1-1-2016

Understanding the Exotic Distributions of Invasive Species

Tyler Evan Schartel

Recommended Citation

Schartel, Tyler Evan, "Understanding the Exotic Distributions of Invasive Species" (2016). *Theses and Dissertations*. 4838.
https://scholarsjunction.msstate.edu/td/4838

Understanding the exotic distributions of invasive species

By

Tyler Evan Schartel

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Biological Sciences
in the Department of Biological Sciences

Mississippi State, Mississippi

December 2016

Understanding the exotic distributions of invasive species

By

Tyler Evan Schartel

Approved:

Christopher P. Brooks
(Major Professor)

Richard L. Brown
(Committee Member)

Brian A. Counterman
(Committee Member)

Gary N. Ervin
(Committee Member)

Brian J. Rude
(Committee Member)

Lisa Wallace
(Committee Member)

Mark E. Welch
(Graduate Coordinator)

Rick Travis
Interim Dean
College of Arts & Sciences

Name: Tyler Evan Schartel

Date of Degree: December 9, 2016

Institution: Mississippi State University

Major Field: Biological Sciences

Major Professor: Christopher P. Brooks

Title of Study: Understanding the exotic distributions of invasive species

Pages in Study: 149

Candidate for Degree of Doctor of Philosophy

Invasive species are organisms whose introduction and spread in exotic ranges result in a multitude of ecological impacts. Understanding the factors that constrain the exotic distributions of invasive species is of considerable interest. Biotic associations formed with taxa in the invaded community may be particularly important in shaping invader distributions. These associations emerge from interactions between the traits of the invasive species and some subset of the traits present in the invaded community. Focusing on how organism traits influence the outcomes of biotic interactions may inform predictions of invader distributions. This kind of trait-based approach may be most easily applied to systems where invaders specialize on particular hosts because such associations imply a close correspondence between the traits of the invader and hosts. This dissertation focuses on the South American cactus moth (*Cactoblastis cactorum*, Lepidoptera: Pyralidae), an invasive consumer in North America whose larvae infest prickly-pear cacti (*Opuntia* spp.).

Chapter One is a brief introduction providing background and context to the presented research. In Chapter Two, I quantify *Opuntia* morphological and tissue macronutrient traits hypothesized to correlate with patterns of *C. cactorum* host use.

Tissue macronutrient traits appear important in predicting *C. cactorum* infestation whereas a model containing *Opuntia* morphological traits had poor predictive ability. Chapter Three describes a method that uses host *Opuntia* identity and availability to estimate habitat suitability in order to predict the North American distribution of *C. cactorum*. I then simulate *C. cactorum* dispersal relative to scenarios of habitat suitability and *Opuntia* availability. Chapter Four alters the model in Chapter Three so that habitat suitability for *C. cactorum* is determined by the availability of trait-based groupings of *Opuntia* hosts. I then simulate *C. cactorum* dispersal via a different method from that described in Chapter Three. In Chapters Three and Four, I evaluate the degree of similarity among model predictions and the relative contribution of modeling constraints in generating variation in this similarity. Chapters Three and Four predictions were most affected by estimates of abiotic suitability and dispersal constraints, respectively. Chapter Five is a short summary of my results and a discussion of their more general applicability.

DEDICATION

This dissertation is dedicated to everyone that has supported me throughout the long journey affectionately known as “graduate school”.

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions of my major advisor, Christopher Brooks, to this work. His guidance and advice have been essential to the planning, execution, and publishing of this research. This research would not have been successful without him.

I would also like to acknowledge the contributions of my advisory committee to this work. My conversations with committee members have been invaluable to refining my research and have helped me place my research into broader context.

Lastly, I would like to acknowledge the contributions of Ronn Altig to this work. Ronn has been a source of constant support during my Ph.D. program, and our conversations have helped motivate me to see this research through to the end.

I am grateful to all of you. Thank you.

TABLE OF CONTENTS

DEDICATION ... ii

ACKNOWLEDGEMENTS ... iii

LIST OF TABLES ... vi

LIST OF FIGURES ... viii

CHAPTER

I. INTRODUCTION ... 1

II. NATIVE HOST TRAITS PREDICT PATTERNS OF INVADER HOST USE ... 6

- Materials and Methods ... 10
  - Study area and cladode sampling procedures ... 10
  - Quantifying *Opuntia* morphological and tissue macronutrient traits ... 11
  - Data analysis ... 12
- Results ... 14
- Discussion ... 15

III. EVALUATING THE RELATIVE IMPORTANCE OF FACTORS GOVERNING INVADER DISTRIBUTIONS ... 27

- Materials and Methods ... 31
  - Study area and occurrence records ... 31
  - Generating Maxent predictions of *C. cactorum* and *Opuntia* taxa distributions ... 32
  - Estimating habitat suitability for *C. cactorum* ... 33
  - Simulating *C. cactorum* dispersal ... 34
  - Evaluating similarity among PAB and Maxent predictions ... 36
- Results ... 38
- Discussion ... 40

IV. A TRAIT-BASED METHOD FOR PREDICTING THE EXOTIC DISTRIBUTIONS OF INVASIVE SPECIES ... 51

- Introduction ... 51

Materials and Methods 55
  Study area and cladode sampling procedures 55
  Quantifying *Opuntia* morphological and tissue macronutrient traits 56
  Generating trait-based *Opuntia* clusters 57
  Generating Maxent predictions and assessing prediction performance 60
  Integrating abiotic suitability and trait-based understanding 62
  Incorporating *C. cactorum* dispersal 63
  Evaluation of trait-based SDM, PAB, and Maxent predictions 64
Results 66
Discussion...69

V. DISSERTATION SUMMARY 86

REFERENCES 90

APPENDIX

A. CHAPTER 2 SUPPLEMENTARY INFORMATION 107

B. CHAPTER 3 SUPPLEMENTARY INFORMATION 112

C. CHAPTER 4 SUPPLEMENTARY INFORMATION 132

LIST OF TABLES

2.1 List of cladodes included in this study. 18

2.2 Morphological traits quantified for *Opuntia* cladodes. 19

2.3 GLMs predicting infestation by *M. prodenialis*. 20

2.4 GLMs predicting infestation by *C. cactorum* 21

2.5 Averaged GLMs predicting infestation by *C. cactorum* or *M. prodenialis*. 22

2.6 Consensus best model for predicting patterns of *M. prodenialis* infestation. 23

2.7 Consensus best model for predicting patterns of *C. cactorum* infestation. 24

3.1 List of 19 BIOCLIM predictor rasters. 45

3.2 List of *Opuntia* taxa modeled. 46

3.3 Dunn's test for multiple comparisons between PAB and Maxent-generated predictions. 47

3.4 Likelihood ratio tests of nested beta regression models. 48

4.1 List of *Opuntia* taxa included in this study. 75

4.2 *Opuntia* morphological traits. 76

4.3 Likelihood ratio tests and PERMANOVA of GLMs evaluating cladode assignment to clusters. 77

4.4 GLMs evaluating if trait-based clusters predict *C. cactorum* infestation. 78

4.5 List of tissue macronutrient-based *Opuntia* clusters modeled. 78

4.6 Dunn test for multiple comparisons between trait-based SDM, PAB, and Maxent predictions. 79

4.7 Results of likelihood ratio test comparing nested beta regression models. 81

LIST OF FIGURES

2.1 Map of *Opuntia* cladode collection sites. 25

2.2 Boxplots of *Opuntia* cladode tissue macronutrient content. 26

3.1 Map of 5214 *Opuntia* (open points) and 233 *C. cactorum* (black points) occurrence records. 48

3.2 Maxent prediction of the North American distribution of *C. cactorum*. 49

3.3 Mosaics of habitat suitability for *C. cactorum*...49

3.4 Boxplots comparing Schoener's D values associated with PAB and Maxent-generated predictions..50

4.1 Map of the southern United States indicating where *Opuntia* cladodes collection sites...81

4.2 Internal (filled black circles) and external (open circles) validation criteria for all number of partition scenarios......................................82

4.3 Map of the study region where 75 cladode samples were collected. 83

4.4 Mosaics of habitat suitability for *C. cactorum*...84

4.5 Boxplots comparing Schoener's D values associated with trait-based SDM, PAB, and Maxent predictions. 85

4.6 Boxplots comparing Schoener's D values associated with trait-based SDM, PAB, and Maxent predictions. 85

B.1 Maxent prediction for *O. humifusa* var. *ammophila*. 121

B.2 Maxent prediction for *O. atrispina*. 122

B.3 Maxent prediction for *O. austrina*. 122

B.4 Maxent prediction for *O. basilaris*..123

B.5 Maxent prediction for *O. camanchica*. 123

B.6 Maxent prediction for *O. ellisiana*. 124

B.7 Maxent prediction for *O. engelmannii* var. *engelmannii*. 124

B.8 Maxent prediction for *O. ficus-indica*. 125

B.9 Maxent prediction for *O. humifusa* var. *humifusa*. 125

B.10 Maxent prediction for *O. engelmannii* var. *lindheimeri*. 126

B.11 Maxent prediction for *O. engelmannii* var. *linguiformis*. 126

B.12 Maxent prediction for *O. macrocentra*. 127

B.13 Maxent prediction for *O. macrorhiza*. 127

B.14 Maxent prediction for *O. microdasys*. 128

B.15 Maxent prediction for *O. phaeacantha*. 128

B.16 Maxent prediction for *O. pottsii* 129

B.17 Maxent prediction for *O. pusilla*. 129

B.18 Maxent prediction for *O. santa-rita*. 130

B.19 Maxent prediction for *O. stricta*. 130

B.20 Maxent prediction for *O. strigil*. 131

B.21 Maxent prediction for *O. undulata*. 131

C.1 Boxplot comparing tissue macronutrient content of *Opuntia* tissues across 8 *Opuntia* taxa. 145

C.2 Dendogram depicting the relationships between *Opuntia* morphological and tissue macronutrient traits. 146

C.3 Maxent prediction for the first tissue macronutrient-based *Opuntia* cluster. 147

C.4 Maxent prediction for the second tissue macronutrient-based *Opuntia* cluster. 147

C.5 Maxent prediction for the third tissue macronutrient-based *Opuntia* cluster. 148

C.6 Maxent prediction for the fourth tissue macronutrient-based *Opuntia* cluster. 148

C.7 Maxent prediction for the fifth tissue macronutrient-based *Opuntia* cluster. 149

CHAPTER I

INTRODUCTION

Ecologists have long pursued the study of species' niches in an effort to understand the mechanisms governing the limits of species' distributions (Grinnell 1917, Elton 1927, Hutchinson 1957, MacArthur 1958, Connell 1961, Guisan and Zimmermann 2000**,** Chase and Liebhold 2003, Soberon and Peterson 2005, Pagel and Schurr 2012, Schurr et al. 2012). Chase and Liebhold (2003) define a species' niche as the set of environmental conditions that allow a species to persist (e.g., Grinnellian niche; Grinnell 1917), as well as the effect of the species on its local environment (e.g., Eltonian niche; Elton 1927). The concept of a species' niche has subsequently inspired many attempts to characterize species niches through correlations between large-scale abiotic conditions and a species' known occurrence (e.g., species distribution models, environmental niche models, invasive species distribution models). Yet, large-scale correlative techniques often neglect fine-scale biotic factors influencing a species' local population dynamics and persistence. As a result, much attention has focused on integrating large-scale correlative techniques with understanding of the fine-scale biotic factors that influence local patterns in species distributions (Peterson 2003, Soberon and Nakamura 2009, Brooks et al. 2012, Kissling et al. 2012, Wisz et al. 2013).

Methods of predicting species' distributions that reconcile large-scale correlative techniques with fine-scale understanding of biotic factors will need to address what

Simon Levin (1992) called the central problem in ecology – how do patterns observed at one scale emerge from patterns and processes operating at both larger and smaller scales? Levin argues that the solution will require "…study of how pattern and variability change with the scale of description, and the development of laws for simplification, aggregation, and scaling." The scale-dependent nature of linkages among large-scale environmental conditions and local biotic interactions present considerable difficulties to both overcoming Levin's problem and constructing mechanistic models to predict species' distributions.

Focusing on organism traits and how they influence the outcomes of biotic interactions may be useful in relating local biotic factors to broader scale patterns in species' niches and distributions. This is because individuals survive and reproduce based on the adaptive fit of their traits (Webb et al. 2010, Verberk et al. 2013) to local environmental selective pressures (McGill et al. 2006, Westoby and Wright 2006, Webb et al. 2010). However, trait-based approaches have generally been unsuccessful in characterizing the niches and informing predictions of the distributions of invasive species. This is because ecological niche shifts (Brooks et al. 2012) or rapid evolutionary changes in invader traits (Lee 2002, Sax et al. 2007, Keller and Taylor 2008) can prevent predictive success based on individual traits (Peterson 2003, Wiens and Graham 2005, Strubbe et al. 2013).

Effective distribution models rely on consistent relationships between a species' occurrence and selected predictor variables. In this respect, the conservation of invader traits may aid in the development of predictive distribution models. Yet, the extent to which an invader's traits are conserved or differ between its native and exotic ranges is often unknown. Factors that strongly restrict a species' niche, such as invader obligate specialization on hosts or resources, can limit the number of confounding factors and thereby simplify the construction of predictive models. This is especially the case when trait-based approaches are applied to biological invasions because the traits of specialist invaders that influence their biotic interactions with taxa in their native range are expected to closely correspond, or have tight ecological fit (Janzen 1985, Agosta 2006), with the traits of the taxa associated with in invaded communities.

The following dissertation uses a conceptual framework proposed by Catford et al. (2009) to characterize the interactions and predict the exotic distributions of an invasive consumer, the South American cactus moth (*Cactoblastis cactorum*), that specializes on its host plants, prickly-pear cacti (*Opuntia* spp. in tribe Opuntieae). This framework proposes that biological invasions are governed, and invader distributions constrained, by three main groups of factors: 1) the number of invaders introduced to and dispersing across the exotic range (propagule pressure); 2) invader physiological tolerances to abiotic conditions; and 3) invader gain or loss of biotic interactions (Catford et al. 2009). Each dissertation chapter deals with either an individual main factor or seeks to characterize the effect of all three main factors on *C. cactorum* distributions. As a result, this dissertation represents an initial step towards obtaining trait-based understanding of the biotic associations and North American distribution of *C. cactorum*.

