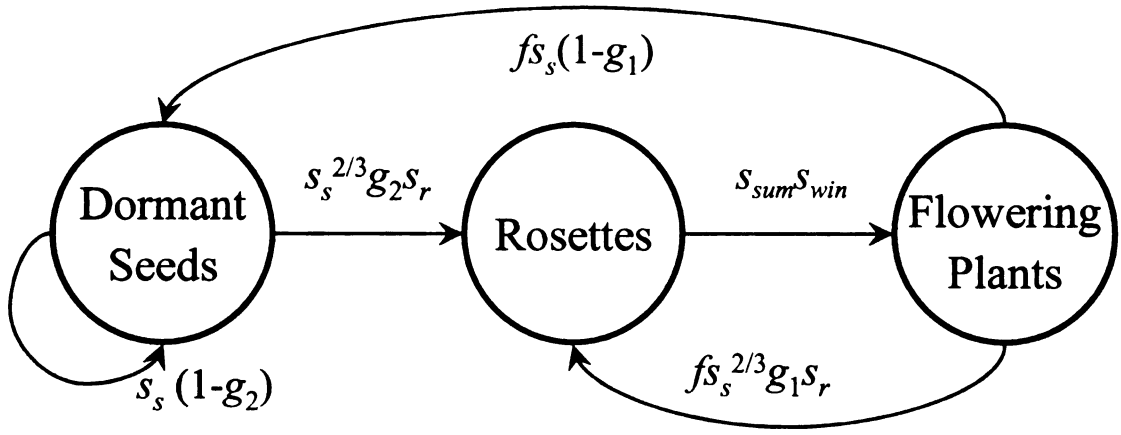


Figure 2.2. Schematic diagram of *A. petiolata* life cycle. Arrows represent one-year transitions from June to June and are comprised of multiple lower level demographic transitions. These are abbreviated as follows: g_1, germination of new seeds within one year of seed set; g_2, germination of dormant seeds from the soil seed bank; s_r survival of newly emerged rosettes to the rosette stage in June; s_{sum}, summer survival of new rosettes from June until late October; s_{win}, winter survival of rosettes to the flowering stage from October until June; f, fecundity (seeds/plant); s_s, survival of dormant seeds in the soil seed bank. Because seed survival (s_s) was measured over a full year but is used twice as an 8 month, sub-annual transition, s_s is raised to the two-thirds power in the seed to rosette and the flowering plant to rosette transitions to scale its affect.

Figure 2.3. Observed mean *A. petiolata* demographic rates over the study period. Light gray lines follow mean values within individual sites, averaged across quadrats. Heavy black lines are mean values (±SEM) of these site means. Year indicates the year during with each measurement was begun. For example, winter survival 2005 was measured from fall of 2005 until June of 2006. The three years constitute the three "cohorts" of plants as grouped in the study, i.e. 2006 germination 1 and 2005 summer survival are grouped together. *Values of germination 2 from in 2009 are estimations (Appendix 2.B.v).

Germination 1
Germination 2
Seedling Survival
Fecundity (seeds/plant)
Summer Rosette Survival
Winter Rosette Survival
Dormant Seed Survival
Year
2006
2007
2008
2009*
2005

Figure 2.4. Frequency distributions of *A. petiolata* demographic rates from this study. Data shown are raw quadrat level observations. Rosette to flowering plant survival (s_{rf}) is calculated as r_{sum}*r_{win} for comparison with previous studies that did not split summer and winter survival. Letters beneath histograms show observations from previous studies. References: *a* (Pardini et al. 2008); *b* (Pardini et al. 2009); *c* (Anderson et al. 1996); *d* (Meekins and McCarthy 2002); e (Drayton and Primack 1999); *f* (Nuzzo and Blossey unpublished data); *g* (Baskin and Baskin 1992); and *h* (Cavers et al. 1979). Overlapping observations are shown as: *i* (Anderson et al. 1996, Meekins and McCarthy 2002); *j* (Drayton and Primack 1999, Pardini et al. 2009, Nuzzo and Blossey unpublished data); *k* (Cavers et al. 1979, Meekins and McCarthy 2002).

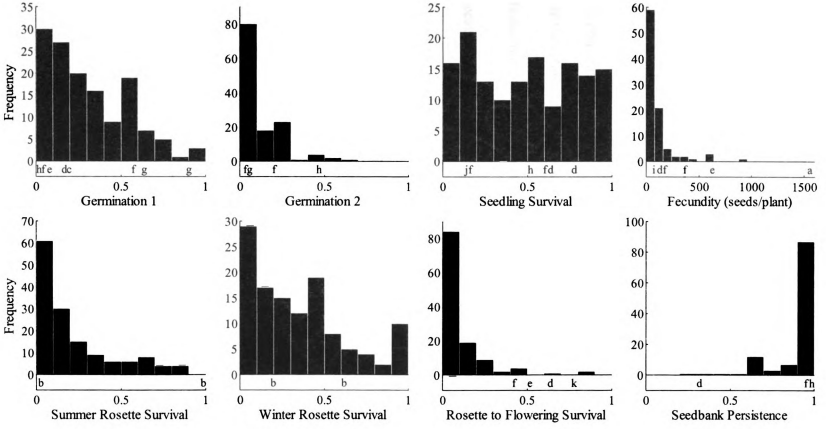

Frequency
Germination 1
Germination 2
Seedling Survival
Fecundity (seeds/plant)
Summer Rosette Survival
Winter Rosette Survival
Rosette to Flowering Survival
Seedbank Persistence

Figure 2.5. Predicted versus observed values of *A. petiolata* seed bank viability (s_s), germination after 1 winter (g_1), seedling to rosette survival (s_r), summer (s_{sum}) and winter (s_{win}) rosette survival, *per capita* fecundity (f), and germination of dormant seeds from the seed bank after two winters (g_2). Fitted values are least squares means estimates for each site by year combination predicted from the hierarchical GLMs and GLMMs. Observed values are simple arithmetic means except f. Silique counts were first averaged across plants within quadrats, then across quadrats for each site and year, and finally scaled to show the estimated number of seeds per plant. Fitted f, values were modeled as silique counts and then scaled to generate seed estimates. Only the first two years of g_2 are shown, because the third year was estimated from years one and two. Unadjusted R^2 are shown because the site by years means are shown (marginal model predictions), whereas the model was fit with random quadrat effects. All R^2s and Adjusted R^2 in this and subsequent figures are calculated in the original data scale.

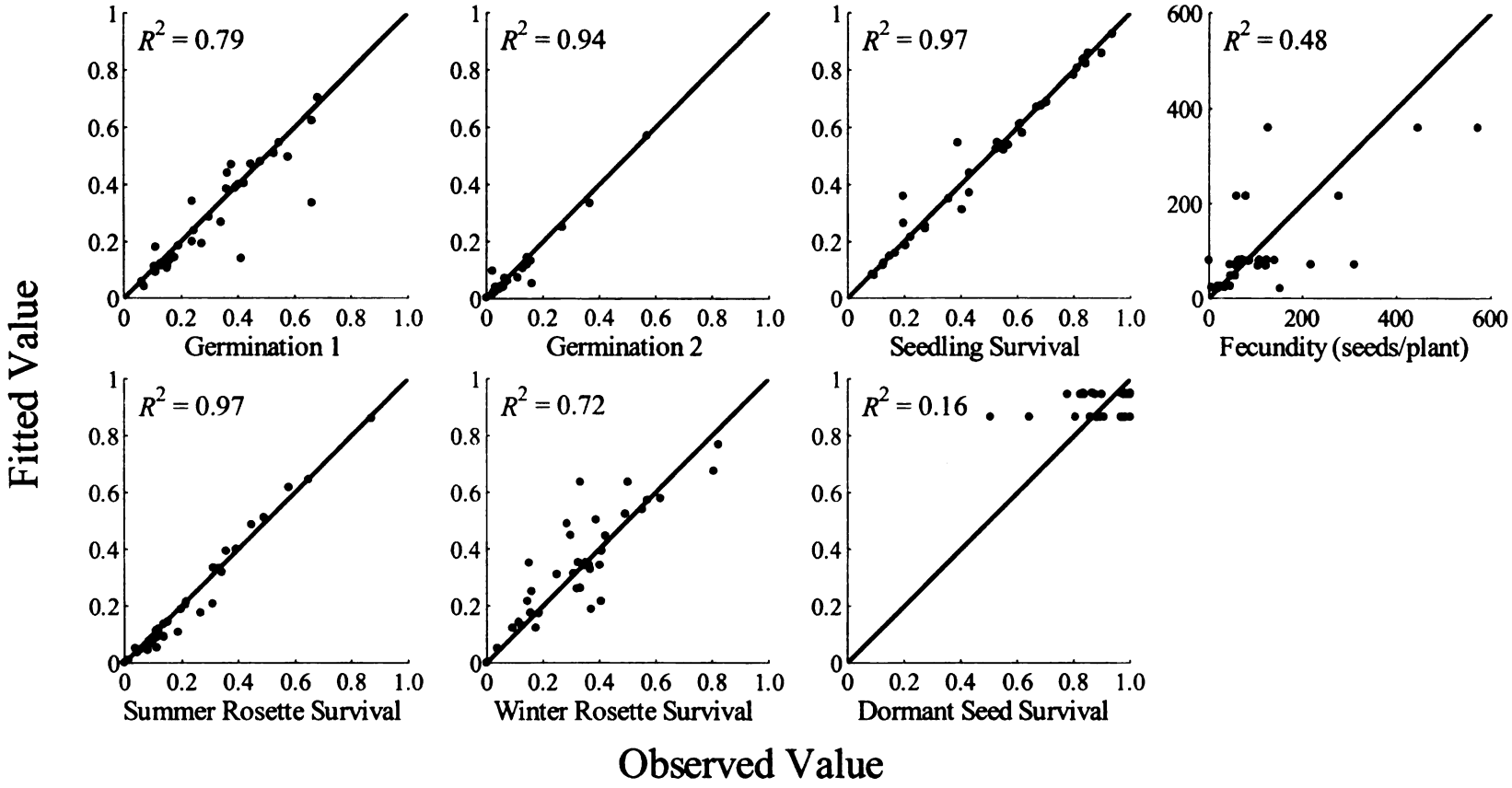

Fitted Value
$R^2 = 0.79$
Germination 1
$R^2 = 0.94$
Germination 2
$R^2 = 0.97$
Seedling Survival
$R^2 = 0.48$
Fecundity (seeds/plant)
$R^2 = 0.97$
Summer Rosette Survival
$R^2 = 0.72$
Winter Rosette Survival
$R^2 = 0.16$
Dormant Seed Survival
Observed Value

Figure 2.6. Pred

scatterplots sho

was overlaid w

predicted proba

(i.e. either 0

versus the actu

observations t

category had

probabilities

mean rosette

observed prob

observation b

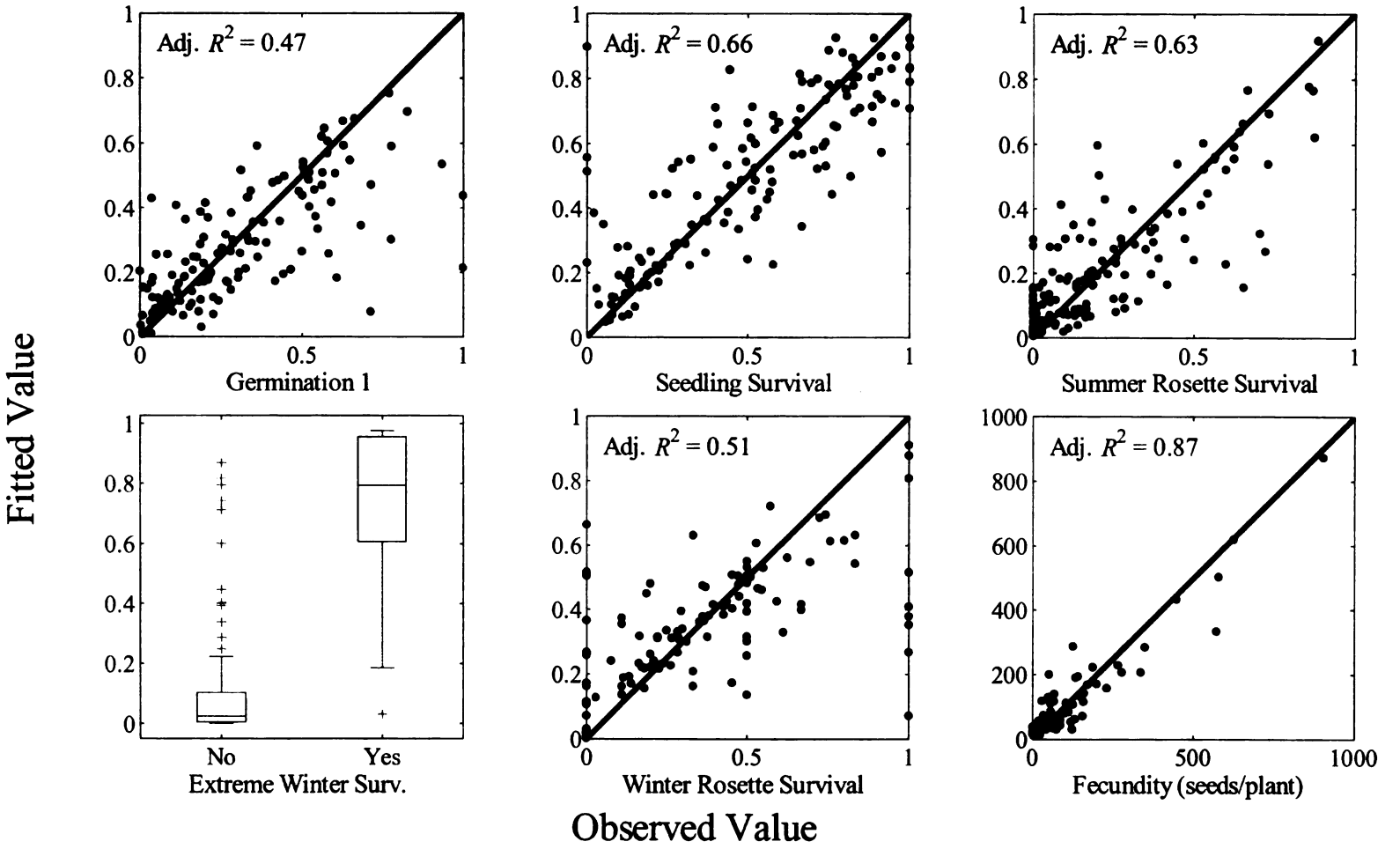
outcome will

observed proba

obtained from

regression fun

Figure 2.6. Predicted versus observed values from environmental models. The five scatterplots show the demographic rates generated from the BLUPs versus the observed rates, overlaid with a 1:1 reference line. The vertical axis of the box plot shows the predicted probability of a winter rosette survival observation resulting in an extreme value, either 0 or 100% survival, versus a non-extreme value from the $r_{win}1$ binary GLM versus the actual outcome. Observed extreme values were coded as "Yes" and include all observations that had either zero or 100% percent survival. Observations in the "No" category had intermediate survival rates. A well fitting model should predict higher probabilities for the "Yes" group and lower the probabilities for the "No" group. The winter rosette survival plot shows the combined results of models $r_{win}2a$ and $r_{win}2b$. For observed probabilities of 0 and 1 the predicted values represent the probability of an observation being 1, versus the alternative of it being 0, conditional on knowing that the outcome will be one of these two extreme values. The remaining points are predicted survival probabilities as in the upper row of plots. Mean per capita fecundity was estimated from the predicted and observed numbers of siliques using the breakpoint regression function in Appendix 2.B.iv.


Fitted Value
Adj. $R^2 = 0.47$
Germination 1
Adj. $R^2 = 0.66$
Seedling Survival
Adj. $R^2 = 0.63$
Summer Rosette Survival
No
Yes
Extreme Winter Surv.
Adj. $R^2 = 0.51$
Winter Rosette Survival
Adj. $R^2 = 0.87$
Fecundity (seeds/plant)
Observed Value

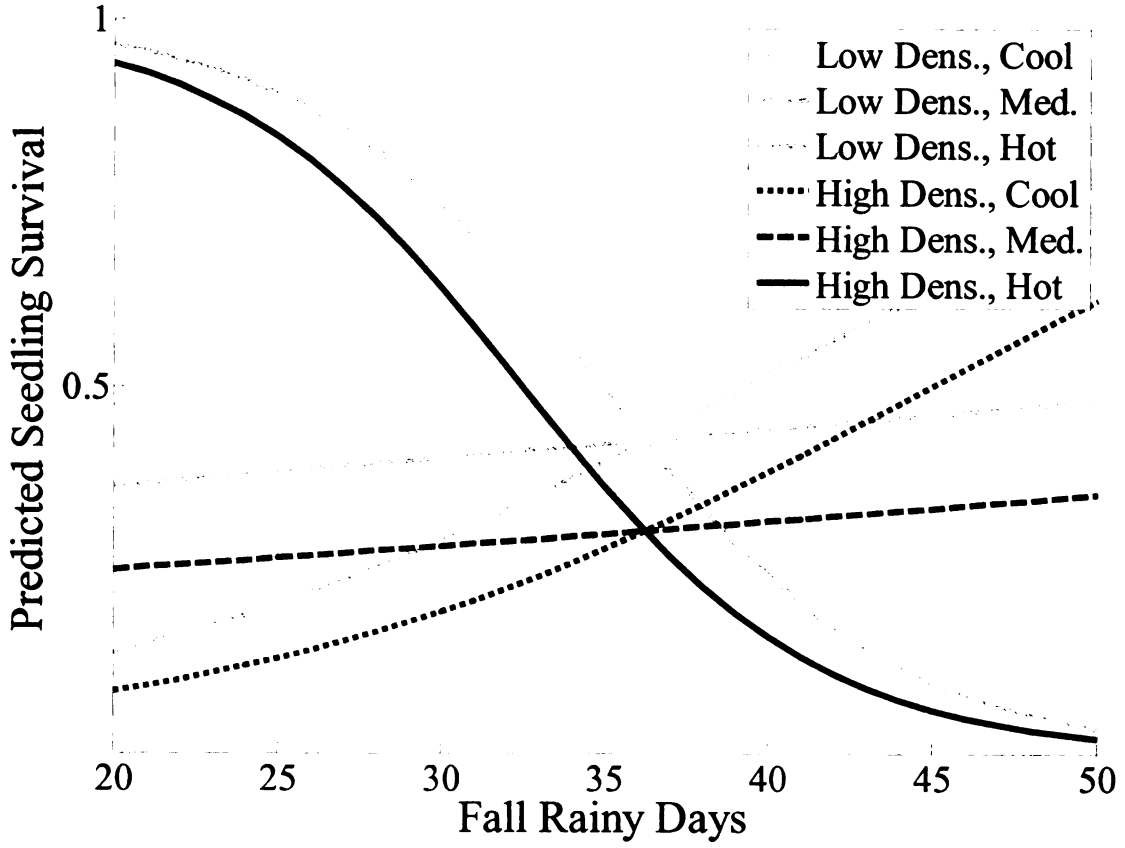


Figure 2.7. Predicted seedling survival is conditioned by interactions among multiple extrinsic and intrinsic factors. Here, the expected seedling survival probability is calculated from the best supported seedling model across the observed range of fall rainy days ($rain_{fall}$), during the year before germination at the minimum (cool), mean (med.), and maximum (hot) number of observed hot days during the summer before germination. The figure illustrates the interaction between population density, spring temperature (hot_{sum}), and fall rain at population densities of 500 (gray) and 5000 (black) seedlings m^{-2}. Mean observed values were used for all other variables. The change in slope from negative to positive results from the interaction between $rain_{fall}$ and hot_{sum}.

APPENDIX 2.A: SOIL DATA

T

Site	Soil USCS	Sand	Silt	WHC	OC	P	K	Mg	Ca	Mn	Zn

Table 2.A.1 Mean soil parameters ± SEM. Site abbreviations in Table 2.1

Site	Soil H2O	Sand	Silt	W.H.C. -1/3 bar	OC	P	K	Mg	Ca	soilpH	NO3	InorgN
B	33.8 (4.8)	43 (2.4)	38 (1.8)	19 (1)	31.8 (3.1)	10.8 (0.6)	5.7 (0.5)	8 (1.8)	96.5 (23.6)	270 (68.9)	1387.5 (263.3)	7.5 (0.3)
ELF	27.2 (1.8)	63 (1.3)	29 (1.3)	8 (1.2)	19.5 (1.1)	6.5 (0.6)	4.8 (0.6)	32.8 (7.3)	60.3 (11.4)	110 (20.7)	962.5 (142)	5.7 (0.4)
F	35.6 (1.7)	43 (4.3)	42.3 (2.4)	14.7 (1.9)	30.5 (2.3)	9.7 (0.8)	5.6 (0.4)	22 (1.9)	124.3 (17.6)	248.3 (28.3)	1516.7 (117.9)	6.8 (0.1)
HL	25.3 (1.3)	27.5 (12.2)	51.5 (10.8)	21 (2.4)	34.1 (1.8)	8.6 (0.6)	4.5 (0.7)	10.8 (0.5)	63.8 (15)	255 (76.7)	1287.5 (293.2)	6.2 (0.3)
HR	27.5 (7.4)	29.8 (0.9)	45.3 (3.3)	25 (3.1)	35.2 (2.8)	16.7 (1.8)	8.5 (1)	39 (6.3)	257 (40.2)	606.3 (39.4)	2200 (167.1)	7.2 (0.2)
HSP	6 (1.4)	89.8 (1.3)	8.8 (1.2)	1.5 (0.5)	4.2 (1.1)	2.3 (0.7)	2.4 (0.5)	13.3 (2.6)	24.3 (0.9)	146.3 (23.6)	925 (143.6)	7.1 (0.2)
IP	30.8 (3.7)	31.3 (9.9)	43.3 (6.1)	25.5 (3.9)	33.6 (0.7)	13.8 (0.5)	6.9 (0.5)	29 (5.9)	163.8 (28.6)	233.8 (29.5)	1562.5 (198.3)	5.3 (0.3)
IR	28.3 (8.3)	64.3 (8.6)	22.5 (3.9)	13.3 (4.8)	24.6 (6.4)	11.2 (3.8)	7.6 (2.3)	14 (3.7)	88 (12.7)	240 (56.2)	1600 (307.5)	6.6 (0.2)
JP	14.2 (1.5)	84 (2.4)	10.5 (1.5)	5.5 (1)	11.3 (2.2)	3.5 (0.3)	3.4 (0.5)	50 (10.3)	43 (6)	77.5 (14.8)	500 (97.9)	5.4 (0.3)
RF	19.1 (1.4)	77.5 (4.8)	17.5 (3.9)	5 (1)	13 (2.2)	4.8 (0.4)	4.4 (0.4)	39.8 (5.1)	49.8 (6.2)	67.5 (14.8)	487.5 (68.8)	4.9 (0.4)
RL	30.6 (8.3)	68.5 (2.6)	22.5 (2.2)	9 (1.3)	29.6 (4)	11.1 (4.3)	7.6 (2.9)	9.8 (2.3)	45.5 (5.2)	271.3 (99.2)	1450 (558.6)	6.1 (0.4)
SH	19 (1.3)	72.5 (1.7)	19.5 (1.7)	8 (0)	20 (1.1)	4.6 (0.5)	3.4 (0.4)	16 (0.9)	79.5 (7.7)	98.8 (3.8)	687.5 (23.9)	5.7 (0.1)

APPENDIX 2.B: STATISTICAL MODEL FITTING NOTES

2.B.i. Model Fitting

Maximum likelihood (ML) and the Laplace approximation of the maximum likelihood were used to fit GLMs and GLMMs for model selection, respectively. Akaike's Information Criterion with a correction factor for small sample bias (AICc) was then used rank models by their relative support from the data as described by Anderson (2008). AICc is based on Kullback-Leibler information and is an estimate of the distance, in arbitrary units of "information", from a model to the full reality of the processes or factors driving a system. Models which better describe the variation in the data have lower AICc scores and are interpreted to more closely approximate reality. Competing models of each vital rate were compared with ΔAIC (Δ_i), calculated for each model i as the difference between its AICc and that of the best supported model. Thus, the best supported model always has a Δ_i of 0. Models with Δ_i less than approximately 10 are considered plausible alternatives to the best model (Anderson 2008). The Akaike weight (w_i) is the probability that model i is actually the best model within the set. The w_is are estimated from the Δ_is and sum to 1 within each set of competing models.

Models of s_{win}, f, g_2, and s_s were evaluated using AICc. Models of g_1, s_r, and s_{sum} were overdispersed. In models with binomial errors the expected variance is a function of the mean, μ, expressed as $\mu(1-\mu)$. Overdispersion occurs when the observed variance exceeds this expectation. This can indicate that important covariates are missing from the model, poor model specification, or an inappropriate error distribution (Crawley 2007). Although overdispersion often has little effect on parameter estimates (e.g. slopes, intercepts), it can cause underestimation of parameter standard errors (Joe and Zhu 2005)

and the

on the (

way as

Pearson

estimat

for mod

(PQL) a

distribu

overdisp

estimate

best sup

have wer

overdisp

be interp

using Q

although

distribut

which pr

residual

are model

and thereby inflate the type I error rate. Selection among overdispersed models was based on the Quasi-Akaike Information Criterion (QAICc), which is interpreted in the same way as AICc but also includes a scale parameter in its calculation, □, calculated as the Pearson's χ^2/df statistic (Bolker et al. 2009). For each overdispersed vital rate □ was estimated from the full model and applied to the QAICc calculation of all reduced models for model selection.

Overdispersed hierarchical models were then refit using pseudo quasi-likelihood (PQL) estimation where possible. This allowed using an overdispersed quasibinomial distribution by fitting multiplicative scale parameter □ to the variance structure of the overdispersed models using the random _residual_ statement in GLIMMIX. Parameter estimates and standard errors for g_1 and s_r are derived from these re-fit PQL models. The best supported hierarchical model of s_{sum}, would not converge using PQL. Parameters for s_{sum} were estimated using the Laplace approximation and are uncorrected for overdispersion. The standard errors for this model may be underestimations and should be interpreted with caution.

The overdispersed environmental models of g_1, s_r and s_{sum} were fit and ranked using QAICc scores calculated from the Laplace-approximated maximum likelihoods. Although I then refit the best supported model for each vital rate with a quasibinomial distribution that corrects for overdispersion, SAS can only fit this distribution using PQL, which produces less accurate parameter estimates than the Laplace approximation. The residual diagnostics of the quasibinomial environmental models were poor, and the fit of the model predictions to the observed data was much worse than the overdispersed

Laplace estimated models. All final environmental models therefore were fitted with the Laplace approximation (= better parameter estimates) but were uncorrected for overdispersion (= possible underestimation of parameter standard errors). Type-3 F-tests for parameter estimates may have inflated risks of type-1 errors and should be interpreted conservatively. The significance of random quadrat effects was tested for each model with a likelihood ratio test. The residual deviance explained by the full GLMM was compared to the residual deviance of a reduced GLM with no random terms. The difference in deviance explained was compared to a mixture of χ^2 distributions with 1 degree of freedom using the covtest / glm option in GLIMMIX to test the null hypothesis that the random effects did not improve the amount of deviance explained by the model.

2.B.ii. "Pretending variables"

Anderson (2008) describes a scenario that arises in model selection in which two competing models spuriously appear to have nearly equal support from the data. In these cases, the ln(l) scores of both models will be nearly identical, and the second model will differ from the better supported model only by having one or more additional independent variables that parameters do not explain any additional variance in the response. Because AIC scores are based on the ln(l) scores plus a penalty for the number of parameters estimated, the difference in AIC between these two models will be equal to the penalty imposed for the extra parameters. The Δi of the competing model may suggest at first that it is a reasonable alternative to the reduced model. However, if the AIC differ only by the penalty for fitting extra parameters but the ln(l) are equal, the extra variables do not actually improve the model fit. Anderson (2008) refers to this as the

"pretending variable" phenomenon, where a parameter with an effect size of zero is "pretending" to improve model fit. Models with pretending variables are identified in the text when they occur.

2.B.iii. Conditional Modeling of Seed Survival

Because the seed survival date contained a high proportion of ones, the data were analyzed in two steps. First, the probability π of observing 100% seed viability in a sample was modeled as a Bernoulli process: each quadrat-level estimate of s_s with 100% seed viability was coded as a one, and all other observation were coded as zeros. The recoded binary data were then fit with a set of binary GLMs with different competing fixed effects structures. Second, I modeled the subset of observations with values less than one using a beta distribution. Thus the process for simulating n new observations is to first generate a set of n random Bernoulli variables with probability π (determined in the binary GLM) of success (i.e. having a 100% viability rate). If s random trials are successes, then the remaining n - s trials have probabilities < 1 and are simulated from the fitted beta distribution. Tests of this simulation process faithfully reproduce the observed mixture distribution. This approach is similar to Cunningham and Lindenmayer's (2005) solution for modeling zero-inflated distributions.

2.B.iv. Fecundity estimation

Measuring per capita fecundity directly in the field was not feasible. Instead, I nondestructively counted the number of siliques on each surviving mature plant in the rosette quadrats. The number of siliques was then multiplied by the number of seeds per silique to estimate total per capita seed production.

The number of seeds per silique increased with plant size. To estimate the shape of this function, a set of 145 test plants was destructively harvested and the number of siliques and seeds per plant were counted. Test plants were collected from seven field sites in southern Michigan and Illinois in 2004 and 2005. Individual test plants ranged from having 0-266 siliques and from 0-3864 seeds.

Three plants had zero siliques and hence had zero seeds. The 142 remaining plants were analyzed to determine the relationship between total seeds per plant and total siliques per plant. The number of seeds and siliques per plant were each $\log_{10}$ transformed to stabilize the variance. Exploratory graphical analysis suggested a linear relationship with a break in the slope close to ten siliques per plant. I created a program in PROC NLMIXED in SAS version 9.2 (SAS Institute 2008) to fit a breakpoint linear regression which estimated the location of the break in slope as well as the two slopes and the intercept.

The fitted model had a maximum likelihood estimated breakpoint of 8 siliques per plant (Table 2.B.1) and had an adjusted R^2 of 0.975 (Figure 2.B.1). The estimated number of seeds per plant, after the back-transformation, in plants with 1-8 seeds is:

$$\log_{10}(seeds_{siliques\leq 8}) = \beta_0 + \beta_1 \times \log_{10}(siliques)$$
$$= 0.5860 + 1.5022 \times \log_{10}(siliques)$$
$$seeds_{siliques\leq 8} = 10^{0.5860+1.5022\times\log_{10}(siliques)}$$

where β_0 and β_1 are the slope and intercept estimates. Seeds in plants with more than 8 siliques are estimated as:

$$\log_{10}(seeds_{siliques>8}) = ((\beta_0 + \beta_1 \times \log_{10}(break)) + \beta_2 \times ...$$
$$((\log_{10}(siliques) - \log_{10}(break - 1))$$
$$= ((0.5860 + 1.5022 \times \log_{10}(8)) + 1.027 \times ...$$
$$((\log_{10}(siliques) - \log_{10}(8 - 1))$$
$$= (0.5860 + 1.5022 \times 0.9031) + 1.027 \times ...$$
$$(\log_{10}(siliques) - 0.8451)$$
$$= 1.9426 + 1.027 \times (\log_{10}(siliques) - 0.8451)$$
$$seeds_{siliques>8} = 10^{1.9426+1.027\times(\log_{10}(siliques)-0.8451)}$$

where *break* is the estimated breakpoint, β_2 is the slope, intercept is the maximum value predicted by the equation for smaller plants. Using this as the intercept in the second equation and subtracting $\log_{10}(break + 1)$ from the silique count ensures that the two functions transition smoothly into each other.

The Michigan plants used in this analysis were previously used to generate the seed-silique function in Evans and Landis (2007). This re-analysis includes additional test plants from Illinois and accounts for the nonlinearity between silique number and per capita fecundity in small plants. Complete SAS code and data for the analysis is provided below.

```
data allo;  input seeds siliques; cards;
0     0         61    4         118   10        181   18        656   49
0     0         45    5         137   10        201   18        786   50
0     0         51    5         146   10        138   19        869   50
1     1         53    5         155   10        222   19        816   52
1     1         55    5         148   10        270   20        817   54
2     1         57    5         135   11        281   20        850   54
2     1         75    5         140   11        264   21        661   55
3     1         49    6         165   11        296   21        891   56
4     1         52    6         181   11        309   21        874   64
5     1         67    6         183   11        257   22        811   65
7     1         93    6         131   12        357   25        1094  73
8     1         43    6         166   12        449   25        1031  75
6     2         38    7         175   12        328   25        839   76
10    2         44    7         201   12        368   26        1145  77
12    2         56    7         111   13        47    27        1339  79
17    2         73    7         200   13        346   27        1125  84
18    2         82    7         149   14        393   27        469   86
22    2         87    7         151   14        401   27        1527  91
23    2         94    7         208   14        341   30        1923  127
24    2         121   7         163   14        336   34        2091  151
20    3         15    8         199   15        373   35        2762  173
21    3         70    8         205   15        506   39        2724  209
21    3         83    8         214   16        598   41        3786  248
23    3         89    8         243   16        627   42        3718  249
19    4         97    8         139   17        565   43        3864  266
34    4         67    8         193   17        683   44        ;
37    4         86    8         237   17        772   45
40    4         104   9         308   17        663   46
42    4         107   9         231   17        685   46
47    4         116   10        61    18        623   48
data allo; set allo;
logsil=log10(siliques);
logseed=log10(seeds);
run;

title 'Fit Breakpoint Regression';
* drop plants with zero siliques;
proc nlmixed
data=allo(where=(siliques>0));

* range of breakpoints to try;
parms break 2 to 200;
if siliques <= break then
      * regression for smaller
plants;
    eta = b01 + b1*logsil;
else
      * regression for larger
plants;
    eta = (b01 + b1*log10(break))+
       b2*(logsil-log10(break-1));

* fit model;
model logseed ~ normal(eta,V);
* generate fitted values;
predict eta out=pred;
run;

* backtransform predicted values;
* calculate residuals;
data pred; set pred;
backpred = 10**pred;
resid = pred-logseed;
run;

* plot results;
title 'Predicted and Observed
Values versus Siliques';
axis1 order=(0 to 4500 by 500);
symbol1 color=black height=2;
symbol2 color=red value=dot;
proc gplot data=pred;
  plot seeds*siliques/
vaxis=axis1;
  plot2 backpred*siliques /
vaxis=axis1;
run; quit;

title 'Residual Plot';
proc gplot data=pred;
  plot resid*logseed;
run;quit;
```

2.B.v. Estimation of 2007 *G2*

Germination of *A. petiolata* seeds from the soil seed bank (g_2) after two winters of dormancy was measured for the first two experimental cohorts (2005-2006, and 2006-2007). No measurements of g_2 were made for the 2007-2008 cohort, which would have been sown in 2007 and germinated in 2009. g_2 for this third cohort was estimated from the relationship between g_1 and g_2 in previous years. A generalized linear mixed model was fit to the first two years of data as:

$$g_{2(y+1)} = g_{1(y+1)} + seeds_{2(y+2)} + \varepsilon_{site} \sim N(0,\sigma^2)$$

with binomially distributed errors, where y is the year that seeds were sown, g_1 is the proportion of seeds that germinated after one winter, $seeds_2$ is the estimated number of ungerminated seeds remaining in the soil in year $y+1$, estimated as the number of seeds added in year y – the number of seeds that germinated in year $y+1$, g_2 is the proportion of $seeds_2$ that germinated in year $y+2$, and ε_{site} is the random error due to site.