The second chapter of this dissertation uses a trait-based approach to investigate how organism traits may influence invader gain or loss of biotic interactions. Chapter Two is motivated by the question: what traits of *Opuntia* influence *C. cactorum*-*Opuntia* associations? It is hypothesized that *Opuntia* morphological and tissue macronutrient traits correlate with patterns of *C. cactorum* host use. The morphological and tissue macronutrient traits of several *Opuntia* taxa are quantified and then evaluated for correlations with patterns of *C. cactorum* host use.

The third chapter of this dissertation focuses on characterizing the intersection of the three main factors that influence biological invasions and constrain invader distributions. Chapter Three is motivated by the question: how does incorporating propagule pressure, abiotic conditions, and biotic interactions affect predictions of the North American distribution of *C. cactorum*? It is hypothesized that predictions generated via the modeling method described in this chapter will differ from Maxent (Phillips et al. 2006) predictions generated using just environmental conditions. The method described in this dissertation chapter uses host *Opuntia* identity and availability to estimate scenarios of habitat suitability for *C. cactorum*. This method then simulates *C. cactorum* dispersal relative to scenarios of habitat suitability and *Opuntia* availability. The degree of similarity (or overlap) among predictions generated by the method described in this chapter, as well as the relative contribution of modeling constraints in generating variation in overlap among predictions, is evaluated.

The fourth chapter of this dissertation seeks to reconcile trait-based understanding of *C. cactorum-Opuntia* associations with methods for predicting invader distributions. Chapter Four is motivated by the question: how does incorporating trait-based understanding of *C. cactorum-Opuntia* associations influence predictions of the North American distribution of *C. cactorum*? It is hypothesized that predictions informed by trait-based understanding of *C. cactorum-Opuntia* associations would differ from predictions generated by Maxent and the modeling method described in Chapter Three.

The method described in Chapter Four uses trait-based *Opuntia* groupings to estimate habitat suitability for *C. cactorum.* Dispersal of *C. cactorum* is then simulated across tissue macronutrient-based maps of habitat suitability via a different method than used in Chapter Three. The degree of similarity (or overlap) among predictions generated by the method described in Chapter Four, as well as predictions generated in Chapter Three, is evaluated. Lastly, the relative contribution of modeling constraints in generating variation in overlap among Chapter Four predictions is evaluated.

CHAPTER II

NATIVE HOST TRAITS PREDICT PATTERNS OF INVADER HOST USE

A major goal of invasion ecology is to understand the factors that influence the geographic distributions of invaders in exotic ranges. Invader niches and distributions can be constrained at multiple scales by abiotic and biotic factors (Grinnell 1917, Elton 1927, Hutchinson 1957, MacArthur 1958, Chase and Leibold 2003). The traits possessed by an invader, as well as the traits of the species it associates with in the exotic community and landscape, can mediate the abiotic and biotic factors that constrain invader distributions.

Invaders are often introduced into communities that differ from those in the native range, and experience abiotic conditions that are novel (Torchin et al. 2003, Colautti et al. 2004, Jiménez-Valverde and Peterson 2011). The ability of the invaders to establish in such conditions has prompted discussion of whether invader traits are conserved or can rapidly evolve following displacement (Wiens and Graham 2005, Sax et al. 2007, Keller and Taylor 2008, Brooks et al. 2012, Petitpierre et al. 2012, Strubbe et al. 2013, Stigall 2014). Rapid evolutionary changes can alter both the invader's niche and its ability to successfully colonize poorer quality (relative to those occupied in the native range) habitats (Lockwood et al. 2005, Simberloff 2009, Guisan et al. 2014). Alternatively, invader niches and traits can be conserved following displacement (Wiens and Graham 2005, Wiens et al. 2010), and changes in occupied habitat may be explained by shifts in the abiotic conditions where suitable biotic associations are available (Brooks et al.

2012). Invader establishment and spread are challenging to predict because of the difficulties in distinguishing between instances of evolutionary change or ecological shifts (Brooks et al. 2012, Petitpierre et al. 2012, Strubbe et al. 2013, Stigall 2014). As a result, the extent to which invader niches and traits are conserved may influence the ability to predict invader distributions.

The gain or loss of biotic interactions between the invader and members of the invaded community is influenced by the traits of both the invader and native taxa (Sih et al. 2010, Guisan et al. 2014, Tingley et al. 2014). There is growing evidence that organism traits mediate intra- and inter-specific biotic interactions in native communities (Werner and Peacor 2003, Stang et al. 2006, Verberk et al. 2013, Leach et al. 2015). Both pieces of evidence are consistent with Darwin's naturalization hypothesis (Darwin 1859, Daehler 2001, Shea and Chesson 2002, Mitchell et al. 2006). Instances where environmental conditions are suitable, but there is low similarity in traits between invader native and exotic communities, may result in ecological shifts that cause an invader to occupy different habitats than those occupied in its native range (Brooks et al. 2012). As a result, focusing on the traits of taxa that associate with the invader in its native and exotic ranges may identify traits that mediate the invader's biotic interactions.

Systems where invaders specialize and form few obligate associations with their hosts or resources may facilitate the identification of traits important to invader occurrence and persistence. This is because few, direct obligate associations between an invader and its hosts or resources may reduce the number of biotic associations that directly influence the invader's occurrence and persistence. In this respect, the South American cactus moth, *Cactoblastis cactorum* (Lepidoptera: Pyralidae), is an excellent study organism for evaluating how invader establishment and spread can be influenced by host or resource traits. The natural history of this moth species is described elsewhere (Dodd 1940, Zimmerman et al. 2004), but its larvae often infest and feed on prickly-pear cacti (*Opuntia* spp.). Used as a biological control of pest *Opuntia* in Australia (Dodd 1940), South Africa (Pettey 1948), and the Caribbean (Simmonds and Bennet 1966), *C. cactorum* was unintentionally introduced to North America as early as the 1980's (Habeck and Bennet 1990, Dickel 1991). Following its introductions to North America (Simonsen et al. 2008), *C. cactorum* has since spread north along the Atlantic coast to Charleston, South Carolina (Hight et al. 2002) and west along the Gulf coast to parishes southwest of New Orleans, Louisiana (Rose 2009).

Propagule pressure likely played an enormous role in *C. cactorum* establishment and spread when it was used as a biological control in Australia (A.P. Dodd estimated in his notebooks that he released more than three billion eggs), but the factors driving the North American invasion of *C. cactorum* are less clear. Models generated using abiotic conditions from the native range of *C. cactorum* successfully recaptured its native distribution, but failed to accurately predict its North American distribution (Brooks et al. 2012). There is considerably less microsatellite diversity in North American *C. cactorum* populations than exists in the native range (Marsico et al. 2011), and environmental conditions at sites that were invaded in Australia, South Africa and North America were widely divergent relative to those in the native range (Brooks et al. 2012). Taken together, it is unlikely that there is sufficient genetic variation across exotic populations to allow for large evolutionary changes in the physiological tolerances of *C. cactorum* to conditions in Australia, South Africa and the southern United States (Marsico et al. 2011, Brooks et al. 2014). Instead, Brooks et al. (2012, 2014) suggest that the observed shift in the abiotic conditions occupied by *C. cactorum* in North America is likely driven by an ecological shift in host availability relative to abiotic conditions.

The identity and availability of *Opuntia* hosts appear to influence both the native and North American distributions of *C. cactorum* (Brooks et al. 2012, Sauby et al. 2012, Brooks et al. 2014). As a result, focusing on host *Opuntia* traits could be useful in determining suitable habitat for *C. cactorum*. A trait-based approach is adopted here for several reasons. First, previous research indicates host plant nutritional quality can affect larval lepidopteran survival and performance (Awmack and Leather 2002). In addition, *Opuntia* morphological traits appear to affect *C. cactorum* selection of and oviposition on *Opuntia* hosts (Myers et al. 1981, Robertson 1987, Jezorek et al. 2010, Sauby et al. 2012). Lastly, taxonomic relationships among *Opuntia* have been difficult to resolve as a result of extensive hybridization and polyploidy (see Majure et al. 2012a, 2012b). The lack of taxonomic resolution among *Opuntia* has complicated efforts to elucidate general patterns in *C. cactorum*-*Opuntia* associations relative to *Opuntia* taxonomic identity.

This investigation represents a step towards identifying *Opuntia* traits that mediate the biotic interactions thought to be responsible for shaping the realized niche of *C. cactorum* in North America. It is hypothesized that *Opuntia* morphological and tissue macronutrient (i.e., crude fiber, lipid, and crude protein) traits are correlated with patterns of *C. cactorum* host use. This work primarily focuses on *C. cactorum*, but the degree to which the traits of *C. cactorum* overlap with closely related taxa native to North America may influence the ability of *C. cactorum* to establish and spread in this range (naturalization hypothesis; Darwin 1859). As a result, the same suites of *Opuntia* traits were examined for correlations with patterns of host *Opuntia* use by a native cactus moth species, *Melitara prodenialis*. In general, our results indicate that *Opuntia* tissue macronutrient traits may be useful in predicting patterns of *C. cactorum*-*Opuntia* infestation.

Materials and Methods

Study area and cladode sampling procedures

Collections of *Opuntia* cladodes were made between spring 2012 to summer 2014 across 94 unique GPS locations in South Carolina and Florida (Figure 2.1). This geographic area includes much of the existing North American ranges of *M. prodenialis* and *C. cactorum*, as well as the ranges of multiple North American *Opuntia* taxa (Anderson 2001, Rebman and Pinkava 2001, Stuppy 2002, Hunt 2006, Majure et al. 2012a, 2012b). Sites were visited from August-October, November-February, and March-May in order to 1) capture spatiotemporal variation in *M. prodenialis* and *C. cactorum* infestation, 2) sample *Opuntia* taxa that persist throughout South Carolina and Florida, and 3) collect *Opuntia* cladodes exposed to heterogeneous environmental conditions that may influence tissue macronutrient content.

All sites were visited, and cladodes collected, in the morning or mid-afternoon. A time-constrained search of 1 person-hour was conducted at the start of each visit to a site in order to detect cactus moth infestation. During a search, cladodes with suspected cactus moth infestation were removed from the plant and set aside for further inspection. Cladodes with suspected cactus moth infestation were dissected after the time-constrained search to confirm infestation status and moth identity. Any *C. cactorum* found were collected and preserved in 90% ethanol. Uninfested cladodes were collected from both infested and uninfested plants after the time-constrained search and cladode dissection. No more than 3 terminal cladodes were collected from any individual *Opuntia* plant, but the total number of cladodes collected per site varied relative to the number of plants present. The GPS location for each cladode collection was recorded along with information on host *Opuntia* height and growth form.

Quantifying *Opuntia* morphological and tissue macronutrient traits

A total of 273 cladodes were collected from 148 plants across all locations from spring 2012 to summer 2014 (Table 2.1). Pictures were taken of each cladode collection at a resolution of no less than 3264 x 2448 pixels. Majure and Ervin (2007) and Majure et al. (2012b) were used to identify collected *Opuntia* cladodes.

Morphological traits were quantified for a total of 236 cladodes. Thirty-two of these were collected from *Opuntia* plants infested with *M. prodenialis* and 22 from plants infested with *C. cactorum*. Data on cladode morphology were collected in a laboratory setting (Table 2.2). Morphological traits included; spines per areole, spine color, cladode shape, plant growth form, plant height, spine shape, spine persistence, spine pattern, mean number of spines per areole, mean length of up to 10 spines, median length of up to 10 spines, and length of the longest spine. Measurements of spine length were made from digital photographs using ImageJ (Rasband 1997). Measurements in ImageJ were calibrated using the known length of a whiteboard included in each photograph. Morphological traits were chosen for their ease of measurement and because previous studies suggest they are correlated with infestation by *C. cactorum* (Myers et al. 1981, Robertson 1987, Jezorek et al. 2010, Sauby et al. 2012). Yet, it is possible that the morphological traits used here may not be the specific traits that female *C. cactorum* use to identify *Opuntia* hosts on which they can successfully oviposit and their larvae infest.

Proximate analysis was used to quantify the macronutrient (i.e., crude fiber, lipid, and crude protein) content of the remaining 37 cladodes. Of the cladodes used to quantify *Opuntia* tissue nutritional content, 4 and 10 cladodes were collected from *Opuntia* plants infested with *M. prodenialis* and *C. cactorum*, respectively. The internal tissue (chlorenchyma, vascular tissue and medullar parenchyma) of each cladode was removed, frozen at -80°C, and then freeze-dried at -45°C and 133 x 10^{-3} mbar Torr until sample mass was constant. Procedural guidelines from the Association of Official Analytic Chemists (AOAC; AOAC 2012) were followed to quantify the crude fiber, lipid, and crude protein content of these freeze-dried tissues. Crude fiber was measured by H_2SO_4 and NaOH extraction (AOAC Official Method 962.09), % lipid was measured by petroleum ether extraction (AOAC Official Method 920.39), and % crude protein was measured via nitrogen analysis (AOAC Official Method 984.13).

Data analysis

All data analyses were performed in the R statistical language, version 3.2.3 (Appendix A, R Development Core Team 2016). Model averaging and an information theoretic approach (Akaike's Information Criterion [AIC]; Burnham and Anderson 1998, 2004) were used to identify which *Opuntia* morphological or tissue macronutrient traits best predicted cactus moth infestation. Generalized linear models (GLMs) were generated and fit using either solely *Opuntia* morphological or tissue macronutrient traits. All GLMs considered infestation by either *M. prodenialis* or *C. cactorum* as the binomially-distributed response variable.

An exhaustive search of all main predictors, as well as one-way and two-way predictor interactions, was conducted for GLMs including solely *Opuntia* tissue macronutrient or morphological traits (package glmulti; Calcagno and de Mazancourt 2010). The fit of subsequent models was assessed using AIC. A subset of fitted GLMs containing solely *Opuntia* tissue macronutrient traits and within ΔAIC of 2 of the model with the lowest AIC value were used to generate averaged models predicting either *M. prodenialis* or *C. cactorum* infestation. Theses averaged models were also used to calculate model-averaged coefficient estimates, unconditional variance, predictor importance (number of models within ΔAIC of 2 including each predictor), and the 95% confidence interval around each model-averaged coefficient estimates. Fit of the averaged models predicting either *M. prodenialis* or *C. cactorum* infestation was assessed by calculating each model's coefficient of determination (R^2).