Because a number of data points were suspected of being inaccurate (e.g. were vandalized, damaged by animals, etc...) pooled parameter estimates were calculated from the remaining good observations from each site within each year. This was done by adding the number of individuals from all good quadrats at within a given time interval and then calculating summary values from these. For example, if three of four g_1 quadrats were deemed useable, pooled g_1 was calculated as the total number of seedling that emerged from the three good quadrats divided by the total number of seeds added to the three good quadrats. Using pooled values precludes analyses of density dependence, as

different individual quadrats had different densities of seeds and seedlings. The primary analyses of g_1 and g_2 in this study were conducted on the raw, quadrat level data, not the pooled values. Analyses of g_2 excluded the pooled values estimated here for 2009.

The fit of the model to the data is illustrated in Figure 2.B.2, and the parameter estimates and model fit statistics are given in Table 2.B.2. One observation (Homer Lake 2006) was an outlier and was omitted from model parameterization

Table 2.B.1 Parameter estimates from breakpoint linear regression of $\log_{10}$(seeds/plant) versus $\log_{10}$(siliques/plant). The breakpoint is the number of [untransformed] siliques above which the slope of the regression changes. The slope for plants with 1-8 siliques is given by β_1, while the slope for plants with more than 8 siliques is given by β_2.

Parameter	Estimate	*Standard Error*	*DF*	t Value	Pr > |t|
breakpoint	8	3.259	142	2.5	0.0
Intercept	0.586	0.052	142	11.3	<.0001
β_1	1.5022	0.090	142	16.7	<.0001
β_2	1.027	0.048	142	21.4	<.0001
V	0.03133	0.004	142	8.3	<.0001

Table

proba

Type-

Param
Interce
S
acade
Residu

Table 2.B.2. Parameter estimates from generalized linear mixed model of germination probability of dormant *A. petiolata* seeds with binomial errors and random site effects. Type-3 F-tests of parameter significance and parameter standard errors are given.

Parameter	Test	Estimate	SE
Intercept	na	6.6141	0.5907
g_1	($F_{1,9} = 141.37$, $P < 0.0001$)	-3.5725	0.3005
$seeds_2$	($F_{1,9} = 276.07$, $P < 0.0001$)	-1.3836	0.08327
Random Site	($\chi 2 = 363.54$, $df = 1$, $P < 0.0001$)	1.0847	0.2381

Seeds per Plant

Figure

dots) a

the cha

line).

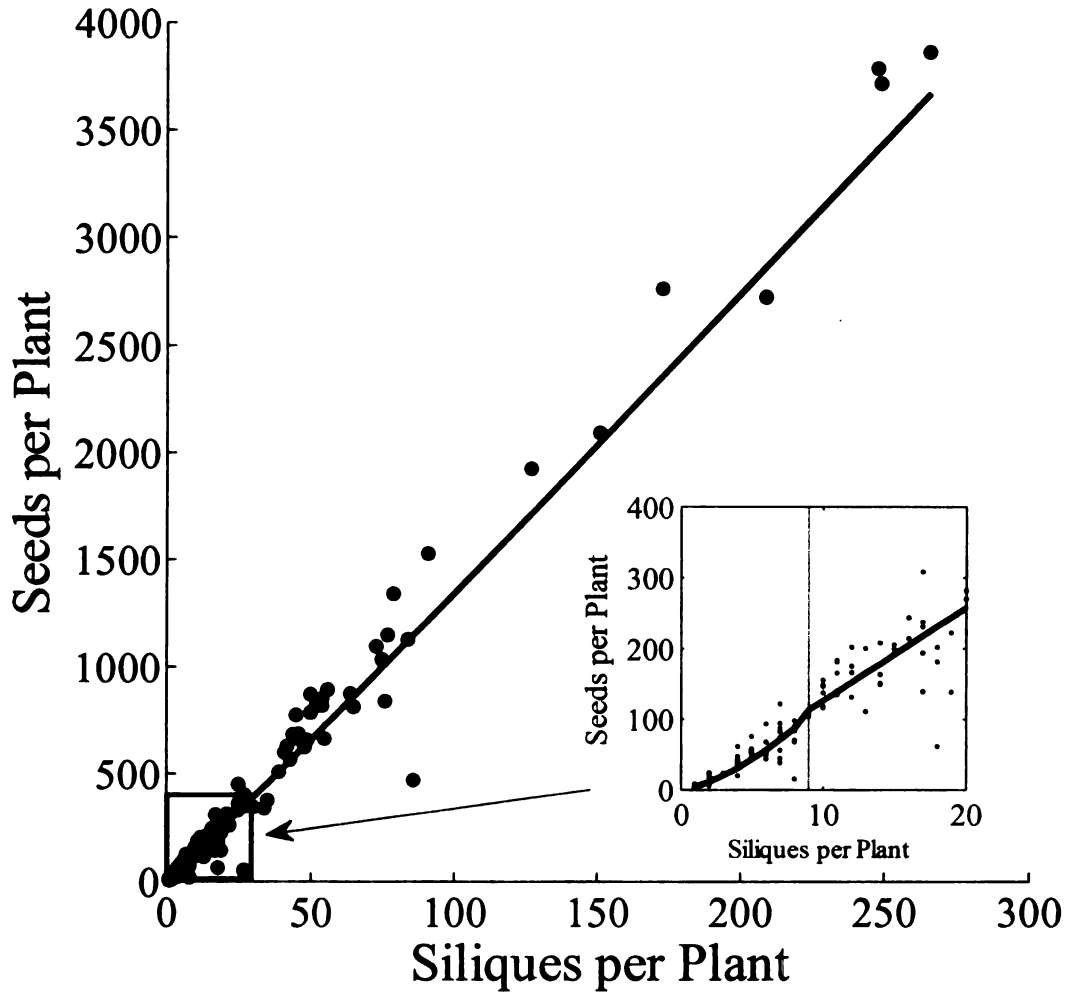


Figure 2.B.1. Number of seeds versus number of siliques in 142 *A. petiolata* test (black dots) and fitted values from breakpoint regression (black line). Inset detail graph shows the change in slope above the breakpoint value of 8 siliques per plant (vertical dashed line).

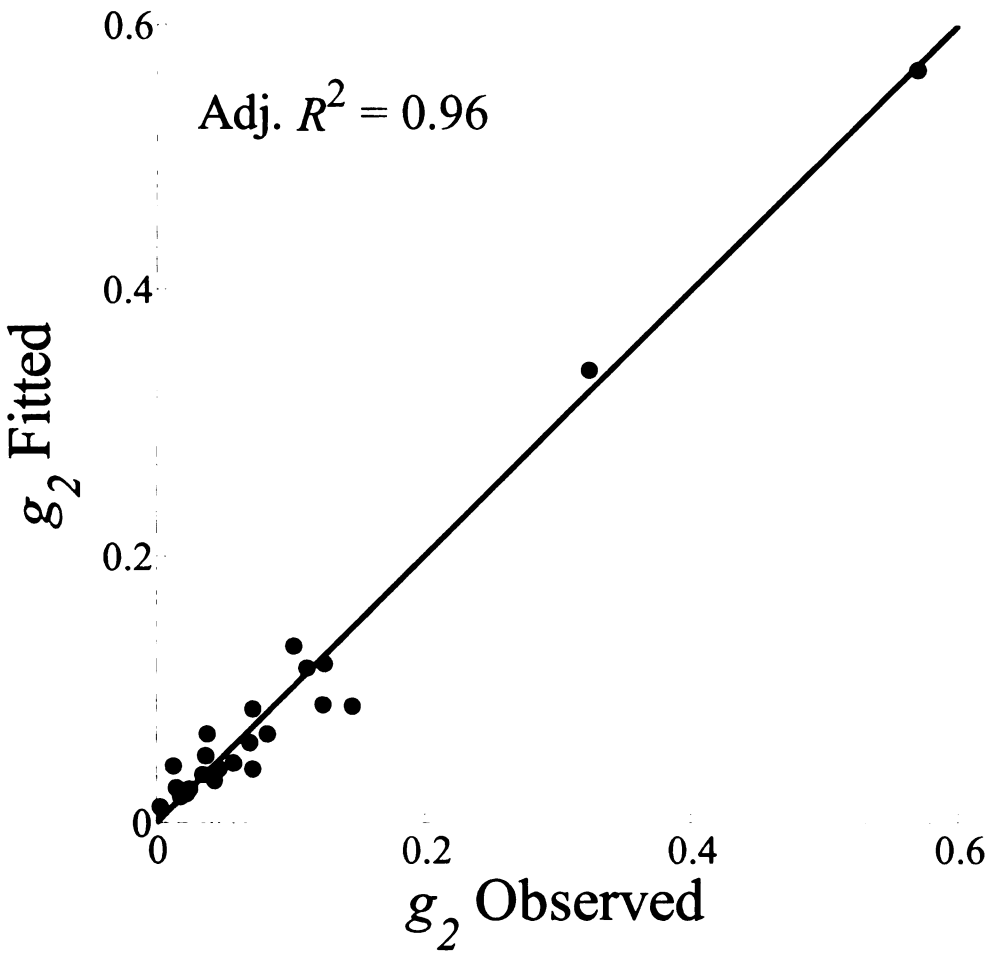


Figure 2.B.2. Fitted versus observed values of pooled *A. petiolata* germination rates from dormant seed (g_2) overlaid with 1:1 reference line. The first two years of data, which were used to fit the model, are shown. This relationship was applied to the third year of data to estimate the unknown values of g_2. Model was fit with binomially distributed errors and g_1 and $seeds_2$ as fixed effects plus random site effects.

APPENDIX 2.C: ENVIRONMENTAL MODEL PREDICTIONS

Figure 2.C.1. Observed *A. petiolata* demographic rates versus up to five independent variables from the best supported environmental models.

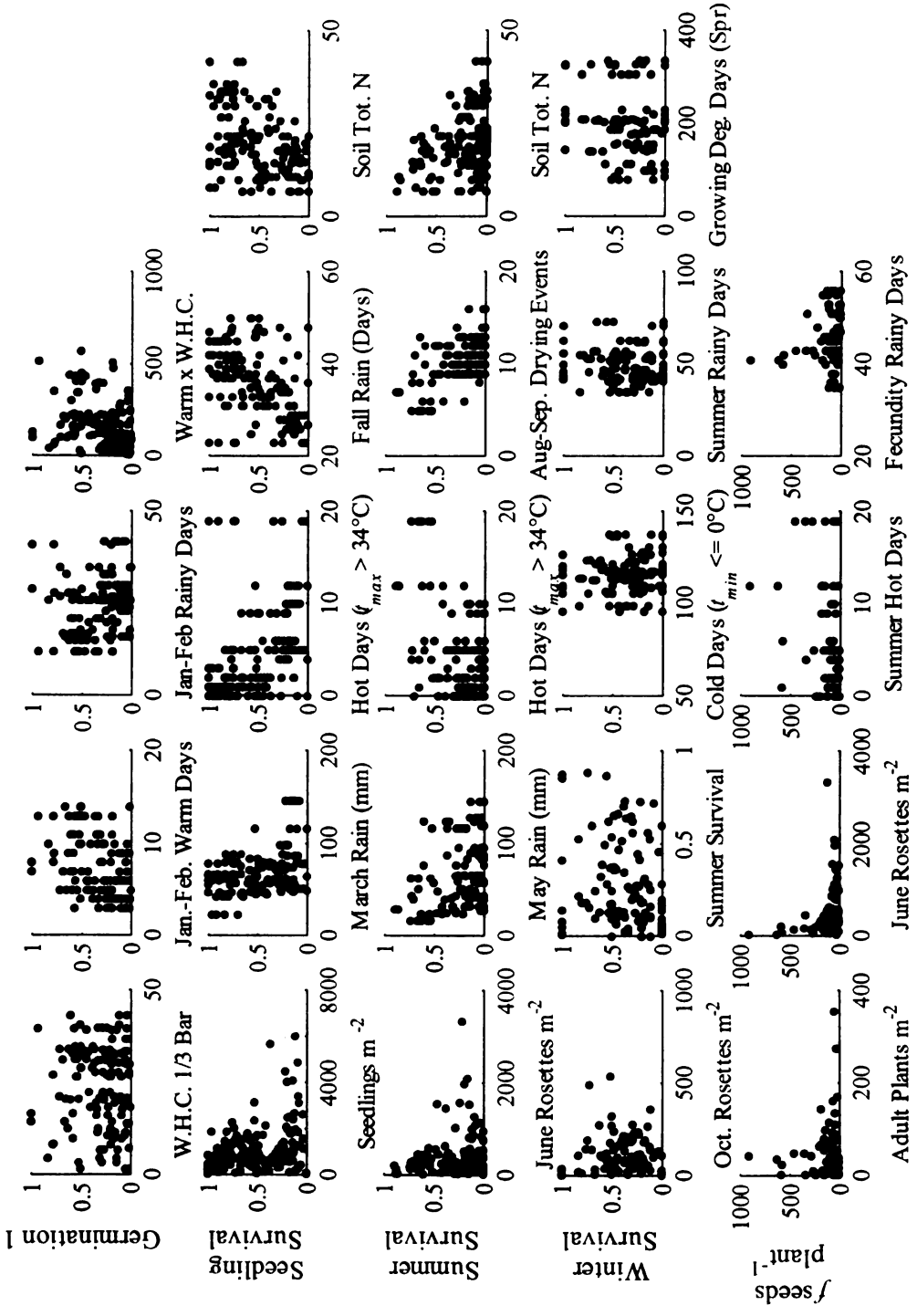
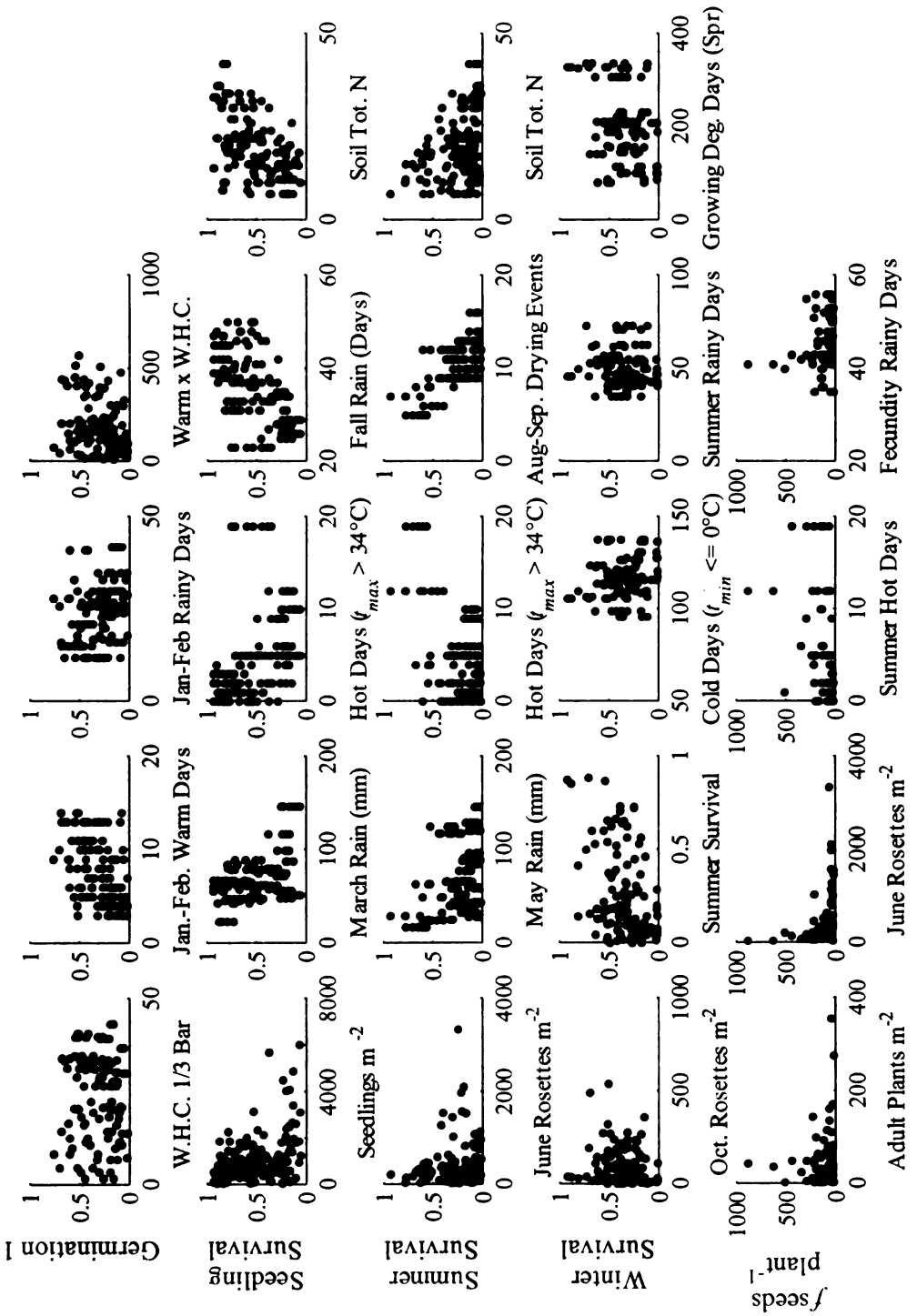


Figure 2.C.2. Predicted *A. petiolata* demographic rates versus the same independent variables in Figure D1 from the best supported environmental models. Predicted values were generated using the BLUPs, and thus include the random quadrat effects.



APPENDIX 2.D: CORRECTIONS TO MEEKINS AND MCCARTHY (2002) MATRIX MODEL

The calculations of λ for *A. petiolata* in Meekins and McCarthy (2002) are problematic for two reasons. First, the life history is incorrect. Seeds of *A. petiolata* either germinate in the spring after they were produced or after a more prolonged period of dormancy in the seed bank as described by Meekins and McCarthy (2002) in their introduction. Meekins and McCarthy's model omits this first pathway, whereby seeds are produced in June, germinate the following spring, and are rosettes by June. This life history pathway should appear as a non-zero entry in the flowering plant (FP) to summer rosette (RS) transition of their **A** matrix. An identical error was made in another earlier study of *A. petiolata* demography (Drayton and Primack 1999) and was later identified and corrected by Rejmanek (2000).

The second problem pertains more generally to transition matrix models. All upper level transitions in a matrix population model must represent equal-length time intervals equal to the projection interval of the model (Caswell 2001). Meekins and McCarthy's (2002) model violates this assumption. For example, they show the seedling to rosette transition taking as long as the seed to seed transition in their figure 7. It is unclear from their methods what time interval the seed to seed transition represents (May to October or October to May?). In recalculating λ from their data I assume that the seed to seed transitions as given are correct and represent one full year of seed survival in the seed bank. Lower level transitions of variable time-lengths can be combined multiplicatively into equal length upper level transitions to meet the assumption of equal time intervals.

I reanalyzed the data in Table 3 of Meekins and McCarthy (2002), combining their mean lower-level transition estimates into a 3x3 annual projection matrix based on

the life cycle in Figure 1 of Davis et al. (2006). Only one germination rate estimate was provided in Meekins and McCarthy, whereas Davis et al. used different germination rates for newly shed seeds and seeds that had been dormant in the seed bank. Here I assume that both of these rates are the same. Additionally, the annual transitions from seed to rosette and from flowering plant to seed in the corrected model include sub-annual seed dormancy transitions. To account for this I use the square root of Meekins and McCarthy's annual seed to seed transition assuming that seed survival accounts for half of the annual projection interval. Estimates of λ in the original paper were 1.451, 1.375, and 1.256 in their quadrats that had 16, 40, and 80 rosettes m^{-2}, respectively, during the first year of their study. Re-calculated values of λ are 2.14, 1.84, and 1.43.

A printout of a spreadsheet is provided in Appendix 2.EAppendix 2.E which shows how Meekins and McCarthy's transitions were combined for this re-analysis. Estimates of λ were calculated using the free PopTools plug-in for Excel (Hood 2006).

APPENDIX 2.E: MEEKINS AND MCCARTHY (2002) CALCULATIONS

Meekins and McCarthy Low Density Matrix

	S	SL	RS	RF	2R	FP
S	0.301	0	0	0	0	102.14
SL	0.158	0	0	0	0	0
RS	0	0.935	0	0	0	0
RF	0	0	0.672	0	0	0
2R	0	0	0	0.787	0	0
FP	0	0	0	0	0.925	0

Adjusted Matrix

	s	r	p
s	0.253	0	25.887
r	0.081	0	8.279
p	0	0.489	0

Rosette Density	M&M λ	Corrected λ
16	1.451	2.143

Meekins and McCarthy Medium Density Matrix

	S	SL	RS	RF	2R	FP
S	0.311	0	0	0	0	62.997
SL	0.153	0	0	0	0	0
RS	0	0.972	0	0	0	0
RF	0	0	0.724	0	0	0
2R	0	0	0	0.802	0	0
FP	0	0	0	0	0.96	0

Adjusted Matrix

	s	r	p
s	0.263	0	16.594
r	0.083	0	5.225
p	0	0.557	0

Rosette Density	M&M λ	Corrected λ
40	1.375	1.843

Meekins and McCarthy High Density Matrix

	S	SL	RS	RF	2R	FP
S	0.316	0	0	0	0	58.766
SL	0.147	0	0	0	0	0
RS	0	0.958	0	0	0	0
RF	0	0	0.569	0	0	0
2R	0	0	0	0.663	0	0
FP	0	0	0	0	0.940	0

Adjusted Matrix

	s	r	p
s	0.27	0	15.84
r	0.079	0	4.652
p	0	0.355	0

Rosette Density	M&M λ	Corrected λ
80	1.256	1.426

Meekins and McCarthy Abbreviations

- S Seeds
- SL Seedlings March
- RS Rosettes in June
- RF Rosettes in October
- 2R Second Year Rosettes Flowering
- FP Plants

Adjusted Matrix Abbreviations

- s Seeds in June
- r Rosettes in June
- p Flowering Plants in June

CHAPTER 3: COMMENT ON "COMPLEX POPULATION DYNAMICS AND CONTROL OF THE INVASIVE BIENNIAL *ALLIARIA PETIOLATA* (GARLIC MUSTARD)

INTRODUCTION

Demographic models are powerful tools for understanding the dynamics of natural populations. As analytical methods and computing power become more accessible, population models are increasingly being used to guide or interpret the development of population management plans. Recently, two studies of the invasive weed *Alliaria petiolata* (garlic mustard, Brassicaceae [M. Bieb] Cavara and Grande) have been published which did just that (Davis et al. 2006, Pardini et al. 2009). Each identified optimal stages in the plant's life history to target with management actions, but they reach contrasting conclusions. Pardini et al.'s (2009) stage-classified model, which included density dependent rosette mortality and fecundity, predicted that management of newly recruited rosettes would have to reduce survival by 95% to be effective. The authors also predicted that rosette management could be counterproductive and increase population density if managers killed fewer than 95% of individuals. This result differs from the elasticity analysis of Davis et al.'s (2006) linear, deterministic model, which indicated that rosette survival was one of three lower level demographic transitions, along with fecundity and seedling survival, that was likely to have the greatest effect on population growth rate (λ). Several critical errors in the construction and parameterization of the model by Pardini et al. (2009, henceforth PDCK after the authors' initials) result in management recommendations are incorrect. Because the published recommendations are potentially already being implemented by some natural resource managers (Adam Davis, personal communication), it is important that these problems in the model be corrected.

The problems stem from omission of demographic transitions in the *A. petiolata* life cycle, errors in the statistical estimation of parameters in two density dependent functions, incorrect implementation of density dependent functions in the model, and questionable interpretation of the model output itself. I will briefly discuss each of these and show the effects of sequentially correcting them. A revised version of their model based on information extracted from the published article makes different management predictions than the original model. This new model indicates that any management that decreases rosette or adult survival probability or decreases per capita fecundity will reduce population sizes. Further, the reworked model predicts stable population dynamics rather than the chaotic or cyclical dynamics predicted by the original model. Readers should consult PDCK or Davis et al. (2006) for more background and details on model specifics.

Study System

Alliaria petiolata is a biennial forb and a frequent invader of forested habitats in North America. Its life history is well documented (e.g. Cavers et al. 1979, Nuzzo 2000, Rodgers et al. 2008). The life cycle diagram in Figure 2.E.1 shows the life history stages present at an annual census point in mid-June just prior to seed set. These are: dormant seeds in the soil that are one or more years old (S), new rosettes which germinated several months earlier in spring (R), and second year flowering plants referred to as "adults" (A). The arrows represent transitions individuals can make over the course of one full year and are comprised of one or more sub-annual transitions: per capita fecundity (f), germination probabilities of new (g_1) and dormant (g_2) seeds, and seed (v), seedling (s_1),

summer rosette (s_2), and winter rosette (s_3) survival probabilities. The parameters c_1 and c_2 are varied from between 0:1 to simulate the effects of management as a proportional reduction in rosette or adult density. Notation alert: in this chapter I follow the abbreviation scheme for demographic transitions from PDCK to facilitate direct comparison with the original paper by PDCK.

Management Inference

Pardini et al.'s management conclusions were drawn from bifurcation plots which show the equilibrium sizes of *A. petiolata* populations (vertical axis) across a range of simulated management efficacy (horizontal axis). I have reproduced PDCK's bifurcation plots using Equation 5 from their paper as the starting point for the reanalysis. The reproduced original result is shown without modification in Figure 2.E.2*a, b*. I then incrementally modify the bifurcation plots as each successive correction is made to the model. In each plot either c_1 or c_2 is varied from 0:1 by 0.001 while the other is fixed at 1. Populations were projected for 2000 years and the first 500 years of population sizes were discarded to eliminate transient dynamics. Population size (adults plus rosettes) was then plotted against c_1 or c_2.

ANALYSES

Life History

In Figure 2.E.1 I show three corrections made to PDCK's published life cycle diagram. As presented in PDCK, the model applies a seed survival probability

(abbreviated v in their paper) to newly produced seeds (Figure 1, PDCK) but not to older seeds in the seed bank. The only source of seed loss from the population after the first winter in PDCK's *A. petiolata* life cycle is by germination, shown as the seed to seed loop in the life cycle diagram (Figure 2.E.1). This loop is parameterized in their model as $1\text{-}g_2$, where g_2 is the probability of an older, dormant seed in the seed bank germinating during any given year. This means that the number of seeds remaining in the soil every year is the proportion that did not germinate. Seeds in the soil seed bank as modeled in PDCK are therefore effectively immortal. However, seeds are lost from the seed bank to mortality as well. Accounting for this requires modification of two transitions. First, it is necessary to multiply the *S* to *S* transition by an annual seed survival probability, abbreviated s_s in Davis et al. (2006) and v in PDCK, to account for seed mortality or loss from the seed bank by paths other than germination (Drayton and Primack 1999, Rejmanek 2000, Meekins and McCarthy 2002, Davis et al. 2006). Second, the transition from dormant seeds in June of one year to rosettes in the next must also account for losses via seed mortality prior to germination. This partial-year seed mortality over 8 months (2/3 of the 1 year model projection interval) is first scaled by raising it to the two-thirds power ($v^{2/3}$) and then multiplied by the existing *S* to *R* transition as $v^{2/3}g_2s_r$. Finally, because the losses from seed mortality during the *P* to *R* transition are only a partial year lower-level transition, this parameter is also rescaled as $v^{2/3}$.

PDCK's model included seed mortality (v) in the A to S (a_{13}; matrix notation refers to analogous transitions in Davis et al. 2006 and in Chapter 4) and A to R (a_{23}) transitions. Although Pardini et al. state that omitting seed mortality from the *S* to *S*

transition (a_{11}) changes the bifurcation result (i.e. the effect of simulated management) only qualitatively (which is true), they also omitted it from the S to R (a_{12}) transition. Adding full and partial year seed mortality to the base model as described above and shown in Figure 2.E.1 changes the model predictions substantially (Figure 2.E.2*c, d*). Unmanaged populations exhibit chaotic instead of cyclic dynamics. More importantly, whereas PDCK concluded that proportionally raising rosette mortality between 0-95% caused maximum population size to increase when v was omitted from several transitions, restoring v to these transitions results in population suppression as rosette mortality is increased to about 47.5%. Reducing per capita fecundity, which PDCK describe as reducing adult plant density (discussed later), up to about 39% results in an increase in population size relative to unmanaged populations. Thus, it is important that seed bank mortality not be omitted from the model.

Parameterization of Density Dependent Functions

Summer Survival, s_2: The function for density dependent summer rosette survival (s_2) was estimated by PDCK using a multiple logistic regression (PDCK Equation 2) of survival versus the sum of adult and rosette density ($T=A+R$) and the product of adult and rosette densities ($U=A*R$). The authors developed this function by comparing regression models with different combinations of A, R, T, and U as predictors. Parameter significance (i.e. P-value) was used to decide which parameters were retained in the final model. The major issue with the statistical analysis using the above predictors is that it ignores muticolinearity between U and T, which violates the assumption of independence

among predictors in the multiple regression. Additionally, the inclusion of an interaction term U without the main effects A and R can result in biased estimates of model parameters. These specific issues set up a scenario that favors spuriously identifying variables as important if P-values are used to select parameters. The resulting density dependence function used in PDCK's population model does not fit the data well and predicts survival probabilities that do not make biological sense (Appendix 3.A, Figure 3.A.1). In the absence of adults, rosette survival probability asymptotically approaches 100% as rosette density increases. For example, this function predicts 100% rosette survival probability (to within the machine precision limits of my computer) when rosette density is 1000 m^{-2} and adult density is 1 m^{-2}. This is incongruous with other estimates of *A. petiolata* summer rosette survival rates between 10-40% (Anderson et al. 1996, Byers and Quinn 1998, Evans Chapter 4) The details of these issues are discussed at greater length in Appendix 3.A and demonstrated in a MATLAB script (The MathWorks 2008) in Appendix 3.F.

I re-parameterized the summer survival function by fitting a logistic regression of survival probability to T and to $\log_e(T+1)$. Data were obtained from Figure 2a in PDCK by extracting x and y pixel coordinates from each data point in the published PDF graph using photo editing software. Because it was not possible to determine values of R or A from the graph, models including A, R, or U were not evaluated. The model which used $\log_e(T+1)$ as the covariate was better supported by the data (ΔAICc = 1.9), although not overwhelmingly. While this function fits the data reasonably well (Appendix 3.A, Figure 3.A.1), alternative functions should be re-fit using the complete dataset and evaluated

using an information criterion such as AICc rather than parameter significance (Appendix 3.A). The re-fit function is:

$$s_{2(t)} = \frac{1}{1+e^{-(\beta_{0s}+\beta_{1s}(A_t+R_t+1))}} \tag{3.1}$$

where β_0s and β_1s are the summer intercept and slope, respectively, and the covariate is the sum of rosette and adult densities at the beginning of summer in year t. Note that in the back transformation of a logistic regression the negative of the exponentiated term (the linear predictor) is taken (Neter et al. 1996). It is not clear whether PDCK did this in their original model of s_2 (Appendix 3.A). Parameters for the refit function are given in Table 3.B.1 (Appendix 3.B).

Winter Survival, s_3: The density dependent rosette survival function was estimated using a linear regression of $\log_e$(survival) on $\log_e(R_t + 1)$. The back transformed survival rate at time t was estimated from rosette density as

$$s_{3(t)} = e^{\beta_{1w}\log_e(R_t+1)} \tag{3.2}$$

where β_{1w} is a slope parameter estimate for winter survival.

There are two problems with this analysis. First, it appears that the authors may have inadvertently dropped a number of observations of 0% survival from their analysis by using a log transformation of the survival response. In their Figure 2b, the only observations with either 0% survival or 100% survival were those with only a single rosette (singletons). Dropping all of the 0% observations by log transforming them [$\log_e(0) = -\infty$] leaves all remaining estimates of singleton survival as 100% (PDCK

Figure 2b, inset). By excluding all of the observations with 0% survival from the analysis, the survival probability of populations at very low densities will be overestimated by the fitted regression model. Second, because PDCK did not estimate an intercept in the model, fixing it at zero on the log-transformed scale, the intercept is fixed at unity (s_3 = 100% survival) on the back-transformed scale. Winter survival probability therefore always goes to one as population size approaches zero because $e^{\beta \log_e(0+1)} = 1$ for any β. This results in a deterministic overestimation of *A. petiolata*'s ability to persist or invade when rare. While in the "real world" the survival probability of less than one individual is meaningless, fractional individuals exist in the "model world" and model predictions can be significantly impacted by this type of error.

Variance stabilizing transformations can be problematic when applied in models used for making predictions, as demonstrated above, and the model-predictions do not always make sense when back-transformed to the original data scale. For example a linear model of arcsin-square root transformed proportion data could predict negative values that cannot be back-transformed to proportions in a meaningful way. When the goal of analysis is to interpret pattern, working in a transformed scale can be the most straightforward option for meeting the assumption of normality in a linear model. However, caution is required when the goal is to make predictions in the original data scale. A better solution is to use a generalized linear model (GLM) or a generalized linear mixed model (GLMM) with an embedded link function appropriate to the distribution of the data (Bolker et al. 2009).

I fit a logistic regression of winter survival probability versus $\log_e(R_t + 1)$ to data obtained from Figures 2b and 2c using photo editing software (Appendix 3.C). This

model is appropriate for survival data and fits the data satisfactorily (Figure 3.C.1). Substituting the new density dependence functions for summer and winter rosette survival changes the population dynamics. The bifurcation analysis of this corrected model indicates a single stable equilibrium at any management efficacy (Figure 2.E.2*e,f*). Any management which reduces either rosette survival or fecundity is predicted to reduce population size. This is an important difference from PDCK's prediction that populations would exhibit cyclical or chaotic dynamics, or that management could cause an increase in population size.

Implementation of Density Dependence

The two rosette density dependence functions (summer and winter) each reference the same rosette density (R_t) in their calculations. This is incorrect, as it is not how the functions were parameterized. Calculation of the winter survival rate must be based on the density of rosettes at the beginning of winter. This requires creating an intermediate age class of winter rosettes, *Rw*, that is passed from the summer survival function to the winter survival function. This change also requires distinguishing between management of rosettes in summer versus winter. The new calculation of A_{t+1}, simulating reduction of winter rosette survival (c_2), is given in two equations:

$$Rw_t = R_t \frac{1}{1+e^{-(\beta_{s0}+\beta_{s1}\log_e(A_t+R_t+1))}} \tag{3.3a}$$

$$A_{t+1} = (1-c_2)Rw_t \frac{1}{1+e^{-(\beta_{w0}+\beta_{w1}\log_e(Rw_t+1)}} \tag{3.3b}$$

With these changes incorporated, the strength of winter density dependent mortality is decreased because summer mortality is accounted for, and the predicted population size

at equilibrium is greater (Figure 2.E.2*g*, *h*). The model predicts that any management that reduces rosette survival or fecundity will reduce population size.

Interpretation of Model Output

In the analyses of *A. petiolata*'s response to management, PDCK refer to management of rosettes versus management of adults. They simulate these scenarios by including an extra parameter in the model which proportionally reduces one of the demographic rates. The equation set below is modified from PDCK equation 5 to include all of the changes described in the previous sections and was used to produce Figure 2.E.2*g* and *h*:

$$S_{t+1} = v(1-g_1)(1-c_1)A_t e^{\beta_{0f}+\beta_{1f}A_t} + S_t v(1-g_2) \tag{3.4a}$$

$$R_{t+1} = v^{2/3} g_1 s_1 (1-c_1) A_t e^{\beta_{0f}+\beta_{1f}A_t} + v^{2/3} g_2 s_1 S_t \tag{3.4b}$$

$$Rw_t = R_t \frac{1}{1+e^{-(\beta_{0s}+\beta_{1s}\log_e(A_t+R_t+1))}} \tag{3.4c}$$

$$A_{t+1} = (1-c_2)Rw_t \frac{1}{1+e^{-(\beta_{0w}+\beta_{1w}\log_e(Rw_t+1)}} \tag{3.4d}$$

where $e^{\beta_{0f}+\beta_{1f}A_t}$ is the density dependent fecundity function .

The effect on population dynamics of managing a particular life history stage depends on when the management is applied. I refer to c_1 as a simulation of managing fecundity rather than adult density, as PDCK describe it, because it is only inserted in the equation set where it affects the number of plants contributing new seeds to the rosette stage and to the seed bank without affecting the relationship of adult density to other life

history stages. In this respect, it is identical to the simulation of reducing per capita fecundity in Davis et al.'s (2006) linear model of *A. petiolata*.

Things are more complicated in a nonlinear model. To simulate the effect of increasing adult mortality, one must be explicit about where and when the impacts of reducing adult density occur. Adult management as modeled in PDCK is analogous to hand pulling a fraction of the adult population as they senesce, just before seed dispersal. Management could be applied at different times, though. For example, reduction of adult plant density earlier in the spring could affect per capita fecundity as well as summer rosette survival. To simulate this, adult density (A_t) should be adjusted by (1-c_1) in the density dependence functions for per capita fecundity in the $A \rightarrow S$ and $A \rightarrow R$ transitions, and in the summer rosette survival function. Similarly, managing rosettes early in winter could reduce the strength of density feedbacks on rosette survival. Making these changes further alters the population dynamics of the model system by decreasing the population's responsiveness to rosette management (Appendix 3.D). The types of management likely to be implemented should therefore be considered. The effects of management timing relative to sub-annual transitions and subsequent impacts on density dependent functions can then be explored by varying where the management parameters are inserted into the equation set.

Finally, the effects of density dependent seedling mortality should also be explored. PDCK cite the low survival rate of seedlings to justify not including density dependent rosette mortality in their model. *Alliaria petiolata* seedling survival probabilities are variable, though, ranging from 0.13 to 0.63 (Table 1 in Davis et al. 2006) and higher (Chapter 2). Adding a density dependent seedling survival function to

Equation set 3.4 has a large effect on the stable equilibrium population size as the strength of the density feedback is increased (Appendix 3.E). More importantly, it greatly changes how populations respond to management Appendix 3.E, Figure 3.E.2).

Conclusions

The study of *A. petiolata* by Pardini et al. (2009) presented a model of *A. petiolata* population dynamics based on a study of a single cohort of plants in 40 1 m^2 sampling quadrats at a single study site. The authors quantified density dependence in summer and winter rosette survival and per capita fecundity, but did not to quantify density dependent seedling survival, which may be important in this species given its variable population density (Trimbur 1973, Nuzzo 2000, Evans Chapter 2, Chapter 5). Using this model, the authors recommended that management of *A. petiolata* should be focused on reducing adult density. Because their model predicted that incomplete removal of rosettes could cause increases in population size, PDCK also recommended that management of rosettes should be avoided. Given the limited geographical and temporal scope of the data, the numerous quantitative errors in formulating the model, and the known spatiotemporal variability in *A. petiolata* demographic rates, the management recommendations made by Pardini et al. (2009) should not be applied at this point or until the modeling framework is examined more thoroughly for the inconsistencies reported here. The suggestion that rosette management could increase population density is particularly concerning. This should include: 1) a recalculation of the density dependence functions for winter and summer rosette survival using appropriate statistical models and parameter selection tools, 2) consideration of how

management might actually be implemented in the field and testing models which reflect this, 3) experimentation with density dependent seedling survival, and 4) revisiting the analyses of the sensitivity of population dynamics to variation in the lower level demographic rates which was included in a supplement to the original paper. Because some of the proposed management strategies for *A. petiolata* affect both rosette survival and fecundity (Davis et al. 2006), the authors should also look at the population responses to combinatorial variation in simulated management these two demographic transitions.

In contrast to PDCK's conclusions, the population dynamics of *A. petiolata* appear to be driven by stable equilibria in simulations of both unmanaged and managed populations based in the final model parameterized from PDCK's data (Figure 2.E.2g, h). There is also compelling evidence for the importance of modeling seed bank mortality. Most reconfigurations of the model predicted that management of either rosettes or adults in the study population should be an effective strategy. However, it will be necessary to think more carefully about how and where in the model to apply the simulated management if the goal is to simulate ways that populations might be managed in the field. While the predictions generated in this re-analysis represent an improvement over the original model, they should not be taken as conclusive. More and better long term empirical data are needed on the seed bank dynamics of *A. petiolata* and density feedbacks on seeding survival and other transitions. Because of the tremendous spatiotemporal variability in *A. petiolata* demographic rates, more general management recommendations will need to come from a larger set of demographic data that has greater coverage in space and time. The results of Pardini et al., once the model is corrected, should be expected as a subset of the many possible models of *A. petiolata*.

Figures: Chapter 3

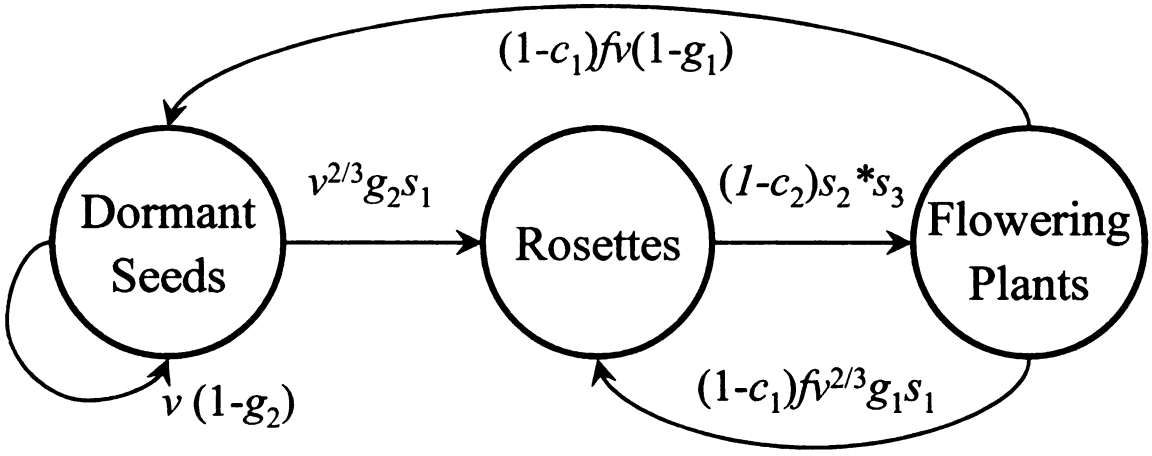


Figure 2.E.1. Life cycle diagram of *A. petiolata* showing the upper level transitions (arrows) and the lower level transitions which comprise them (equations). Symbols are defined in the text. This version differs from Figure 2.2 in that it follows the notation of Pardini et al. (2009) and includes parameters for simulation of management.

Figure 2.E.2. Bifurcations of *A. petiolata* populations with simulated management of rosette survival (left hand panels) and simulated management of per capita fecundity (right hand panels) using equation set 4. In each case, the unmanaged population is represented on the far left where additional mortality or reduction in fecundity is zero, and the dashed reference line shows the maximum equilibrium size of the unmanaged population. Points above this line indicate potential increases in population size relative to the unmanaged case, while those below it indicate population decline. Individual panels show (*a*,*b*) the original model as parameterized and published in Pardini et al. (2009), (*c*,*d*) the effects of adding seed bank morality to the *S* to *S* and *S* to *R* transitions and rescaling sub-annual seed mortality as described in the text, (*e*,*f*), the result of changing the function for winter rosette density dependence in *a-d* to a logistic curved response, and (*g*,*h*) the result of correcting the sign of the summer rosette density dependence function used in *a-f*.

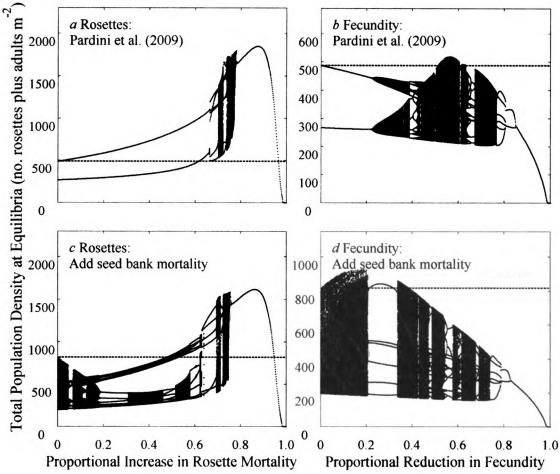
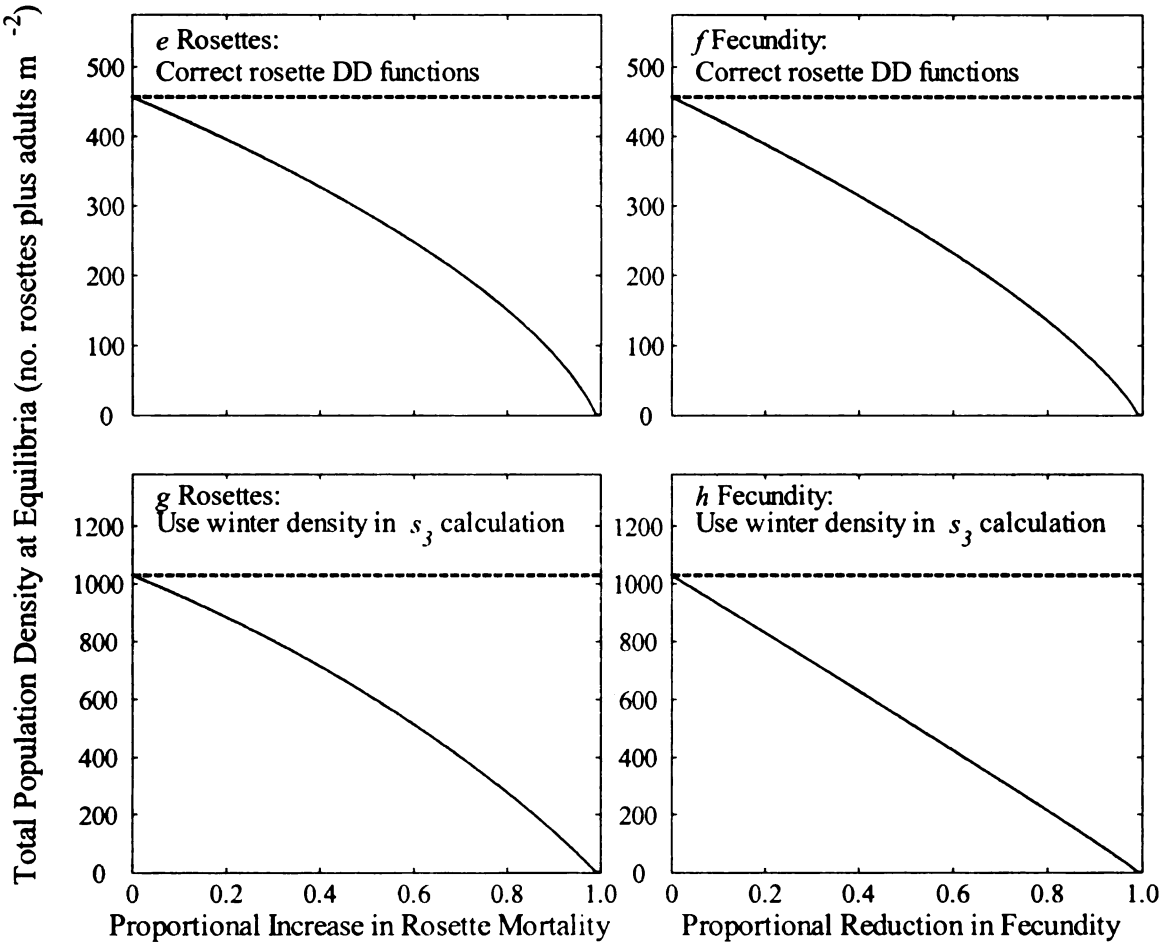

a Rosettes:
Pardini et al. (2009)
b Fecundity:
Pardini et al. (2009)
c Rosettes:
Add seed bank mortality
d Fecundity:
Add seed bank mortality
Total Population Density at Equilibria (no. rosettes plus adults m-2)
Proportional Increase in Rosette Mortality
Proportional Reduction in Fecundity
0
500
1000
1500
2000
100
200
300
400
600
800
0.2
0.4
0.6
0.8
1.0

Figure 2.E.2 **cont'd.**



APPENDIX 3.A: SUMMER DENSITY DEPENDENCE AND MULTICOLLINEARITY

3.A.i. Summer Density Dependence Function (s_2)

The function for density dependent summer rosette survival (s_2) in Pardini et al. (2009, henceforth abbreviated 'PDCK') does not fit the data which were used to parameterize it, which are shown in Figure 2a of PDCK. The function is given in PDCK Equation 2. Parameters were selected for inclusion in the model based on significance (i.e. *P*-values). The authors may have been lead to conclude that this function fit their data well based on its parameters' *P*-values and its R^2 value (0.79). Because the distribution of plant densities used to parameterize the function were clustered, rather than continuous (PDCK, Figure 2a), the function they used only intersects the two clusters of data by chance (Figure 3.A.1, dashed line). If there had been more continuous variation in *A. petiolata* rosette density, it is unlikely that this model would have garnered much support via R^2. The original function for summer survival in PDCK is given as:

$$s_2 = \frac{1}{1 + e^{\beta_0 + \beta_1(A+R) + \beta_2(A*R)}} \qquad (3.A.1)$$

where e is the base of the natural logarithm, β_0 is the intercept, β_1 and β_2 are the slopes of the covariates, and A and R are the densities of adults and rosettes, respectively, at the beginning of summer. In PDCK, the sum $A+R$ is abbreviated T, and the product $A*R$ is abbreviated U. These are expanded here for clarity.

Because the lower-order terms for A and R were dropped, predicted rosette survival probability is symmetrical with respect to A and R. For example, this means that s_2 is the same whether there are 100 rosettes and 5 adults, or 5 rosettes and 100 adults. A surface plot of the predicted rosette survival rates versus adult and rosette densities

illustrates some of the problems with this function (Figure 3.A.2). When adult density is zero, survival probability increases asymptotically, approaching 100% survival as rosette density increases. This does not correspond with the data plotted in PDCK Figure 2a or with the known biology of *A. petiolata*. The symmetry of the function means that the same is true when rosette density is extremely low and adult density is increased. Survival only decreases when there are both rosettes and adults present, dropping most rapidly when $A = R$. This apparent Allee effect that occurs at low densities or when the ratio of A:R is uneven occurs because the sign of coefficient for T is actually positive, although it is shown as negative in PDCK Table 1. The reason for this and consequences if this assumption is incorrect are explained below in the section on back transforming logistic regressions.

There are two related statistical problems with this function. The first is the inclusion of a term for the interaction U without its constituent lower-order effects A and R. Used in this way, the interaction lacks biological meaning, although it has a large effect on model predictions.

The additive term T implies that the linear effect of rosettes on survival is the same as that of adults, and thus, that the slopes and intercepts of s_2 with respect to A and R are equal. If this is true, how do they interact? If there is a significant interaction between A and R, the main effects for these lower-order terms should be retained in the model even if they are not significant. Removing them from the regression can cause a shift in the function relative to the origin and alter its curvature. Depending on the signs and magnitudes of the terms in the model, this will result in over or under estimation of the intercept and of the remaining slope parameters. Unless there is compelling evidence

to do otherwise, lower order terms should always be retained if they are included in higher order terms (see chapter 12.3 in Draper and Smith 1998). Without them, the interaction term lacks meaning, and the regression model predictions are likely to be flawed.

3.A.ii. Multicollinearity and Parameter Bias

Second, the additive and multiplicative density terms T and U should not both be included in the model. Given two random variables x and y, $(x + y)$ and $(x * y)$ will often be strongly, positively correlated, even if x and y are uncorrelated. If z is a function of $(x + y)$, it will spuriously appear to be a function of $(x * y)$ as well. If z is an additive function of x and y, it can also appear to be a function of $(x*y)$ and/or $(x+y)$, depending on the signs, magnitudes, and correlations among x and y. When there is strong correlation among predictor variables in a regression model, there can be multiple solutions to the regression (parameter estimates) that fit equally well. In such cases where there is multicolinearity among the predictors, the parameter estimates will not be unique, and alternate solutions to the model will generate different, unrelated predictions. Neter et al. (1996) discuss this issue with useful examples in Chapter 7.7 of their text. Model selection based on variable significance (P-values) will not be able to distinguish effectively among these models, because the parameter estimates will be significantly different from zero, even though they are unrelated to the process driving variation in z. These issues are explored and illustrated more fully in a supplemental MATLAB program (Appendix 3.F).

The results of one simulation from Appendix 3.F illustrating the correlations among x, y, and z, are shown in Figure 3.A.3 for those without access to MATLAB. In this figure, x and y are randomly generated variables with means 3 and 5, respectively, standard deviations of 1, and 100 'observations' each. The random variable z is a linear function of x and y, calculated with Model Function A in Appendix 3.F as:

$$z = 5 + 2x + 6y + \varepsilon \quad \varepsilon \sim N(0,1) \tag{3.A.2}$$

where ε is added as random error. Three linear models were then fit to the data with different combinations of parameters:

$$\begin{aligned}
&\text{M1} \quad z = \beta_0 + \beta_1 x + \beta_2 y \\
&\text{M2} \quad z = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy \\
&\text{M3} \quad z = \beta_0 + \beta_1 (x + y) \\
&\text{M4} \quad z = \beta_0 + \beta_1 xy \\
&\text{M5} \quad z = \beta_0 + \beta_1 (x + y) + \beta_2 xy
\end{aligned}$$

Models were then compared using parameter significance as well as AICc. AICc is based on Kullback-Leibler information and is a relative measure of the distance, in arbitrary units of 'information', from a given model to the full reality of the underlying processes driving the variance in the data (Anderson 2008). The model with the lowest AICc score is thus the closest to describing the truth among a set of competing models and is considered the best model. ΔAICc is calculated as the difference between each model's AICc and that of the best model ($\Delta AICc_{best} = 0$). The Akaike weight (w) for each model is the probability that it is the best among a set of competing models, and is calculated from the ΔAICc scores. Anderson (2008) offers a very user friendly introduction to the subject which is more accessible than the more comprehensive volume by Burnham and Anderson (2002). The simulation was run in MATLAB (The

MathWorks 2008), and the analysis was conducted in PROC GLIMMIX in SAS (SAS Institute 2008).

The results of this exercise, including parameter estimates, *P*-values, and Akaike's Information Criterion adjusted for sample size (AICc) are given in Table 3.A.1. The analysis of the simulation highlights the potential pitfalls of model selection based on parameter significance alone. All parameters were highly significant ($P < 0.0001$) in all models except for the interaction term in M2, which was correctly estimated as non-significant. The interaction term was only identified as significant when the lower order terms for x and y were dropped (M4, M5). Model selection based on parameter significance can therefore easily lead to retention of spuriously significant terms. In models where this happened (M3-M5), the parameter estimates were highly inaccurate. The true intercept was 2. In M3, it was estimated as 9.5, in M4 it was estimated as 29.9, and in M5, which has the same form as PDCK's summer survival function, it was estimated as -12.3. The other parameters in these models were not part of the process underlying the variance in z and so have no 'true' values, although the estimates of these parameters vary substantially across models. The parameter estimates in the true model, M1, were reasonably close to the true parameter values, as were those in M2.

Comparison of models using AICc correctly ranked M1 as the best model, given the data. As 90.4% of the Akaike weight was assigned to M1, there is a 90.4% probability that M1 is the best model among the five. M2 was ranked second with 9.6% of the Akaike weights. All other models had approximately 0% probability of being the best model and thus are discarded as unlikely.

In summary, when an interaction term is retained, its lower-order terms should also be retained because the function can become shifted relative to the origin (Draper and Smith 1998) and the remaining parameter estimates will likely be inaccurate if lower order terms are dropped. If U is kept in the model, A and R should be kept as well, even if they are not significant. Because of the high probability that U and T are strongly correlated, if U is included in the model, T should not be. If T is included, it should be the only covariate, as A, R, and U cannot be included with it. Model selection based on parameter significance can be misleading and result in incorrect acceptance of an inappropriate model. Model selection based on information theory represents a better alternative. One very effective way of assessing whether a fitted model is appropriate or not is to plot its predictions across a range of values and judge whether it reflects the patterns in the data. If it does not, then the model likely mischaracterizes the underlying processes.

Notes on Logistic Regression Back-Transformation

A logistic regression model linearizes the relationship between the independent variables, which can be categorical or continuous, and the dependent response variable, which is the probability of observing p "successes" (e.g. survivors) out of q trials (e.g. initial plants). The linearized form of the logistic model is

$$\log_e\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \ldots \beta_n x_n \qquad (3.A.3)$$

where the right hand side of the equation is collectively called the linear predictor, $x_1, \ldots$ x_n is the set of n predictor variables, $\beta_0, \beta_1, \ldots \beta_n$ are the estimated intercept and slope

parameters, π is the response probability of interest, and the left hand side of the equation is the logit of π (Neter et al. 1996). Solving Equation 3.A.3 for π in terms of the x_i by taking the antilog of both sides allows predictions of survival probabilities to be made from the fitted model if the x_i are known:

$$\pi = \frac{e^{\beta_0+\beta_1 x_1+\ldots\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+\ldots\beta_n x_n}} \quad (3.A.4a)$$

$$= \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\ldots\beta_n x_n)}} \quad (3.A.4b)$$

These are equivalent expressions. PDCK used the form in Equation 3.A.4b in their summer rosette density dependence function, reprinted here as Equation 3.A.1. Because they do not indicate taking the negative of the exponentiated linear predictor, the implication is that they either took the negative of the individual terms before presenting them in Table 1 of their paper, or that they omitted this step. Assuming the former, this means that the 'true' signs of all the coefficients for s_2 are reversed from their signs on the linear scale as printed in PDCK Table 1, the slope of T is positive, and the slope of U is negative. This assumption was made in plotting Figure 3.A.2 and in reproducing the original bifurcation plots shown in Figures 2a and 2b of the main paper (this paper). If this assumption is wrong, failure to take the negative of the linear predictor in the back-transformation would reverse the sign of the density dependent function for s_2 and would result in an additional set of problems. To visualize what this would do to the function, stand on your head while looking at Figure 3.A.2.

Table 3.A.1 Analysis of simulation results. A random variable z was created as a function of x and y using the formula $z = 5 + 2x + 6y$ + random error. Z was modeled as a function of five different combinations of x and y. M1 represents the 'true' form of the function used to generate z. Parameter estimates from models without the main effects x and y are poor, although they have significant P-values. ΔAICc values correctly identify the M1 as the model with the greatest support from the data. M1 received > 90% of the Akaike weights (w) indicate a 90.4% probability that it is the best supported model among the five competing models.

Model	Effect	Estimate	SE	*df*	*t*-Value	*P*-value	-2LL	AICc	ΔAICc	*w*
M1	Intercept	5.701	0.582	97	9.80	<.0001	294.9	303.3	0.0	0.904
	x	1.865	0.097	97	19.29	<.0001				
	y	5.934	0.095	97	62.18	<.0001				
M2	Intercept	4.603	1.451	96	3.17	0.002	297.2	307.8	4.5	0.096
	x	2.248	0.473	96	4.75	<.0001				
	y	6.151	0.280	96	22.00	<.0001				
	x * *y*	-0.076	0.092	96	-0.83	0.4105				
M3	Intercept	9.557	1.860	98	5.14	<.0001	526.1	532.3	229.0	<0.001
	x + *y*	3.928	0.229	98	17.18	<.0001				
M4	Intercept	29.877	1.295	98	23.08	<.0001	603.4	609.6	306.3	<0.001
	x * *y*	0.744	0.080	98	9.27	<.0001				
M5	Intercept	-12.331	1.952	97	-6.32	<.0001	424.7	433.1	129.8	<0.001
	x + *y*	9.138	0.407	97	22.44	<.0001				
	x * *y*	-1.326	0.098	97	-13.56	<.0001				

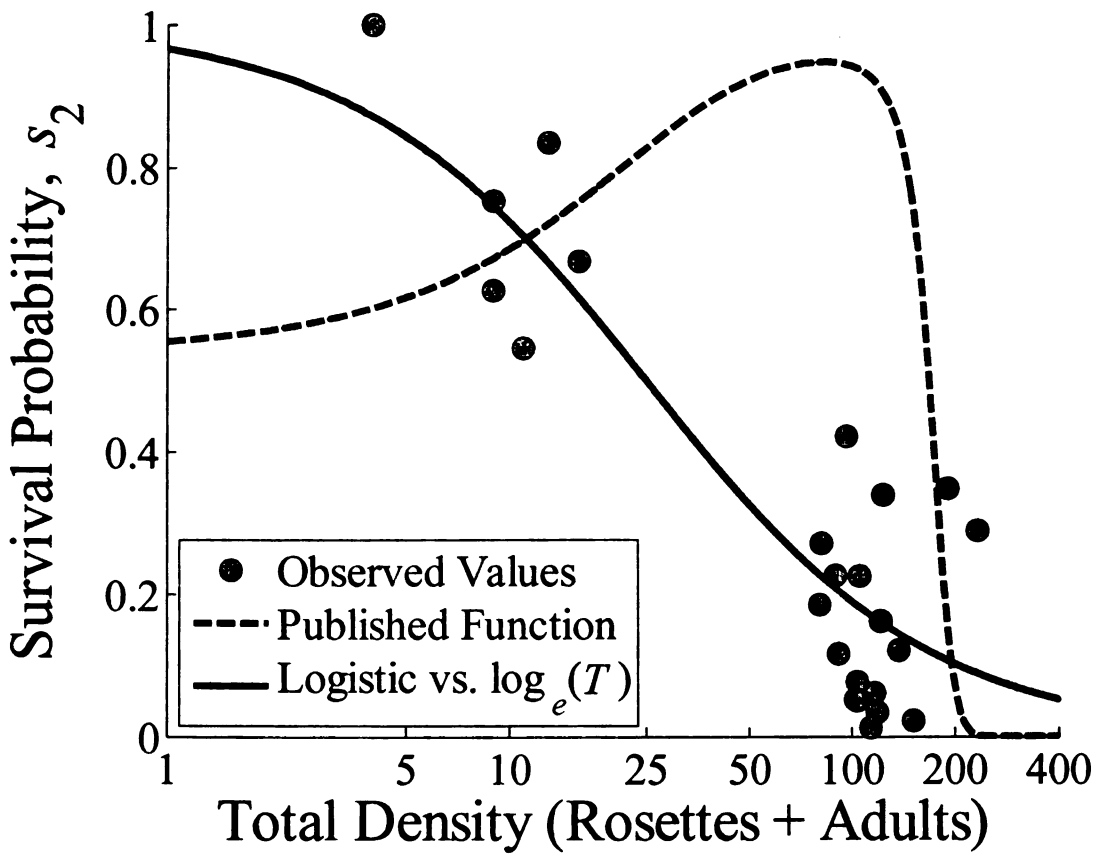


Figure 3.A.1. The function used to characterize density dependent mortality of *A. petiolata* in summer did not fit the data well. Data extracted from Pardini et al. (2009) are shown (dots) against the published logistic regression function evaluated at $R = A$ across the range of total plant density from the study (dashed line). Data extracted from Figure 2A in Pardini et al. using photo editing software were used to refit a new logistic regression function of survival probability vs. $\log_e(A + R)$, shown as a solid line. This new function is used in the modified model in the main text. Because A and R could not be extracted from the figure, lower-order terms and interactions could not be fit.

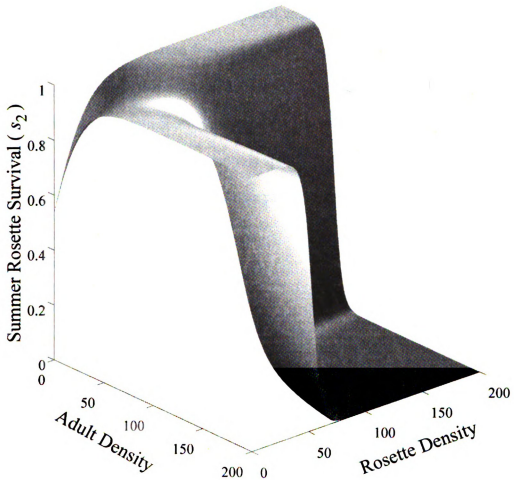
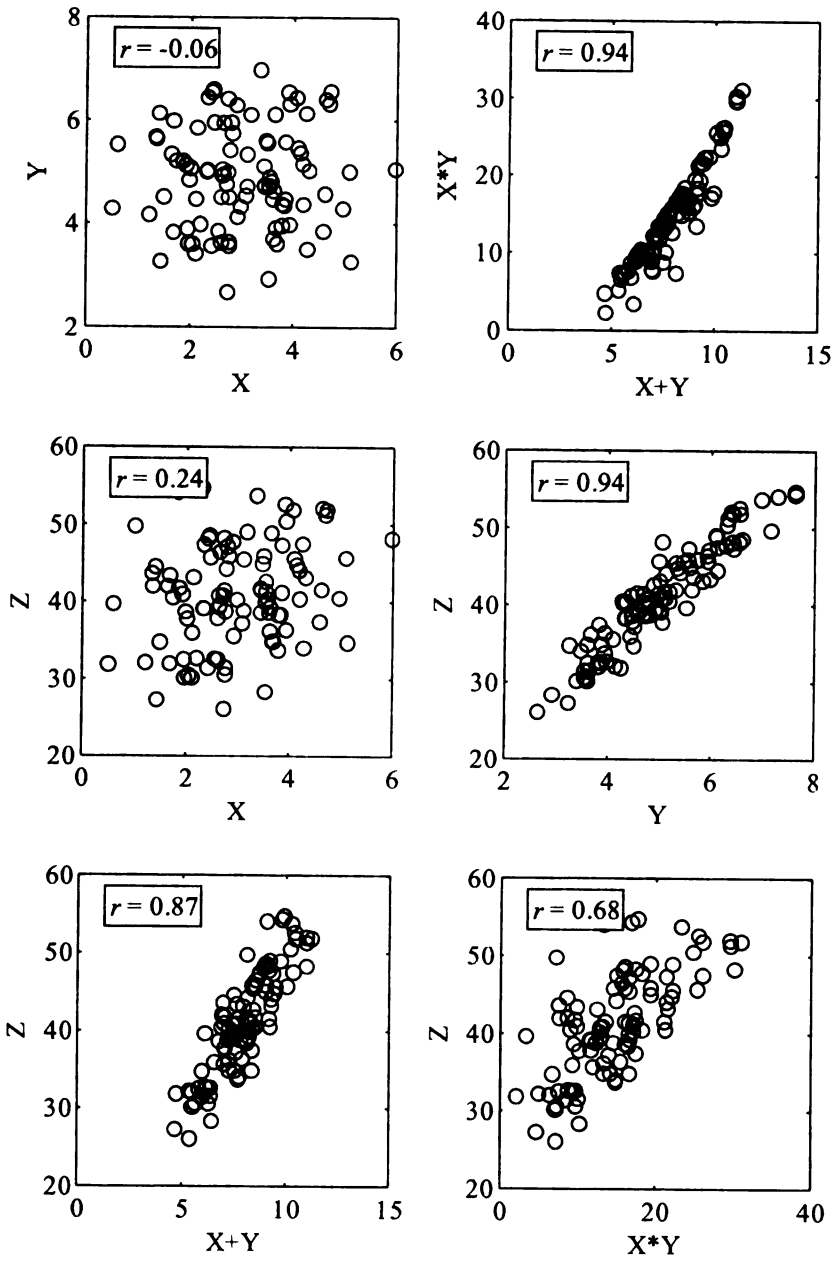


Figure 3.A.2. Surface plot of the published function for density dependent summer rosette survival (s_2) from Pardini et al. (2009) evaluated at all combinations of $A = 0{:}200$ and $R = 0{:}200$.

Figure 3.A.3. A simulation illustrating how the sum ($x + y$) and product ($x * y$) of two uncorrelated random variables can be highly correlated with each other. The variables x and y are each comprised of 100 random draws from normal distributions with means 3 and 5, respectively, and standard deviations of 1. The variable z was calculated as $z = 5 + 2x + 6y$ + random error. Correlation coefficients (r) are shown for each relationship. Although z is only an additive function of x and y, it appears to be strongly correlated with both their sum and product. The simulated data plotted here were also used in the example analysis presented in Table 3.A.1.


r = -0.06
Y
X
r = 0.94
X*Y
X+Y
r = 0.24
Z
X
r = 0.94
Z
Y
r = 0.87
Z
X+Y
r = 0.68
Z
X*Y

APPENDIX 3.B: REFIT DENSITY DEPENDENCE EQUATIONS

Table 3.B.1. Parameter estimates and fit statistics from refit summer (s_2) and winter (s_3) *A. petiolata* density dependence functions. Data were extracted from figures 2a and 2b of Pardini et al. (2009) *A. petiolata* winter survival data.

Model	Parameter	Test	Estimate	SE
s_2	Intercept	($t_{22} = 1.81$, $P = 0.0840$)	3.5626	1.9682
	$\log_e$(rosettes+1)	($t_{22} = -2.28$, $P = 0.0325$)	-1.0850	0.4753
s_3	Intercept	($t_{32} = 3.837$, $P < 0.0001$)	1.1928	0.3108
	$\log_e$(rosettes+1)	($t_{32} = -7.344$, $P < 0.0001$)	-0.4752	0.0647

APPENDIX 3.C: WINTER DENSITY DEPENDENCE

Extraction of winter rosette survival data

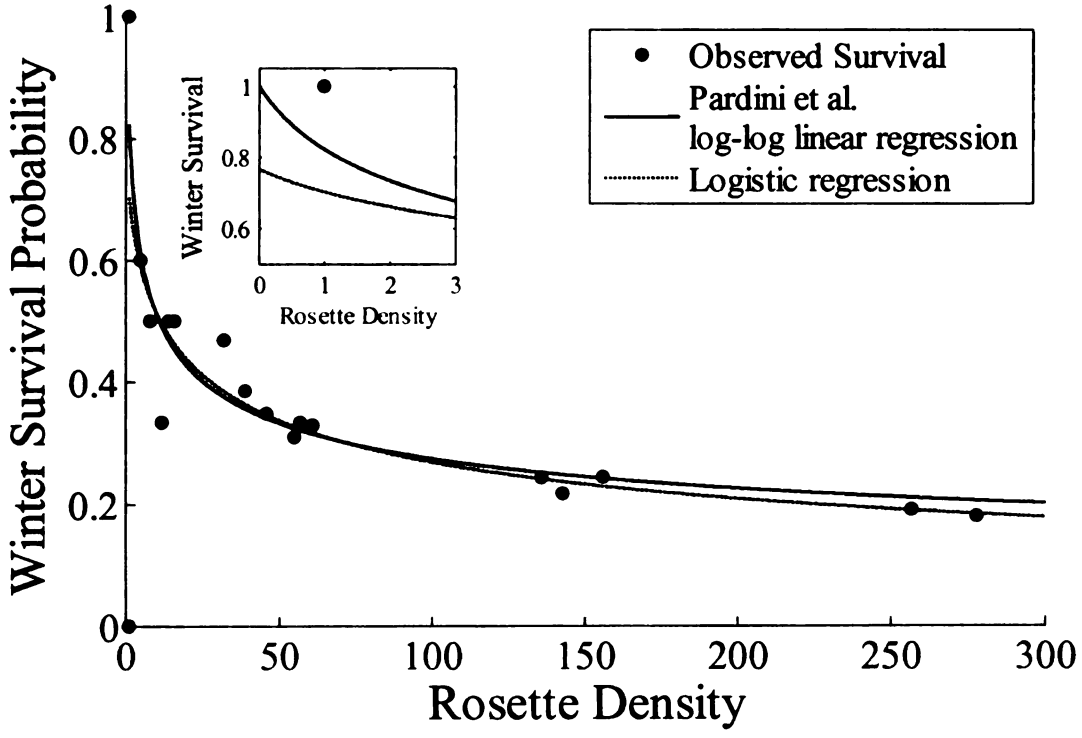
Several problems were noted in Pardini et al.'s (2009) function for density dependent survival of *Alliaria petiolata* rosette in winter (s_3). The most significant of these was the omission of an intercept, which forces survival probability to 100% as densities approach 0. Data to re-parameterize this survival function were obtained from Figures 2b and 2c in PDCK by extracting x and y pixel coordinates from each data point in the published PDF file using photo editing software. Because they reported sampling winter survival in 34 quadrats but only 18 points are visible in their Figure 2b, it was inferred that the remaining 16 quadrats had rosette densities of 1 plant m^{-2} and obscure each other in the figure. The number of survivors in these 18 singleton quadrats was obtained from the number of fecundity estimates made in plots with a single plant in their figure 2c. None of the other non-singleton plots had survival rates low enough to leave them as singletons the following spring. Thus, of these 18 singleton quadrats, 12 must have had 100% survival and six had 0% survival. Data extracted from PDCK figure 2b reveal that there were approximately 1340 rosettes used in this calculation. It seems that PDCK mistakenly stated that 1795 rosettes were used in calculating winter survival (PDCK, pg. 390). This corresponds closely with the 1346 rosettes they report marking for measurements of summer survival. Thus, the number of plants indicated in the text for summer and winter rosette survival must be reversed in the text on page 390 of their paper.