A genetic algorithm was used to explore potential combinations of individual predictors, as well as one-way and two-way predictor interactions, in GLMs including solely *Opuntia* morphological traits (package glmulti; Calcagno and de Mazancourt 2010). Genetic algorithms provide an efficient way to explore large candidate sets of models (Trevino and Falciani 2006, Orestes Cerdeira et al. 2009, Calcagno and de Mazancourt 2010) because the algorithm initially selects a combination of predictor variables (and their interactions) randomly and then proceeds through model fitting and selection by successively fitting better models (as measured by AIC). A total of 5 replicate runs of the genetic algorithm were conducted for *M. prodenialis* or *C. cactorum* infestation. Collinearity between predictor variables was evaluated prior to running the genetic algorithm, and interactions between highly correlated terms were explicitly excluded from consideration in each replicate run. A consensus best model (package glmulti; Calcagno and de Mazancourt 2010) was then generated for each cactus moth species. The consensus best models for *M. prodenialis* and *C. cactorum* were then evaluated to determine if any morphological traits, or the model itself, were correlated with infestation by their respective cactus moth species.

Results

Substantial variation in tissue crude fiber, crude protein, and lipid content was observed within and among *Opuntia* taxa included in this investigation (Figure 2.2). A total of 11 and 21 GLMs with equal predictive power were generated to predict *M. prodenialis* and *C. cactorum* infestation, respectively (Tables 2.3 and 2.4). The averaged models were positively correlated with *M. prodenialis* ($R^2 = 0.905$) and *C. cactorum* ($R^2 = 0.555$) infestation (Table 2.5). Crude protein and the interaction of crude fiber and crude protein were the most important variables in predicting *M. prodenialis* infestation (Table 2.5). Alternatively, crude fiber and the interaction of crude fiber and crude protein were most important to predicting infestation by *C. cactorum* (Table 2.5).

Cactus moth infestation was poorly predicted by models containing solely *Opuntia* morphological traits. The consensus best models for predicting *M. prodenialis* (Adjusted R-squared: -0.06567, $F_{18,217}$=0.03456, P = 1; Table 2.6) and *C. cactorum* (Adjusted R-squared: -0.0807, $F_{18,217}$=0.02212, P = 1; Table 2.7) infestation were not correlated with data on cactus moth infestation.

Discussion

The data suggest that infestation by cactus moths is correlated with, and can be predicted by, the nutritional content of *Opuntia* tissues. Specialist lepidopterans appear particularly sensitive to the nutrient content of their host plants and diets (Moore 1985, Genc and Nation 2004). The crude protein content of *Opuntia* tissues appears to be important for influencing *M. prodenialis* infestation. Increasing dietary crude protein content can increase larval lepidopteran nitrogen content but simultaneously decrease larval fat content (Karowe and Martin 1989). Lipid reserves acquired during larval feeding and growth can subsequently affect adult fecundity and fitness (Boggs and Freeman 2005, Colasurdo et al. 2009), but additional work is needed to verify this occurs with *M. prodenialis*. Alternatively, crude fiber content and the ratio of crude fiber to crude protein in *Opuntia* tissues appear important for influencing infestation by *C. cactorum*. Increasing tissue fiber content (and thus tissue toughness) can reduce tissue palatability and digestibility for herbivores (Grubb 1986, Hanley et al. 2007) as well as increase the time required to mechanically process plant tissues (Laca et al. 2001). In this system, neonate larval *C. cactorum* survival is negatively affected by increasing *Opuntia* epidermal toughness (Jezorek et al. 2010), and more neonate larvae are required to successfully penetrate and infest tougher *Opuntia* cladodes (Varone et al. 2012). In addition, the nutritional tradeoff between the crude fiber and sugar contents of plant tissues (Brokensha 1996) can suppress larval lepidopteran feeding (Beck 1956, Bartelt et al. 1990). Adding sugars to meridic diets can increase larval *C. cactorum* survival (Carpenter and Hight 2012), but the extent to which *C. cactorum* is affected by the nutritional tradeoff between dietary crude fiber and sugar is not understood.

Cactus moth infestation was poorly predicted by models containing solely *Opuntia* morphological traits. As a result, it is difficult to conclude if individual predictors, or interactions among predictors, significantly correlated with cactus moth infestation are indeed biologically important. Host *Opuntia* spine characteristics were significantly correlated with infestation by *M. prodenialis*, but the extent to which host *Opuntia* morphology influences plant choice and infestation by *M. prodenialis* is currently unknown. Results were also consistent with studies of *C. cactorum* that suggest *Opuntia* height or plant size (Myers et al. 1981, Robertson 1987, Sauby et al. 2012) and spine characteristics (Jezorek et al. 2010) may be important. Yet, the observed non-existent correlations between host *Opuntia* morphology and infestation by either cactus moth species suggest that previous relationships with traits other than plant size do not reflect a direct influence of morphological traits on host use. In this respect, plant size (height or biomass) may be better predictors of *C. cactorum* host use than traits of individual cladodes. Alternatively, the morphological traits used in this investigation may not include those traits that influence cactus moth, particularly *C. cactorum*, infestation success. It currently is impossible to distinguish between these explanations given the available data.

Invasive species can drive native populations extinct (Rose et al. 2005), influence the evolutionary pathways of native species (Mooney and Cleland 2001), and disrupt the function of the communities and ecosystems they invade (Dukes and Mooney 2004, Kenis et al. 2008, Gallardo et al. 2016). Catford et al. (2009) suggested that invasion success occurs at the intersection of three broad factors; 1) propagule pressure; 2) abiotic conditions; and 3) the gain and/or loss of biotic interactions. The importance of each factor in defining a specific invader's distribution is often unclear.

It has previously been hypothesized that novel biotic interactions are important in shaping the realized niche and North American distribution of *C. cactorum* (Torchin et al. 2003, Mitchell et al. 2006, Brooks et al. 2012, 2014). A potentially important interaction for *C. cactorum* may be the extent to which it competes with *M. prodenialis*. This expectation is consistent with Darwin's naturalization hypothesis (Darwin 1859). In this respect, *C. cactorum* and *M. prodenialis* may possess similar suites of traits that interface with *Opuntia* traits. Similarity in the traits of *M. prodenialis* and *C. cactorum* may subsequently affect competition between these cactus moth species. The degree to which competition occurs between these species is not well understood, and it is currently impossible to test this hypothesis given the data collected in this investigation.

The research presented here also suggests the potential for *Opuntia* tissue macronutrient concentrations to affect *C. cactorum-Opuntia* associations. Variation in tissue macronutrient content among *Opuntia* hosts, particularly in areas of greater *Opuntia* diversity, could affect the future spread of *C. cactorum*. But, the nutritional profiles of most North American *Opuntia* have not been characterized, and manipulative experiments have not been conducted. An analysis of host traits in Argentine *Opuntia* (those in the native range of *C. cactorum*) may allow for the detection of additional, trait-based *C. cactorum-Opuntia* associations. Trait-based clusters of Argentine and North American *Opuntia* correlated with *C. cactorum* infestation may then be compared to suggest *Opuntia* traits that may be mediating *C. cactorum-Opuntia* associations.

Table 2.1 List of cladodes included in this study.

Opuntia taxa	N	Morphological	*M. prodenialis*	*C. cactorum*	Tissue Macronutrient	*M. prodenialis*	*C. cactorum*
O. humifusa var. *ammophila*	19	13	0	1	6	0	3
O. humifusa var. *humifusa*	179	162	30	19	17	2	5
O. pusilla	23	18	1	0	5	0	0
O. stricta	52	43	1	2	9	2	2
	273	236	32	22	37	4	10

Table 2.2 Morphological traits quantified for *Opuntia* cladodes.

Trait	Type of Variable	Number of trait states	Trait states/units of measurement
Spines per areole	Categorical	19	0, 0 or 1, 0 to 2, 0 to 3, 0 to 4, 0 to 5, 1, 1 or 2, 1 to 3,1 to 4, 1 to 5, 1 to 6, 2, 2 or 3, 2 to 3, 2 to 4, 2 to 5, 3 to 5, 3 to 6
Spine color	Categorical	17	black to white tip, black to yellow tip, brown, brown-yellow, grey, grey and red, grey-yellow, none, purple, purple to white tip, red and white bands, red to white tip, red to yellow tip, white, yellow, yellow-white
Cladode shape	Categorical	4	elliptic, lanceolate, obovate, orbicular
Plant growth form	Categorical	2	erect, sprawling
Plant height	Categorical	2	< 1m, > 1m
Spine shape	Categorical	3	curved, none, straight
Spine persistence	Categorical	2	none, persistent
Spine pattern	Categorical	3	birds-foot, none, other
Mean spines per areole	Numerical	Discrete	# spines per areole
Mean length of up to 10 spines	Numerical	Continuous	cm
Median length of up to 10 spines	Numerical	Continuous	cm
Length of longest spine	Numerical	Continuous	cm

Table 2.3 GLMs predicting infestation by *M. prodenialis*.

Model	Model	AIC	Weight
1	Melitara ~ 1 + crude protein	27.256	0.169
2	Melitara ~ 1	27.348	0.162
3	Melitara ~ 1 + crude fiber:crude protein	27.930	0.121
4	Melitara ~ 1 + crude protein + crude fiber:crude protein	28.937	0.072
5	Melitara ~ 1 + crude protein + lipid:crude protein	28.949	0.073
6	Melitara ~ 1 + crude protein + crude fiber	28.949	0.073
7	Melitara ~ 1 + crude fiber:crude protein + lipid:crude fiber	29.028	0.070
8	Melitara ~ 1 + crude fiber	29.071	0.068
9	Melitara ~ 1 + lipid:crude protein	29.117	0.067
10	Melitara ~ 1 + crude protein + lipid	29.201	0.064
11	Melitara ~ 1 + crude protein + lipid:crude fiber	29.256	0.062

Table 2.4 GLMs predicting infestation by *C. cactorum*

Model #	Model	AIC	Weight
1	Cacto ~ 1 + crude fiber + lipid:crude protein	29.741	0.089
2	Cacto ~ 1 + crude fiber:crude protein + lipid:crude fiber	29.866	0.084
3	Cacto ~ 1 + crude fiber + lipid	29.881	0.083
4	Cacto ~ 1 + lipid + crude fiber:crude protein	30.043	0.077
5	Cacto ~ 1 + crude fiber + lipid:crude fiber	30.272	0.065
6	Cacto ~ 1 + crude fiber:crude protein	31.076	0.046
7	Cacto ~ 1 + crude protein + lipid + crude fiber:crude protein	31.099	0.045
8	Cacto ~ 1 + crude fiber + lipid + crude fiber:crude protein	31.186	0.432
9	Cacto ~ 1 + lipid:crude fiber	31.234	0.042
10	Cacto ~ 1 + lipid + crude fiber:crude protein + crude lipid:crude protein	31.290	0.041
11	Cacto ~ 1 + crude protein + crude fiber + lipid	31.398	0.039
12	Cacto ~ 1 + crude protein + crude fiber:crude protein	31.510	0.037
13	Cacto ~ 1 + crude protein + crude fiber:crude protein + lipid:crude protein	31.526	0.036
14	Cacto ~ 1 + crude protein + crude fiber:crude protein + lipid:crude fiber	31.540	0.036
15	Cacto ~ 1 + crude fiber	31.546	0.036
16	Cacto ~ 1 + crude fiber + lipid + lipid:crude protein	31.580	0.036
17	Cacto ~ 1 + crude fiber + crude fiber:crude protein + lipid:crude fiber	31.645	0.034
18	Cacto ~ 1 + crude fiber:crude protein + lipid:crude protein	31.701	0.033
19	Cacto ~ 1 + crude fiber + crude fiber:crude protein + lipid:crude protein	31.717	0.033
20	Cacto ~ 1 + crude fiber + lipid:crude protein + lipid:crude fiber	31.719	0.033
21	Cacto ~ 1 + protein + crude fiber + lipid:crude protein	31.740	0.0327

Table 2.5 Averaged GLMs predicting infestation by *C. cactorum* or *M. prodenialis*.

Moth species	Averaged model (R^2)	Model terms	Estimate	Uncond. Var.	No. models	Import.	+/- (alpha=0.05)
C. cactorum	CC ~ lipid + crude fiber + crude protein + crude fiber:lipid + crude protein:lipid + crude fiber:crude protein (0.555)	crude protein	0.078	0.046	6	0.226	0.435
		crude fiber:lipid	-0.013	0.0006	6	0.294	0.048
		crude protein:lipid	-0.010	0.0006	8	0.334	0.050
		lipid	-0.083	0.018	7	0.363	0.274
		crude fiber	-0.316	0.180	11	0.524	0.863
		crude fiber:crude protein	-0.080	0.011	12	0.545	0.211
		(Intercept)	2.486	2.681	21	1.0000	3.329
M. prodenialis	MePr ~ lipid + crude fiber + crude protein + crude fiber:lipid + crude protein:lipid + crude fiber:crude protein (0.905)	lipid	0.002	0.0001	1	0.064	0.022
		crude fiber:lipid	0.002	2.73E-05	2	0.132	0.011
		crude protein:lipid	0.001	2.67E-05	2	0.139	0.011
		crude fiber	-0.010	0.0008	2	0.141	0.059
		crude fiber:crude protein	-0.014	0.0008	3	0.263	0.057
		crude protein	-0.312	0.200	6	0.513	0.908
		(Intercept)	-0.676	2.709	11	1.0000	3.342

The coefficient of determination (R^2) is provided with each averaged model. Coefficient estimates, unconditional variance, number of models (out of the best-performing) including the predictor, predictor importance, and the 95% confidence interval were calculated for each predictor in each averaged model.

Table 2.6 Consensus best model for predicting patterns of *M. prodenialis* infestation.