Re-running PDCK's linear regression of $\log_e$(survival probability) on $\log_e$(rosette density+1) in SAS (SAS Institute 2008) using the extracted data reproduced their original result very closely. This estimate of β_{1w} was -0.2804, compared with their estimate of -

0.2890 (PDCK Table 1), and confirms that the data extraction was successful. As expected, the six quadrats with 0% survival were dropped from the analysis by taking the natural logarithm of 0, which is -∞. Model fit to the untransformed data was good with the 0% survival observations excluded (*Adj.* $R^2 = 0.86$) but not when the dropped observations were included (*Adj.* $R^2 = 0.09$). Adding a small fraction (10^{-7}) to observations with 0% survival resulted in poor model fit (*Adj.* $R^2 = 0.018$), and fitting an intercept predicted survival probabilities greater than one.

A new logistic regression of survival probability as a function of $\log_e$(rosette density +1) was fit to the extracted data in PROC GLIMMIX in SAS (SAS Institute 2008). The parameter estimates for the new function are presented in Table 3.B.1 (Appendix 3.B). The refit winter survival function fits the data similarly across most of the range of observed densities (Figure 3.C.1). The functions diverge most prominently at low densities as they approach the *y* axis (Figure 3.C.1, detail). The original function increases and intercepts the axis at 1, while the *y* intercept of the back-transformed refit function is 0.767. Incorporating this new function for s_3 into the population model further changes the model's management predictions.

Figure 3.C.1. Density dependence of winter rosette survival probability using data extracted from Pardini et al. (2009). Pardini et al. conducted a linear regression of $\log_e$(winter survival) on $\log_e$(rosettes + 1) with the intercept fixed at 1. Thus the model predicts that survival probability approaches 100% as population size approaches 0. An alternative to this is a logistic regression of survival probability on $\log_e$(rosettes + 1). This model has a lower y intercept and is less prone to overestimating survival of low density populations. The inset detail shows the shapes of the two functions at the y intercept.



APPENDIX 3.D: APPLY SEEDLING DENSITY DEPENDENT MORTALITY

The implementation of simulated *A. petiolata* rosette and adult morality in Pardini et al (2009) was limited to overall survival functions and was not incorporated into the density dependence functions that comprised the upper level demographic rates. Here, rosette and adult mortality are applied to simulate: 1) the effect of increasing rosette mortality during the winter survival period and 2) increasing adult mortality early in the spring. Equation set 3.4 in the main paper is used as the starting point. The new equation set is:

$$S_{t+1} = v(1-g_1)(1-c_1)A_t e^{\beta_{0f}+\beta_{1f}A_t} + S_t v(1-g_2) \quad (3.D.1a)$$

$$R_{t+1} = v^{2/3} g_1 s_1 (1-c_1) A_t e^{\beta_{0f}+\beta_{1f}A_t} + v^{2/3} g_2 s_1 S_t \quad (3.D.1b)$$

$$Rw_t = R_t \frac{1}{1+e^{-(\beta_{0s}+\beta_{1s}\log_e(((1-c_1)A_t+R_t)+1))}} \quad (3.D.1c)$$

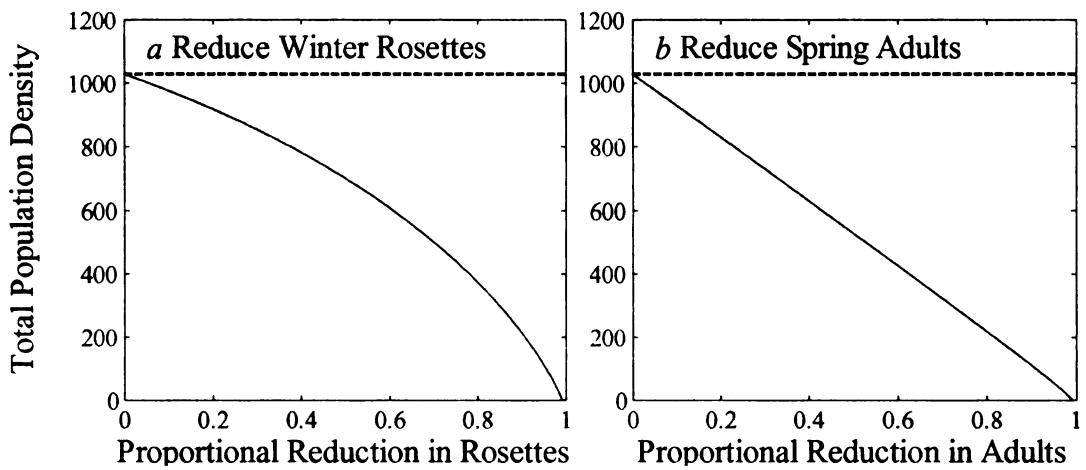
$$A_{t+1} = (1-c_2)Rw_t \frac{1}{1+e^{-(\beta_{0w}+\beta_{1w}\log_e((1-c2)Rw_t+1)}} \quad (3.D.1d)$$

where increasing the parameter c_1 reduces adults survival proportionally and c_2 reduces rosettes survival. In scenario 1, management of rosette density is simulated to occur early enough in the winter that it decreases the negative feedback of rosette density on the remaining survivors. In scenario 2, management of adult density is simulated to occur early enough in the spring that the negative effects of adults on rosette survival are decreased, although not early enough to reduce per capita fecundity. To run scenario 1, c_2 is varied from 0:1 and c_1 is fixed at 1 (Figure 3.D.1a). To explore the effects of rosette management the reverse is done (Figure 3.D.1b).

The results of these simulations are qualitatively similar to the results of the model in the main paper (Figure 2.E.2g, h). Comparison of the curvature of the equilibrium density curve in Figure 3.D.1a with Figure 2.E.2g in the main text shows that the changes made

in the density dependent functions here slow a population's response to management. This is evidenced by the increased curvature in Figure 3.D.1a.

Figure 3.D.1 Bifurcations of *A. petiolata* populations with (*a*) simulated management of rosettes in early winter and (*b*) simulated management of adults in early spring using equation set C. In each case, the unmanaged population is represented on the far left where additional mortality is zero, and the dashed horizontal reference line shows the maximum equilibrium size of the unmanaged population. As all point fall below this line, any increase in morality results in decreased population size relative to the unmanaged case.



APPENDIX 3.E: APPLICATION OF SEEDLING DENSITY DEPENDENCE

Alliaria petiolata populations frequently reach high initial seedling densities (Davis et al. 2006) and have variable seedling mortality rates (2006). To explore the possible affects of density dependent seedling mortality on *A. petiolata*'s population dynamics and response to management, a density dependent function for seedling survival can be added to the model developed in Appendix 3.D of this paper. A sub-annual seedling age class (*Sdl*) is created to which a logistic density dependent function then applied to estimate the number of rosettes at time $t+1$ (R_{t+1}) that survive from the seedling stage at time t. Note that the number of seedlings at time t is calculated from the number of seeds germinating from the soil at time t and the number of new seeds contributed from adult plants at time t, thus making the seedling stage temporally intermediate to time t and time $t+1$:

$$S_{t+1} = v(1-g_1)(1-c_1)A_t e^{\beta_{0f}+\beta_{1f}A_t} + S_t v(1-g_2) \quad \text{(3.E.1a)}$$

$$Sdl_t = v^{2/3} g_1 s_1 (1-c_1) A_t e^{\beta_{0f}+\beta_{1f}A_t} + v^{2/3} g_2 S_t \quad \text{(3.E.1b)}$$

$$R_{t+1} = Sdl_t \frac{1}{1+e^{-(\beta_{0sdl}+\beta_{1sdl}\log_e(Sdl_t+1))}} \quad \text{(3.E.1c)}$$

$$Rw_t = R_t \frac{1}{1+e^{-(\beta_{0s}+\beta_{2s}(1-c_1)A_t R_t+\beta_{2s}(A_t(1-c_1)+R_t))}} \quad \text{(3.E.1d)}$$

$$A_{t+1} = (1-c_2)Rw_t \frac{1}{1+e^{-(\beta_{0w}+\beta_{1w}\log_e(Rw_t(1-c2)+1)}} \quad \text{(3.E.1e)}$$

By setting β_{0sdl} to 0.5, the maximum seedling survival rate is 0.6225 when the strength of density dependence is zero. This is approximately equal to the highest rate compiled from the literature by Davis et al. . As β_{1sdl} is varied from 0: -1, the strength of the density dependent feedback on seedling survival probability increases. In the absence of management (fixing c_1 and c_2 at 0) the maximum density of the unmanaged population

decreases sharply as β_{1sdl} increases, although the population dynamics remain stable (Figure 3.E.1).

To explore the effects of managing rosettes or adults across this range of increasing seedling density dependence, c_1 and c_2 were varied from 0:1 as β_{1sdl} was varied from 0:-1 (Figure 3.E.2). The lines in Figure 3.E.2 were created by running a 2000 year simulation from Equation set F with each combination of c_1 or c_2 as β_{1sdl} was varied from 0 to 1 and discarding the first 500 years of population sizes to eliminate transient dynamics. For each level of seedling density dependence (β_{1sdl}), the proportion of the unmanaged population size from each simulation was calculated by dividing the maximum population size at each value of along the x axis by the maximum population size when $x = 0$. Hence, all lines cross the y axis at 1. In each case, increasing the strength of seedling density dependence (as β_{1sdl} becomes more negative) reduces the responsiveness of the population to management. At the strongest level of seedling density dependence simulated ($\beta_{1sdl} = -1$), increasing rosette management actually causes an increase in population size, whereas increasing adult management still suppresses population size.

Seedling density could negatively affect the responsiveness of the population to rosette management if it is strong enough. Dismissing its potential effects on population dynamics an management may be unwarranted.

Figure 3.E.1 Bifurcation of unmanaged *A. petiolata* populations with varied strength of seedling density dependence modeled from equation set F. Variability in the strength of density dependence was expressed by changing the slope parameter in the logistic function for seedling survival to the rosette stage (β_{1sdl}) from 0 to -1. The case of no density dependence is shown on the far left where $\beta_{1sdl} = 0$ and seedling survival probability is fixed by the intercept parameter at 0.6225. The dashed horizontal reference line shows the maximum equilibrium size of the population with no density dependence.

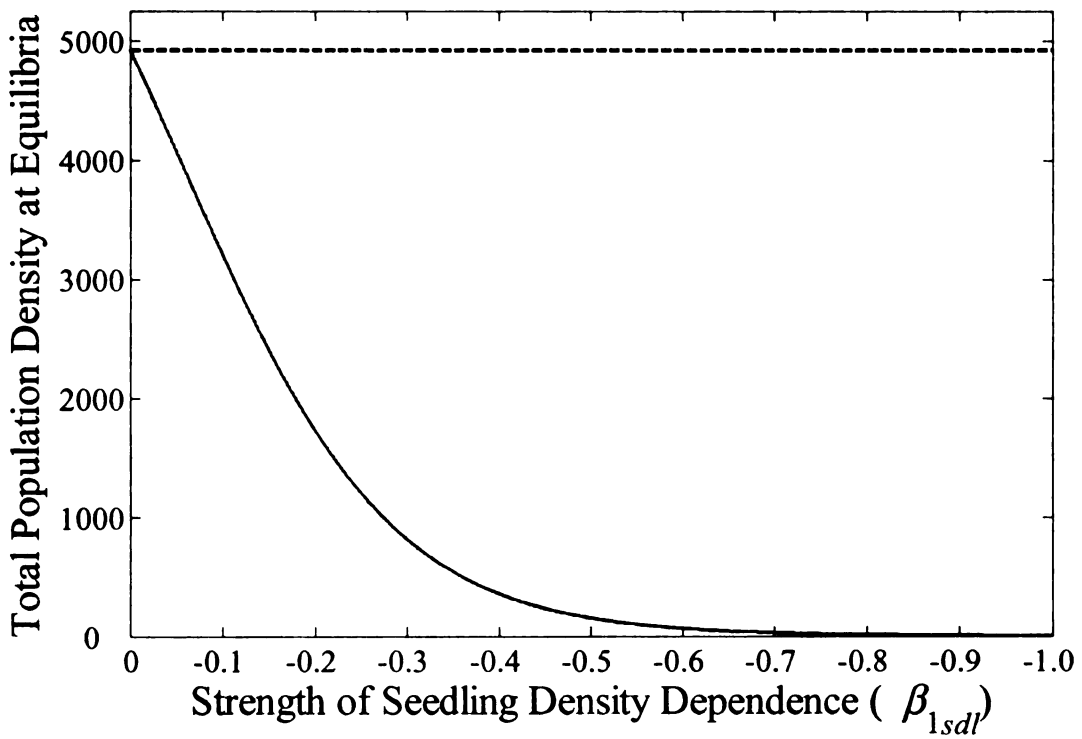
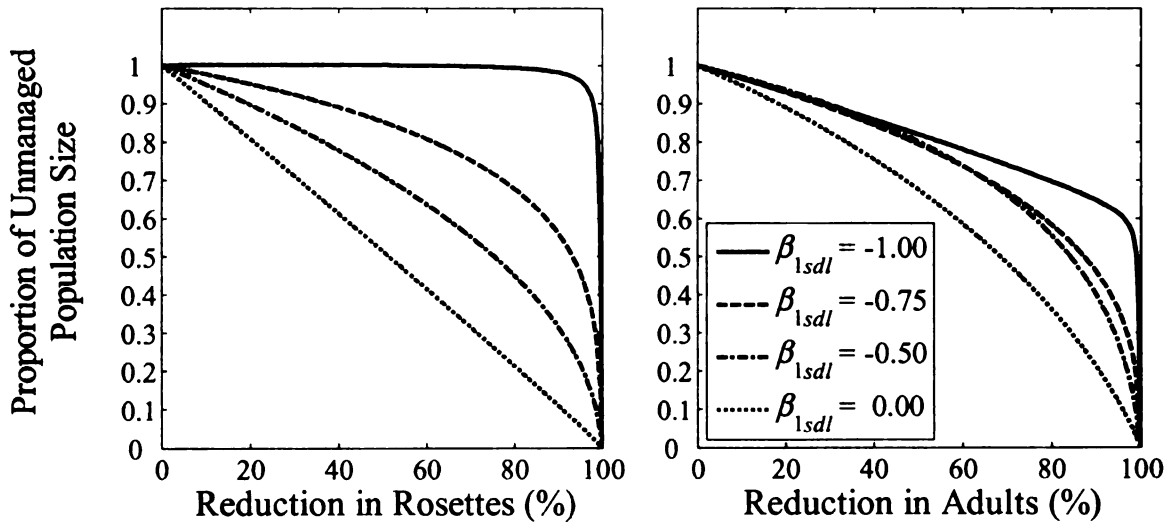


Figure 3.E.2. Proportional change in *A. petiolata* equilibrium population densities across a range of varied strength of seedling density dependence and simulated management of rosettes (left) and simulated management of Adults (right). The lines show the proportion of the unmanaged population size for a given strength of seedling density dependence as the strength of rosette or adult management is increased. Strength of seedling density dependence increases as values of β_{1sdl} become more negative (see Figure 3.E.1). For example, when there is no seedling density dependence, ($\beta_{1sdl} = 0$, dotted line), increasing rosette mortality causes an approximately linear proportional decrease in maximum population size, relative to the size of an unmanaged population with no density dependence.



APPENDIX 3.F: MULTICOLLINEARITY SIMULATION:

MATLAB COMPUTER CODE

```
% Appendix B from J. Evans: Response to Pardini et al. (2009)
% Multicollinearity Simulation
% Created with MATLAB 7.6.0.324a (R2008a)
% To run this program, open a new editor in MATLAB and paste the entire
% script below into it. Save the file as '\Multicollinearity_Demo.m'.
% The product of two random variables, x and y, will be correlated even
% when x and y are independent. If z is an additive function of x and y
% (Model Function A below) or a linear function of the sum of x and y
% (x+y), (Model Function B below), z will also be strongly correlated with
% their product (x*y). This can lend false support to regression models
% with an interaction term for x*y if the additive terms for x and y are
% excluded.
%
% The simulations below illustrate this. First run the script with model
% function A. The figure will show the relationships among x, y, z, x+y,
% and x*y, including correlation coefficients. If the MATLAB statistics
% toolbox is installed, four linear regressions will be run relating z to
% different combinations of x and y. Because we know the relationship of z
% to x and y, we can judge which correlations are 'true' and which are
% spurious.
%
% When z is a linear function of x and y with no interaction (Model
% Function A), both (x+y) and (x*y) are spuriously picked up as significant
% predictors of z if the main effects of x and y are excluded from the
% model, as seen below in the output of regressions (m3) and (m4). If the
% main effects are retained (m5), the regression correctly identifies the
% interaction  term as non-significant. The same is true when z is a
% function of (x+y) (Model Function B).
%
% When there is a significant interaction between x and y, substituting the
% sum of x and y (x+y) for the two independent main effects of x and y
% results in incorrect parameter estimation. Run Model Function C and
% compare the parameter estimates for the intercepts and interaction terms
% in regression analyses (m3), (m4) and (m5). When the regression includes
% both main effects and the interaction (m5), the parameter estimates are
% accurate, even thought the P-values for the main effects are not always
% significant. However, substituting (x+y) for the main effects results in
% underestimation of the interaction coefficient and gross underestimation
% of the intercept (m3). Excluding just the main effects x and y results in
% overestimation of the intercept and the interaction term (m4).
%
% This is true also when the main effects are small but the interaction is
% significant and is one of the principal reasons to not drop non-significant
% main effects. Run Model Function D. It is identical in form to C, but the
% slopes for the main effects of x and y are an order of magnitude smaller
% than in C. Compare the regression output from (m3), (m4) and (m5). In
% (m5), when the main effects are retained, the parameter estimates are
% close to their 'true' values. In (m3), where the main effects are
% replaced with the sum of x and y, the parameter estimate for the
% interaction is reasonably close to its true value of 2, although
% underestimated. The intercept is also underestimated and, because we know
% that z is a function of x and y, not (x+y), the result of a highly
% significant fit of (m3) to the data is spurious, due to the correlation
% between (x+y) and (x*y). In (m4), where the main effects for x and y are
% dropped, both remaining coefficients are overestimated. However, without
% the main effects, the increase in z as either x or y is increased will be
```

```
% an underestimate.
%
% Therefore, it is important to retain the main effects in the model, even
% though they could be non-significant. Information-theoretic statistics
% such as Akaike's Information Criteria (AICc) may provide clearer guidance
% for model selection than variable significance alone.
%
%
% Version Date 6/27/2009 Jeffrey Evans
%
%
%WWWWWWWWWWWWWWWWWWWWWWWWWW
%% Begin Simulation
clear all; close all;

% The simplest way to run the code is to set 'modelFun' below and press F5.
% Figure will be created, and regression output will appear in the command
% window if the statistics toolbox is installed.

% Model function A: z = 5 + 2x +6y;
% Model function B: z = 5 +3(x+y);
% Model function C: z = 5 + 2x + 6y + 2(x*y);
% Model function D: z = 5 + 0.2x + 0.6y + 2(x*y);
% Choose which model function to use by typing 'A', 'B','C', or 'D' between single quotes:
modelFun = 'b';

% Use a large sample size in simulation The sample size is set by
% changing the value of 'reps'.
reps = 100;

% generate two random, normally distrusted variables, x and y. x has a
% mean 3 and standard deviation of 1. y has a mean 5 and standard
% deviation of 1.
x = random('normal',3,1,reps,1);
y = random('normal',5,1,reps,1);


% Select among functional relationships between z, x, and y The random
% number generation for each parameter simulates the effects of process
% error in the relationship of z to x and y.
if ~isempty(strmatch('A',modelFun))
   z = 5 + 2.*x+6.*y + normrnd(0,1,reps,1);
   t = 'z = 5 + 2x +6y'; % title for figure

elseif ~isempty(strmatch('B',modelFun)) || ~isempty(strmatch('b',modelFun))
   z =5 + 3.*(x+y) + normrnd(0,1,reps,1);
   t = 'z = 5 +3(x+y)'; % title for figure

elseif strmatch('C',modelFun) % with significant interaction term
   z = 5 + 2.*x+ 6.*y + 2.*(x.*y) + normrnd(0,1,reps,1);
   t = 'z = 5 + 2x + 6y + 2(x*y)'; % title for figure

elseif strmatch('D',modelFun) % with small main effects and significant interaction term
   z = 5 + 0.2.*x + 0.6.*y + 2.*(x.*y) + normrnd(0,1,reps,1);
   t = 'z = 5 + 0.2x + 0.6y + 2(x*y)'; % title for figure
end
```

```
%% Create Scatterplots

% Set size and aspect ratio of figure. Defaults will be restored if entire
% code is run (by pressing F5) or when MATLAB is closed and restarted.
scrsz = (get(0,'ScreenSize'));
set(0,'DefaultFigurePosition',[50 50 scrsz(4)*.5 scrsz(4)*.75]);

figure(1);

% plot y vs. x
subplot(3,2,1);
plot(x,y,'ok')
set(gca,'FontName','Times')
xlabel('X','FontName','Times');ylabel('Y','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(x,y),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% plot  x*y vs x+y
subplot(3,2,2);
plot(x+y,x.*y,'ok');
set(gca,'FontName','Times')
xlabel('X+Y','FontName','Times');ylabel('X*Y','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(x+y,x.*y),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% plot z vs. x
subplot(3,2,3);
plot(x,z,'ok');
set(gca,'FontName','Times')
xlabel('X','FontName','Times');ylabel('Z','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(x,z),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% plot z vs. y
subplot(3,2,4);
plot(y,z,'ok');
set(gca,'FontName','Times')
xlabel('Y','FontName','Times');ylabel('Z','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(y,z),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% plot z vs. x+y
subplot(3,2,5);
```

```
plot(x+y,z,'ok');
set(gca,'FontName','Times')
xlabel('X+Y','FontName','Times');ylabel('Z','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(x+y,z),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% plot z vs. x*y
subplot(3,2,6);
plot(x.*y,z,'ok');
set(gca,'FontName','Times')
xlabel('X*Y','FontName','Times');ylabel('Z','FontName','Times');
text(0.1,0.9,['\rm\itr\rm = ' num2str(corr(x.*y,z),'%3.2f\n')],...
   'Units','Normalized',...
   'EdgeColor','k','BackgroundColor','w',...
   'FontName','Times') % overlay correlation coefficient
axis square

% Set title for figure. Code modified from 'suplabel' by Ben Barrowes %
% Copyright information at bottom of script.
h=axes('Units','Normal','Position',[.08 .08 .87 .87],'Visible','off');
set(get(h,'Title'),'Visible','on')
title(t,'FontSize',14);
t_ax=get(h,'Title');
set(t_ax,'VerticalAlignment','middle','FontName','Times')


%VVVVVVVVVVVVVVVVVVVVVVVVVVVV


%% Compare Regressions
% Note: This section requires that the MATLAB statistics toolbox be
% installed.

% Run four linear regressions of z vs. different combinations of x and y.
% The four regression models are:

% (m1) z = b0 + b1x + b2y (m2) z = b0 + b1(x+y) (m3) z = b0 + b1(x+y) +
% b2(x*y) (m4) z = b0 + b2(x*y) (m5) z = b0 + b1x + b2y +b3(x*y)

% where b0 is the intercept and b1, b2, and b3 are slope estimates

% The regression parameter estimates and associated p-values will be output
% in the MATLAB command window.

[b1 dev stats1] = glmfit([x y],z,'normal');
[b2 dev stats2] = glmfit([x+y],z,'normal'); %#ok<NBRAK>
[b3 dev stats3] = glmfit([x+y x.*y],z,'normal');
[b4 dev stats4] = glmfit([x.*y],z,'normal'); %#ok<NBRAK>
[b5 dev stats5] = glmfit([x y x.*y],z,'normal');

out1 = cell(4,3);
out1{1,1} ='Parameter';
out1{2,1} ='Intercept'; out1{3,1} ='X'; out1{4,1} ='Y';
```

```
out1{1,2} = 'Coefficient';
out1{2,2} = b1(1); out1{3,2} = b1(2); out1{4,2} = b1(3);
out1{1,3} = 'P-Value';
out1{2,3} = stats1.p(1); out1{3,3} = stats1.p(2); out1{4,3} = stats1.p(3);
disp('(m1): z = b0 + b1x + b2y'); disp(out1)

out2 = cell(3,3);
out2{1,1} ='Parameter';
out2{2,1} ='Intercept'; out2{3,1} ='X+Y';
out2{1,2} = 'Coefficient';
out2{2,2} = b2(1); out2{3,2} = b2(2);
out2{1,3} = 'P-Value';
out2{2,3} = stats2.p(1); out2{3,3} = stats2.p(2);
disp('(m2): z = b0 + b1(x+y)'); disp(out2)

out3 = cell(4,3);
out3{1,1} ='Parameter';
out3{2,1} ='Intercept'; out3{3,1} ='X+Y'; out3{4,1} ='X*Y';
out3{1,2} = 'Coefficient';
out3{2,2} = b3(1); out3{3,2} = b3(2); out3{4,2} = b3(3);
out3{1,3} = 'P-Value';
out3{2,3} = stats3.p(1); out3{3,3} = stats3.p(2); out3{4,3} = stats3.p(3);
disp(sprintf('M3 is analogous to the summer rosette\n survival function in Pardini et al. (2009)'));
disp('(m3): z = b0 + b1(x+y) + b2(x*y)');disp(out3)

out4 = cell(3,3);
out4{1,1} ='Parameter'; out4{2,1} ='Intercept'; out4{3,1} ='X*Y';
out4{1,2} = 'Coefficient';
out4{2,2} = b4(1); out4{3,2} = b4(2);
out4{1,3} = 'P-Value';
out4{2,3} = stats4.p(1); out4{3,3} = stats4.p(2);
disp('(m4): z = b0 + b2(x*y)'); disp(out4)

out5 = cell(5,3);
out5{1,1} ='Parameter'; out5{2,1} ='Intercept'; out5{3,1} ='X';
out5{4,1} ='Y';out5{5,1} ='X*Y'; out5{1,2} = 'Coefficient';
out5{2,2} = b5(1); out5{3,2} = b5(2); out5{4,2} = b5(3); out5{5,2} = b5(4);
out5{1,3} = 'P-Value';
out5{2,3} = stats5.p(1); out5{3,3} = stats5.p(2); out5{4,3} = stats5.p(3); out5{5,3} = stats5.p(4);
disp('(m5): z = b0 + b1x + b2y +b3(x*y)'); disp(out5)


% Restore default screen size settings
set(0,'DefaultFigurePosition',scrsz);


%% Copyright Information for redistribution of 'suplabel' code
% Used to place equation on figure with subplots.
%
% Code available at Mathworks file exchange:
% http://www.mathworks.com/matlabcentral/fileexchange/7772
%
% Copyright (c) 2004, Ben Barrowes All rights reserved.
%
% Redistribution and use in source and binary forms, with or without
% modification, are permitted provided that the following conditions are
```

```
% met:
%
%    * Redistributions of source code must retain the above copyright
%      notice, this list of conditions and the following disclaimer.
%    * Redistributions in binary form must reproduce the above copyright
%      notice, this list of conditions and the following disclaimer in the
%      documentation and/or other materials provided with the distribution
%
% THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS
"AS
% IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO,
% THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR
% PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR
% CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL,
% EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO,
% PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR
% PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY
OF
% LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING
% NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS
% SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.
```

CHAPTER 4: THE PREDICTED RESPONSE OF *ALLIARIA PETIOLATA* TO BIOLOGICAL CONTROL: LINEAR DETERMINISTIC DEMOGRAPHIC MODELS

ABSTRACT

Temporal and spatial variation in the demographic rates of invasive weed species can affect the outcome of management. When classical biological control is applied as a management tool, the introduction of ineffective agents can increase the risk of impacts to non-target species. Building from an earlier published model of the invasive weed *Alliaria petiolata* (garlic mustard), I evaluated the sensitivities and elasticities of the asymptotic population growth rate (λ) to both annual and sub-annual demographic transition probabilities. The model was parameterized using data from twelve sites in Michigan and Illinois, USA, and over three plant generations. Annual estimates of λ ranged from 0.48 to 5.88 across all sites and years. Population growth rates within sites were highly variable, ranging from 0.80 to 5.88 within one site. I used a megamatrix model to summarize variation in growth within sites. Site growth rates (λ_M) ranged from 0.83 to 3.54 with an arithmatic mean of 1.90. Sensitivities and elasticities varied with λ, indicating that the transitions with the largest impacts on population growth differ for growing and declining populations. This makes management options dependent on population growth rate. Rosette survival (summer and winter) consistently emerged as the transition with the greatest effects on λ in populations with positive growth, as did germination of new seeds and fecundity. This result is consistent with previous predictions by Davis et al. (2006) that rosette survival should be targeted by management. My model raises the caveat that rosette survival is only a target transition when growth is positive; its proportional effect on λ decreases as λ decreases. I predict a lower probability of successfully managing *A. petiolata* with biocontrol than Davis et al. My simulations predict that the root-crown mining weevil *Ceutorhynchus scrobicollis* is

likely to control up 5 of 12 populations, while introducing a second species could extend control to as many 9 populations, although the probability of success is very low (< 0.1) at 4 of these 9 sites. Better data on the distribution of agent impacts is necessary to refine these predictions.

INTRODUCTION

Demographic analyses are increasingly being used to interpret the ecology of invasive plant species and to guide the development of population management plans (e.g. Vitousek and Walker 1989, Byers and Quinn 1998, Shea and Kelly 1998, Parker 2000, Sakai et al. 2001, Buckley et al. 2003a, b, Koop and Horvitz 2005, Shea et al. 2005). McEvoy and Coombs (1999) proposed that demographic analyses of weed populations could contribute especially to improving success rate of weed biological control and to reducing possible negative effects on native communities. Some of the safety concerns about weed biocontrol relate to the common practice of releasing multiple agent species per target as a means of finding one that successfully suppresses the target (the "lottery model", McEvoy and Coombs 1999, McEvoy and Coombs 2000, Denoth et al. 2002). There are two principal concerns with this approach. First, biocontrol agent species which establish but do not suppress their hosts are more likely to become abundant in the absence of negative feedbacks from their host plants. Indeed, the establishment of one or more ineffective biocontrol agents is a common occurrence (McEvoy and Coombs 1999, Denoth et al. 2002). This increases the risk of the agent indirectly affecting non-target species at higher trophic levels by altering food web structures (Pearson and Callaway 2003, Ortega et al. 2004, Pearson and Callaway 2005). To minimize this risk, it is desirable to release the smallest number of agents necessary to suppress the target weed. Second, testing the feeding preferences and host specificity of agents that do not ultimately control the target is an unprofitable investment of monetary resources and scientist time. Testing and introduction of each agent species cost $460,000

U.S. dollars in 1997 (McFadyen 1998 $611,155 in 2009, http://www.bls.gov/data/inflation_calculator.htm).

Using sensitivity and elasticity analyses of matrix population models, life history transitions in the target species' life cycle that have a large effect on the population growth rate (λ) can be easily identified (Caswell 2001) before biocontrol agents are released. This information can then be applied as a filter to the biocontrol agent selection process. By increasing emphasis on species that disrupt these critical transitions, biological control researchers can conduct a more directed search for agents that are most likely to have a large impact on the target plant (McEvoy and Coombs 1999). This could serve to reduce the total number biocontrol agent species introduced, thereby reducing the risk of non-target impacts, as well as the costs of testing ineffective agents.

Davis et al. (2006) were the first to apply predictive plant demographic models to the development of a new biocontrol program before the introduction of any biocontrol agents. A biological control program for the invasive biennial forb *Alliaria petiolata* (garlic mustard, Brassicaceae [M. Bieb] Cavara and Grande) has been in development since 1998 (Blossey et al. 2001b) with efforts focused on testing the host specificity of four European weevils in the genus *Ceutorhynchus* (Coleoptera: Curculionidae) that affect *A. petiolata* at different stages in its life cycle. Davis et al. (2006) used a linear, deterministic matrix population model of *A. petiolata* to estimate the probability of successfully controlling it with either one or two biocontrol agent species. Their model was parameterized using the ranges of reported *A. petiolata* demographic rates from multiple published sources. It generated predictions about the efficacy required of biocontrol agents that attack different stages in the life cycle to suppress the target

species, defined here as the reduction of λ to 1.0 or less, This is the level at which individuals exactly replace themselves. Comparing their model predictions to data on the impacts of the candidate biocontrol agents from host specificity trials in Europe, Davis et al. concluded that the majority (72%) of simulated populations were likely to be controlled effectively by a single agent that simultaneously reduces per capita reproductive output of mature plants by 49% and survival of rosettes to the flowering stage by 43%.

One other demographic model of *A. petiolata* has been published which made predictions about which life history stages should be targeted for management (Pardini et al. 2009). Unfortunately, a number of biological and statistical problems in the conception and implementation of this model as published render its predictions about *A. petiolata* population dynamics and management questionable. An extensive discussion of Pardini et al.'s model is presented in Chapter 3.

The approach used by Davis et al. (2006) considers the potential effects of managing one or two life history transitions simultaneously. This is likely a reflection of the kinds of impacts that would be seen if either one or two of the candidate biocontrol agents were introduced and performed at its maximum possible efficacy. However, the Davis model only considered the asymptotic dynamics of individual populations by holding the transition probabilities within populations static through time. Although they looked at the projected responses of *A. petiolata* across a range of demographic rates, their analysis did not account for the possibility that demographic rates vary through time within populations. Because natural populations can be highly variable (in population structure, dynamics, etc...) across locations and over time, it is important to consider how

the variation in a population's demographic rates over time will affect its long-term response to perturbations (i.e. management). The transition matrices in Davis et al.'s analysis were assembled by varying the demographic parameters individually using a single natural population as a starting point. Because of this, the matrices they used do not necessarily reflect the correlations among parameters likely to be found in natural populations.

I have previously shown that *A. petiolata* demographic rates are highly variable across its range within the North Central Region of the United States (2006, Chapter 2) This variation is highly structured across populations and over multiple plant generations as a function of population density, climate, and soil attributes. This finding suggested that a static matrix modeling approach in which the transition probabilities are fixed through time likely does not accurately represent *A. petiolata*'s population dynamics. In this paper, I build on the matrix models of *A. petiolata* by Davis et al. (2006) and extend them using data from multiple natural populations of *A. petiolata*. My goal was to assess variation in population growth rates and population structure of *A. petiolata* across its demographic parameter space. Here, I examine the effects of variation in the observed *A. petiolata* demographic rates on estimated population growth rates (λ), the sensitivities and elasticities of λ to variation in transition probabilities among the study populations, and the projected response of populations to different combinations of simulated biological control management. I do this using linear, deterministic matrix population models and with "megamatrix" models (sensu Horvitz and Schemske 1986, Pascaraella and Horvitz, 1998) which account for temporal variation in vital rates within sites by cross-classifying individuals by life stage and habitat condition.

METHODS

Study System and Data Collection

Alliaria petiolata is a common invasive weed of forested ecosystems across much of North America (USDA-NRCS 2007). Its general biology and ecology are well studied and have been summarized in several reviews (Cavers et al. 1979, Nuzzo 2000, Rodgers et al. 2008). I will briefly review its biology as recorded in Michigan and Illinois. All demographic statistics refer to data from Chapter 2 unless noted otherwise. Seeds germinate in March or April in Michigan and Illinois producing dense carpets of seedlings with densities (seedlings m^{-2}) ranging 75-6025 (mean = 1171, median = 878). Seedling survival (annual proportion surviving in 0.16-0.25 m^{-2} plots) is variable (range = 0-1, mean = 0.486) and negatively density dependent. Seedlings that survive form low rosettes of basal leaves by June. Rosettes persist over the winter as green plants. All surviving rosettes bolt and flower by May of their second year. Between 0 and 20% (median = 1.5%) of seedlings that emerge in the spring survive to flower. The majority of individuals self-pollinate (Durka et al. 2005). Seeds are produced in slender fruits (siliques) along the upper stem. The number of seeds per silique increases nonlinearly with plant size, reaching an asymptote of around 14 seeds per silique in large plants in my system (see Appendix 2.B.iv). All second year plants senesce by August and set their seeds over the summer as the siliques dry and become brittle. Individual plants produced 0-1705 (mean = 85.7 seeds $plant^{-1}$) seeds (Chapter 2). Seeds require cold stratification to germinate (Baskin and Baskin 1992, Anderson et al. 1996, Raghu and Post 2008). Between 6.1-68.2% (mean =30.1%) of newly shed seeds germinate after one winter,

although there is strong inter-annual variation in this rate. Seeds which do not germinate can remain dormant for twelve or more years in the soil (Nuzzo and Blossey, unpublished data). Annual seed survival probability is high (mean = 0.919), while germination probabilities of dormant seeds after two winters ranged from 0.001-0.569 (mean = 0.110). These each contribute to the longevity of *A. petiolata*'s seed bank.

The search for biocontrol agents for *A. petiolata* was initiated in 1998. From an initial pool of over 70 phytophagous insects known to feed on *A. petiolata* in Europe, four weevil species in the genus *Ceutorhynchus* were considered the most promising and have been studied in feeding preference trials (Blossey et al. 2001b, Davis et al. 2006, Gerber et al. 2007b, Gerber et al. 2007a, 2008a). The four weevils each complete their life cycle on *A. petiolata*. All feed on *A. petiolata* foliage as adults and decrease fecundity by reducing photosynthetic capacity. As larvae they affect *A. petiolata* at different stages. The multivoltine root crown feeder *C. scrobicollis* aestivates during summer in Europe and oviposits on rosette leaves from mid-September through April while mean temperatures are above -5°C (Gerber et al. 2007a). Larvae hatch from eggs laid in buds, leaves, and petioles and mine in the root crown over the winter. In spring, *C. scrobicollis* larvae leave *A. petiolata* and pupate in the soil, emerging as adults in May-June (Gerber et al. 2007a). The univoltine stem-borers *C. alliariae* and *C. robertii* each oviposit on *A. petiolata* in early spring (Gerber et al. 2008a). Larvae mine within stems and petioles until late spring and pupate in the soil. Adults emerge in June and July, feed on *A. petiolata* foliage, and overwinter in the litter. Finally, *C. constrictus* feed on developing seeds in June and July as larvae. Each larva consumes 2-3 seeds on average before

migrating to the soil to pupate (Gerber et al. 2008b). *Ceutorhynchus constrictus* overwinter as pupae and emerge as adults in the spring.

Host specificity testing of *C. scrobicollis* and *C. alliariae* are the closest to completion and these species are the most likely of the four to be introduced first (Hinz and Gerber 2005). The impacts of these two species varied across several feeding trials conducted in greenhouse and common garden experiments from zero effect on seed production by either agent (Gerber et al. 2007a, 2008a), zero effect on rosette survival by *C. alliariae* (Gerber et al. 2007b) and 7% reduction of rosette survival by *C. scrobicollis* (data from Hinz and Gerber 2001). By itself, *C. scrobicollis* has caused up to a 49% reduction in fecundity and a 45% increase in rosette mortality (Gerber et al. 2002). The combined action of the two agents caused up to 82% and 50% reductions in rosette survival and fecundity, respectively (Gerber et al. 2002) Given the small sample sizes used and wide range of impacts caused by these species in different experiments, it is hard to know exactly what effects they might have on *A. petiolata* if released in North America or what kinds of error terms to associate with their predicted efficacies. This uncertainty makes demographic models potentially useful in estimating impacts by allowing simulation of different scenarios.

Demographic Data

Plant demographic data were collected from 12 *A. petiolata* populations in southern Michigan ($n = 7$) and Illinois ($n = 5$) between 2005 and 2008 (Table 2.1). Site selection was based on accessibility and on the presence of established, unmanaged *A. petiolata* populations. Demographic rates were measured during three sampling periods

denoted here by the year in which the sampling period began (*year* = 2005, 2006, and 2007). Beginning in 2005, four replicate measures of each demographic rate were made at each site during each sampling period. Four sampling areas were established within each site (n = 48 experimental units total): two at the forest edge and two in the forest interior. Analysis of spatiotemporal variation in demographic rates indicated no differences between edge and interior quadrats for any demographic rate, so all replicates were treated equally Chapter 2.

Survival rates were recorded in three quadrats within each sampling area. Seedling survival to the rosette stage (s_r, seedling survival) was recorded in 20 x 20 cm (IL) and 25 x 25 cm (MI) quadrats between late February and June of each year. Davis et al. (2006) modeled rosette survival from June to June as a single survival parameter abbreviated s_{rf}. I split this transition into two lower-level demographic rates to improve the temporal resolution of the analysis and model. Rosette survival from June to October (s_{sum}, summer survival) and from October to June the following year (s_{win}, winter survival) were recorded in 40 x 40 cm (IL) and 50 x 50 cm (MI) quadrats. The positions of individual rosettes within each quadrat were marked on transparent plastic data sheets in June each year. Survivors were located in October and again the following June using the transparency as a map. S_{sum} and s_{win} can be multiplied together to generate s_{rf} in the same temporal currency of Davis et al.'s parameter. Fecundity was estimated from all survivors in the rosette quadrate in June of the second year by counting the number of siliques per plant and scaling by the number of seeds per silique estimated from nearby plants not included in the demographic censuses (Appendix 2.B.iv).

Germination of one-year-old and two-year-old seeds and seed survival from June to June were followed in 20 x 20 cm stainless steel wire mesh trays buried at the soil surface. Each experimental seed tray was paired with a control tray that did not receive seeds. This yielded estimates of the probability of germination of new seeds after one winter (g_1) and the probability of germination of older dormant seeds from the soil seed bank (g_2). The annual survival probability of dormant seeds (s_s) was also measured from these seed trays. Greater detail on all sampling procedures can be found in Evans et al. (unpublished manuscript). Mean values of each demographic rate were calculated within each site during each year for use in the present analysis (n = 36 site x year mean rates). Replicate measures of some demographic rates were lost to vandalism or animal disturbance. In other cases, demographic rates could not be measured because there were zero plants at the beginning of a transition interval (e.g. if all rosettes died in summer, winter rosette survival could not be estimated). In each of these cases, the replicate was omitted from the calculation of the mean transition rate.

Matrix Model Construction

Davis et al. (2006) described the construction of a matrix model for *A. petiolata* which I reiterate here with some modifications and apply to the parameterized matrices from the 12 study sites. The life cycle of *A. petiolata* in the northern continental United States is described in terms of the three life history stages present in mid-summer just before seed set: dormant seeds in the soil seed bank (S), first year rosettes (R), and flowering second-year plants (P) about to set seed (Figure 4.1). In this life cycle diagram, each arrow represents a transition that an individual can make between two stage classes

over the course of a full year (e.g. $S \rightarrow R$, $R \rightarrow P$). Following the terminology in Caswell (2001), transitions which occur over one full projection interval of a model (one year, in this model) are referred to as upper-level transitions. Transitions that do not occur in North American *A. petiolata* populations were assumed to have zero probability of occurring and are not represented in Figure 4.1. Newly shed seeds produced by flowering plants in June of year 1 can either germinate after one winter and become rosettes in year 2 with some probability ($P \rightarrow R$ transition) or they can remain dormant and enter the soil seed bank in year 2 ($P \rightarrow S$). Note that the $P \rightarrow R$ transition only occurs via seed production. Flowering plants do not perenniate or become rosettes. Dormant seeds in the seed bank in year 1 can either germinate in the spring to become rosettes in year 2 ($S \rightarrow R$) or can remain dormant for another year ($S \rightarrow S$). There is no stasis in the rosette stage. Because *A. petiolata* does not flower in the year that it germinates, rosettes in year 1 either transition to the flowering plant stage in year 2 ($R \rightarrow P$) or die.

The upper-level transitions are quantified as the products of sub-annual, lower-level transitions. For *A. petiolata*, the five nonzero upper-level transitions are composed from combinations of seven lower-level transitions which I measured in the field and appear as the equations alongside each arrow. The lower-level demographic rates quantify the probabilities associated with each lower level transition. Three changes were made from the model in Davis et al. (2006). First, the *R* to *P* transition (s_{rf} in Davis et al.) was split into the lower-level transitions s_{sum} and s_{win} because summer and winter survival were measured in separate quadrats in Michigan in 2005. Summer survival was measured in the smaller seedling quadrats that year, and summer and winter survival was measured in the larger rosette quadrats thereafter. Second, because seed survival (s_s) was

quantified in the field over a full year but is used twice as a sub-annual transition ($P \rightarrow R$, $S \rightarrow R$), I scaled the observed value proportionally to account for s_s over 8 months in these transitions by raising it to the two-thirds power as $s_s^{2/3}$. Third, following Davis et al., the variables c_1 and c_2, which simulate rosette mortality and fecundity reduction, are shown as $(1\text{-}c_n)$ to express them in terms of survival. The transition probabilities in the life cycle diagram make up the entries in the projection matrix **A**. The **A** matrix has one row and one column for each life history stage in the population (Figure 4.1). The 3 x 3 **A** matrix for *A. petiolata* has rows 1, 2, and 3 and columns 1, 2, and 3 for seeds, rosettes, and flowering plants, respectively. The individual elements in the **A** matrix are "the a_{ij}" (the entry in the i^{th} row and the j^{th} column number) and represent probabilities of an individual in life stage j at time t transitioning to stage i at time $t + 1$ (Caswell 2001). Thus a_{32} is the probability of individuals that are rosettes (stage 2) at time t surviving to becoming flowering plants (stage 3) at time $t + 1$, while a_{23} corresponds with the transition from flowering plants at time t to rosettes at time $t + 1$. One **A** matrix was parameterized for each site during each year using the mean lower-level parameters, yielding 3 matrices per site, for a total of 36 A matrices

The numbers of individuals in each stage class are tracked in a column vector, **n**, which has one row for every life history stage. The entries in rows 1-3 ($\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$) represent the numbers of seeds, rosettes, flowering plants, respectively, in the population at a given time. Given a population vector $\mathbf{n}_t$, the population size can be projected

forward to time $t + 1$ using matrix multiplication to multiply the transition matrix by the population vector as

$$\mathbf{n}_{t+1} = \mathbf{A}_{sy}\mathbf{n}_t \tag{4.1}$$

where $\mathbf{n}_{t+1}$ and $\mathbf{n}_t$ are the population vector at times $t + 1$ and t, respectively, and $\mathbf{A}_{sy}$ is the projection matrix for site s in year y. The population growth rate (λ) was calculated as the dominant eigenvalue of the **A** matrix (Caswell 2001). When $\lambda > 1$, the population size increases, whereas $\lambda < 1$ indicates population decline. In terms of demography, the goal of *A. petiolata* management is to persistently drive λ below 1. The sensitivities and elasticities of λ to perturbation of the a_{ij} were used to analyze which transitions have the greatest impact on λ. The sensitivity of each matrix element is the partial first derivative of λ with respect to a_{ij}. The sensitivities are the local slope of λ at a given value of a_{ij} (Caswell 2001) and are compiled in the sensitivity matrix **S** as:

$$s_{ij} = \frac{\partial \lambda}{\partial a_{ij}} \tag{4.2}$$

Elasticities of λ to perturbation of the a_{ij} scale the sensitivities (s_{ij}) by their proportional contributions to λ. For each matrix $\mathbf{A}_{sy}$ the elasticity matrix **E** contains the elasticities of lambda to the a_{ij}. These are calculated as

$$e_{ij} = \frac{a_{ij}}{\lambda}\frac{\partial \lambda}{\partial a_{ij}} \tag{4.3}$$

where e_{ij} is the elasticity of λ to a proportional change in matrix element a_{ij}, $\partial\lambda/\partial a_{ij}$ is the sensitivity of λ to an additive perturbation of a_{ij}, calculated as the partial derivative of λ

with respect to a_{ij}, and the e_{ij} sum to 1 within each matrix **E** (Caswell 2001). Because management of *A. petiolata* is likely to be applied at the level of the lower-level transitions, which are components of the a_{ij} (e.g. management of fecundity, management of seedling survival, etc...), it is useful to analyze how λ responds to perturbations of the lower-level transitions. The sensitivities of λ to the lower-level transitions are:

$$S_x = \frac{\partial \lambda}{\partial x} = \sum_{ij} \frac{\partial \lambda}{\partial a_{ij}} \frac{a_{ij}}{\partial x} \tag{4.4}$$

and the elasticities of λ to the lower-level transitions are:

$$E_x = \frac{x}{\lambda} \frac{\partial \lambda}{\partial x} = \frac{x}{\lambda} \sum_{ij} \frac{\partial \lambda}{\partial a_{ij}} \frac{a_{ij}}{\partial x} \tag{4.5}$$

where E_x is the elasticity of λ to a proportional perturbation of the lower level transition x (Caswell 2001).

I used two previously described methods (Davis et al. 2006) to visualize the potential impacts of the proposed biological control agents on *A. petiolata*. I first calculated the sensitivity and elasticity of λ to each of a) the five non-zero upper level transitions (a_{11}, a_{21}, a_{32}, a_{23}, and a_{13},) and b) the seven lower-level transitions (s_s, s_r, g_1, g_2, s_{sum}, s_{win}, and f) for each of the 36 **A** matrices. The sensitivity and elasticity of lambda to each of these transitions was then plotted against the observed transition values. For example, the elasticities (E_x) of λ to each of the lower-level transitions were plotted against g_1 to show how the rankings of the E_x change as g_1 changes. This is, effectively, a set of empirical second derivatives of lambda with respect to each transition. It illustrates how the response of lambda to a proportional perturbation in one

of the lower-level transitions changes across the observed range of each transition. This is used to interpret the relative potential magnitudes of the response of λ to management of each lower-level transition. Rankings of the sensitivities and elasticities were compared among populations with low and high growth rates. Pearson correlation coefficients between the elasticities and λ were used in some cases to help elucidate these patterns. The natural logarithm of λ was used in these calculations to meet the assumption of normality. Because lower-level elasticities do not sum to 1 as the upper-level statistics do, they are not interpretable as proportional contributions to λ in the same way (Caswell 2001).

In the second set of analyses, I used a megamatrix model (Pascarella and Horvitz 1998) for preliminary exploration of the effects of Markovian environmental variation on projected biocontrol outcomes. Note that the megamatrix analysis with uncorrelated environments is equivalent to a conventional analysis of the weighted mean matrix $\bar{\mathbf{A}}$ created from the means of the a_{ij} across the three years within a site (Tuljapurkar et al. 2003). I chose to pursue the megamatrix approach because it set up a flexible analytical framework that can be adapted in future models as additional environmental data becomes available, and because of difficulty in estimating appropriate weights associated with the mean **A** matrix. In a megamatrix model, individual plants are cross-classified by life stage and by the state of the environment, defined here by the three annual **A** matrices within each site. The three annual **A** matrices parameterized for each site characterize a portion of the local variability in *A. petiolata* demographic rates which is driven by a combination of intrinsic and extrinsic biotic and abiotic factors (see Chapter 2). For the purpose of this analysis I assume these are representative of the true variation in **A** for

each site. For the megamatrix analysis, I create a 3 x 3 matrix **P** of environmental transition probabilities with one row and one column for each of three environmental states (**A** matrices) parameterized for each site. The **P** matrix characterizes how the environment changes over time. Each entry p_{kl} in **P** is the probability that the environment in state l at time t will transition to state k at time t+1, where the time interval t to t+1 is the same as the projection interval of the population matrix **A** and the probabilities within each row of **P** sum to 1. The m x n population projection matrix **A** is combined multiplicatively with the p x q environmental transition matrix **P** to create the mp x nq megamatrix **M**, whose entries quantify the probability that an individual in stage j in environment l at time t transitions to stage i in environment k at time t+1. See Pascarella and Horvitz (1998 Appendix B) for details on the parameterization of **M**. The dominant eigenvalue of **M**, λ_M, is the asymptotic growth rate of the population as the environment changes according to **P**. In the analysis presented here, I assumed that the environments were uncorrelated, where the probabilities p_{kl} in **P** were equal for all k,l and generate a stochastic sequence of environments. The 3 x 3 environmental matrix is

$$\mathbf{P} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \tag{4.6}$$

and was used to generate the megamatrix **M**.

For each site I created zero growth isoclines (contours of $\lambda = 1$) from the megamatrix by calculating λ_M for all factorial combinations of c_1 and c_2 between 0 to 1

in increments of 0.001, and then creating a contour plot of $\lambda_M = 1$ versus c_1 and c_2. These were interpreted by overlaying them on plots showing the expected range of biocontrol agent impacts on *A. petiolata* rosette survival and fecundity estimated from common garden and greenhouse studies during agent host specificity testing (Gerber et al. 2007b, Gerber et al. 2007a, 2008a, E. Gerber and H. Hinz, personal communication). Sites whose zero growth isoclines intersected the plotted range of biocontrol impacts can potentially be controlled. The empirical probability distributions of agent impacts on *A. petiolata* are unknown. We only know that the range of impacts each agent has is from zero to an observed maximum. The probability of successfully controlling *A. petiolata* at each site was assessed from the plots assuming a uniform probability distribution for agent impacts (i.e. an agent is equally likely to have any impact from zero up to its maximum). For each site, the ranges of possible impacts from *C. scrobicollis* alone and from *C. scrobicollis* with *C. alliariae* were each plotted as shaded rectangles, and the zero growth isoclines were plotted over them. The probability of success was then estimated as the proportion of the shaded area above the zero growth isocline.

I also considered the annual variation in the response to biocontrol, although management recommendations are based on the megamatrix results. For each site I created zero growth isoclines (contours of $\lambda = 1$) from the three annual **A** matrices, and varied rosette mortality and per capita fecundity to simulate the effects of the proposed biocontrol agents. The annual zero growth isoclines were created by solving the characteristic equation of each matrix for c_1, the simulated reduction in per capita fecundity, in terms of the observed demographic rates, λ, and c_2, the simulated reduction in per capita rosette survival:

$$c_1 = \lambda^3 - s_s(1-g_2)\lambda^2 \div \left[c_2\left(-s_s(1-g_2)fs_s{}^{2/3}g_1s_r\right)+ \lambda c_2\left(fs_s{}^{2/3}g_1s_rs_{sum}s_{win}\right)+c_2\left(fs_s(1-g_1)s_s{}^{2/3}g_2s_rs_{sum}s_{win}\right)\right] \quad (4.7)$$

Equation 4.7 is based on Equation 4 in Davis et al. (2006) but incorporates changes in the calculation of vital rates described above. Details on its derivation are available in the Appendix in Davis et al. (2006). By setting λ to 1 and varying c_2 from 0 to 1, I was able to calculate combinations of c_1 and c_2 that correspond with stable population size and plot zero growth isoclines for each annual matrix within each site.

RESULTS

Population growth rates

Years: There was considerable temporal variation in population growth rates (λ) within each site. Over the three years of study, λ varied among the twelve sites from 0.48 to 5.88 (Table 4.2). Some sites consistently had low annual population growth rates. Two (HR, RL) never reached a λ of 1. Five sites had positive growth during a single year only, four during two years only, and just one site, HSP, had positive growth during all three years. Within five sites (41%) λ was greatest during the first year of the study and decreased in years two and three. Temporal variation in λ across all sites is summarized by comparing the column arithmatic means of the annual λ estimates in Table 4.1. Overall, λ was greatest in 2005 and was nearly identical thereafter, dropping from 1.825 to 1.014 and 1.019 successively.

Sites: Spatial variation in λ across sites is evaluated with λ_M, the megamatrix λ for each site (Table 4.2). Estimates of λ_M ranged from 0.932 (HR) to 3.543 (B) and were

frequently greater than even the highest annual λ within a site Figure 4.2. For example, the highest annual λ in the JP site was 1.595 (2006), while λ_M was 2.298. There were no consistent geographic patterns in the spatial distribution of λ within years, although the two sites with the highest λ_M were both in southern Illinois. The overall mean of the12 λ_M estimates was 1.897, meaning that on average the *A. petiolata* study sites supported rapid positive growth. This number is within the range of published estimates of λ for *A. petiolata* from previous studies which vary from 1.2 (V. Nuzzo and B. Blossey, unpublished data in Davis et al. 2006) to 4.4 (Drayton and Primack 1999, Rejmanek 2000). Only one site had negative growth (HR). The two sites with the highest site mean λ, both in Illinois, each had a $\lambda > 5$ during 2005 (Table 4.2), and in one of these sites (B) λ dropped to below 1 during the following two years.

Sensitivity and Elasticity Analyses

Upper-Level Transitions: Across the study populations, management which reduces the transition of rosettes to the flowering stage ($R \rightarrow P$) is expected to cause the greatest absolute reduction in λ. The sensitivity of λ to perturbation of the a_{ij} varied in rank order among the 36 **A** matrices. These are shown in Figure 4.3 (upper panel) sorted from left to right by s_{32}, the sensitivity of λ to a_{32} ($R \rightarrow P$). The population growth rate λ most frequently had the greatest sensitivity to perturbation of a_{32} ($R \rightarrow P$) followed by a_{21} ($S \rightarrow R$). This is a biologically intuitive result, as a_{32} is the proportion of *A. petiolata* individuals that survive to reproductive maturity. Raising a_{21} similarly increases the rate

at which dormant seeds exit the seed bank to become rosettes and can progress toward reproductive maturity. Sensitivity of λ to a_{21} was only large when s_{32} was small. The mean sensitivity to a_{32} was over twice that of a_{21} (5.7 vs. 2.4), and the sensitivities of λ to all other upper-level transitions were comparatively low.

In contrast, the elasticity (e_{ij}) of λ to a_{11} ($S \rightarrow S$) was, on average, much greater than the other elasticities. However, e_{11} was strongly, negatively correlated with $\log_e(\lambda)$ (log transformed to meet assumption of normality) across the 36 **A** matrices ($r = -0.9118$); e_{11} was large almost exclusively among poorly performing populations with λ near or below 1. This occurred because a_{11} was largely driven by seed survival (s_S), which had relatively little variation and was almost universally high, while there was much more variation in the other a_{ij} and lower-level transitions (Table 4.1). The elasticity of λ to a_{11} becomes inflated when reductions in the other more variable transition probabilities cause λ to decrease. When this happens, a_{11}/λ, which is multiplied by the sensitivity of λ to a_{11} (s_{11}) to calculate e_{11} (Equation 4.3), becomes disproportionately large and drives up e_{11}. Thus, although a_{11} appears to rise in importance in a number of cases, it does so somewhat passively by not changing in synch with other transitions and only makes large contributions to λ in populations where λ is already low (Figure 4.4, left). Russ Forest (2007) is the most extreme example of this, where there was 100% summer mortality in all quadrats which forced winter survival and fecundity to 0. The A and E matrices for this site in 2007 were:

$$\mathbf{A} = \begin{bmatrix} 0.7688 & 0 & 0 \\ 0.0388 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.8)$$

and

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.9)$$

Because the individuals transitioning to the rosette stage in a_{21} ($S \rightarrow R$) at time t have 0 survival probability thereafter (because a_{32}, $R \rightarrow P$, = 0), a proportional change in a_{21} has no effect on λ. Only individuals in a_{11} ($S \rightarrow S$) contribute to the next time step, and $\lambda = a_{11} = 0.7688$. This was not the most robust population.

After a_{11}, a_{32} ($R \rightarrow P$) made the next largest contributions to λ (Figure 4.3, bottom). The elasticity of λ to a_{32} ($e32$) was positively correlated with $\log_e(\lambda)$ (r = 0.8152) and was greater than all other elasticities in the populations and years with the highest λ (Figure 4.4, left). Based on this, the transition from rosettes to the flowering stage is the most important management target among the fastest growing and most problematic populations. Transition a_{23} ($P \rightarrow R$) caused the next largest proportional response in λ followed by a_{13} ($P \rightarrow S$) and a_{21}, ($S \rightarrow R$), which had approximately equal e_{ij} (Figure 4.3).

Lower-Level Transitions: The population growth rate λ had the greatest elasticity to perturbation of dormant seed survival (s_S) in most sites and most years (Figure 4.5). Elasticities of λ to g_1 and to f, s_r, s_{sum}, and s_{win} were positively associated with increases in each of those five lower-level demographic rates, and were most strongly correlated with increasing s_{sum} and g_1. The elasticities of λ to f, s_r, s_{sum}, and s_{win} (abbreviated E_4) are shown with a single line (solid line, Figure 4.5) because their values are equal across the range of the data. E_4 exceeded E_{ss} at the higher levels of s_{sum}, (dash-dot line, Figure 4.5). Seed survival (s_S) had low variability (Table 4.1) and was only weakly correlated with $\log_e(\lambda)$ ($r = 0.2838$), while s_{sum} was positively correlated with $\log_e(\lambda)$ ($r = 0.7059$), and s_S and s_{sum} were uncorrelated ($r = 0.0256$). This and the plot of the elasticities versus λ (lower right, Figure 4.5) suggest that this reversal in ranking parallels what was seen in the upper-level elasticities (Figure 4.4, right). It results not from a tradeoff between s_S and s_{sum}, but rather from the response of λ to increasing s_{sum} as s_S remains stable and large in the background. The higher λ becomes, the greater its proportional response is to perturbation of all transitions except s_S and g_2. The sensitivities of λ to the lower level transitions reveal that this is accompanied by a larger increase in the magnitude of λ's response to additive perturbation of s_{win}, g_1, and s_{sum} (Figure 4.6). Thus, the more rapidly *A. petiolata* grows, the more vulnerable it becomes [demographically] to management of rosette survival and germination. Management should be preferentially directed at these transitions.

Zero Growth Isoclines

Alliaria petiolata showed a range of responses to the simulation of management (Figure 4.7). In Figure 4.7, the larger shaded area represents the ranges of increased rosette mortality (x-axis, 0-50%) and reduction of fecundity (y-axis, 0-82%) induced in *A. petiolata* during several studies of the candidate biocontrol agents *C. scrobicollis* and *C. alliariae* (Gerber et al. 2002, Gerber et al. 2007b, Gerber et al. 2007a, 2008a, E. Gerber personal communication). The smaller shaded area shows the range of observed impacts of *C. scrobicollis,* acting by itself, on fecundity (0-49%) and rosette survival (0-45%) across all experiments. Zero growth isoclines which intersect the shaded area or fall below the origin (and thus are not visible) indicate combinations of *A. petiolata* demographic rates and agent efficacy which suppress λ below 1. Predicted management outcomes are based on interpretation of the megamatrix zero growth isocline for each site. Annual zero growth isoclines are shown in Appendix 4.B

Nine of the twelve (75%) *A. petiolata* study populations are potentially controllable if *C. scrobicollis* and *C. alliariae* are released together. This assumes that management efficacy is at the extreme high end of the expected range for both agents (top right corner of shaded area, Figure 4.7). Of these eight sites, four have a predicted probability of being successfully controlled below 0.1 (Table 4.3). Fewer sites are predicted to be controlled if the agents perform below this level. If *C. scrobicollis* is released alone and performs at its maximum observed efficacy, *A. petiolata* could be suppressed at up to five sites (41%; HR, F, SH, RL, and RF). One of the five sites (RL) is predicted to have a 0.007 probability of being successfully controlled. However, if *C. scrobicollis* only reduces rosette survival by 7% (Hinz and Gerber 2001) and has no

effect on fecundity (Gerber et al. 2007a), this number drops to one site (HR), which is not an increase over the number of unmanaged sites with $\lambda_M < 1$.

The annual variation in demographic rates within each site caused considerable differences in agent efficacy required to suppress growth during individual years. Annual zero growth isoclines are shown in Appendix 4.B. At each site, at least one of the three annual zero growth isoclines crossed the shaded agent-impacts area. In all sites except HSP, which had positive growth each year, at least one isocline also falls below the origin of the figure.

DISCUSSION

Improving the predictability of weed biological control remains a common goal of weed scientists, biocontrol practitioners, and conservation managers (McEvoy and Coombs 1999, Louda et al. 2003b, Pearson and Callaway 2003, Briese 2006, Davis et al. 2006, Raghu and van Klinken 2006). Because biological control management of *A. petiolata* is a longer-term endeavor, management decisions should not be made based on data from individual years. Four demographic models of *A. petiolata* have been published previously (Drayton and Primack 1999, but see Rejmanek 2000, Meekins and McCarthy 2002, but see Evans et al. unpublished manuscript Appendix D, Davis et al. 2006, Pardini et al. 2009, but see Evans unpublished manuscript), two of which made specific management recommendations. Of these four previous models, three had significant errors that required later correction (hence the "but sees"). Drayton and Primack's model incorrectly characterized *A. petiolata*'s life history by omitting a_{23} ($P \rightarrow R$). Meekins and McCarthy used sub-annual transitions as upper-level transitions, thus introducing errors

in the model's projection interval. Pardini et al.'s model incorporated errors from the parameterization of several density dependent functions and omitted seed bank mortality from two transitions (a_{11} and a_{21}), which together sabotaged their management recommendations.

Davis et al. explored a parameter space centered on the demographic rates of a natural *A. petiolata* population (Meekins and McCarthy 2002) and made a number of specific predictions about the expected variability of successful management. One of their main conclusions was that the elasticity structure, at least in terms of relative rankings of the elasticities, was fairly constant as each vital rate was sequentially varied from 0 to 1 while the others were held fixed (Figure 3 in Davis et al. 2006). As they acknowledge, this method did not account for correlations or covariance among the vital rates likely to occur in natural populations. Because the matrices in my analysis were parameterized from natural populations and were fixed, the analysis implicitly accounts for these correlations and avoids having to make assumptions about the relationships among the a_{ij} (van Tienderen 1995).

In this study, I have shown that *A. petiolata*'s demographic parameter space is, in fact, larger and more varied than what Davis et al. modeled or than previous reports in the literature indicated. One of my most important findings is that, contrary to Davis et al.'s conclusion, neither the magnitudes nor rankings of either the lower or upper level elasticities are constant in space or time. We know that spatial variation in the importance of individual life history pathways can drive significant differences in population dynamics and response to management among isolated populations (Shea et al. 2005) and even within populations (Parker 2000). I show here that temporal variation in life history

within populations can also contribute significantly to variation in long-term population performance. *Alliaria petiolata* populations fluctuate over time through periods of rapid growth and expansion and periods of survival. As these fluctuations occur, the population dynamics change from being driven primarily by growth and reproduction, to being driven by the persistence of the seed bank. This is evidenced by the varying elasticity structures within populations over time. When λ was low, the survival of dormant seeds in the seed bank (s_s and a_{11}) had the largest proportional effect on growth, although the absolute magnitude (sensitivity) of this effect was small relative to other transitions. In contrast, when λ was high, the transition from rosettes to flowering plants (a_{32}) had the largest proportional effect on λ (Figure 4.5, bottom right panel). This effect was largely dependent on variation in summer rosette survival, which was the largest demographic bottleneck in the system. Other lower-level transitions with large proportional effects when λ was large were germination of new seeds, seedling survival, fecundity, and winter survival. The sensitivity analysis highlights the importance of winter and summer rosette management among populations where λ is large. Collectively, the elasticity and sensitivity analyses indicate that management which reduces the survival of rosettes to the flowering stage (a_{23}) or which reduces the transition from the flowering to rosette stages (a_{32}), either via reducing fecundity, seedling survival, or germination of new seeds, will have the greatest impact on λ during years with the highest growth rates.

The high variability in λ both among and within populations could also have significant consequences for the outcome of biological control. As a population experiences a sequence of years with alternating positive and negative growth, the

rankings and magnitudes of the elasticities and sensitivities can change. For example the 'Bob' site (Peoria) had a λ of 5.78 in 2005 and 0.79 in 2006. In 2005,the lower level elasticity of λ to s_S (E_{ss}) was 0.29, while the elasticity to E_4, which includes s_{sum} and s_{win}, was 0.4989. The following year, the rankings were reversed, and E_{ss} was 0.9666 while E_4 dropped to 0.0235. Practically speaking, this means that when a transition probability is driven to a small value relative to other transitions, a proportional change in its value would have to be very large to have a meaningful impact on λ. This could help *A. petiolata* buffer against biocontrol management of the transitions in E_4 such as rosette survival and fecundity if their effects on λ consistently diminish when populations are suppressed. If so, it would allow *A. petiolata* a means of recovering from biocontrol suppression after its growth rate is initially reduced. This would be exacerbated if the biocontrol agents' efficiency is negatively dependent on *A. petiolata* density, for example if they spend more time searching for the plant at low densities. This highlights one of the limitations of using static, linear population models, as there are no built-in feedbacks of density on the vital rates or the elasticity of λ to them. Future nonlinear models will need to explore this possible feedback mechanism in greater detail.