Model (AIC)	Term	Estimate	Std. Error	Z value	Pr(>|z|)
Melitara ~ Spine pattern + Cladode color + Cladode shape + Plant height + Spine persistence + Mean spines per areole + Median length + Mean length of 10 spines + Longest spine length + Spine pattern:Spine persistence + Median length of 10 spines:Longest spine length + Mean length of 10 spines:Longest spine length + Plant height:Mean spines per areole (183.87)	(Intercept)	-34.43	2375.7	-0.014	0.99
	Spine pattern - other	1.9	6713.75	0	0.99
	Cladode color - green	15.64	1590.48	0.1	0.99
	Cladode color - yellow-green	-0.84	3520.66	0	0.99
	Cladode shape - obovate	17.19	1764.75	0.01	0.99
	Cladode shape - orbicular	-0.11	3024.21	0	0.99
	Plant height >1m	-3.22	1.75	-1.84	0.067
	Spine persistence - persistent	6.15	2.45	2.51	0.012
	Mean spines per areole	1.50	0.87	1.72	0.086
	Median length of 10 spines	5.09	9.91	0.51	0.61
	Mean length of 10 spines	-5.67	10.48	-0.51	0.59
	Longest spine length	-4.54	1.76	-2.59	0.0097
	Spine pattern - other:Spine persistence - persistent	-4.51	6713.75	-0.001	0.99
	Median length of 10 spines:Longest spine length	-1.63	3.76	-0.43	0.67
	Mean length of 10 spines:Longest spine length	3.26	3.81	0.86	0.39
	Plant height >1m:Mean spines per areole	3.72	4.21	0.89	0.38

Model fit: Adjusted R-squared: -0.06567, $F_{18,217} = 0.03456$, P = 1

Table 2.7 Consensus best model for predicting patterns of *C. cactorum* infestation.

Model (AIC)	Term	Estimate	Std. Error	Z value	Pr(>|z|)
Cacto ~ Spines per areole + Cladode color + Cladode shape + Plant height + Mean spines per areole + Median length of 10 spines + Mean length of 10 spines + Longest spine length + Median length of 10 spines:Longest spine length + Cladode shape:Mean spines per areole + Plant height:Median length of 10 spines (154.52)	(Intercept)	-17.70	2826.23	-0.01	1.00
	Spines per areole - 0 or 1	0.96	2.20	0.44	0.66
	Spines per areole - 0 to 2	-1.89	3.27	-0.58	0.56
	Spines per areole - 0 to 3	-15.33	2634.74	-0.01	1.00
	Spines per areole - 1 or 2	-10.69	6059.31	0.00	1.00
	Spines per areole - 1 to 3	1.27	10754.02	0.00	1.00
	Cladode color - green	17.41	2826.23	0.01	1.00
	Cladode color - yellow-green	-0.46	5968.55	0.00	1.00
	Cladode shape - obovate	-1.79	1.34	-1.34	0.18
	Cladode shape - orbicular	-24.93	4650.14	-0.01	1.00
	Plant height >1m	-1.54	1.47	-1.05	0.30
	Mean spines per areole	-94.02	25234.08	0.00	1.00
	Median length of 10 spines	11.96	7.31	1.64	0.10
	Mean length of 10 spines	-13.69	7.04	-1.95	0.05
	Longest spine length	-1.61	2.51	-0.64	0.52
	Median length of 10 spines:Longest spine length	1.49	0.76	1.95	0.05
	Cladode shape - obovate:Mean spines per areole	85.76	25234.08	0.00	1.00
	Cladode shape - orbicular:Mean spines per areole	240.87	45113.88	0.01	1.00
	Plant height >1m:Median length of 10 spines	6.44	3.00	2.15	0.03
	Model fit: Adjusted R-squared: -0.0807, $F_{18,217} = 0.02212$, P = 1				

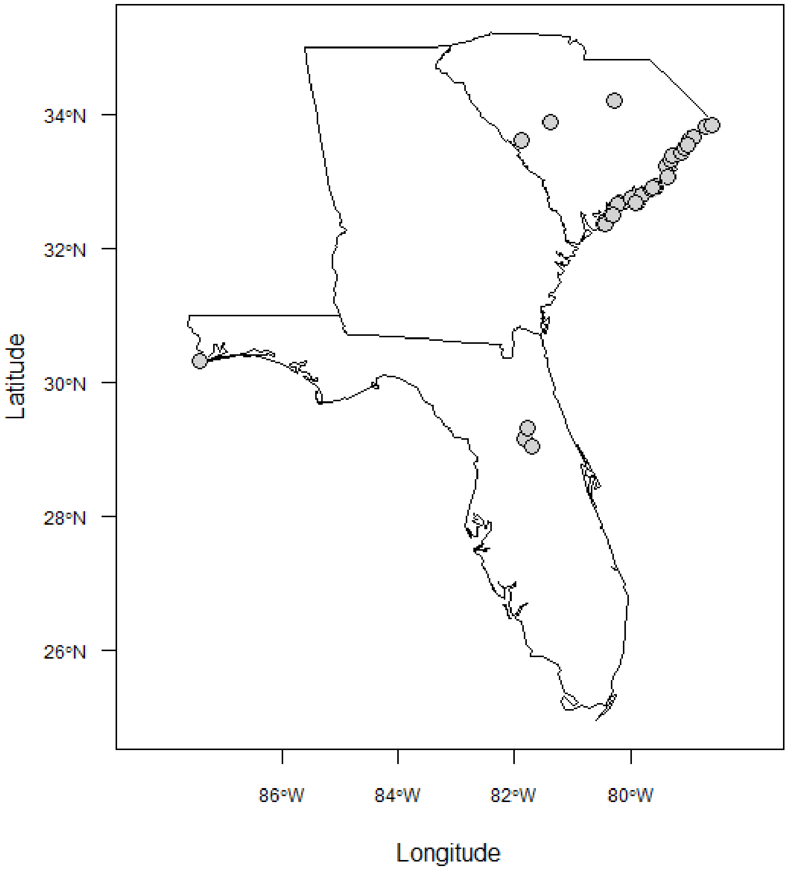

Figure 2.1 Map of *Opuntia* cladode collection sites.

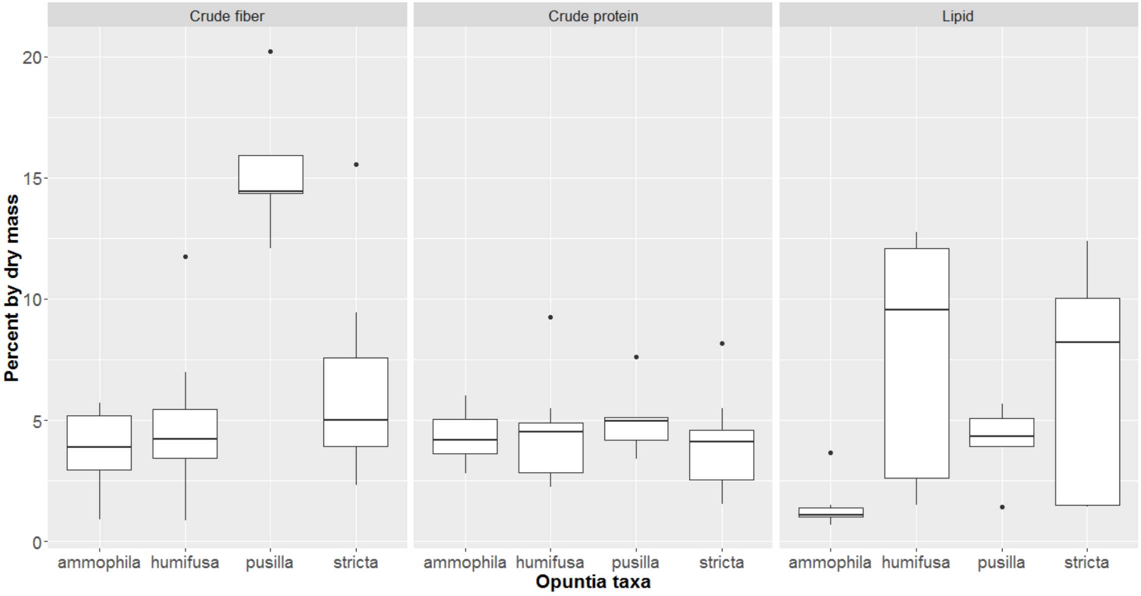

Figure 2.2 Boxplots of *Opuntia* cladode tissue macronutrient content.

CHAPTER III

EVALUATING THE RELATIVE IMPORTANCE OF FACTORS GOVERNING INVADER DISTRIBUTIONS

Biological invasions have increased in frequency throughout the 20th and 21st centuries (Simberloff et al. 2013). Consequently, an increasing number of biological invasions pose threats to communities and ecosystems worldwide. The threats posed by biological invasions to invaded communities and ecosystems make clear the need for both understanding of the factors that govern biological invasions and techniques for predicting invader distributions.

Catford et al. (2009) suggest that biological invasions are governed, and invader distributions ultimately shaped, by three broad groups of factors: 1) the number of invaders introduced to and dispersing across the exotic range (propagule pressure); 2) invader physiological tolerances to abiotic conditions; and 3) invader gain or loss of biotic interactions. Two approaches to characterizing how conditions in an exotic range influence invader distributions appear feasible. The first approach focuses on understanding the individual contributions of propagule pressure, abiotic conditions, or biotic interactions to shaping an invader's distribution. Considering only one broad factor governing invader distributions can simplify the generation and interpretation of model predictions (Kearney 2006, Morin and Lechowicz 2008, Paine 2010, Pigot and Tobias 2013). Yet, this approach can affect the extent to which methods of modeling species distributions can be extrapolated to new regions or different suites of constraining conditions (Beaumont et al. 2009, Brooks et al. 2012, Wang and Jackson 2014, Bradley et al. 2015). Alternatively, propagule pressure, abiotic conditions, and biotic interactions can be simultaneously considered when predicting an invader's distribution. Methods for modeling species distributions that adopt this approach are uncommon. However, considering all three major factors simultaneously may allow researchers to discern the relative importance of each factor in constraining invader distributions by evaluating variation among ensuing predictions.

Predictions of invader distributions that simultaneously consider propagule pressure, abiotic conditions, and biotic interactions may be most easily developed in systems where invaders form few obligate associations with hosts or resources. This is because obligate associations between an invader and its hosts or resources imply decreased complexity in the network of biotic interactions that directly influences the invader's occurrence and persistence. It follows that abiotic suitability for the invader in an exotic range can be estimated from locations where the invader has formed and maintains obligate biotic associations (Sih et al. 2010, Guisan et al. 2014, Tingley et al. 2014). Lastly, invader spread is influenced by invader ability to form and maintain obligate associations when host or resource availability is heterogeneous (Davis et al. 2000). Predictions of invader distributions should therefore consider both invader dispersal abilities and factors that influence the number of dispersing individuals (Lonsdale 1999, Williamson 1999).

The South American cactus moth, *Cactoblastis cactorum* (Lepidoptera: Pyralidae), is a well-known invader whose larvae are dependent on host plants in the tribe

Opuntieae for survival. Although successfully introduced as a biological control of pest prickly-pear cacti (*Opuntia* spp.) in Australia (Dodd 1940), introductions of *C. cactorum* into South Africa (Pettey 1948) and the Caribbean (Simmonds and Bennet 1966) led to its unintentional introduction to the United States. First detected in the Florida Keys in 1989 (Habeck and Bennet 1990, Dickel 1991), field surveys have determined *C. cactorum* has spread north to Charleston, South Carolina (Hight et al. 2002) and west to parishes west of New Orleans, Louisiana (Rose 2009, Rose et al. 2011).

The factors driving the North American invasion of *C. cactorum* are not clear. The North American distribution of *C. cactorum* appears to be relatively stable (Hight et al. 2002, Rose et al. 2009, 2011). Previous work using Maxent (Maxent; Phillips et al. 2006) and abiotic conditions from the native range of *C. cactorum* (Argentina) generated predictions that successfully recaptured the native, but not North American, distribution of *C. cactorum* (Brooks et al. 2012). Also, estimates of *C. cactorum* genetic diversity in North America indicate low diversity that may decrease the potential for evolutionary change in *C. cactorum* physiological tolerances to North American abiotic conditions (Marsico et al. 2011). Consequently, it is thought that the North American invasion of *C. cactorum* is being driven by biotic interactions.

Host *Opuntia* identity and availability appear to play important roles in conserving *C. cactorum* host preferences and forming novel *C. cactorum-Opuntia* associations (Dodd 1940, Marsico et al. 2011, Brooks et al. 2012, Sauby et al. 2012). It is estimated that approximately 30 *Opuntia* taxa occur in the United States, but this number is debatable due to issues with hybridization and polyploidy among *Opuntia* (Anderson 2001, Rebman and Pinkava 2001, Stuppy 2002, Hunt 2006, Majure et al. 2012a, 2012b).

Several North American *Opuntia* taxa are known hosts for *C. cactorum* (Johnson and Stiling 1998, Soberon et al. 2001, Sauby et al. 2012). As a result, concern has arisen over the potential for *C. cactorum* to invade regions of greater *Opuntia* diversity in the southwestern United States and Mexico (Soberon et al. 2001). Current understanding of North American *Opuntia* ranges indicates considerable variation in range size (Powell et al. 2008), but abiotic constraints on many (if not all) *Opuntia* ranges are not well understood. Knowledge of the geographic ranges of *Opuntia* hosts may provide information on *Opuntia* availability to *C. cactorum* that can be used to inform predictions of the North American distribution of *C. cactorum*.

This work describes the development of a method that incorporates propagule pressure, abiotic conditions, and biotic interactions to predict the North American distribution of *C. cactorum*. Predictions of species distributions are often appropriately called species distribution models (SDMs). The predictions generated by the method used in this investigation are referred to as PAB predictions to distinguish these predictions from other SDMs as a result of their inclusion of propagule pressure (P), abiotic conditions (A), and biotic interactions (B). It was hypothesized that PAB predictions would differ from predictions generated using Maxent (Phillips et al. 2006). It was expected that the geographic extent of PAB predictions would be more constrained than the geographic extent of Maxent-generated predictions. The degree of similarity (or in other words, overlap) between PAB and Maxent predictions, as well as the relative contribution of propagule pressure, abiotic suitability, and biotic interactions in generating overlap among PAB predictions, is quantified. Results indicate that abiotic suitability is more important than biotic interactions and propagule pressure in generating overlap among PAB predictions. The work described here provides a foundation to incorporate additional biological complexity in order to improve predictions of the North American distribution of *C. cactorum* and, more generally, to implement similar methods to characterize the distributions of other species of interest.