The two candidate biocontrol agent species, *C. scrobicollis* and *C. alliariae,* are well matched to *A. petiolata* in terms of affecting it at the appropriate life stages. Because it can affect both rosette survival and fecundity simultaneously, the root mining weevil *C. scrobicollis* has the potential to suppress *A. petiolata* in as many as five of the 12 study populations (41%) if it performs at its maximum efficacy. Using the same basic approach applied here, Davis et al. (2006) concluded that *C. scrobicollis* would control *A. petiolata* in 13 of the 18 (72%) scenarios they modeled. They also concluded that introducing the

weevil *C. alliariae* could broaden the impacts to include control *A. petiolata* in of 16 of the 18 (88%) simulations. It is important to note that these numbers represent the best possible outcome, given the models in Davis et al. and in this study, and assume that the biocontrol agents perform at the maximum efficacy recorded from among at least four different feeding and impacts trials. In reality, the impacts of either agent will likely be lower than the maximum and, thus, those numbers are likely overly optimistic. Given these caveats, both studies agree that *C. scrobicollis* is likely to suppress at least some *A. petiolata* populations if released on its own. Davis et al. found that introducing a second agent, *C. alliariae*, could extend control to an additional 16% of the scenarios modeled. My analysis predicts potentially higher returns from the release of a second agent by extending the possibility of control to an additional 33% of the study populations. Whether either species is capable of consistently delivering a strong enough impact on rosette mortality or reproduction to suppress populations remains unclear from currently available data. Without knowing what the mean or distribution are of the agents' impacts below the reported maxima, it is not possible to assess the degree of uncertainty in the probability of success. The estimated probabilities of successfully managing *A. petiolata* presented here should be interpreted with this understanding. Better data on the distributions of candidate agent impacts are needed from well replicated field trials and should be sought. These should include information on insect demographic rates and density dependence, as well as the insects' demographic responses to *A. petiolata* densities and environmental conditions. From these, a coupled plant-herbivore model could be developed which links the demography of *A. petiolata* with the biology of the biocontrol agents. This would allow generation of much more refined predictions.

The lack of ability or desire to make these kinds of evaluations prior to releasing biological control agents has been criticized as a major historical shortcoming of weed biocontrol programs. However, the last decade since McEvoy and Coombs (1999) has seen the beginnings of a cultural shift in the philosophy of weed biocontrol. The number of studies on agent impacts and target plant demography continues to grow, and biological control is transforming into a more rigorous practice guided by scientific principles and founded on ecological theory.

Tables: Chapter 4

Table 4.1. Mean observed *A. petiolata* demographic rates by site, calculated for each of three years.

Site	State	Year	g_1	g_2	s_r	s_{sum}	s_{win}	f	s_s
B	IL	2005	0.682	0.049	0.843	0.648	0.369	276.857	0.823
		2006	0.239	0.031	0.149	0.119	0.500	59.027	0.806
		2007	0.070	0.071	0.094	0.119	0.321	79.527	0.975
ELF	MI	2005	0.445	0.019	0.357	0.342	0.408	46.170	0.988
		2006	0.111	0.068	0.702	0.213	0.617	48.576	1.0
		2007	0.190	0.267	0.832	0.095	0.037	56.877	1.0
F	IL	2005	0.360	0.143	0.167	0.578	0.492	86.446	1.0
		2006	0.129	0.019	0.125	0.151	0.333	83.543	0.642
		2007	0.061	0.072	0.087	0.136	0.175	106.959	1.0
HL	IL	2005	0.403	0.021	0.198	0.872	0.823	445.234	1.0
		2006	0.245	0.269	0.273	0.138	0.286	572.051	0.880
		2007	0.389	0.032	0.429	0.054	0.250	126.388	0.969
HR	IL	2005	0.577	0.112	0.222	0.492	0.406	23.301	0.778
		2006	0.110	0.048	0.429	0.094	0.389	46.540	0.505
		2007	0.166	0.039	0.195	0.187	0.125	20.079	0.979
HSP	MI	2005	0.529	0.021	0.528	0.216	0.185	59.891	1.0
		2006	0.106	0.161	0.935	0.073	0.403	105.165	0.975
		2007	0.238	0.266	0.685	0.311	0.325	122.102	0.979
IP	IL	2005	0.546	0.075	0.128	0.447	0.308	108.276	0.839
		2006	0.395	0.367	0.205	0.267	0.146	71.698	0.894
		2007	0.480	0.288	0.403	0.085	0.806	124.029	1.0
IR	MI	2005	0.421	0.001	0.273	0.393	0.352	44.425	0.875
		2006	0.273	0.047	0.798	0.048	0.370	310.587	0.882
		2007	0.411	0.207	0.388	0.081	0.156	217.960	0.867
JP	MI	2005	0.661	0.062	0.606	0.309	0.350	68.905	1.0
		2006	0.177	0.144	0.617	0.357	0.552	63.669	0.983
		2007	0.660	0.440	0.523	0.111	0.092	56.877	1.0
RF	MI	2005	0.362	0.064	0.609	0.036	0.114	140.325	0.900
		2006	0.130	0.129	0.533	0.331	0.421	63.173	0.859
		2007	0.109	0.077	0.566	0.0	0.0	0.0	0.833
RL	MI	2005	0.163	0.031	0.835	0.067	0.161	15.849	1.0
		2006	0.151	0.157	0.810	0.013	0.333	152.414	0.969
		2007	0.340	0.189	0.899	0.113	0.152	20.079	1.0
SH	MI	2005	0.297	0.026	0.550	0.197	0.569	22.737	0.977
		2006	0.152	0.569	0.667	0.044	0.298	5.717	0.907
		2007	0.377	0.169	0.849	0.111	0.365	34.090	1.0

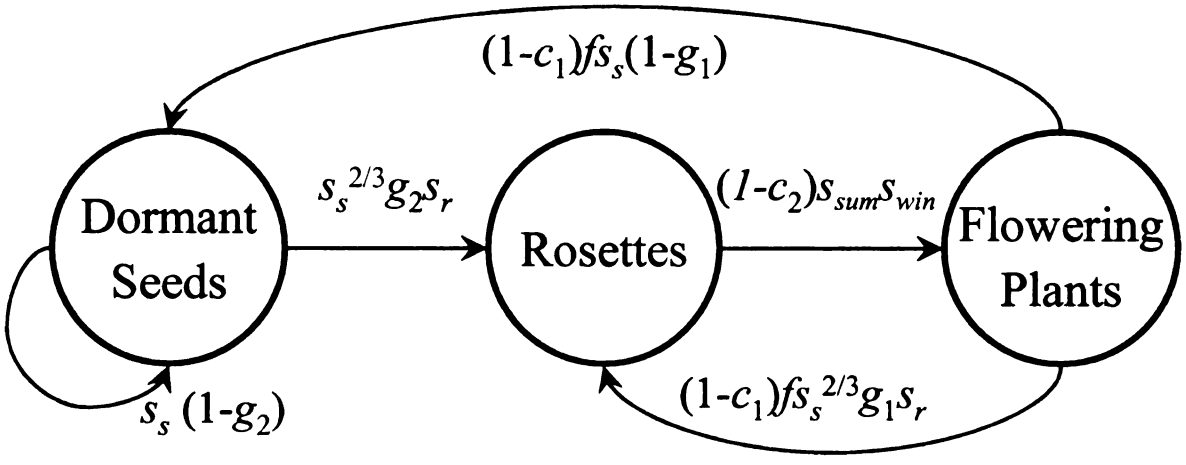
Table 4.2. Annual estimates of λ for each site (λ_{year}). These were summarized within each site as the megamatrix λ (λ_M), the population growth rate of the average population within each site calculated from the megamatrix **M** with uncorrelated environments. The mean growth rate across sites within each year and across all λ_M is given in the bottom row. The arithmetic mean λ_M is the average λ across the study system.

Site	State	λ_{2005}	λ_{2006}	λ_{2007}	λ_M
B	IL	5.883	0.798	0.927	3.543
ELF	MI	1.097	1.212	0.793	1.586
F	IL	1.453	0.642	0.945	1.120
HL	IL	5.061	1.586	0.959	3.405
HR	IL	0.855	0.530	0.945	0.832
HSP	MI	1.014	1.184	1.856	2.082
IP	IL	1.064	0.809	1.509	1.887
IR	MI	0.880	1.207	0.922	2.370
JP	MI	1.763	1.595	0.709	2.298
RF	MI	0.862	1.212	0.769	1.285
RL	MI	0.973	0.908	0.870	1.233
SH	MI	0.995	0.483	1.029	1.123
Column Means:		1.825	1.014	1.019	1.897

Table 4.3. The probability of successfully controlling *A. petiolata* at each site was graphically estimated as the proportion of the two shaded areas above the zero growth isoclines in Figure 4.7. The sizes of the shaded areas for *C. scrobicollis* alone (smaller rectangle in Figure 4.7) and with *C. alliariae* (larger rectangle) were 13899 and 26082 pixels, respectively. This analysis assumes a uniform probability distribution for agent impacts.

		Shaded Pixels Above Isocline		Probability of Control	
Site	State	*C. scrobicollis*	*C. alliariae* + *C. scrobicollis*	*C. scrobicollis*	*C. alliariae* + *C. scrobicollis*
B	IL	0	0	0	0
ELF	MI	0	408	0	0.016
F	IL	8074	20173	0.581	0.773
HL	IL	0	0	0	0
HR	IL	13899	26082	1.000	1.000
HSP	MI	0	0	0	0
IP	IL	0	2327	0	0.089
IR	MI	0	208	0	0.008
JP	MI	0	103	0	0.004
RF	MI	2469	13827	0.178	0.530
RL	MI	98	8056	0.007	0.309
SH	MI	8951	21050	0.644	0.807
			Mean:	0.201	0.295

Figures: Chapter 4



$$\mathbf{A} = \begin{bmatrix} a_{11} = s_s(1-g_2) & a_{12} = 0 & a_{13} = (1-c_1)fs_s(1-g_1) \\ a_{21} = s_s^{2/3}g_2s_r & a_{22} = 0 & a_{32} = (1-c_1)fs_s^{2/3}g_1s_r \\ a_{31} = 0 & a_{32} = (1\text{-}c_2)s_{sum}s_{win} & a_{33} = 0 \end{bmatrix}$$

Figure 4.1. Alliaria *petiolata* life cycle diagram and corresponding projection matrix **A**. Arrows represent one-year transitions from June of year t to June of year t+1 and are comprised of sub-annual, lower level demographic transitions: per capita fecundity (f), germination probabilities of new seeds within one year of seed set (g_1) and of dormant seeds from the soil seed bank (g_2), and seedling (s_r), summer rosette (s_{sum}), winter rosette (s_{win}), and dormant seed (s_s) survival probabilities. Matrix rows correspond with the life history stage individuals are transitioning from time t, where rows 1, 2, and 3 (from left) are seeds, rosettes, and flowering plants. Columns indicate life history stages individuals are transitioning to in time t+1. For example matrix element a_{32} represents the transition from rosettes (column 2) to flowering plants (row 3) in matrix A. Following Davis et al. (2006), the variables c_1 and c_2 simulate rosette mortality and fecundity reduction due to biocontrol, and are shown as (1-c_n) to express them in terms of survival.

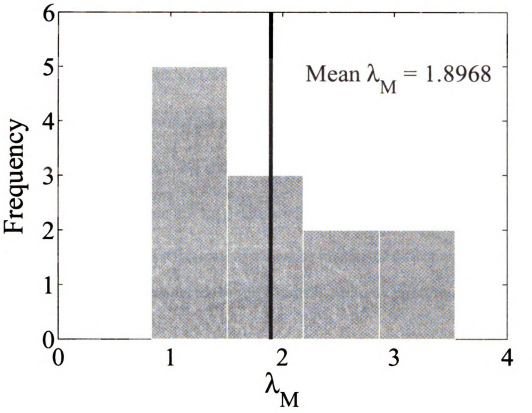
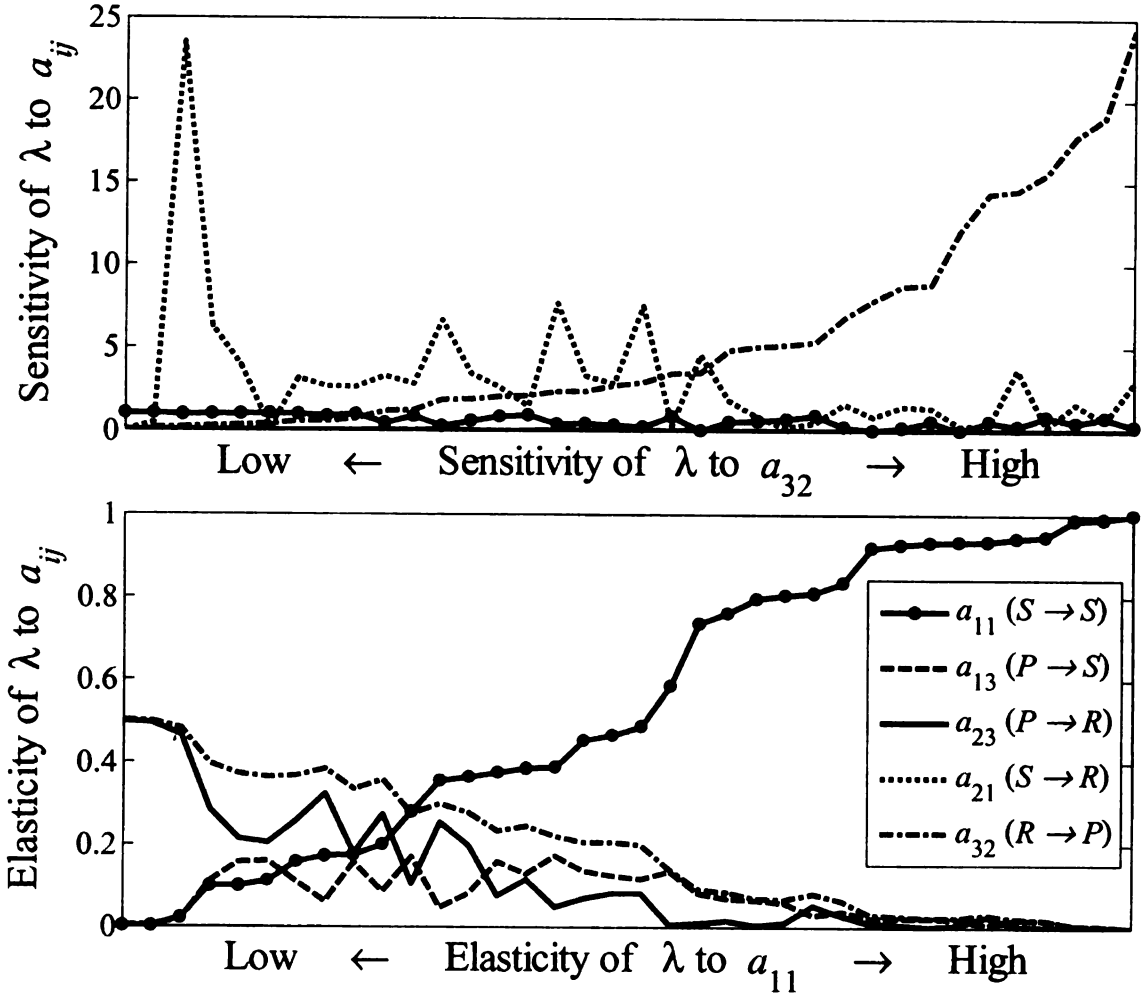


Figure 4.2. Frequency distribution of *A. petiolata* megamatrix population growth rate across all sites (n = 12). The black line indicates the mean value of λ_M across sites.

Figure 4.3. Sensitivities (top) and elasticities (bottom) of λ to perturbation of A matrix elements a_{ij}. Sensitivities from the combined 36 site by year matrices are sorted from left to right by sensitivity matrix element s_{32} ($R \rightarrow P$), and elasticities are sorted by elasticity matrix element e_{11} ($S \rightarrow S$). The sensitivities of λ to a_{23} and a_{13} are small and lie along the horzontal axis. The elasticities of λ to a_{13} and a_{21} are equal; the former obscures the latter in the lower plot. NOTE: The dotted line (a_{11}) does *not* indicate a different density of data points than the other four line styles and should not be interpreted differently. It is used because the software which produced the graph (MATLAB) only prints four basic line styles.


Sensitivity of λ to a_{ij}
25
20
15
10
5
0
Low ← Sensitivity of λ to a_{32} → High
Elasticity of λ to a_{ij}
1
0.8
0.6
0.4
0.2
0
a_{11} ($S \rightarrow S$)
a_{13} ($P \rightarrow S$)
a_{23} ($P \rightarrow R$)
a_{21} ($S \rightarrow R$)
a_{32} ($R \rightarrow P$)
Low ← Elasticity of λ to a_{11} → High

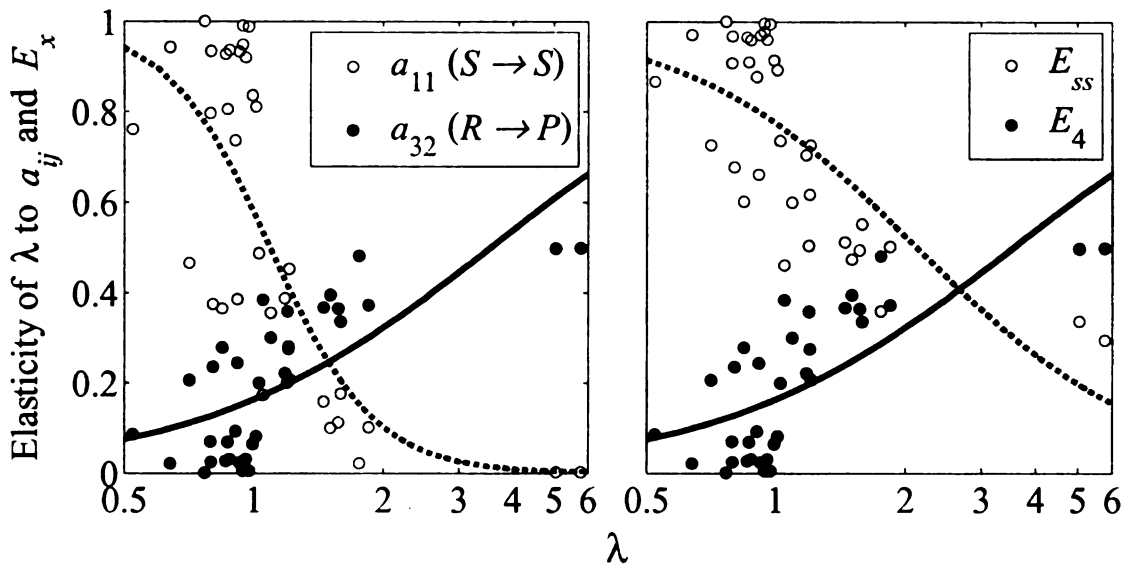


Figure 4.4. Elasticities (e_{ij}) of λ to A matrix elements a_{11} ($S \rightarrow S$) and a_{32} ($R \rightarrow P$) (left) and to lower level transitions seed (s_s) survival and E_4 (right) versus the population growth rate λ. The elasticities of λ to f, s_r, s_{sum}, and s_{win} were identical and are represented together with one line, labeled E_4. E_4 and e_{32} are equal as well. Among populations with high growth rates, λ has the greatest elasticity to a_{32} and E_4. Logistic regression lines are overlaid to show trends.

Figure 4.5. Empirical second derivatives of λ. Observed variation in elasticities of *Alliaria petiolata* population growth versus observed variation in the lower demographic transitions. In each panel the elasticities of λ to each lower-level demographic rate from the 36 site by year matrices is plotted against the observed value of the transition indicated on the horizontal axis. The eight panels show the same empirical data sorted differently. The data are sorted from left to right by the transition labeled on the horizontal axis. In the lower right panel the x axis is shown on a log scale, although the x axis tick marks are back-transformed to the original scale. Contours were smoothed using a LOESS smoother (Burkey 2009) to clarify the relationships among variables. This is why the values in each panel appear different even though they show the same data. The raw, unsmoothed plots are presented in Appendix A.

Elasticity of λ to variation in x_{ij}
1.0
0.8
0.6
0.4
0.2
0
Germination 1
Germination 2
Seedling Survival
Summer Rosette Survival
Winter Survival
Per Capita Fecundity
Seedbank Survival
ss
sr, sum,
win, f
gl
g2
λ
0.5
1
2
3
4
5
6

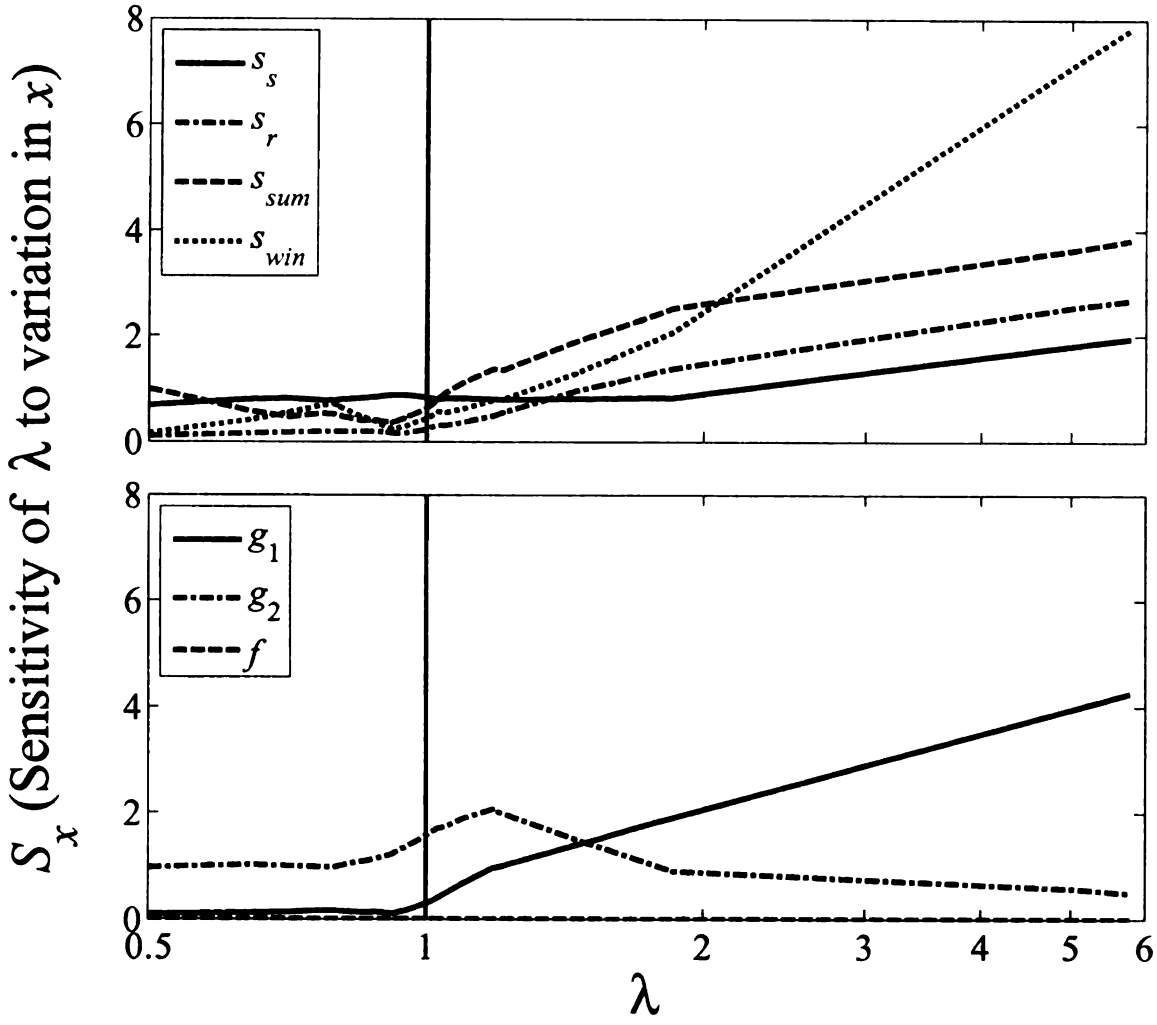
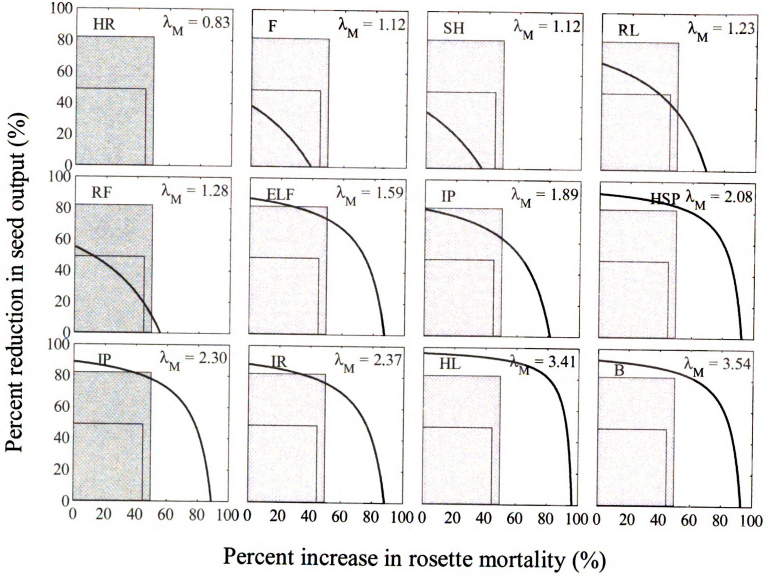


Figure 4.6. Sensitivities of λ to lower-level *A. petiolata* transitions x versus λ for 36 individual site-years. As λ increases, the rankings and magnitudes of the S_x change to favor management of rosettes in winter or summer and germination of new seeds. The vertical lines differentiate between expanding populations (right) and declining populations (left). The seven sensitivity contours are split into two panels to improve readability. LOWESS smoothing was applied to each line to for the same purpose.

Figure 4.7. Zero growth isoclines (contours of $\lambda = 1$, indicating stable *A. petiolata* population size) for each study population assuming a random sequence of the three environments (years) using a megamatrix model. The x and y axes in each plot represent the efficacy of rosette or fecundity management simulated by increasing the variables c_2 or c_1, respectively, from 0 to 1. The larger shaded area shows the observed range of seed reduction and rosette mortality caused by the combined actions of the root and stem mining weevils *Ceutorhynchus scrobicollis* and *C. alliariae* (see Table 2 in Davis et al., 2006). The smaller rectangle within the shaded area indicates the range of observed impacts of *C. scrobicollis* alone, where the upper right corner is the maximum observed impact of *C. scrobicollis*. Contours which intersect the larger shaded area represent populations which could theoretically be controlled by *C. scrobicollis* and *C. alliariae* together at a given level of agent efficacy (e.g. IP would only be controlled if actual agent efficacy is at the high end of the observed range). Contours which intersect the smaller shaded area could theoretically be controlled by *C. scrobicollis* alone. The line from HR does not appear in the plot because its $\lambda_M < 1$.Impacts of *C. alliariae* alone are not shown because it has not been shown to reduce fecundity under realistic field conditions.


HR λ_M = 0.83
F λ_M = 1.12
SH λ_M = 1.12
RL λ_M = 1.23
RF λ_M = 1.28
ELF λ_M = 1.59
IP λ_M = 1.89
HSP λ_M = 2.08
IP λ_M = 2.30
IR λ_M = 2.37
HL λ_M = 3.41
B λ_M = 3.54
Percent reduction in seed output (%)
Percent increase in rosette mortality (%)
0 20 40 60 80 100

APPENDIX 4.A: RAW ZERO GROWTH CONTOURS

Figure 4.A.1 Empirical elasticity contours without smoothing function. Observed variation in elasticities of *Alliaria petiolata* population growth versus observed variation in the lower demographic transitions. In each panel the elasticities of λ to each lower-level demographic rate from the 36 site by year matrices is plotted against the observed value of the transition indicated on the horizontal axis. The eight panels show the same empirical data sorted differently. The data are sorted from left to right by the transition labeled on the horizontal axis. In the lower right panel the x axis is shown on a log scale, although the x axis tick marks are back-transformed to the original scale.

Elasticity of λ to variation in x_{ij}
Germination 1
Winter Survival
Germination 2
Per Capita Fecundity
Seedling Survival
Seedbank Survival
Summer Rosette Survival
λ
ss
sr, sum, win, f
g1
g2

APPENDIX 4.B: ANNUAL ZERO GROWTH CONTOURS

Figure 4.B.1. Annual Zero growth isoclines (contours of $\lambda = 1$) for each site. The megamatrix result used in the primary analysis is also shown. Within each site, at least one of the three annual zero growth isoclines crossed the shaded agent-impacts area. In all sites except HSP, which had positive growth each year, at least one isocline falls below the origin of the figure.

B
ELF
F
HL
HR
HSF
IP
IR
JP
RF
RL
SH
Megamatrix
2005
2006
2007
0
20
40
60
80
100
1000
Percent reduction in seed output (%)
Percent increase in rosette mortality (%)

CHAPTER 5: DEMOGRAPHIC VARIANCE ACROSS GRADIENTS OF POPULATION DENSITY IN *ALLIARIA PETIOLATA*

ABSTRACT

Demographic models of the invasive plant *Alliaria petiolata* (garlic mustard, Brassicaceae) have been developed previously to inform the selection of effective biological control agents. Understanding how density dependent feedbacks affect vital rates and vary spatiotemporally among populations may further the refinement of these model predictions and lend an important regional perspective to future management models. From 2004 to 2007, I studied the survival and reproduction of *A. petiolata* in 13 natural populations in Illinois and Michigan, USA, across a natural density gradient spanning over three orders of magnitude. Using mixed effects models I quantified the contributions of population density, region, cohort, and site effects to variability in demographic rates. Seedling and summer survival and fecundity parameters varied significantly between Michigan and Illinois and between cohorts, but winter survival did not. In contrast to other studies of *A. petiolata*, only mean seedling survival and fecundity were negatively correlated with initial density, while summer and winter rosette mean survival were not. Instead, I found that the residual variance in each analysis was negatively dependent on population density, even when the mean response to density was not significant. By explicitly modeling the residual variance in each vital rate as a function of plant density, I show that the demographic variance and stochasticity are themselves density dependent functions; the maximum possible value of each demographic rate decreases as density increases, but vital rates vary below these density dependent upper bounds. Thus, survival and fecundity can be broadly distributed at low densities but funnel into narrow distributions with lower maxima at high densities. The scaling of demographic stochasticity with population density has potentially important

consequences for populations of management concern, as small populations may become more susceptible to local extinctions. The robustness of predictive weed management models may therefore be improved by incorporating elements of environmental and density dependent demographic stochasticity into their designs.

INTRODUCTION

Density dependent processes and demographic variability are important in natural populations (Horvitz and Schemske 1995, Turchin 1999). Failure to incorporate them properly into predictive management models could lead to inaccurate characterization of the demography of the study species. Density feedbacks can vary in importance (Elliott and Hurley 1998) and even direction (Goldberg et al. 2001) among life history stages, which suggests that linear population models may oversimplify complex biological processes. Similarly, a single species can exhibit life history differences across its range, in different habitats, or over time that could affect how it responds to management (Parker 2000, Shea et al. 2005). For species with large geographic distributions, this means either that different management approaches need to be implemented in different areas, or that a single strategy must be applied which maximizes the outcome at the greatest number of locations.

When biological control is considered for management of invasive weeds, this latter, probabilistic approach may be the most appropriate option. Because herbivorous insect agents with strong dispersal capabilities are not likely to stay where they are locally released (Simberloff and Stiling 1996, Stiling et al. 2004, Louda et al. 2005b), weed biocontrol management decisions should ideally be made on a scale equivalent to the potential distribution of the proposed biocontrol agent(s) (Louda et al. 2003a). Modeling a target species' biology should ideally account for demographic variability across the species' range that could affect the outcome of management, including possible variability in density dependence. The critical question then is how to best

account for this variability and create a model of invasive species control that is robust in its predictions across a wide range of environments.

I considered this problem in an ongoing study (Davis et al. 2006) of the invasive weed *Alliaria petiolata* (garlic mustard, Brassicaceae [M. Bieb] Cavara and Grande) for which biological control has not yet been implemented but is likely forthcoming (Gerber et al. 2008b). This species has been the subject of several previous studies that quantified various aspects of its demography (Cavers et al. 1979, Baskin and Baskin 1992, Anderson et al. 1996, Drayton and Primack 1999) and indicated negative density dependent feedbacks in survival and reproduction (Meekins and McCarthy 2000, Pardini et al. 2009). Data from Meekins and McCarthy (2002) also shows density dependence of *A. petiolata* population growth rates (λ), although the estimates of λ in the original paper were calculated incorrectly (see Appendix 2.D). These studies provide valuable estimates of *A. petiolata* transition probabilities across a large geographical area in eastern North America (Kentucky to Ontario), but none accounts for the variability in demographic rates and population density feedbacks on vital rates at large spatial or temporal scales.

Here, I present a re-examination of density dependence in the *A. petiolata* demographic data described in previous chapters. Using these data, I asked (1) if there were differences in demographic rates between Michigan and Illinois populations, (2) whether populations were variable from year to year, and (3) whether there were density dependent feedbacks that contribute to this variability. I used mixed effects models to explore the spatial and temporal variability in survival and reproduction and evaluate the importance of density dependent feedbacks across the observed gradient of population density measured in my study system. This study differs from the analysis in Chapter 2 in

the way that residual variance was modeled and interpreted. While the generalized linear models and generalized linear mixed models parameterized in Chapter 2 can be used to predict survival rates, they have limited flexibility in the way they handle residual variance and overdispersed data. The linear mixed models explored here offer a different view of the data structure and highlight some of the most interesting aspects of *A. petiolata*'s demography.

METHODS

Sampling

Study site descriptions and sampling methods are presented in Chapters 2 and 4. Note that as in previous chapters, demographic rates were grouped into "cohorts" for analysis based on the June to June annual projection interval used in previous *A. petiolata* studies (Davis et al. 2006). This required grouping rosettes from one year in the same cohort as seedlings that germinated the following year. For example rosette survival in the 2005-2006 cohort was measured from June 2005 until June 2006 when those plants matured and produced seeds. Seedling survival was measured during the spring of 2006.

Statistical Analyses

The main objective of this study was to evaluate the contributions of location, time, and density to differences in demographic rates between *A. petiolata* populations. I constructed linear mixed effects models with a 2 by 3 factorial fixed effects structure to analyze the variability in seedling, summer, and winter survival and in per capita fecundity. Fixed factors in the analyses included region (Illinois and Michigan), cohort

(2005-2006, 2006-2007, 2007-2008), and initial *A. petiolata* density to test for density dependence. Site nested within year was included as a random factor assuming that each site is a random sample from the population of all possible study sites and therefore contributes to the observed variation in the data. The nesting structure accounts for the covariance of observations across sites within each year. To avoid introducing multiplicative errors into the fecundity analysis, I modeled the number of siliques per plant as the response rather than the estimated number of seeds per plant. However, seeds per plant and siliques per plant are strongly correlated (Appendix 2.B.iv), so model predictions will be applicable to seed production. In the three analyses of survival, each observation was weighted by the number of plants used in its calculation. For example, a survival rate estimated from a sample of 100 plants would be more heavily weighted than one estimated from 50 plants. Data used in the analysis of fecundity were not weighted because these were exhaustive counts of the true number of siliques on each plant in each sampling quadrat.

Maximum seedling density and June and fall rosette densities (plants m^{-2}) were used as density covariates in analyses of seedling, summer, and winter survival, respectively. Fall rosette density was used as the covariate in the analysis of fecundity. All survival probability response variables were arcsine square root transformed, and silique counts used in the fecundity analysis were square root transformed to meet the assumptions of the analyses.

Initial statistical analyses indicated that there was overdispersion in the survival data (i.e. the residual variance was greater than expected under a binomial distribution, where the variance is a function of the mean (Littell et al. 2006)) even after applying

variance stabilizing transformations. The fecundity data were also overdispersed. Two alternative structures for the residual variance were fitted in each analysis to correct for this: the power of mean (POM) function and an exponential function in which the residual variance is a function of the density covariate. Denominator degrees of freedom were approximated using the Satterthwaite method (Littell et al. 2006). Each analysis was run with fixed effects only, fixed effects plus random site within year effects, and finally with all fixed and random effects including the alternative residual variance structures. The best-fit model was selected using the information-theoretic methods (Littell et al. 2006).

RESULTS

Alliaria petiolata initial population densities varied approximately 100-fold for each life stage. Maximum seedling densities ranged from 75 to 6025 plants m^{-2} (Figure 5.1). Summer and fall rosette densities used as covariates in the summer and winter survival analyses ranged from 31 to 3344 plants m^{-2} and from 4 to 540 plants m^{-2}, respectively (Figure 5.1). Bayesian Information Criteria (BIC) (Littell et al. 2006) indicated that, for each of the three survival transitions analyzed, the data best supported the full model which included random effects terms and the variance modeled as a function of density (Table 5.1). There was equivalent support for the POM and density dependent variance functions in the fecundity analysis ($\Delta BIC = 0.4$). For ease of biological interpretation, I present the results of the density dependent variance model of fecundity. Model fit statistics are presented in Table 5.2 including slope and variance parameter estimates for covariates. Many of the highest starting densities in the summer

survival analysis occurred in the smaller quadrats that were monitored through the summer in Michigan during 2005. Re-running the model without these data points did not change sign or significance of the density covariate ($F_{1, 32} = 4.13$, $P = 0.0504$) nor the general conclusions about the exponential variance term; the results presented include these data.

Fixed Effects

Mean seedling survival ($F_{1, 27.3} = 43.05$, $P < 0.0001$) varied between Michigan and Illinois (Table 5.2), but not between cohorts. Summer survival varied both between regions ($F_{1, 34.9} = 5.32$, $P = 0.0272$) and cohorts ($F_{2, 32.7} = 18.0$, $P < 0.0001$). The region*cohort interaction in the summer survival analysis was significant as well ($F_{2, 32.2} = 5.59$, $P = 0.0082$), indicating that year to year differences in mean summer survival were not consistent between Michigan and Illinois. Mean winter survival rates were not significantly different between regions or cohort groups (Table 5.2), though mean fecundity varied between regions ($F_{1, 26.4} = 5.76$, $P = 0.0237$).

Density dependent effects on survival were strongest early in the life cycle and for seed production. Maximum seedling density negatively affected seedling survival (slope = -0.00004, $F_{1, 45.2} = 8.88$, $P = 0.0046$). The effect of rosette density on summer survival was also negative (slope = -0.00002) but not significant ($F_{1, 18.8} = 0.61$, $P = 0.4456$), and reversed to significantly positive (slope = 0.0006, $F_{1, 42.4} = 6.61$, $P = 0.0137$) in winter. This reversal was driven by two influential experimental units (quadrats) with high

leverage during the winter that had both high starting densities and high survival rates. Dropping these data points from the analysis changed the sign of the slope to negative (slope = -0.00018) but non significant ($F_{1, 57.2} = 0.24$, $P = 0.6241$). Because the interpretation of the slope of the density covariate for winter rosette survival was entirely dependent on these two observations, I chose to drop them from further analysis. Fecundity responded negatively to initial *A. petiolata* density (slope = -0.0055, $F_{1, 595}$, $P < 0.0001$) and varied between regions as well ($F_{1, 26.4} = 5.76$, $P = 0.0237$), but not between cohorts ($F_{2, 25.9} = 0.85$, $P = 0.4378$). Although the absolute values of the slope estimates are small, the ranges of the dependent variables were also small relative to the independent variables. Thus, shallow slopes such as these can be biologically meaningful. All three-way interactions (region*cohort*density) were non-significant and were dropped from the models.

Random Effects

The random effects of sites nested within cohorts were significant for all response variables (Table 5.2), indicating that plants growing together at the same site likely experienced more similar biotic and abiotic conditions than those growing at different sites and in different years. This effect was least significant during the winter, suggesting that winter survival is less dependent on local site conditions.

Perhaps more significantly, the residual variance in each vital rate decreased as a function of initial density. The slope estimates for the variance terms were negative and significant in all analyses: seedling survival (slope = -0.0007, $Z = -4.81$, $P < 0.0001$),

survival in summer (slope = -0.001, Z = -2.84, P = 0.0045) and winter (slope = -0.0096, Z = -6.11, P = < 0.0001), and fecundity (slope = -0.0049, Z = -6.81, P < 0.0001). The consequence of this is reflected in the decreasing ranges of the response variables with increasing initial density. At low population densities, survival probability estimates in each model ranged from 0 to 1, and fecundity ranged from 0 to over 120 siliques plant^{-1} (Figure 5.1). As the population density increased, the variation around the mean survival probabilities and fecundity declined exponentially (Figure 5.2), such that at high plant densities survival probabilities were rarely greater than 0.5, and fecundity rarely exceeded 20 siliques plant^{-1}. This pattern is evident in the data as the decreasing upper boundaries and the more subtly increasing lower boundaries of the parameter space across the gradient of increasing density in the raw data (Figure 5.1) and in the prediction intervals generated from the parameterized models (Figure 5.2-5.5). Thus, although the density covariate slopes of the summer and winter survival analyses were not different from zero, survival rates did respond to initial plant density.

DISCUSSION

Population density is one of several important factors influencing the survival and reproduction of *A. petiolata* across the 12 Illinois and Michigan study sites. These effects appear to be at least partially seasonal. Negative density dependent feedbacks were demonstrated in all transitions except winter using generalized linear mixed models with additional covariates in Chapter 2. Using simpler linear mixed models of transformed data here, the main effects of population density appear more variable. Negative density dependent feedbacks were strongest during the spring recruitment phase (seedling to

rosette). During this period, seedling populations reached high initial densities in some locations and then thinned over the remainder of the plant's life cycle until senescence in the summer of the second year. It was only during this recruitment stage, when population densities reached almost 7000 m^{-2}, that there was a significant negative effect of plant density on mean survival (Figure 5.2, Table 5.2). Though still negative, this relationship weakened through the summer (Figure 5.3) and winter (Figure 5.4), when the highest starting density was only 360 plants m^{-2} (Table 5.2). Winter rosette density had a negative effect on seed production during the spring of the second year in *A. petiolata*'s life cycle. This switching of the covariate slope from negative to zero may reflect the importance of intraspecific competitive or interference interactions early in the life cycle later becoming subordinate to abiotic or stochastic factors once self thinning of seedlings has increased the spacing between individuals and reduced the likelihood of further competitive interactions (Goldberg et al. 2001). Alternately, because fecundity was density dependent as well, it could indicate that spring is a critical period in the life cycle during which differences in resource availability or growing conditions have the potential to interact with population density.

The failure to detect density dependence in summer and winter rosette survival rates may be either exclusion of other important covariates or that natural population densities were too low after thinning of seedlings in spring to have further effects (sensu Smith et al. 2003). Biased selection of low density rosette populations could also obscure density dependence, especially if its effects are subtle or absent at these low densities. Quadrat placement was not biased against high densities. If it had been biased, I would not have captured the exponentially declining variance of survival and fecundity as

density increased. Thus, the data show that mean demographic rates are relatively constant across the naturally occurring density gradient after recruitment with only the variance exhibiting a strong negative density response.

The Importance of Variance

The slope of the density covariate in the fixed effects is an estimate of the relationship between initial density and mean survival or reproductive rates. When this slope is not significantly different from zero, the conventional interpretation is that the response variable and covariate are unrelated. However, I find that population density is strongly connected with the variance of demographic rates in *A. petiolata*, even when the main effect of density is not significant. This was true of the density dependent models presented in Chapter 2 as well, although it was not possible to explicitly account for the density-variance relationship within the generalized linear modeling framework used there. Modeling the variance as a random exponential function of the density covariate reveals a decreasing upper bound to the distribution of lower level demographic rates as density increases even in the absence of an overall mean trend. For example, the slope of the summer survival density covariate was not significantly different from zero. Mean summer survival was nearly the same in both the high and low density plots: the mean summer survival rates of the lower and upper 50^{th} percentiles of quadrats were 0.207 and 0.224, respectively. But summer survival in low density quadrats ranged from 0 to 1, while there were no measurements of summer survival greater than 0.5 at densities above ~600 plants m^{-2} (Figure 5.1). Maximum survival probability is limited or capped at any given density but can vary below that maximum. This limit decreases as density

increases. Vital rates show a broad distribution in the range of values they can attain at low densities where abiotic and stochastic factors are likely most limiting and are funneled into a more restricted range of values as density increases and crowding and competition dominate. Predicting population behaviors will therefore be less precise at low densities.

Ecological data from natural populations are characteristically noisy. Unmeasured factors such as habitat quality, microsite conditions, and environmental and demographic stochasticity undoubtedly contribute to the observed variance in demographic rates. Decoupling these from the effects of density requires controlled experiments. What observational studies such as this one can provide are important estimates of the actual distribution of demographic parameters among natural populations and over time. For application to a regional management strategy like biological control, understanding this variability in vital rates will be essential to determining what proportion of populations across the landscape will respond to a given control strategy. Linking results from observational studies of natural populations with experiments conducted under known gradients may be useful for further partitioning the observed variance into density dependent, environmental, and stochastic components.

My mixed model approach to estimating demographic rates in *A. petiolata* suggests a significant negative feedback of population density on survival probability and fecundity across multiple populations and corroborates findings of other studies of this species (Meekins and McCarthy 2000, Smith et al. 2003, Winterer et al. 2005, Rebek and O'Neil 2006). More interestingly, though, I did not see strong effects of plant density on mean demographic rates after the recruitment phase. Rather, density dependence in this

system was best characterized by the negative relationship between demographic variance and population density for all demographic parameters. Density dependence is an important factor structuring *A. petiolata* populations, but is only important relative to other factors at high densities. The estimated ranges of demographic parameters changed from spanning the entire range of possible values at low densities to having increasingly restricted distributions at higher densities. Thus, the demographic rates in these models effectively change from behaving like stochastic variables to behaving in a more deterministic fashion with increasing density, and density dependence becomes the most important factor limiting the response variable.

Demographic variability is predicted to be dependent on population size when there is density dependence in survival and reproductive rates (Engen et al. 1998). This has been observed empirically in a songbird population (Saether et al. 1998) as well as in experimental populations of *Daphnia magna* in which demographic stochasticity appeared to have density dependent components (Drake 2005). Incorporating density dependent demographic stochasticity into a stochastic population model greatly improved Drake's ability to accurately predict extinction rates of the experimental populations. This suggests future models of *A. petiolata* biocontrol that incorporate density dependent stochasticity as well as site and regional demographic differences will provide more accurate predictions of population responses to biocontrol. Graphical analysis of the data shows strong density dependent mortality in some sites during certain years, but not others. Density feedbacks are therefore important in this system, but might better be thought of as effects that rise and fall in importance relative to other mortality factors across space and time.

It is uncertain exactly how these elements will interact in future *A. petiolata* biocontrol models. There is a tension between the generally stabilizing effects of density dependent survival and reproduction (e.g. Saether et al. 1998, Buckley et al. 2001) and the destabilizing effects and increased extinction probabilities resulting from density dependent stochasticity (Drake 2005). Conceivably, these processes could interact to enhance biocontrol effectiveness. An effective herbivore that reduces a vital rate such as survival could increase the relative importance of stochasticity. Mortality induced in early life history stages will truncate the distribution of population densities in later, non-density dependent stages. By driving populations towards densities that are more susceptible to demographic stochasticity, a biocontrol agent could thus increase a population's probability of stochastic extinction. Complex dynamics between herbivore and plant densities could complicate this relationship.

Populations of *A. petiolata* exhibit spatiotemporal variability in demographic rates. Dynamic stochastic feedbacks on demographic rates clearly play an important role in this variability and should not be overlooked. Because populations whose demographics are dissimilar may not succumb a common control tactic (Parker 2000, Shea et al. 2005), incorporating these estimates of the spatiotemporal variance and density dependence of demographic rates and stochasticity in the *A. petiolata* system into predictive management models could potentially indicate whether different management strategies will be required in different areas. The role of demographic variance scaling across population density gradients should be explored further in future models of *A. petiolata* population dynamics and biological control. The models presented here are useful for interpreting patterns in these data. However, because the data were transformed

before analysis, the models cannot be used in a predictive context. Doing so will require development of new generalized linear mixed models that can accommodate modeling the residual variance as a function of a covariate. .

Tables: Chapter 5

Table 5.1 Comparisons of fixed- and mixed effects models of density dependent survival and fecundity in *A. petiolata.* Seedling survival was calculated as the number of rosettes in June divided by the maximum number of seedlings observed during the spring germination period. Summer and winter survival were calculated as the proportion of rosettes surviving from June to October and from October to the following June, respectively. All models included fixed effects terms for initial *A. petiolata* density (plants m^{-2}), state, year, and state by year interactions. Mixed effects models also included random effects terms for site within year. In the mixed models, the variance structure was modeled normally, as an exponential function of the density covariate, and using the power of the mean (POM) structure. AIC and BIC are the Akaike and Bayesian Information Criterion, respectively. Δ AIC and Δ BIC indicate the difference in fit between each model and the fit of the best fit model within each set of analyses. Table 5.1.

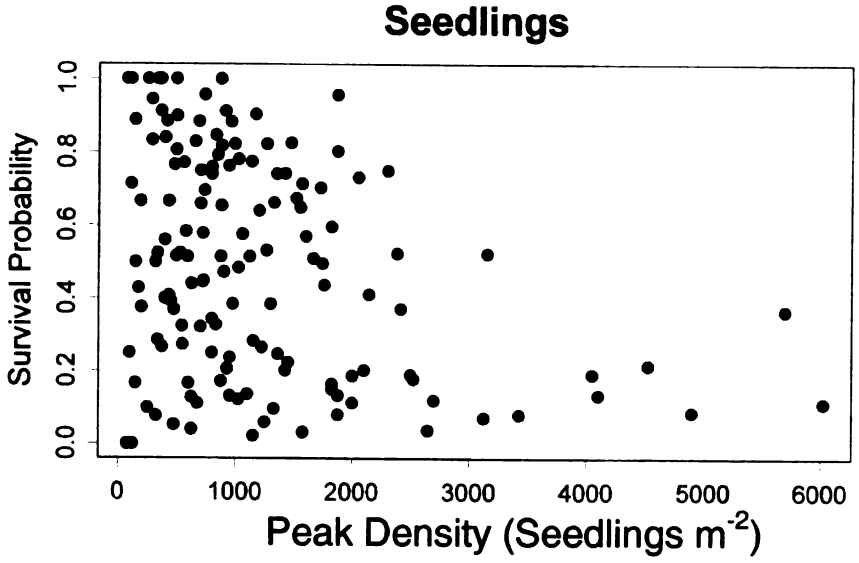
Transition	Model, Covariance Structure	-2 Res. Log likelihood	AIC	Δ AIC	BIC	Δ BIC
Seedling Survival	Mixed, *exp*(seedlings)	35.4	41.4	0.0	46.2	0.0
	Mixed, POM	50.1	56.1	14.7	60.8	14.6
	Mixed, normal	54.0	58.0	16.6	61.1	14.9
	Fixed, normal	80.1	82.2	40.9	85.1	38.9
Summer Survival	Mixed, *exp*(June rosettes)	24.6	30.6	0.0	35.3	0.0
	Mixed, POM	31.3	37.3	6.7	42.0	6.7
	Mixed, normal	44.9	48.9	18.3	49.8	14.5
	Fixed, normal	46.0	48.0	17.4	50.9	15.6
Winter Survival	Mixed, *exp*(Oct. rosettes)	130.5	136.5	0.0	141.1	0.0
	Mixed, normal	159.2	161.2	24.7	162.8	21.7
	Fixed, normal	159.2	161.2	24.7	163.9	22.8
Fecundity	Mixed, POM	2687.2	2693.2	0.0	2697.7	0.0
	Mixed, *exp*(Oct. rosettes)	2687.5	2693.5	0.3	2698.1	0.4
	Mixed, normal	2733.7	2737.7	44.5	2740.8	43.1
	Fixed, normal	2815.7	2817.7	124.5	2822.4	124.7

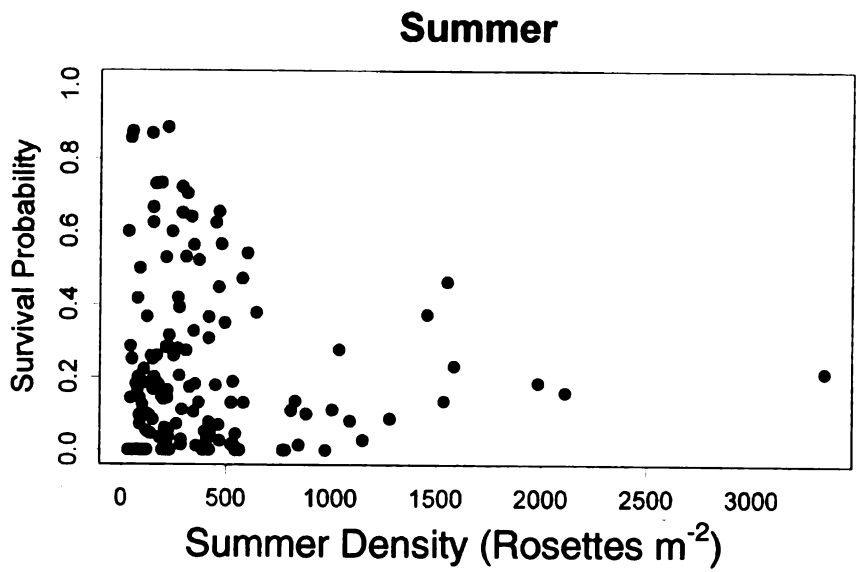
Table 5.2 Type III *F*-tests of fixed effects and estimates of covariance parameters from analyses of arcsine square root transformed seedling, summer, and winter survival probabilities in *A. petiolata*. In each model, plant density (# m^{-2}) is the number of *A. petiolata* plants at the start of the interval over which survival was measured, except for seedling survival, where seedling density is the maximum density of seedlings observed during the spring. The estimated density effects are slopes associated with each density covariate. Random effects are indicated in italics and were evaluated with *Z* tests. Congratulations! You're almost done reading this thing. I will buy a beer for any committee member who finds this text and mentions it at my defense.

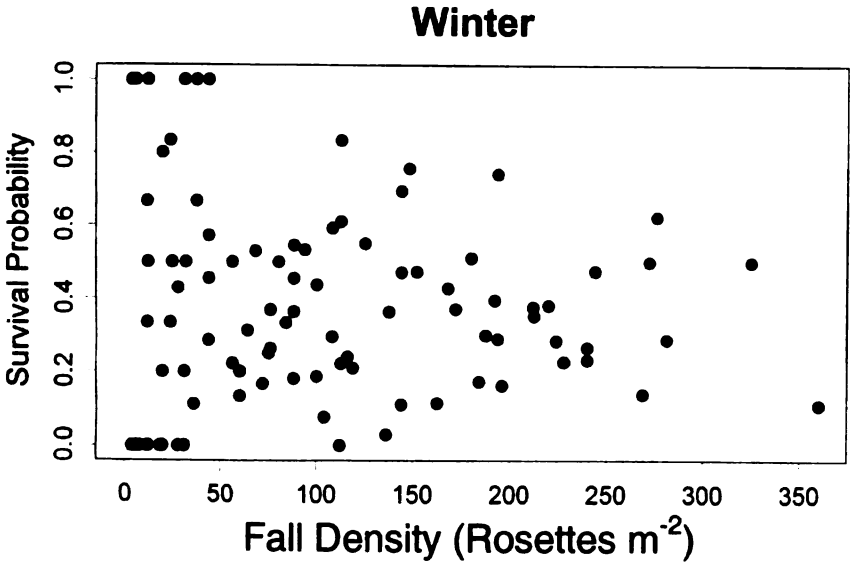
Transition	Effect	Estimate	Num DF	Den DF	Test Statistic	*P*
Seedling Survival	Region		1	27.3	43.05	<.0001
	Cohort		2	26.7	0.58	0.5664
	Region*Cohort		2	27	1.76	0.1910
	Maximum Seedling Density	-0.00004	1	45.2	8.88	0.0046
	Site(Cohort)	0.0287			2.67	0.0038
	exp(Seedling Density)	-0.0007			-4.81	<.0001
	Residual	0.0902			4.88	<.0001
Summer Survival	Region		1	34.9	5.32	0.0272
	Cohort		2	32.7	18.03	<.0001
	Region*Cohort		2	32.2	5.59	0.0082
	Summer Density	-0.00002	1	18.8	0.61	0.4456
	Site(Cohort)	0.0198			2.46	0.0069
	exp(Summer Density)	-0.0011			-2.84	0.0045
	Residual	0.0634			5.24	<.0001
Winter Survival	Region		1	46.9	0.03	0.8588
	Cohort		2	49.7	1.25	0.2946
	Region*Cohort		2	38	1.04	0.3628
	Winter Density	-0.00018	1	57.2	0.24	0.6241
	Site(Cohort)	0.0175			1.7100	0.0435
	exp(Winter Density)	-0.0096			-6.1100	<.0001
	Residual	0.2904			5.4700	<.0001
Fecundity	Region		1	26.4	5.76	0.0237
	Cohort		2	25.9	0.85	0.4378
	Region*Cohort		2	23.9	0.18	0.8392
	Winter Density	-0.00551	1	595	44.89	<.0001
	Site(Cohort)	0.4183			2.12	0.0172
	exp(Winter Density)	-0.0049			-6.81	<.0001
	Residual	3.71			7.53	<.0001

Figures: Chapter 5

Figure 5.1 Untransformed survival and fecundity data plotted against plant density (*A. petiolata* plants m^{-2}) at the beginning of the time interval over which survival was evaluated. Seedling density is the maximum density observed during the spring, summer and winter starting rosette densities were measured in June and October, respectively. Each survival parameter estimate shown is a quadrat mean value, calculated as the proportion of plants marked at the beginning of the time interval that were still alive at the end of the interval. Fecundity is plotted against the density of rosettes during the previous October and is shown as the estimated number of seeds produced by each individual plant within each quadrat rather than quadrat mean values. All plants in a quadrat experience the same starting density and thus align in vertical bands by quadrat in the figure. Note the differences in the x axes. The maximum plant density steadily decreases as *A. petiolata* populations thin over the course of the growing season.


Seedlings
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
2000
3000
4000
5000
6000
Peak Density (Seedlings m-2)


Summer
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
2500
3000
Summer Density (Rosettes m-2)


Winter
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
250
300
350
Fall Density (Rosettes m-2)

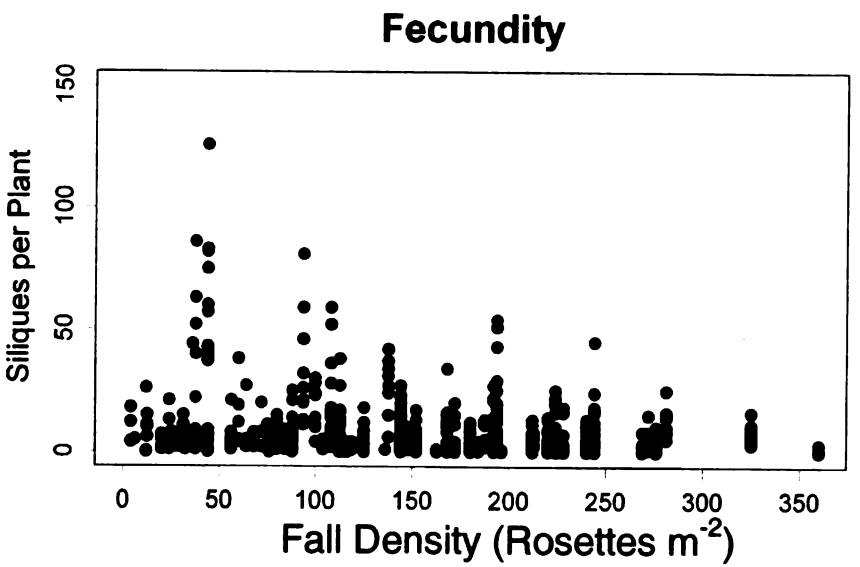

Fecundity
Siliques per Plant
0
50
100
150
0
50
100
150
200
250
300
350
Fall Density (Rosettes m-2)

Figure 5.2. Observed seedling survival data (dots) and initial seedling density (# m^{-2}) showing mean predicted trend (solid line) and lower and upper 95% prediction intervals (dashed lines) for each cohort in Illinois and Michigan based on model results. Dashed lines show 5^{th} and 95^{th} percentiles of the distribution of survival rates based on simulated observations generated from the statistical model which was parameterized from the data. Each distribution is based on 100,000 simulated data sets. Trend lines have been smoothed to illustrate the overall shape of the relationship between density and the probability of survival. Analyses were performed on arcsin-square root transformed survival probabilities and are shown back-transformed to the original scale. Predicted survival rates with negative values were rounded to zero prior to back-transformation to better illustrate the prediction intervals.

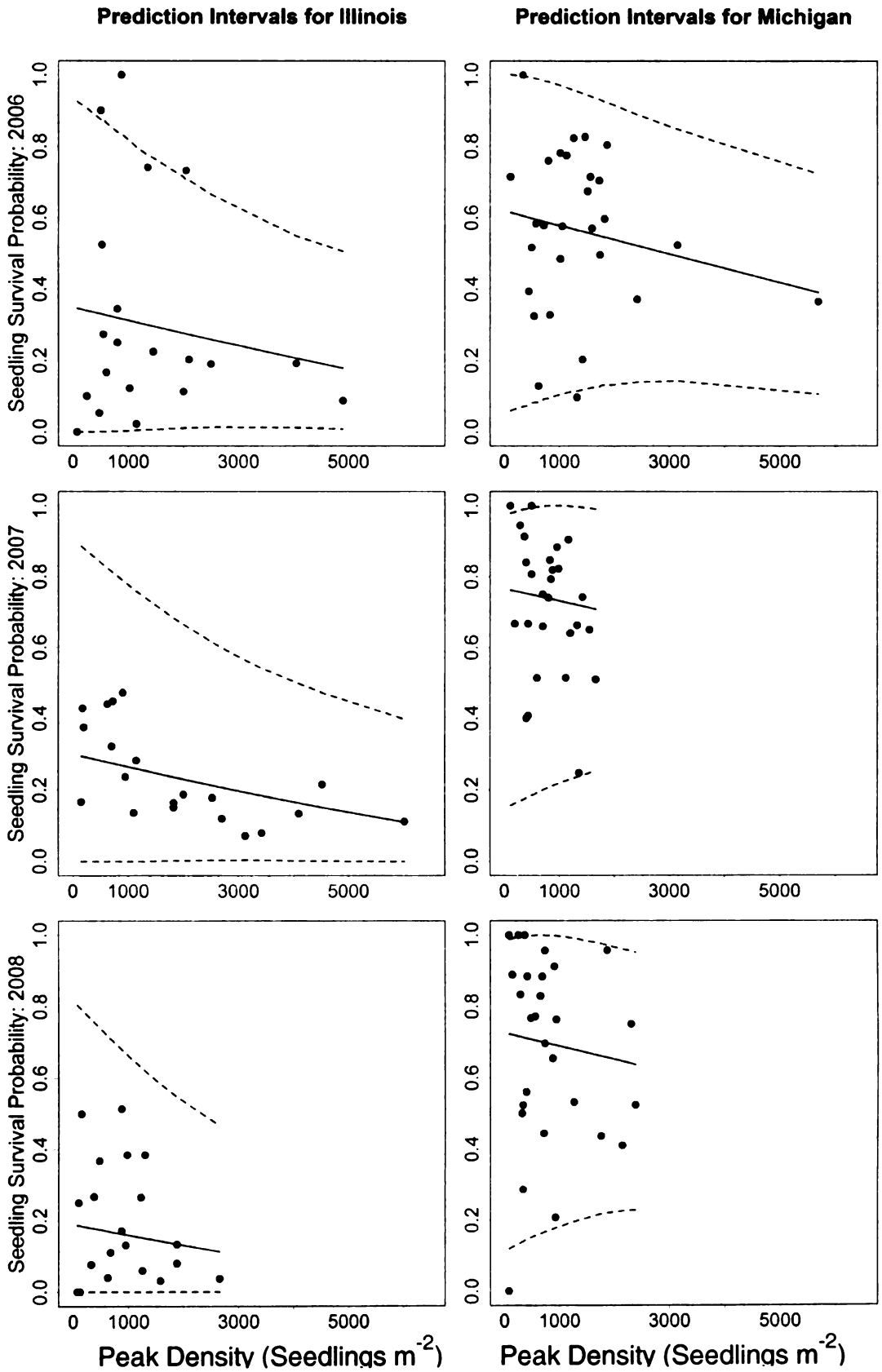

Prediction Intervals for Illinois
Prediction Intervals for Michigan
Seedling Survival Probability: 2006
Seedling Survival Probability: 2007
Seedling Survival Probability: 2008
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
3000
5000
Peak Density (Seedlings m-2)
Peak Density (Seedlings m-2)

Figure 5.3. Observed summer survival data (dots) and initial rosette density (# m^{-2}) in June showing mean predicted trend (solid line) and lower and upper 95% prediction intervals (dashed lines) for each cohort in Illinois and Michigan based on model results. See Figure 5.2 caption for more details. Note the scale of the X axis for Michigan 2005.

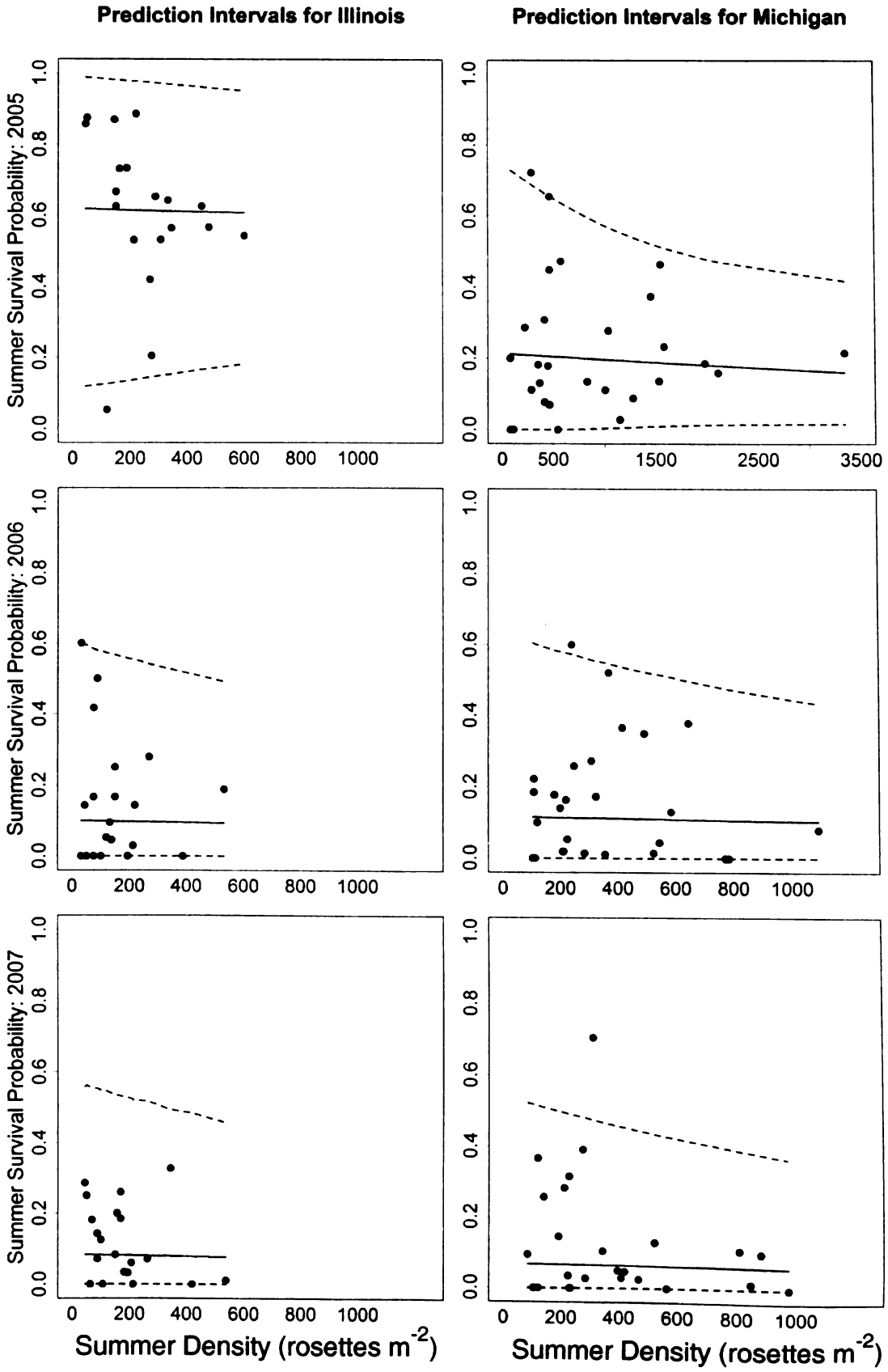

Prediction Intervals for Illinois
Prediction Intervals for Michigan
Summer Survival Probability: 2005
Summer Survival Probability: 2006
Summer Survival Probability: 2007
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
0
500
1500
2500
3500
Summer Density (rosettes m-2)
Summer Density (rosettes m-2)

Figure 5.4. Observed winter survival data (dots) and initial rosette density (# m^{-2}) in fall showing mean predicted trend (solid line) and lower and upper 95% prediction intervals (dashed lines) for each cohort in Illinois and Michigan based on model results. See Figure 5.2 caption for more details.