Materials and Methods

Study area and occurrence records

A total of 5214 *Opuntia* occurrence records from 4737 unique GPS locations, and 233 *C. cactorum* occurrence records from 227 unique GPS locations, were compiled for the current study (Figure 3.1). The region sampled spans from Arizona to South Carolina and includes the known North American distribution of *C. cactorum* and the ranges of many North American *Opuntia* taxa. Occurrence records for *C. cactorum* and the *Opuntia* taxa included in this study were collected during field study by members of the Brooks lab (2008, 2012-13) and supplemented with records from the Cactus Moth Detection and Monitoring Network (CMDMN, 1990 - 2012). GPS coordinates were collected for each *Opuntia* and *C. cactorum* occurrence record. Researchers associated with the CMDMN identified *Opuntia* to taxa when occurrence records were collected. For samples collected by members of the Brooks lab, Powell and Weedin (2004) and Powell et al. (2008) were used to identify *Opuntia* from Arizona, New Mexico, and Texas based on plant morphology. Majure and Ervin (2007) and Majure et al. (2012b) were used to identify *Opuntia* from South Carolina and Florida.

Generating Maxent predictions of *C. cactorum* and *Opuntia* taxa distributions

All model simulations and statistical analyses were conducted in the R statistical language (Appendix B, R Development Core Team 2016). Maxent (Phillips et al. 2006) was used to generate predictions of the North American distributions of *C. cactorum* and each *Opuntia* taxon for which occurrence records were compiled (package dismo; Hijmans 2013). The WorldClim database (http://www.worldclim.org; Hijmans et al. 2005) was used to obtain 19 BIOCLIM rasters at 30 arc-second (~1km^2) resolution to act as predictive environmental variables (Table 3.1). All *C. cactorum* or *Opuntia* taxa occurrence records for which associated BIOCLIM information could be extracted were used to generate each Maxent prediction. All BIOCLIM layers were utilized when generating Maxent predictions in order to produce the most conservative Maxent predictions possible.

The occurrence records for *C. cactorum* and each *Opuntia* taxon were partitioned into 5 groups via *k*-fold partitioning (package dismo; Hijmans 2013). A total of 10,000 background points were then randomly selected from within the geographic boundary of the study region for each taxon modeled. The occurrence records for *C. cactorum* or each *Opuntia* taxon, as well as the randomly selected background points generated for each taxon modeled, were used to train and test Maxent predictions. Only Maxent's logistic output for each taxon modeled was used in this investigation. All Maxent predictions were projected across the extent of the study region at a resolution of 30 arc-seconds.

Threshold-dependent metrics informed by the threshold value that maximized Maxent sensitivity and specificity were used to evaluate Maxent's ability to accurately predict abiotic suitability for each taxon modeled. This threshold value was used because it balances Maxent's ability to predict a species' presence and absence (Liu et al. 2005, 2013) and thus discern suitable from unsuitable habitat. The threshold-dependent performance metrics used here included: model omission rate (proportion of true occurrences misidentified by the defined threshold), sensitivity (proportion of actual positives identified as such), specificity (proportion of actual negatives identified as such), proportion of presence and absence points correctly identified, Cohen's kappa (Cohen 1960), and the true skill statistic (TSS; Allouche et al. 2006). Both Cohen's kappa and the TSS normalize overall model accuracy (the number of cells where an organism is correctly classified as present or absent) by model accuracy that might have occurred due to chance. Values of Cohen's kappa and the TSS can range from -1 to 1; a value of 1 indicates perfect agreement between model accuracy and accuracy expected due to chance whereas negative values indicate that the model's predictions are no better than random chance (Cohen 1960, Allouche et al. 2006). Threshold-independent metrics also were used to assess Maxent's performance; the threshold-independent metric used was the area under receiver operator curve (AUC).

Estimating habitat suitability for *C. cactorum*

Habitat suitability for *C. cactorum* was estimated by combining the Maxent predictions of abiotic suitability for *Opuntia* taxa relative to scenarios considering host *Opuntia* availability to *C. cactorum*. Scenarios of *Opuntia* availability were generated with consideration given to whether host *Opuntia* were known to be hosts for *C. cactorum*. The first scenario of host *Opuntia* availability used the Maxent predictions of eastern *Opuntia* taxa known to be hosts for *C. cactorum*, including; *O. humifusa* var. *ammophila*, *O. austrina*, *O. engelmannii* var. *engelmannii*, *O. ficus-indica*, *O. humifusa* var. *humifusa*, *O. engelmannii* var. *lindheimeri*, *O. macrorhiza, O. pusilla*, and *O. stricta* (Johnson and Stiling 1998, Soberon et al. 2001, Sauby et al. 2012). The second scenario of host *Opuntia* availability used the Maxent predictions of all *Opuntia* taxa modeled.

The Maxent predictions for each *Opuntia* taxon in each host availability scenario were stacked together (package raster; Hijmans et al. 2011). Three different estimates of habitat suitability were then calculated based on the mean, median, or maximum value of each raster cell across all stacked *Opuntia* predictions in both host *Opuntia* availability scenarios. In total, 6 habitat suitability maps (2 scenarios of host *Opuntia* availability, 3 estimates of habitat suitability based on abiotic suitability for *Opuntia*) were generated. All maps of habitat suitability for *C. cactorum* were projected across the same spatial extent as the taxon-level Maxent predictions and at a resolution of 30 arc-seconds.

Simulating *C. cactorum* dispersal

The R package MIGCLIM (Engler and Guisan 2009, Engler et al. 2012) was used to simulate *C. cactorum* dispersal relative to host *Opuntia* availability and estimates of habitat suitability for *C. cactorum*. Several dispersal scenarios, each with multiple constraints, were generated. The first dispersal scenario solely considered local dispersal. Local *C. cactorum* dispersal in North America is not well understood, so historical information regarding *C. cactorum* dispersal in Australia (Dodd 1940) was used to impose local dispersal constraints up to distances of 4km (5 raster cells). Variation in *C. cactorum* local dispersal abilities was generated by drawing 10 different dispersal kernels (labelled kernels 1-10) from exponential distributions over a distance of approximately 4km. Local dispersal kernels primarily differed in regards to the probabilities associated with very local dispersal (2km, up to 3 raster cells); for example, kernels 9 and 10 had greater probabilities of very local dispersal than all other dispersal kernels. In simulations involving solely local dispersal constraints, the total probability of dispersing across the 5 raster cells in each local dispersal kernel was re-scaled to sum to a value of 1.

The remaining dispersal scenarios considered both local and long-distance dispersal (LDD) constraints. In general, the abilities of *C. cactorum* to disperse over long distances in field conditions are not well understood, but Dodd (1940) anecdotally observed a female that dispersed at least 24km. This dispersal distance (24km, 30 raster cells), as well as a distance of 12km (15 raster cells; approximately the longest total distance flown by a male *C. cactorum* in laboratory settings; Sarvary et al. 2008a), were used as the maximum distances over which individual moths could potentially disperse. It is inherently difficult to quantify the probability of LDD for many organisms as LDD events can require specific conditions and are typically rare (Hengeveld 1994, Buchan and Padilla 1999). Few adult *C. cactorum* appear to disperse over distances greater than 2.5km (Hight et al. 2005, Sarvary et al. 2008a, 2008b), but the probability with which *C. cactorum* LDD occurs under field conditions is currently unknown. As a result, two scenarios were generated where *C. cactorum* LDD occurred with arbitrarily chosen probabilities of 0.01 or 0.001. Varying both the maximum distance and the probability of LDD occurring allowed for evaluation of the degree to which this SDM method was sensitive to modifications in dispersal parameters. Scenarios involving LDD also involved the process of local dispersal. Local dispersal kernels were therefore re-scaled so the total probability of local dispersal summed to 0.99 or 0.999.

C. cactorum is known to be multivoltine in warmer climates (Hight et al. 2003) and can have up to three dispersal flights per year in the southeastern United States

(Simonson et al. 2005). Based on this information, all dispersal simulations were run using a total of 300 dispersal steps, or the equivalent of 100 years. The minimum number of replicate simulations needed for each combination of abiotic suitability, host *Opuntia* availability, and dispersal constraints was determined by evaluating variation in the number of cells occupied across replicates. Scenarios including 10, 25, and 50 replicates were examined. The number of cells occupied across 10 replicates was substantially lower than the number of cells occupied across 25 or 50 replicates. As a result, 25 replicate simulations were generated for each combination of modeling constraints (i.e., *Opuntia* availability, habitat suitability values, and dispersal scenario).

Evaluating similarity among PAB and Maxent predictions

The two goals associated with this research were: 1) compare the degree of similarity (or overlap) among PAB and Maxent predictions; and 2) evaluate the relative contribution of propagule pressure, abiotic suitability, and biotic interactions in generating overlap among PAB predictions. Comparing the degree of overlap among PAB and Maxent predictions necessitated that all predictions were on the same value scale. Values of Maxent's logistic output range from 0 to 1 and represent Maxent's estimate of the probability that a species is present at a location given the environment at that location (Elith et al. 2011). PAB predictions were thus re-scaled by classifying all raster cells in each replicate prediction as occupied (value of 1) or unoccupied (value of 0) by *C. cactorum*. Replicates generated for each model scenario (i.e., *Opuntia* availability, habitat suitability values, and dispersal scenario) were then stacked and averaged. Raster cell values in final PAB predictions of the North American distribution of *C. cactorum* ranged from 0 to 1 and represented the proportion of replicates in which any particular cell was occupied by *C. cactorum* at the end of 300 dispersal steps.

Schoener's D (Schoener 1968) was used to evaluate the degree of overlap among predictions of the North American distribution of *C. cactorum* (package ENMeval; Muscarella et al. 2014). Schoener's D permits direct comparisons between predictions generated using different SDM methods (Warren et al. 2008). Values of Schoener's D range from 0, indicating predictions that are highly divergent and do not overlap, to 1, which indicates predictions that overlap exactly. A Kruskal-Wallis analysis of variance was used to evaluate variance in Schoener's D values among PAB and Maxent predictions relative to the modeling constraints of propagule pressure, abiotic conditions, and biotic interactions included in the simulation process. A Dunn's test (Dunn 1964) for multiple comparisons (package dunn.test; Dinno 2014) with a Bonferroni correction was then used to compare pairs of all predictions relative to the constraints of propagule pressure, abiotic conditions, and biotic interactions included in the simulation process.

The relative contribution of abiotic suitability, host *Opuntia* availability, and dispersal constraints in generating overlap among PAB predictions was also evaluated. Discrete probability distributions were fit to Schoener's D values from comparisons among all PAB predictions, and the fit of each probability distribution was assessed via Akaike's Information Criterion (AIC; Burnham and Anderson 2004). A beta distribution best fit all Schoener's D values. Nested beta regression models (package betareg; Cribari-Neto and Zeileis 2010) considering all modeling constraints simultaneously, as well as models including individual constraints, were fit. In all beta regression models, Schoener's D values were used as the response variable whereas abiotic suitability, host

Opuntia availability, and *C. cactorum* dispersal constraints were used as predictor variables. Likelihood ratio tests were used to assess the amount of variation in Schoener's D explained by each of the nested regression models. Predictor significance in each beta regression model was evaluated using a coefficient test (package lmtest; Zeileis and Hothorn 2002).

Results

A total of 21 Maxent predictions were generated for the *Opuntia* taxa included in this study (Table 3.2, Figures B.1-B.21). A Maxent prediction of the North American distribution of *C. cactorum* was also generated (Figure 3.2). The AUC value for the *C. cactorum* prediction was 0.979 and AUC values ranged from 0.786 to 0.997 for *Opuntia* predictions. The threshold that maximized Maxent sensitivity and specificity for *C. cactorum* was 0.28, and thresholds ranged from 0.025 to 0.55 for *Opuntia* predictions. Maxent omission rates never exceeded 21%. Cohen's kappa was 0.345 for the *C. cactorum* prediction and ranged in value from 0.018 to 0.958 for *Opuntia* predictions. Values of the TSS ranged from 0.355 to 0.994 for *Opuntia* predictions, whereas the TSS for the *C. cactorum* prediction was 0.957.

Habitat suitability for *C. cactorum* varied relative to its method of calculation and whether eastern or all *Opuntia* hosts were considered available to *C. cactorum* (Figure 3.3). In general, habitat suitability maps generated using eastern *Opuntia* hosts had more extreme ranges of suitability values than scenarios considering all *Opuntia* hosts. Habitat suitability values ranged from 3.07×10^{-7} to 0.663, 0.0018 to 0.486, and 0.0085 to 0.879 for scenarios considering the median, mean, and maximum values for eastern host

Opuntia. Median, mean, and maximum habitat suitability values ranged from 2.9 x 10^{-6} to 0.413, 0.0038 to 0.347, and 0.023 to 0.927 when all *Opuntia* taxa were considered.

Considerable variation in Schoener's D was observed when evaluating all predictions relative to how constraints on propagule pressure, abiotic conditions, and biotic interactions were imposed during the modeling process (Figure 3.4). A Kruskal-Wallis analysis of variance revealed substantial variation in Schoener's D values among all predictions relative to constraints of abiotic suitability ($\chi^2 = 5476.1$, df = 3, p-value = <0.001), host *Opuntia* availability ($\chi^2 = 2187.2$,df = 2,p-value = <0.001), local dispersal kernel ($\chi^2 = 339.41$,df = 10,p-value = <0.001), long-distance dispersal probability ($\chi^2 = 960.29$,df = 3,p-value = <0.001), and maximum distance dispersed during long-distance dispersal ($\chi^2 = 387.22$, df = 3,p-value = <0.001). When all predictions were compared to each other, it was observed that PAB predictions generally had greater median overlap, and greater variation in that overlap, with all other predictions than did the Maxent prediction. The Dunn's test for multiple comparisons indicated nearly all pairwise comparisons of all predictions were significantly different ($P < 0.05$) with respect to the modeling constraints of interest (Table 3.3). Pairwise comparisons also revealed predictions generated using dispersal kernels 9 and 10 differed significantly ($P < 0.05$) from predictions generated using kernels 1-8.

Likelihood ratio tests of nested beta regression models were used to explore the contribution of abiotic suitability, host *Opuntia* availability, and dispersal conditions to overlap among PAB predictions of the North American distribution of *C. cactorum* (Table 3.4). The full beta regression model that contained all predictor variables explained the greatest amount of variation in Schoener's D values among PAB predictions. How abiotic suitability was estimated, followed by host *Opuntia* availability to *C. cactorum*, accounted for the greatest amount of variation in Schoener's D values among PAB predictions (Table 3.4). The frequency of long-distance dispersal was the most influential dispersal constraint in generating variation in Schoener's D values among PAB predictions.

Discussion

This research focused on generating predictions (PAB predictions) of the North American distribution of *C. cactorum* that included the three major factors that shape invader distributions (propagule pressure, abiotic conditions, and biotic interactions). These PAB predictions were then compared to a prediction generated by Maxent and amongst themselves. The relative contribution of propagule pressure, abiotic conditions, and biotic interactions in generating overlap among PAB predictions was also evaluated. The hypothesis that PAB predictions would differ from predictions generated by Maxent was supported, but the expectation that PAB predictions would be more constrained than a Maxent prediction was not upheld. PAB predictions were generally broader in extent, and extended further inland, than the Maxent prediction. Assessing overlap among PAB and Maxent predictions revealed that PAB predictions generally had greater median overlap, as well as greater variation in overlap, among all predictions than did the Maxent prediction. Abiotic conditions, followed by biotic interactions, were most the influential modeling constraints in generating variation in overlap among PAB predictions. Methods that increase understanding of the influential factor(s) governing similarities among SDM predictions may enable hypothesis testing about specific mechanisms governing biological invasions. Understanding of the effects of specific mechanisms governing biological invasions may then be used to inform efforts to manage invader establishment and spread throughout an exotic range.

The manner in which abiotic suitability was calculated (i.e., maximum, mean, or median suitability) was the most important factor influencing the extent of, and consequently variation in overlap among, PAB predictions. This result is not entirely surprising as abiotic suitability can influence both invader spread across an exotic landscape and predictions of invader exotic distributions (Capinha et al. 2013, Muthukrishnan et al. 2015, Stewart-Koster et al. 2015). Scenarios of maximum suitability considered only one value (i.e., the maximum value of any particular *Opuntia* taxa present in a raster cell) whereas mean and median suitability were aggregate values that considered all *Opuntia* hosts present in a raster cell. Maximum suitability scenarios resulted in the furthest inland and westward spread of *C. cactorum.* Alternatively, mean or median suitability scenarios generated predictions that, to varying degrees, indicated more restricted distributions of *C. cactorum*. A mean suitability scenario, in particular, resulted in an intermediate degree of spread of *C. cactorum*. This intermediate degree of spread in turn resulted in PAB predictions with the most median overlap among all PAB predictions. Conversely, scenarios of median suitability generated more restricted PAB predictions that had the least median overlap with other PAB predictions.

Biotic interactions, specifically host *Opuntia* availability to *C. cactorum*, were the second-most influential factor in generating variation in overlap among PAB predictions. This result appears to corroborate previous work suggesting particular *Opuntia* taxa are important in influencing *C. cactorum* occurrence (Dodd 1940, Marsico et al. 2011, Brooks et al. 2012, Sauby et al. 2012). PAB predictions generated under a scenario of all

Opuntia taxa availability to *C. cactorum* exhibited a greater degree of westward and inland spread of *C. cactorum* than predictions generated under a scenario considering eastern *Opuntia* taxa availability. The observed discrepancies in *C. cactorum* spread between PAB predictions relative to scenarios of host *Opuntia* availability are likely attributable to eastern *Opuntia* taxa having range limits that more closely coincide with the current distribution of *C. cactorum*. Abiotic suitability west of the ranges of these eastern *Opuntia* taxa was less and thus restricted *C. cactorum* spread primarily to the southeastern US. Resource availability in the western portion of the study region also may have affected *C. cactorum* spread and ensuing predictions. Increasing native biodiversity can increase local scale resistance to vector-borne and fungal pathogens, as well as invasive species (Knops et al. 1999, Naeem et al. 2000, Kennedy et al. 2002, Fargione and Tilman 2005, Keesing et al. 2006). In the case of invasive species, extrinsic factors such as heterogeneity in resource availability can generate positive correlations between native biodiversity and invader establishment and spread at regional or community-level scales (Davis et al. 2000, Levine 2000, Byers and Noonburg 2003, Hooper et al. 2005). The western portion of the study region had more *Opuntia* diversity, but aside from the western *Opuntia* taxa included in a study by Jezorek et al. (2010), the suitability of most western *Opuntia* taxa as hosts for *C. cactorum* is not known.

Invader dispersal abilities influence whether an invader can capitalize on resource availability (Davis et al. 2000) or reach unoccupied, but suitable, habitat (Lonsdale 1999, Williamson 1999). The probability of long-distance dispersal was the most important dispersal constraint imposed in generating PAB predictions. Yet, the uncertainty associated with the frequency with which *C. cactorum* disperse long-distances highlights the need for additional research focused on quantifying *C. cactorum* LDD abilities. The local dispersal kernel used in scenarios considering solely local dispersal resulted in substantially different PAB predictions. This is especially true with regards to predictions generated using dispersal kernels 9 and 10 as these two kernels had greater probabilities of dispersing shorter distances (up to 3 raster cells), but lesser probabilities of dispersing longer distances (5 raster cells), relative to kernels 1-8. The results reported here thereby highlight the importance of the shape of a dispersal kernel in dictating local invader spread (Kot et al. 1996). The maximum distance which individuals could disperse was the least important constraint imposed. Predictions where propagules could disperse maximum distances of 12km and 24km did not significantly differ ($P > 0.05$) in their degree of overlap. However, these maximum distances may be conservative estimates as factors such as human-mediated dispersal (Wilson et al. 2009) or inclement weather (Andraca-Gomez et al. 2015) can increase invader dispersal distances. Additional research about *C. cactorum* dispersal under field-conditions in North American would be useful in refining future simulation attempts.

The geographic ranges of important prey or host plant species have been used in SDM methods to integrate biotic interactions and abiotic suitability in order to predict the distributions of organisms in their native communities (Trainor and Schmitz 2014, Trainor et al. 2014, de Araujo et al. 2015). Yet, similar approaches are less commonly used when applying SDM methods to biological invasions (but see Silva et al. 2014, da Silva Doge et al. 2015). Integrating abiotic suitability and biotic interactions in order to estimate habitat suitability for *C. cactorum* highlights a key point: if organism traits mediate biotic interactions (Werner and Peacor 2003, Stang et al. 2006, Verberk et al.

2013, Leach et al. 2015), then the presence or absence of taxa possessing specific suites of traits that mediate interactions with an invader may dictate habitat suitability for the invader. In this respect, scenarios of maximum habitat suitability may be an appropriate constraint when predicting *C. cactorum* distributions if *C. cactorum* perceives habitat suitability relative to the presence of particular *Opuntia* taxa that possess specific suites of traits. Previous work in this system suggests *C. cactorum* occurrence is influenced by the presence of particular *Opuntia* taxa (Dodd 1940, Marsico et al. 2011, Brooks et al. 2012, Sauby et al. 2012), so where these important *Opuntia* taxa occur may also constitute good habitat for *C. cactorum*. Alternatively, aggregate measures of habitat suitability (e.g., mean or median) may be more appropriate constraints if *C. cactorum* perceives habitat suitability as function of all *Opuntia* taxa and traits available.

The over-arching goal of this investigation was to develop an SDM method that integrates propagule pressure, abiotic suitability, and biotic interactions to predict the exotic distributions of invasive species. This work also sought to evaluate the relative importance of all modeling constraints on ensuing predictions. The results presented here indicate both goals were accomplished. However, the accuracy of any SDM predictions can be hindered by limited researcher understanding of invader perception of habitat suitability, which biotic interactions influence invader occurrence in an exotic range, and invader dispersal biology. Researcher lack of understanding about invader life history also hinders the interpretation of SDM predictions, especially with respect to identifying and quantifying what factors are influential in generating SDM predictions and potentially driving specific biological invasions. Future improvements to the method described here should focus on integrating trait-based approaches that characterize how invader traits mechanistically influence the major factors governing invasion.

Table 3.1 List of 19 BIOCLIM predictor rasters.

BIOCLIM Layer	Environmental variable
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality (BIO2/BIO7)*100
BIO4	Temperature Seasonality (standard deviation * 100)
BIO5	Maximum temperature of warmest month
BIO6	Minimum temperature of coldest month
BIO7	Temperature annual range (BIO5 - BIO6)
BIO8	Mean temperature of wettest quarter
BIO9	Mean temperature of driest quarter
BIO10	Mean temperature of warmest quarter
BIO11	Mean temperature of coldest quarter
BIO12	Annual precipitation
BIO13	Precipitation of wettest month
BIO14	Precipitation of driest month
BIO15	Precipitation seasonality (Coefficient of variation)
BIO16	Precipitation of wettest quarter
BIO17	Precipitation of driest quarter
BIO18	Precipitation of warmest quarter
BIO19	Precipitation of coldest quarter

Table 3.2 List of *Opuntia* taxa modeled.

Taxa group	Taxa modeled	#	Thresh	AUC	Omis.	Sens.	Spec.	Prop. correct	K	TSS
Eastern	*O. austrina*	69	0.55	0.993	0	1	0.986	0.987	0.457	0.986
	O. humifusa var. ammophila	122	0.415	0.997	0	1	0.994	0.994	0.79	0.994
	O. humifusa var. humifusa	742	0.33	0.875	0.121	0.88	0.872	0.872	0.381	0.752
	O. engelmannii var. engelmannii	2659	0.34	0.82	0.202	0.798	0.842	0.834	0.549	0.64
	O. engelmannii var. lindheimeri	137	0.43	0.856	0.066	0.934	0.977	0.976	0.509	0.911
	O. ficus-indica	242	0.28	0.908	0.098	0.902	0.914	0.912	0.26	0.816
	O. macrorhiza	74	0.31	0.889	0.054	0.946	0.831	0.832	0.063	0.777
	O. pusilla	197	0.18	0.969	0.01	0.99	0.949	0.949	0.259	0.939
	O. stricta	395	0.26	0.945	0.068	0.932	0.959	0.959	0.958	0.891
Western	*O. atrispina*	34	0.03	0.976	0	1	0.952	0.952	0.12	0.952
	O. basilaris	26	0.08	0.979	0	1	0.957	0.957	0.105	0.957
	O. camanchica	11	0.35	0.906	0.091	0.91	0.903	0.903	0.018	0.813
	O. ellisiana	59	0.28	0.786	0.14	0.86	0.711	0.712	0.022	0.571
	O. engelmannii var. linguiformis	74	0.4	0.854	0.192	0.808	0.901	0.9	0.093	0.709
	O. macrocentra	68	0.26	0.941	0.045	0.955	0.926	0.926	0.136	0.881
	O. microdasys	32	0.19	0.962	0	1	0.924	0.928	0.587	0.924
	O. phaeacantha	122	0.3	0.932	0.09	0.91	0.955	0.954	0.311	0.865
	O. pottsii	18	0.338	0.945	0	1	0.89	0.89	0.028	0.89
	O. santa-rita	30	0.025	0.978	0	1	0.956	0.956	0.115	0.956
	O. strigil	32	0.04	0.984	0	1	0.969	0.969	0.164	0.969
	O. undulata	29	0.22	0.891	0.154	0.846	0.937	0.937	0.06	0.783
Cacto.	*C. cactorum*	234	0.28	0.979	0.012	0.988	0.969	0.97	0.345	0.957

Thresh. is the threshold value that maximized Maxent sensitivity and specificity for each taxon. AUC (area under the receiver operator curve) is a threshold-independent metric. Omis. is the proportion of true occurrences misidentified by the defined threshold. Sens. is the proportion of actual positives identified as such. Spec. is the proportion of actual negatives identified as such. Prop. correct is the proportion of presence and absence points correctly identified. K is Cohen's kappa. TSS is the true skill statistic.

Table 3.3 Dunn’s test for multiple comparisons between PAB and Maxent-generated predictions.

Scenario	Comparisons	Z value	P value
Abiotic suitability	Maximum - Maxent	18.984	< 2.2E-16
	Maximum - Mean	-7.350	5.96E-13
	Maximum - Median	58.453	< 2.2E-16
	Mean - Maxent	-20.019	< 2.2E-16
	Median - Maxent	-10.757	< 2.2E-16
	Mean - Median	65.791	< 2.2E-16
Biotic interactions	All - Eastern	-43.706	< 2.2E-16
	All - Maxent	14.100	< 2.2E-16
	Eastern - Maxent	19.131	< 2.2E-16
LDD Probability	Local - 0.1%	0.458	1
	Local - 1%	-19.178	< 2.2E-16
	Local - Maxent	15.359	< 2.2E-16
	0.1% - 1%	-24.047	< 2.2E-16
	0.1% - Maxent	15.350	< 2.2E-16
	1% - Maxent	18.443	< 2.2E-16
LDD Maximum Distance	12km - 24km	-2.284	0.067
	12km - Local	8.423	< 2.2E-16
	12km - Maxent	16.749	< 2.2E-16
	24km - Local	10.288	< 2.2E-16
	24km - Maxent	17.043	< 2.2E-16
	Local - Maxent	15.359	< 2.2E-16

Reported P values are exact values calculated from the Dunn’s test.

Table 3.4 Likelihood ratio tests of nested beta regression models.