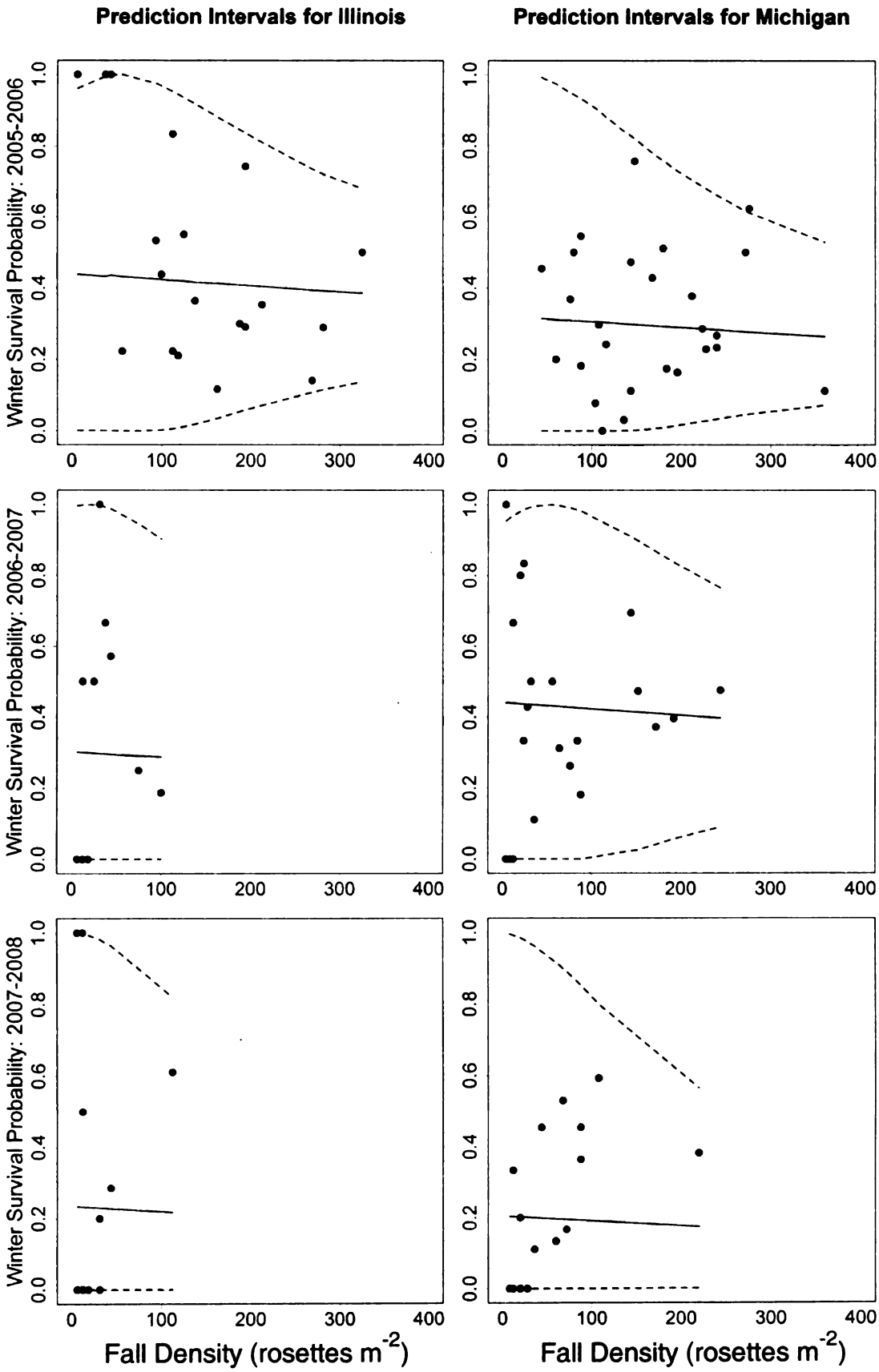
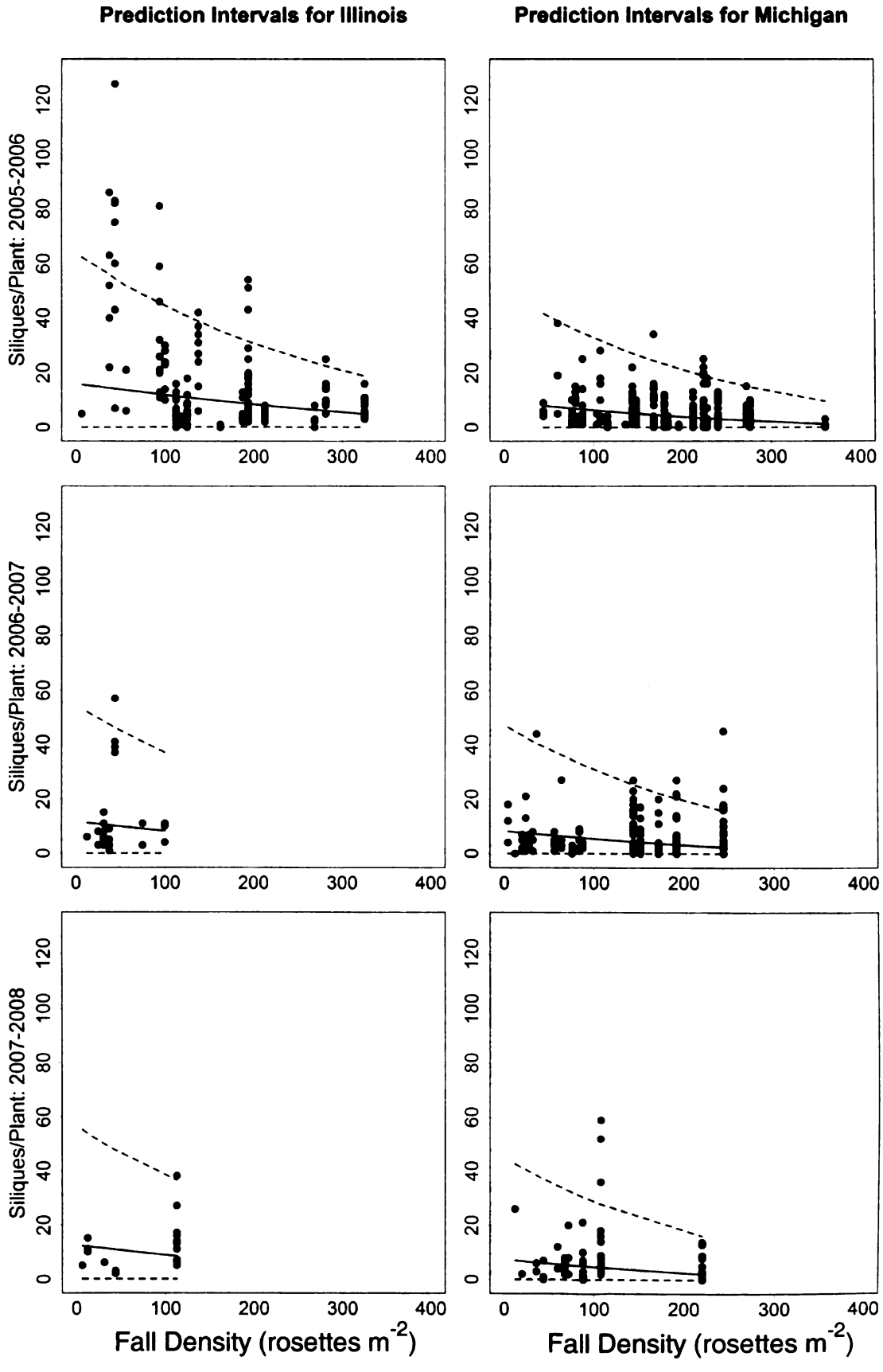

Prediction Intervals for Illinois
Prediction Intervals for Michigan
Winter Survival Probability: 2005-2006
Winter Survival Probability: 2006-2007
Winter Survival Probability: 2007-2008
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
Fall Density (rosettes m-2)
Fall Density (rosettes m-2)

Figure 5.5. Observed fecundity data (dots, siliques plant^{-1}) and initial rosette density (# m^{-2}) in fall showing mean predicted trend (solid line) and lower and upper 95% prediction intervals (dashed lines) for each cohort in Illinois and Michigan based on model results. Each distribution is based on 50,000 simulated data sets. See Figure 5.2 caption for more details.


Prediction Intervals for Illinois
Prediction Intervals for Michigan
Siliques/Plant: 2005-2006
Siliques/Plant: 2006-2007
Siliques/Plant: 2007-2008
0
20
40
60
80
100
120
0
100
200
300
400
Fall Density (rosettes m-2)

LITERATURE CITED

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer.

Anderson, R. C., S. S. Dhillion, and T. M. Kelley. 1996. Aspects of the ecology of an invasive plant, garlic mustard (*Alliaria petiolata*), in central Illinois. Restoration Ecology **4**:181-191.

AOSA. 2000. Tetrazolium testing handbook for agricultural seeds : contribution no. 29 to the Handbook on seed testing. Association of Official Seed Analysts. Tetrazolium Testing Committee., North Brunswick, N.J.

Bais, H. P., R. Vepachedu, S. Gilroy, R. M. Callaway, and J. M. Vivanco. 2003. Allelopathy and exotic plant invasion: From molecules and genes to species interactions. Science **301**:1377-1380.

Bangerter, E. B. 1985. New and interesting records of adventive plants from the Auckland Institute and Museum Herbarium: 11. Records of the Auckland Institute and Museum **22**:41-45.

Baskin, J. M., and C. C. Baskin. 1992. Seed-germination biology of the weedy biennial *Alliaria petiolata*. Natural Areas Journal **12**:191-197.

Bellows, T. S. 1999. Whither hence, Prometheus? Pages 1011-1015 *in* T. S. Bellows and T. W. Fisher, editors. Handbook of Biological Control. Academic Press, New York.

Bierzychudek, P. 1982. The demography of jack-in-the-pulpit, a forest perennial that changes sex. Ecological Monographs **52**:335-351.

Bierzychudek, P. 1999. Looking backwards: Assessing the projections of a transition matrix model. Ecological Applications **9**:1278-1287.

Blossey, B. 1999. Before, during and after: The need for long-term monitoring in invasive plant species management. Biological Invasions **1**:301-311.

Blossey, B., R. Casagrande, L. Tewksbury, D. A. Landis, R. N. Wiedenmann, and D. R. Ellis. 2001a. Nontarget feeding of leaf-beetles introduced to control purple loosestrife (*Lythrum salicaria* L.). Natural Areas Journal **21**:368-377.

Blossey, B., V. Nuzzo, H. Hinz, and E. Gerber. 2001b. Developing biological control of *Alliaria petiolata* (M. Bieb.) Cavara and Grande (garlic mustard). Natural Areas Journal **21**:357-367.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution **24**:127-135.

Brady, N. C., and R. R. Weil. 1996. The nature and properties of soils. 11th edition. Prentice Hall, Upper Saddle River, N.J.

Briese, D. T. 2006. Can an a priori strategy be developed for biological control? The case of *Onopordum* spp. thistles in Australia. Australian Journal of Entomology **45**:317-323.

Buckley, Y. M., D. T. Briese, and M. Rees. 2003a. Demography and management of the invasive plant species *Hypericum perforatum*. I. Using multi-level mixed-effects models for characterizing growth, survival and fecundity in a long-term data set. Journal of Applied Ecology **40**:481-493.

Buckley, Y. M., D. T. Briese, and M. Rees. 2003b. Demography and management of the invasive plant species Hypericum perforatum. II. Construction and use of an individual-based model to predict population dynamics and the effects of management strategies. Journal of Applied Ecology **40**:494-507.

Buckley, Y. M., H. L. Hinz, D. Matthies, and M. Rees. 2001. Interactions between density-dependent processes population dynamics and control of an invasive plant species, *Tripleurospermum perforatum* (scentless chamomile). Ecology Letters **4**:551-558.

Burkey, J. 2009. LOWESS smoothing function for MATLAB, http://www.mathworks.com/matlabcentral/fileexchange/22470.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference : a practical information-theoretic approach. 2nd edition. Springer, New York.

Byers, D. L., and J. A. Quinn. 1998. Demographic variation in *Alliaria petiolata* (Brassicaceae) in four contrasting habitats. Journal of the Torrey Botanical Society **125**:138-149.

Call, L. J., and E. T. Nilsen. 2003. Analysis of spatial patterns and spatial association between the invasive tree-of-heaven (*Ailanthus altissima*) and the native black locust (*Robinia pseudoacacia*). American Midland Naturalist **150**:1-14.

Callaway, R. M., D. Cipollini, K. Barto, G. C. Thelen, S. G. Hallett, D. Prati, K. Stinson, and J. Klironomos. 2008. Novel weapons: Invasive plant suppresses fungal mutualists in America but not in its native Europe. Ecology **89**:1043-1055.

Callaway, R. M., T. H. DeLuca, and W. M. Belliveau. 1999. Biological-control herbivores may increase competitive ability of the noxious weed *Centaurea maculosa*. Ecology **80**:1196-1201.

Carlson, A. M., and D. L. Gorchov. 2004. Effects of herbicide on the invasive biennial *Alliaria petiolata* (garlic mustard) and initial responses of native plants in a southwestern Ohio forest. Restoration Ecology **12**:559-567.

Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation. 2nd edition. Sinauer Associates, Sunderland, MA.

Cavers, P. B., M. I. Heagy, and R. F. Kokron. 1979. Biology of Canadian weeds. 35. *Alliaria petiolata* (M Bieb) Cavara and Grande. Canadian Journal of Plant Science **59**:217-229.

Cory, J. S., and J. H. Myers. 2000. Direct and indirect ecological effects of biological control. Trends in Ecology and Evolution **15**:137-139.

Crawley, M. J. 2007. The R book. Wiley, Chichester, England ; Hoboken, N.J.

Cruden, R. W., A. M. McClain, and G. P. Shrivastava. 1996. Pollination biology and breeding system of *Alliaria petiolata* (Brassicaceae). Bulletin of the Torrey Botanical Club **123**:273-280.

Cunningham, R. B., and D. B. Lindenmayer. 2005. Modeling count data of rare species: Some statistical issues. Ecology **86**:1135-1142.

Davis, A. S., D. A. Landis, V. Nuzzo, B. Blossey, H. Hinz, and E. Gerber. 2006. Demographic models inform biocontrol agent selection for garlic mustard (*Alliaria petiolata*). Ecological Applications **16**:2399-2410.

Denoth, M., L. Frid, and J. H. Myers. 2002. Multiple agents in biological control: improving the odds? Biological Control **24**:20-30.

Dhillion, S. S., and R. C. Anderson. 1999. Growth and photosynthetic response of first-year garlic mustard (*Alliaria petiolata*) to varied irradiance. Journal of the Torrey Botanical Society **126**:9-14.

Drake, J. M. 2005. Density-dependent demographic variation determines extinction rate of experimental populations. Plos Biology **3**:1300-1304.

Draper, N. R., and H. Smith. 1998. Applied regression analysis. 3rd edition. Wiley, New York.

Drayton, B., and R. B. Primack. 1999. Experimental extinction of garlic mustard (*Alliaria petiolata*) populations: implications for weed science and conservation biology. Biological Invasions **1**:159-167.

Durka, W., O. Bossdorf, D. Prati, and H. Auge. 2005. Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Molecular Ecology **14**:1697-1706.

Elliott, J. M., and M. A. Hurley. 1998. Population regulation in adult, but not juvenile, resident trout (*Salmo trutta*) in a lake district stream. Journal of Animal Ecology **67**:280-286.

Engen, S., O. Bakke, and A. Islam. 1998. Demographic and environmental stochasticity - Concepts and definitions. Biometrics **54**:840-846.

Evans, J. A. 2006. Impacts of herbivores and plant communities on establishment and spread of *Alliaria petiolata* (garlic mustard) in Michigan. M.S. Thesis. Michigan State University, East Lansing, MI.

Evans, J. A., A. S. Davis, S. Raghu, A. Ragavendran, D. A. Landis, and D. W. Schemske. unpublished manuscript. The scale and sources of demographic variation in the invasive plant *Alliaria petiolata*.

Evans, J. A., and D. A. Landis. 2007. Pre-release monitoring of *Alliaria petiolata* (garlic mustard) invasions and the impacts of extant natural enemies in southern Michigan forests. Biological Control **42**:300-307.

Gee, G. W., and J. W. Bauder. 1986. Particle-size analysis. Pages 383–411 *in* A. Klute, editor. Methods of soil analysis: Part 1: Physical and mineralogical methods. American Society of Agronomy : Soil Science Society of America, Madison, WI.

Gerber, E., and H. Hinz. 2005. Biology, impact and interactions of potential biocontrol agents on garlic mustard. Pages 4-6 *in* Symposium on the biology, ecology, and management of garlic mustard (*Alliaria petiolata*) and European buckthorn (*Rhamnus cathartica*). USDA Forest Service, University of Minnesotta, St. Paul, MN.

Gerber, E., H. L. Hinz, and B. Blossey. 2007a. Impact of the belowground herbivore and potential biological control agent, *Ceutorhynchus scrobicollis*, on *Alliaria petiolata* performance. Biological Control **42**:355-364.

Gerber, E., H. L. Hinz, and B. Blossey. 2007b. Interaction of specialist root and shoot herbivores of *Alliaria petiolata* and their impact on plant performance and reproduction. Ecological Entomology **32**:357-365.

Gerber, E., H. L. Hinz, and B. Blossey. 2008a. Pre-release impact assessment of two stem-boring weevils proposed as biological control agents for *Alliaria petiolata*. Biological Control **45**:360-367.

Gerber, E., H. L. Hinz, and G. Cortat. 2008b. Biological control of garlic mustard, *Alliaria petiolata*. Annual report 2007. Unpublished Report, CABI Bioscience Switzerland Centre, Delemont, Switzerland.

Gerber, E., H. L. Hinz, N. Guazzone, J. McKenney, S. Michler, and M. Zuefle. 2002. Biological control of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande. Annual report 2001. Unpublished Report, CABI Bioscience Switzerland Centre, Delemont, Switzerland.

Goldberg, D. E., R. Turkington, L. Olsvig-Whittaker, and A. R. Dyer. 2001. Density dependence in an annual plant community: Variation among life history stages. Ecological Monographs **71**:423-446.

Grant, D. W., D. P. C. Peters, G. K. Beck, and H. D. Fraleigh. 2003. Influence of an exotic species, *Acroptilon repens* (L.) DC. on seedling emergence and growth of native grasses. Plant Ecology **166**:157-166.

Gutierrez, A. P., L. E. Caltagirone, and W. Meikle. 1999. Evaluation of results: economics of biological control. Pages 243-252 *in* T. S. Bellows and T. W. Fisher, editors. Handbook of Biological Control. Academic Press, New York.

Harris, P. 1991. Classical biocontrol of weeds - Its definition, selection of effective agents, and administrative political problems. Canadian Entomologist **123**:827-849.

Hinz, H., and E. Gerber. 2005. Summary of host specificity testing conducted in Switzerland and Brassicaceae taxonomy. Pages 12-16 *in* Symposium on the biology, ecology, and management of garlic mustard (*Alliaria petiolata*) and European buckthorn (*Rhamnus cathartica*). USDA Forest Service, University of Minnesotta, St. Paul, MN.

Hinz, H. L., and E. Gerber. 1998. Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande Annual Report 1998. Unpublished Report, CABI Bioscience Switzerland Centre, Delemont, Switzerland.

Hinz, H. L., and E. Gerber. 2001. Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande. Annual Report 2000. Unpublished Report, CABI Bioscience Switzerland Centre, Delemont, Switzerland.

Hoddle, M. S. 2004a. Restoring balance: Using exotic species to control invasive exotic species. Conservation Biology **18**:38-49.

Hoddle, M. S. 2004b. The strength of biological control in the battle against invasive pests: a reply. Conservation Biology **18**:61-64.

Hood, G. M. 2006. PopTools version 2.7. Available free on the internet. URL http://www.cse.csiro.au/poptools/. Albany, Western Australia.

Horvitz, C. C., and D. W. Schemske. 1986. Seed dispersal and environmental heterogeneity in a neotropical herb: a model of poopulation and patch dynamics.

Pages 169-185 *in* A. Estrada and T. H. Fleming, editors. Frugivores and seed dispersal. Dr W. Junk, Boston.

Horvitz, C. C., and D. W. Schemske. 1995. Spatiotemporal variation in demographic transitions of a tropical understory herb - projection matrix analysis. Ecological Monographs **65**:155-192.

Huffaker, C. B., F. J. Simmonds, and J. E. Laing. 1976. Theoretical and empirical basis of biological control. Pages 41-78 *in* C. B. Huffaker and P. S. Messenger, editors. Theory and practice of biological control. Academic Press, New York.

Joe, H., and R. Zhu. 2005. Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution. Biometrical Journal **47**:219-229.

Keeler, M. S., and F. S. Chew. 2008. Escaping an evolutionary trap: preference and performance of a native insect on an exotic invasive host. Oecologia **156**:559-568.

Klute, A. 1986. Water retention: laboratory methods. Pages 635-662 *in* A. Klute, editor. Methods of soil analysis: Part 1: Physical and mineralogical methods. American Society of Agronomy : Soil Science Society of America, Madison, WI.

Koop, A. L., and C. C. Horvitz. 2005. Projection matrix analysis of the demography of an invasive, nonnative shrub (*Ardisia elliptica*). Ecology **86**:2661-2672.

Landis, D. A., D. C. Sebolt, M. J. Haas, and M. Klepinger. 2003. Establishment and impact of *Galerucella calmariensis* L. (Coleoptera : Chrysomelidae) on *Lythrum salicaria* L. and associated plant communities in Michigan. Biological Control **28**:78-91.

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schaebenberger. 2006. SAS for mixed models. 2nd edition. SAS Institute, Cary, N.C.

Louda, S., D. Kendall, J. Connor, and D. Simberloff. 1997. Ecological effects of an insect introduced for the biological control of weeds. Science **277**:1088-1090.

Louda, S. M., A. E. Arnett, T. A. Rand, and F. L. Russell. 2003a. Invasiveness of some biological control insects and adequacy of their ecological risk assessment and regulation. Conservation Biology **17**:73-82.

Louda, S. M., R. W. Pemberton, M. T. Johnson, and P. A. Follett. 2003b. Nontarget effects - The Achilles' Heel of biological control? Retrospective analyses to reduce risk associated with biocontrol introductions. Annual Review of Entomology **48**:365-396.

Louda, S. M., T. A. Rand, A. E. Arnett, A. S. McClay, K. Shea, and A. K. McEachern. 2005a. Evaluation-of ecological risk to populations of a threatened plant from an invasive biocontrol insect. Ecological Applications **15**:234-249.

Louda, S. M., T. A. Rand, F. L. Russell, and A. E. Arnett. 2005b. Assessment of ecological risks in weed biocontrol: Input from retrospective ecological analyses. Biological Control **35**:253-264.

Louda, S. M., and P. Stiling. 2004. The double-edged sword of biological control in conservation and restoration. Conservation Biology **18**:50-53.

Luken, J. O., L. M. Kuddes, and T. C. Tholemeier. 1997. Response of understory species to gap formation and soil disturbance in *Lonicera maackii* thickets. Restoration Ecology **5**:229-235.

McEvoy, P. B. 1996. Host specificity and biological pest control - How well is research on host specificity addressing the potential risks of biological control. Bioscience **46**:401-405.

McEvoy, P. B., and E. M. Coombs. 1999. Biological control of plant invaders: Regional patterns, field experiments, and structured population models. Ecological Applications **9**:387-401.

McEvoy, P. B., and E. M. Coombs. 2000. Why things bite back: unintended consequences of biological weed control. Pages 167-194 *in* P. A. Follett and J. J. Duan, editors. Nontarget effects of biological control. Kluwer Academic Publishers, Boston.

McFadyen, R. E. C. 1998. Biological control of weeds. Annual Review of Entomology **43**:369-393.

Meekins, J. F., and B. C. McCarthy. 1999. Competitive ability of *Alliaria petiolata* (garlic mustard, Brassicaceae), an invasive, nonindigenous forest herb. International Journal of Plant Sciences **160**:743-752.

Meekins, J. F., and B. C. McCarthy. 2000. Responses of the biennial forest herb *Alliaria petiolata* to variation in population density, nutrient addition and light availability. Journal of Ecology **88**:447-463.

Meekins, J. F., and B. C. McCarthy. 2002. Effect of population density on the demography of an invasive plant (*Alliaria petiolata*, Brassicaceae) population in a southeastern Ohio forest. American Midland Naturalist **147**:256-278.

Mulvaney, R. L. 1996. Nitrogen—inorganic forms. Pages 1123-1184 *in* D. L. Sparks, Soil Science Society of America., and American Society of Agronomy., editors. Methods of soil analysis. Part 3, Chemical methods. Soil Science Society of America : American Society of Agronomy, Madison, WI.

Myers, C. V., and R. C. Anderson. 2003. Seasonal variation in photosynthetic rates influences success of an invasive plant, garlic mustard (*Alliaria petiolata*). American Midland Naturalist **150**:231-245.

Myers, C. V., R. C. Anderson, and D. L. Byers. 2005. Influence of shading on the growth and leaf photosynthesis of the invasive non-indigenous plant garlic mustard *Alliaria petiolata* (M. Bieb) Cavara and Grande grown under simulated late-winter to mid-spring conditions. Journal of the Torrey Botanical Society **132**:1-10.

Nelson, D. W., and L. E. Sommers. 1994. Total carbon, organic carbon, and organic matter. Pages 961-1010 *in* R. W. Weaver and Soil Science Society of America., editors. Methods of soil analysis. Part 2, Microbiological and biochemical properties. Soil Science Society of America, Madison, Wis., USA.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear regression models. 3rd edition. Irwin, Chicago, Ill.

Nuzzo, V. A. 1991. Experimental control of garlic mustard [*Alliaria petiolata* (Bieb.) Cavara & Grande] in Northern Illinois using fire, herbicide, and cutting. Natural Areas Journal **11**:158-167.

Nuzzo, V. A. 1993a. Current and historic distribution of garlic mustard (*Alliaria Petiolata)* in Illinois. The Michigan Botanist **32**:23-33.

Nuzzo, V. A. 1993b. Distribution and spread of the invasive biennial *Alliaria petiolata* (garlic mustard) in North America. Pages 137-145 *in* B. N. McKnight, editor.

Biological pollution: The control and impact of invasive exotic species: proceedings of a symposium held at the University Place Conference Center, Indiana University-Purdue University at Indianapolis on October 25 & 26, 1991. Indiana Academy of Science, Indianapolis.

Nuzzo, V. A. 1993c. Natural mortality of garlic mustard (*Alliaria petiolata* (Bieb.) Cavara & Grande) rosettes. Natural Areas Journal **13**:132-133.

Nuzzo, V. A. 1994. Response of garlic mustard (*Alliaria petiolata* Bieb - Cavara and Grande) to summer herbicide treatment. Natural Areas Journal **14**:309-310.

Nuzzo, V. A. 1996. Impact of dormant season herbicide treatment on the alien herb garlic mustard (*Alliaria petiolata* [Bieb.] Cavara and Grande) and groundlayer vegetation. Transactions of the Illinois State Academy of Science **89**:25-36.

Nuzzo, V. A. 1999. Invasion pattern of the herb garlic mustard (*Alliaria petiolata*) in high quality forests. Biological Invasions **1**:169-179.

Nuzzo, V. A. 2000. Element stewardship abstract for *Alliaria petiolata* (*Alliaria officinalis*) garlic mustard. The Nature Conservancy.

Nuzzo, V. A., W. McClain, and T. Strole. 1996. Fire impact on groundlayer flora in a sand forest 1990-1994. American Midland Naturalist **136**:207-221.

Ortega, Y. K., D. E. Pearson, and K. S. McKelvey. 2004. Effects of biological control agents and exotic plant invasion on deer mouse populations. Ecological Applications **14**:241-253.

Pardini, E. A., J. M. Drake, J. M. Chase, and T. M. Knight. 2009. Complex population dynamics and control of the invasive biennial, *Alliaria petiolata* (garlic mustard). Ecological Applications **19**:387-397.

Pardini, E. A., B. J. Teller, and T. M. Knight. 2008. Consequences of density dependence for management of a stage-sructured invasive plant (*Alliaria petiolata*). American Midland Naturalist **160**:310-322.

Parker, I. M. 2000. Invasion dynamics of *Cytisus scoparius*: A matrix model approach. Ecological Applications **10**:726-743.

Pascarella, J. B., and C. C. Horvitz. 1998. Hurricane disturbance and the population dynamics of a tropical understory shrub: Megamatrix elasticity analysis. Ecology **79**:547-563.

Pearson, D., and R. Callaway. 2003. Indirect effects of host-specific biological control agents. Trends in Ecology & Evolution **18**:456-461.

Pearson, D. E., and R. M. Callaway. 2004. Response to Thomas et al: biocontrol and indirect effects. Trends in Ecology & Evolution **19**:62-63.

Pearson, D. E., and R. M. Callaway. 2005. Indirect nontarget effects of host-specific biological control agents: Implications for biological control. Biological Control **35**:288-298.

Pemberton, R. W. 2000. Predictable risk to native plants in weed biological control. Oecologia **125**:489-494.

Pichtel, J., K. Kuroiwa, and H. T. Sawyerr. 2000. Distribution of Pb, Cd and Ba in soils and plants of two contaminated sites. Environmental Pollution **110**:171-178.

Prati, D., and O. Bossdorf. 2004. Allelopathic inhibition of germination by *Alliaria petiolata* (Brassicaceae). American Journal of Botany **91**:285-288.

Pyle, L. L. 1995. Effects of disturbance on herbaceous exotic plant-species on the floodplain of the Potomac River. American Midland Naturalist **134**:244-253.

Raghu, S., and S. Post. 2008. Cold stratification requirements for germination of *Alliaria petiolata*. Invasive Plant Science and Management **1**:315–318.

Raghu, S., and R. D. van Klinken. 2006. Refining the ecological basis for agent selection in weed biological control. Australian Journal of Entomology **45**:251-252.

Rebek, K. A., and R. J. O'Neil. 2006. The effects of natural and manipulated density regimes on *Alliaria petiolata* survival, growth and reproduction. Weed Research **46**:345-352.

Rees, M., and R. L. Hill. 2001. Large-scale disturbances, biological control and the dynamics of gorse populations. Journal of Applied Ecology **38**:364-377.

Rejmanek, M. 2000. On the use and misuse of transition matrices in plant population biology. Biological Invasions **2**:315-317.

Renwick, J. A. A. 2002. The chemical world of crucivores: lures, treats and traps. Entomologia Experimentalis et Applicata **104**:35-42.

Roberts, K. J., and R. C. Anderson. 2001. Effect of garlic mustard *Alliaria petiolata* (Beib. Cavara & Grande) extracts on plants and arbuscular mycorrhizal (AM) fungi. American Midland Naturalist **146**:146-152.

Rodgers, V. L., K. A. Stinson, and A. C. Finzi. 2008. Ready or not, garlic mustard is moving in: *Alliaria petiolata* as a member of eastern North American forests. Bioscience **58**:426-436.

Saether, B.-E., S. Engen, A. Islam, R. McCleery, and C. Perrins. 1998. Environmental stochasticity and extinction risk in a population of a small songbird, the geat tit. The American Naturalist **151**:441-450.

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O'Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics **32**:305-332.

SAS Institute. 2008. The SAS System Version 9.2. Cary, NC.

Shea, K., and D. Kelly. 1998. Estimating biocontrol agent impact with matrix models: *Carduus nutans* in New Zealand. Ecological Applications **8**:824-832.

Shea, K., D. Kelly, A. W. Sheppard, and T. L. Woodburn. 2005. Context-dependent biological control of an invasive thistle. Ecology **86**:3174-3181.

Simberloff, D., and P. Stiling. 1996. How risky is biological control? Ecology **77**:1965-1974.

Skrondal, A., and S. Rabe-Hesketh. 2004. Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. Pages xi, 508 p. Chapman & Hall/CRC, Boca Raton.

Smith, G. R., H. A. Dingfelder, and D. A. Vaala. 2003. Effect of plant size and density on garlic mustard reproduction. Northeastern Naturalist **10**:269-276.

Soltanpour, P. N., G. W. Johnson, S. W. Workman, J. Benton-Jones Jr., and M. R. O. 1996. Inductively coupled plasma spectroscopy. Pages 91-140 *in* D. L. Sparks, Soil Science Society of America., and American Society of Agronomy., editors. Methods of soil analysis. Part 3, Chemical methods. Soil Science Society of America : American Society of Agronomy, Madison, WI.

Stiling, P. 2002. Potential non-target effects of a biological control agent, prickly pear moth, *Cactoblastis cactorum* (Berg) (Lepidoptera: Pyralidae), in North America, and possible management actions. Biological Invasions **4**:273-281.

Stiling, P., D. Moon, and D. Gordon. 2004. Endangered cactus restoration: Mitigating the non-target effects of a biological control agent (*Cactoblastis cactorum*) in Florida. Restoration Ecology **12**:605-610.

Stinson, K., S. Kaufman, L. Durbin, and F. Lowenstein. 2007. Impacts of garlic mustard invasion on a forest understory community. Northeastern Naturalist **14**:73-88.

Stinson, K. A., S. A. Campbell, J. R. Powell, B. E. Wolfe, R. M. Callaway, G. C. Thelen, S. G. Hallett, D. Prati, and J. N. Klironomos. 2006. Invasive plant suppresses the growth of native tree seedlings by disrupting belowground mutualisms. Plos Biology **4**:e140 http://dx.doi.org/110.1371/journal.pbio.0040140.

Susko, D. J., and L. Lovett-Doust. 1999. Effects of resource availability, and fruit and ovule position on components of fecundity in *Alliaria petiolata* (Brassicaceae). New Phytologist **144**:295-306.

Susko, D. J., and L. Lovett-Doust. 2000. Patterns of seed mass variation and their effects on seedling traits in *Alliaria petiolata* (Brassicaceae). American Journal of Botany **87**:56-66.

The MathWorks. 2008. MATLAB. Natick, MA.

The Nature Conservancy. 2002. http://tncweeds.ucdavis.edu/success/ak001.html.

Thomas, G. W. 1996. Soil pH and soil acidity. Pages 475-490 *in* D. L. Sparks, Soil Science Society of America., and American Society of Agronomy., editors.

Methods of soil analysis. Part 3, Chemical methods. Soil Science Society of America : American Society of Agronomy, Madison, WI.

Thomas, M. B., P. Casula, and A. Wilby. 2004. Biological control and indirect effects. Trends in Ecology & Evolution **19**:61-61.

Trimbur, T. J. 1973. The ecological life history of *Allaria officinallis*, a deciduous forest "weed". M.S. Thesis. The Ohio State University, Columbus, OH.

Tuljapurkar, S., C. C. Horvitz, and J. B. Pascarella. 2003. The many growth rates and elasticities of populations in random environments. American Naturalist **162**:489-502.

Turchin, P. 1999. Population regulation: a synthetic view. Oikos **84**:153-159.

USDA-APHIS. 2006. Technical Advisory Group for Biological Control Agents of Weeds home page, http://www.aphis.usda.gov/ppq/permits/tag/. USDA.

USDA-NRCS. 2007. The *PLANTS* Databse, Version 3.5 (http://plants.usda.gov). Data compiled from various sources by Mark W. Skinner. National Plant Data Center, Baton Rouge, LA 70874-4490 USA.

Van Driesche, R. G., and T. S. Bellows. 1996. Biological Control. Chapman & Hall, Detroit.

van Klinken, R. D., and S. Raghu. 2006. A scientific approach to agent selection. Australian Journal of Entomology **45**:253-258.

van Tienderen, P. H. 1995. Life-Cycle Trade-Offs in Matrix Population-Models. Ecology **76**:2482-2489.

Vaughn, S. F., and M. A. Berhow. 1999. Allelochemicals isolated from tissues of the invasive weed garlic mustard (*Alliaria petiolata*). Journal of Chemical Ecology **25**:2495-2504.

Vitousek, P. M., and L. R. Walker. 1989. Biological Invasion by *Myrica faya* in Hawaii - Plant Demography, Nitrogen-Fixation, Ecosystem Effects. Ecological Monographs **59**:247-265.

Wapshere, A. J. 1974. Strategy for evaluating safety of organisms for biological weed control. Annals of Applied Biology **77**:201-211.

Wapshere, A. J. 1989. A testing sequence for reducing rejection of potential biological-control agents for weeds. Annals of Applied Biology **114**:515-526.

Welk, E., K. Schubert, and M. H. Hoffmann. 2002. Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in North America. Diversity and Distributions **8**:219-233.

Weston, L. A., and S. O. Duke. 2003. Weed and crop allelopathy. Critical Reviews in Plant Sciences **22**:367-389.

Whitten, M. J., and M. A. Hoy. 1999. Genetic improvement and other genetic considerations for improving the efficacy and success rate of biological control. Pages 271-296 *in* T. S. Bellows and T. W. Fisher, editors. Handbook of Biological Control. Academic Press, New York.

Williamson, M. 1996. Biological invasions. 1st edition. Chapman & Hall, London ; New York.

Winterer, J., M. C. Walsh, M. Poddar, J. W. Brennan, and S. M. Primak. 2005. Spatial and temporal segregation of juvenile and mature garlic mustard plants (*Alliaria petiolata*) in a central Pennsylvania woodland. American Midland Naturalist **153**:209-216.

Wolfe, B. E., and J. N. Klironomos. 2005. Breaking new ground: Soil communities and exotic plant invasion. Bioscience **55**:477-487.

Wolfe, B. E., V. L. Rodgers, K. A. Stinson, and A. Pringle. 2008. The invasive plant *Alliaria petiolata* (garlic mustard) inhibits ectomycorrhizal fungi in its introduced range. Journal of Ecology **96**:777-783.

Yates, C. N., and S. D. Murphy. 2008. Observations of herbivore attack on garlic mustard (*Alliaria petiolata*) in Southwestern Ontario, Canada. Biological Invasions **10**:757-760.