Model	df	LogLik	Chisq	Pr (>Chisq)
Suit. + Host + LDD Prob. + Disp. method + Disp. kernel	9	7994.9		
Suitability	4	6506.8	2976.3	< 0.001
Host	3	5240.9	2531.8	< 0.001
LDD Probability	3	4632.3	1217.1	< 0.001
Disp. method	4	4343.8	577.1	< 0.001
Disp. kernel	3	4318.8	49.925	< 0.001

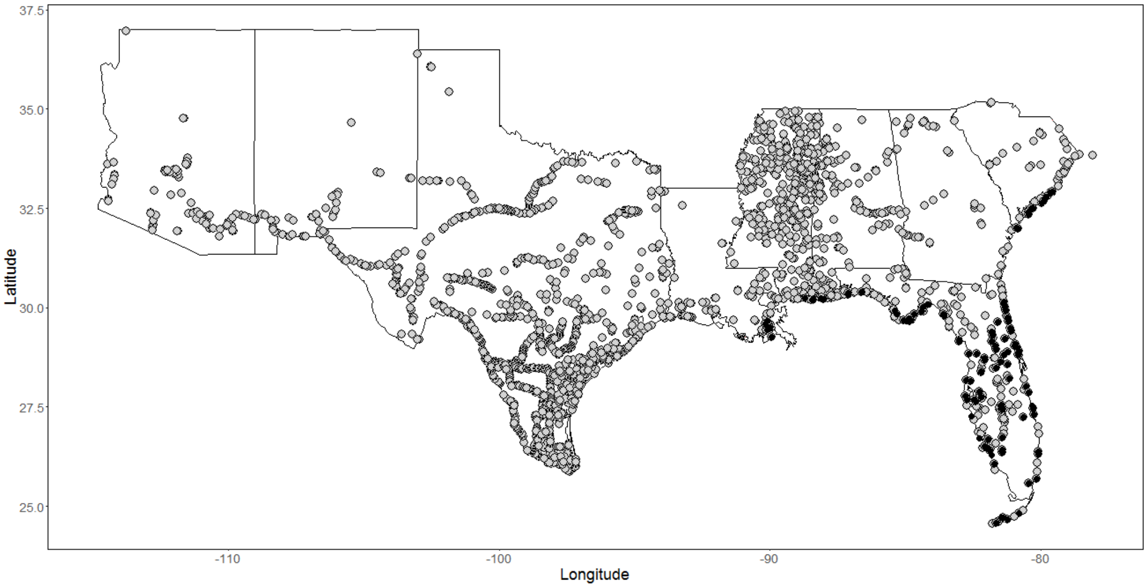

Figure 3.1 Map of 5214 *Opuntia* (open points) and 233 *C. cactorum* (black points) occurrence records.

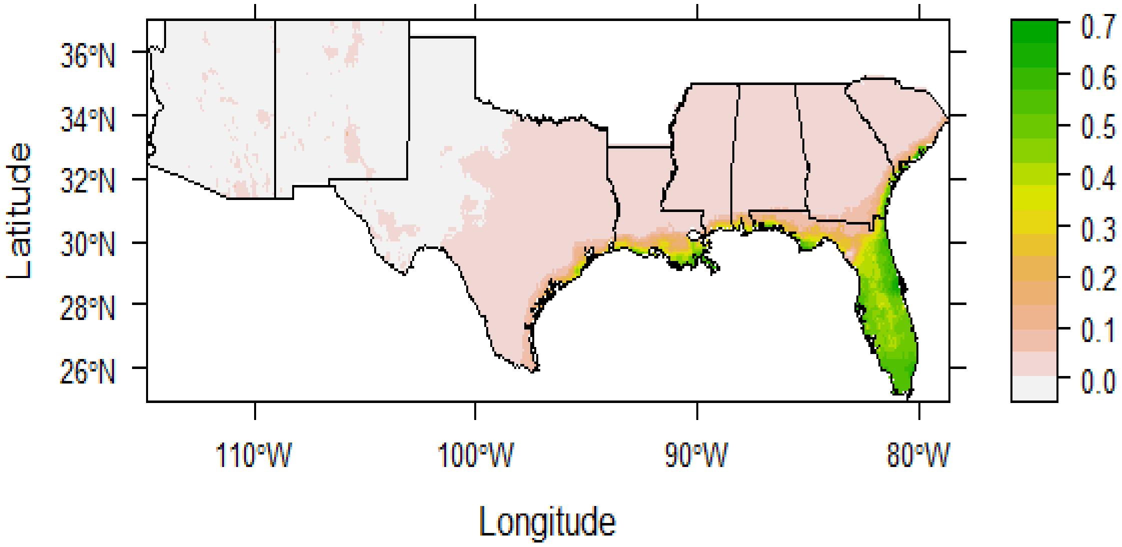

Figure 3.2 Maxent prediction of the North American distribution of *C. cactorum*.

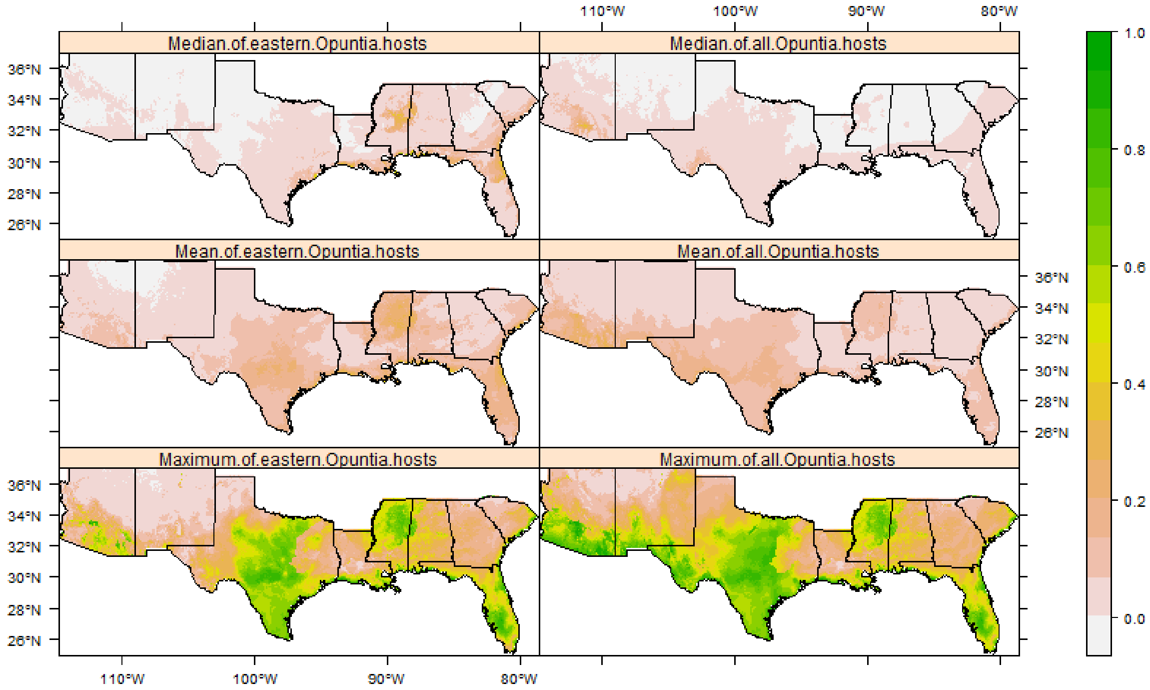

Figure 3.3 Mosaics of habitat suitability for *C. cactorum*.

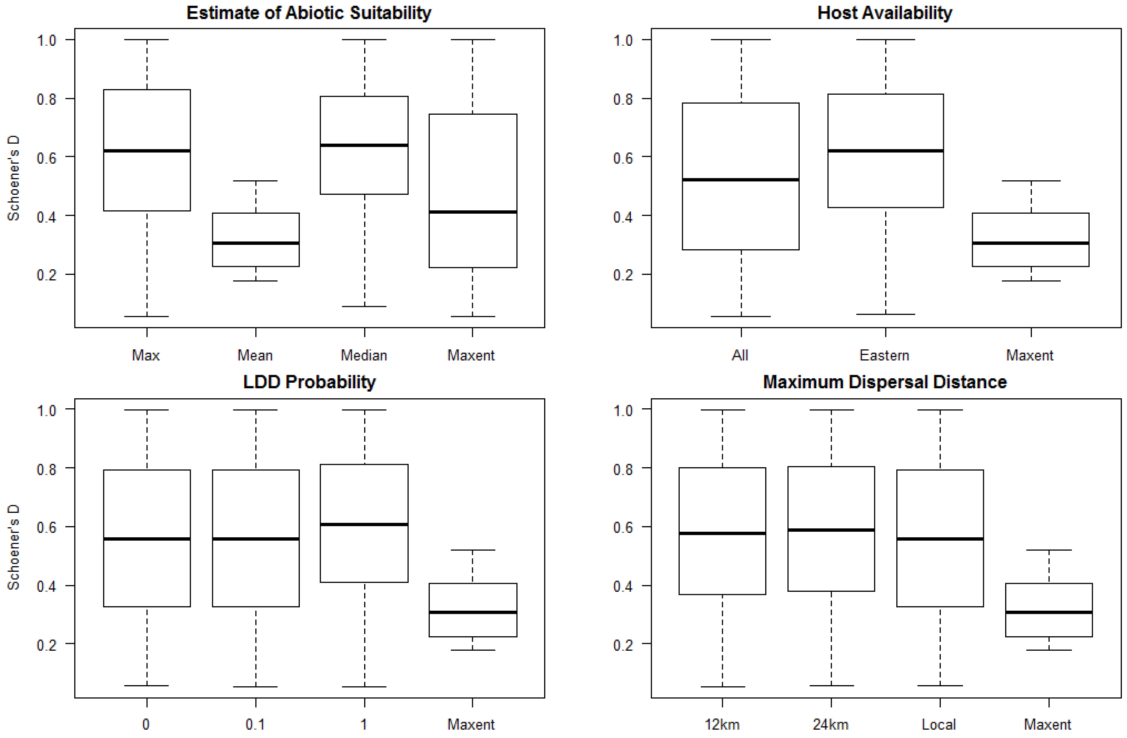

Figure 3.4 Boxplots comparing Schoener's D values associated with PAB and Maxent-generated predictions.

CHAPTER IV

A TRAIT-BASED METHOD FOR PREDICTING THE EXOTIC DISTRIBUTIONS OF INVASIVE SPECIES

Introduction

Trait-based approaches offer considerable promise to identify mechanisms influencing patterns of community assemblage and structure (Keddy 1992, Poff 1997, Lavorel and Garnier 2002, McGill et al. 2006, Verberk et al. 2013). This is because trait-based approaches are founded on the premise that individuals are selected for by local environmental filters relative to the individual's collection of traits (Poff 1997, McGill et al. 2006, Violle et al. 2007, Webb et al. 2010). Surviving individuals subsequently express phenotypes that vary predictably relative to local environmental conditions. It follows that populations separated by large spatial scales, but persisting under similar local environmental conditions, may exhibit similar population-level distributions of phenotypes (Swenson and Enquist 2007, Cornwell and Ackerly 2009, Whitehead et al. 2011). Trait-based approaches may therefore aid in predicting how organism traits affect species occurrences, and thus patterns in species distributions.

Much interest has focused on developing methods to predict the exotic distributions of invasive species (Elith and Leathwick 2009, Vaclavik and Meetemeyer 2009, Gallien et al. 2010, Jones et al. 2010, Jarnevich and Reynolds 2011, Gallien et al. 2012, 2015). Understanding the factors that govern biological invasions is critical for informing methods of developing species distribution models (SDMs). Catford et al. (2009) distilled the factors governing biological invasions into three broad groupings: 1) the number of invaders arriving and dispersing in an exotic range (propagule pressure); 2) invader physiological tolerance to abiotic conditions; and 3) invader gain or loss of biotic interactions in an exotic range. The intersection of these three major groupings defines the conditions that constrain invader distributions in an exotic range (Soberon 2007, Catford et al. 2009). Invader traits (e.g., tolerance to cold, feeding adaptations), or traits of taxa present in invaded communities, can mediate the abiotic and biotic factors that constrain the invader's exotic distributions (Lamouroux et al. 2002, Leps et al. 2006, McGill et al. 2006, Webb et al. 2010, Verberk et al. 2013). Predictions of invader exotic distributions may thus be informed by integrating the traits of the invader or taxa in the invaded community.

Organism traits are not commonly integrated into SDM methods, much less SDM methods applied to biological invasions. Focusing on systems where invaders form a limited number of obligate associations with hosts or resources may benefit the initial development of trait-based SDM methods. This is because invader establishment and persistence is directly linked to invader ability to form associations with hosts or resources in newly-invaded communities that act as surrogates for interactions in the invader's native range. A limited number of obligate interactions between invader and hosts or resources may subsequently aid in identifying specific suites of traits that govern invader associations, and constrain invader distributions, in exotic ranges.

The South American cactus moth, *Cactoblastis cactorum*, is an invasive consumer whose larvae are dependent on their host plants, prickly pear cacti (*Opuntia* spp.).

Collections of *C. cactorum* made in Argentina in early 1926 were successfully introduced into Australia as a means of biological control for pest prickly-pear. Subsequent introductions of *C. cactorum* to South Africa (Pettey 1948) and the Caribbean (Simmonds and Bennett 1966) to control additional pest *Opuntia* species ultimately led to its unintentional introduction into the Florida Keys (Habeck and Bennett 1990, Dickel 1991). Field surveys have since determined *C. cactorum* has spread north to Charleston, South Carolina (Hight et al. 2002) and west to parishes west of New Orleans, Louisiana (Rose 2009, Rose et al. 2011).

Host *Opuntia* identity and availability appear to play important roles in conserving *C. cactorum* host preferences, driving novel associations, and governing *C. cactorum* occurrence in North America (Dodd 1940, Marsico et al. 2011, Brooks et al. 2012, Sauby et al. 2012). It is estimated that approximately 30 *Opuntia* taxa occur in the United States (Anderson 2001, Rebman and Pinkava 2001, Stuppy 2002, Hunt 2006). The potential for most North American *Opuntia* taxa to be used as hosts by *C. cactorum* is not well understood (but see Jezorek et al. 2010), especially in regards to *Opuntia* taxa found in the southwestern US. Infestation of *Opuntia* by *C. cactorum* was found to be correlated with both the tissue macronutrient content of *Opuntia* tissues (Chapter Two). Yet, additional work is needed to identify *C. cactorum* traits that interact with *Opuntia* tissue macronutrient traits so as to develop a mechanistic understanding of *C. cactorum*-*Opuntia* associations.

This research focuses on developing a SDM method that predicts the North American distribution of *C. cactorum*. This method differs from the method described in Chapter Three in that habitat suitability for *C. cactorum* is estimated relative to the availability of trait-based *Opuntia* clusters and additional biological complexity is incorporated when simulating *C. cactorum* dispersal. It was hypothesized that predictions generated via the trait-based SDM method described here would differ from predictions generated by Maxent and a previously described SDM method (Chapter Three). Trait-based SDM predictions were expected to be more constrained than all (PAB and Maxent) predictions generated in Chapter Three. The hypothesis and prediction of this research is addressed by quantifying the degree of similarity (or in other terms, overlap) among trait-based SDM, PAB, and Maxent predictions. The relative importance and contribution of abiotic suitability and dispersal in generating overlap among trait-based SDM predictions is also quantified. Trait-based SDM predictions of the North American distribution of *C. cactorum* were considerably more constrained than PAB predictions generated in Chapter Three. Yet, trait-based SDM predictions including 3 or 4 dispersal events did not differ from Maxent-generated predictions. The number of dispersal events included in the modeling process was the most important factor influencing the overlap among trait-based SDM predictions. Results from this investigation indicate trait-based SDM methods can generate predictions of invader exotic distributions that are consistent with predictions generated by methods that consider invader distributions to be constrained solely by environmental conditions. The importance of the number of dispersal events as a modeling constraint for this trait-based SDM method highlights a need for accurate information regarding invader dispersal in exotic ranges.

Materials and Methods

Study area and cladode sampling procedures

Collections of *Opuntia* cladodes were made between spring 2012 to summer 2014 at 78 geographically distinct sites across the southern and southwestern United States, particularly an area spanning from Arizona to South Carolina (Figure 4.1). This geographic area includes the existing North American range of *C. cactorum*, as well as the ranges of multiple North American *Opuntia* taxa (Anderson 2001, Rebman and Pinkava 2001, Stuppy 2002, Hunt 2006, Majure et al. 2012a, 2012b). Sites were visited from August-October, November-February, and March-May in order to 1) capture spatiotemporal variation in *C. cactorum* infestation, 2) sample *Opuntia* taxa that persist throughout South Carolina and Florida, and 3) collect *Opuntia* cladodes exposed to heterogeneous environmental conditions that may influence tissue macronutrient content.

All sites were visited, and all cladodes collected, in the morning or mid-afternoon. A time-constrained search of 1 person-hour was conducted at the start of each visit to a site in order to detect cactus moth infestation. During a search, cladodes with suspected cactus moth infestation were removed from the plant and set aside for further inspection. Cladodes with suspected cactus moth infestation were dissected after the time-constrained search to confirm infestation status and moth identity. Any *C. cactorum* found were collected and preserved in 90% ethanol. Uninfested cladodes were collected from both infested and uninfested plants after the time-constrained search and cladode dissection. No more than 3 terminal cladodes were collected from any individual *Opuntia* plant, but the total number of cladodes collected per site varied relative to the number of plants present. The GPS location for each cladode collection was recorded along with information on host *Opuntia* height and growth form. All cladodes were collected during early morning to mid-afternoon hours.

A total of 721 cladodes were collected from 364 plants across all sampling sites. Six of these cladodes were excluded from this investigation because they had begun to rot prior to returning to the authors' lab. Pictures were taken of each cladode collection at a resolution of no less than 3264 x 2448 pixels. Powell and Weedin (2004) and Powell et al. (2008) were used to identify *Opuntia* from Arizona, New Mexico, and Texas based on plant morphology. Majure and Ervin (2007) and Majure et al. (2012b) were used to identify *Opuntia* from South Carolina and Florida.

Quantifying *Opuntia* morphological and tissue macronutrient traits

A total of 715 cladodes were utilized in this study (Table 4.1). Both morphological and tissue macronutrient traits of host plants were quantified for 21 of these 715 cladodes. A total of 7 of these 21 cladodes were collected from *Opuntia* hosts infested with *C. cactorum*. Solely morphological traits were quantified for a total of 640 cladodes of which 22 cladodes were collected from *Opuntia* plants infested with *C. cactorum*. Solely tissue macronutrient traits were quantified for a total of 54 cladodes of which 3 cladodes were collected from *Opuntia* plants infested with *C. cactorum*.

All data on cladode morphology were collected in a laboratory setting (Table 4.2). Morphological traits included; spines per areole, spine color, cladode shape, plant growth form, plant height, spine shape, spine persistence, spine pattern, mean number of spines per areole, mean length of up to 10 spines, median length of up to 10 spines, and length of the longest spine. Measurements of spine length were made from digital photographs using ImageJ (Rasband 1997). Measurements in ImageJ were calibrated using the known length of a whiteboard included in each photograph. Morphological traits were chosen for their ease of measurement and because previous studies suggest their importance to *C. cactorum* (Myers et al. 1981, Robertson 1987, Jezorek et al. 2010, Sauby et al. 2012). Yet, it is possible that the morphological traits used here may not be the specific traits that female *C. cactorum* utilize to identify *Opuntia* hosts on which they can successfully oviposit and their larvae infest.

Proximate analysis was used to quantify the macronutrient (i.e., crude fiber, lipid, and crude protein) content of a total of 75 cladodes (54 with just tissue macronutrient traits, 21 with both morphological and tissue macronutrient traits) collected from 73 plants (Figure C.1). The internal tissue (chlorenchyma, vascular tissue and medullar parenchyma) of each cladode was removed, frozen at -80ºC, and then freeze-dried at -45ºC and 133 x 10^{-3} mbar Torr until sample mass was constant. Procedural guidelines from the Association of Official Analytic Chemists (AOAC; AOAC, 2012) were followed to quantify the crude fiber, lipid, and crude protein content of these freeze-dried tissues. Crude fiber was measured by H_2SO_4 and NaOH extraction (AOAC Official Method 962.09), % lipid was measured by petroleum ether extraction (AOAC Official Method 920.39), and % crude protein was measured via nitrogen analysis (AOAC Official Method 984.13).

Generating trait-based *Opuntia* clusters

Hierarchical agglomerative clustering methods were used to determine clusters of *Opuntia* traits related to the likelihood of *C. cactorum* infestation. Clustering was determined using only the 21 cladodes for which both morphological and macronutrient data were available (n = 15 traits [12 morphological, 3 tissue macronutrient]) because data on tissue macronutrients were not available for many of the cladodes. All data analyses were performed in the R statistical language, version 3.2.3 (R Development Core Team 2016).

Bootstrapping was used to generate 100 dendograms representing the relationships between the morphological and tissue macronutrient traits of these 21 cladodes (Figure C.2; package ClustofVar; Chavent et al. 2012). These 100 dendrograms were then split into scenarios of 2 to *n*-1 partitions. A mean adjusted Rand criterion (Rand 1971) was used to evaluate agreements and disagreements in cladode assignment to partition across all 100 dendrograms in each scenario of 2 to *n*-1 partitions. Values of the mean adjusted Rand criterion range from 0 (indicating complete disagreement across dendrograms in cladode assignment to partitions) to 1 (indicating complete agreement across dendrograms in cladode assignment to partitions; Meila 2005, 2007). Transformations of the distributions of the mean adjusted Rand criterion in scenarios of 2 to *n*-1 partitions to meet parametric test assumptions of normality were unsuccessful. A Kruskal-Wallis test with a post-hoc Dunn correction for multiple comparisons (Dunn 1964, package PMCMR; Pohlert 2014) was used to evaluate statistical differences between the mean adjusted Rand criterion of our 2 to *n*-1 partitions. The mean adjusted Rand criterion values for scenarios of 5 to 12 partitions were not statistically different ($P > 0.05$), but were significantly greater ($P < 0.05$) than Rand criterion values in scenarios of 2 to 4 or 13 and 14 partitions. Cladode observations were thus split into groupings (hereafter referred to as clusters) using scenarios of 5 to 12 partitions.

Traits of the 21 samples with both morphological and tissue macronutrient data were used to calculate centroids for each trait relative to the number of partitions (5 to 12) being considered. The Euclidean distance between each cladode sample's value of a specific trait and the 5 to 12 centroids generated for each trait was calculated (package analogue; Simpson et al. 2015). Distances of each cladode sample's individual traits to each trait centroid were summed across the 15 traits considered. This summed Euclidean distance was then standardized by the number of traits associated with each sample (*n* = 15) because Euclidean distance increases with the number of variables or dimensions considered. The minimum standardized Euclidean distance of each cladode sample was used to assign each of the 21 cladodes to a cluster in scenarios of 5 to 12 partitions. This process was then repeated with the 54 cladodes associated with just tissue macronutrient traits and 640 cladodes associated with just morphological traits. The summed Euclidean distances of these 54 and 640 cladode samples were also standardized relative to the number of traits associated with these cladodes (3 and 12, respectively). The minimum standardized Euclidean distance of each cladode sample was then used to assign the 54 tissue macronutrient and 640 morphological cladode samples to clusters in scenarios of 5 to 12 partitions.

Cluster structure in each number of partitions scenario was assessed by first generating dissimilarity matrices of the Euclidean distances between cladodes and then calculating internal (within-cluster sum of squares) and external (entropy) validation criterion (package fpc; Hennig 2010). A small within-cluster sum of squares value indicates clusters are more compact and have low variation among observations contained within clusters. Entropy also evaluates the amount of disorder (i.e., variation) inherent among observations within clusters. An entropy value of 0 indicates no disorder, so lower values indicate less variation among observations and thus stronger cluster definition. Permutational multivariate analysis of variance (PERMANOVA, package vegan; Oksanen et al. 2007) was used to partition variance between clusters (McArdle and Anderson 2001) and evaluate if cladode assignment to cluster predicted dissimilarity between cladode observations.

The contribution of individual morphological and nutritional *Opuntia* traits to assigning samples to clusters was evaluated by considering the fit of several nested GLMs. Models containing all traits, only morphological, and only tissue macronutrient traits were fit for each possible number of partitions. Likelihood ratio tests and ANOVA were used to compare model fit between competing models. To test whether trait-based clusters were predictive of infestation, GLMs were generated where the proportion of hosts within each cluster served as the binomial response and the identity of clusters was used as a predictor. Models of best fit were then analyzed via ANOVA to determine if particular clusters were correlated with cactus moth infestation. For all GLMs, Akaike's Information Criterion (AIC; Burnham and Anderson 2004) was used to evaluate model fit relative to the number of clusters used as a predictor variable. Models within 2 AIC values of each other were considered equally valid solutions.

Generating Maxent predictions and assessing prediction performance

The R statistical language (Appendix C, R Development Core Team 2016) was used to generate Maxent (Phillips et al. 2006) predictions of the distributions of each trait-based *Opuntia* cluster relative to abiotic conditions (package "dismo"; Hijmans et al. 2013). The WorldClim database (http://www.worldclim.org, Hijmans et al. 2005) was used to obtain 19 BIOCLIM layers at 30 arc-second (~1 km^2) resolution to act as predictive environmental variables. All BIOCLIM layers were utilized when generating Maxent predictions.

The occurrence records for all cladodes in each trait-based *Opuntia* cluster were partitioned into 5 groups via *k*-fold partitioning (package dismo; Hijmans 2013). A total of 10,000 background points were then randomly selected from within the geographic boundary of the states included in the study region. The *Opuntia* occurrence records and randomly selected background points were used to train and test Maxent predictions of the range of each trait-based *Opuntia* cluster. Only Maxent's logistic output was used in this investigation, and Maxent predictions for each trait-based *Opuntia* cluster were projected across the extent of the study region at a resolution of 30 arc-seconds. A Maxent prediction of the North American distribution of *C. cactorum* was also generated by following similar methods (Chapter Three).

Maxent's ability to accurately predict abiotic suitability for each trait-based *Opuntia* cluster modeled was evaluated using metrics that are either dependent on thresholds that enable Maxent to discern suitable from unsuitable habitat (threshold-dependent) or independent of these thresholds (threshold-independent). Threshold-dependent metrics were informed by the threshold value that maximized Maxent sensitivity and specificity while modeling each trait-based *Opuntia* cluster. This threshold value was used because it balances Maxent's ability to predict instances of a species' presence and absence (Liu et al. 2005, 2013). The threshold-dependent performance metrics used here included: model omission rate (proportion of true occurrences misidentified by the defined threshold), sensitivity (proportion of actual positives identified as such), specificity (proportion of actual negatives identified as such), proportion of presence and absence points correctly identified, Cohen's kappa (Cohen 1960), and the true skill statistic (TSS; Allouche et al. 2006). Both Cohen's kappa and the TSS normalize overall model accuracy (the number of cells correctly classified as present or absent) by model accuracy that might have occurred due to chance. Values of Cohen's kappa and the TSS can range from -1 to 1; a value of 1 indicates perfect agreement between model accuracy and accuracy expected due to chance whereas negative values indicate that the model's predictions are no better than random chance (Cohen 1960, Allouche et al. 2006). The threshold-independent metric used was the area under receiver operator curve (AUC).

Integrating abiotic suitability and trait-based understanding

Environmental conditions may influence host *Opuntia* availability to *C. cactorum*. Larval *C. cactorum* are dependent on their *Opuntia* hosts for survival, so habitat suitability for *C. cactorum* may be reasonably estimated by considering environmental suitability relative to what *Opuntia* taxa are available to *C. cactorum*. The correlation coefficient between trait-based *Opuntia* clusters and *C. cactorum* infestation was used to weight all cell values in each *Opuntia* cluster's Maxent prediction. The weighted Maxent predictions of each trait-based *Opuntia* cluster were then stacked. Three different estimates of habitat suitability were then calculated based on the mean, median, or maximum value of each raster cell across all weighted and stacked *Opuntia* cluster predictions. Each habitat suitability map for *C. cactorum* was projected across the extent of the study region at a resolution of 30 arc-seconds.

Incorporating *C. cactorum* dispersal

The 233 *C. cactorum* occurrence records compiled for this investigation were examined to determine if each record was associated with BIOCLIM data. Some *C. cactorum* occurrence records were either not associated with, or were only partially associated with, BIOCLIM data. It appears the proximity of some occurrence records to water, especially in instances where *C. cactorum* occurrence records were located on small land masses (e.g., barrier islands along the Gulf Coast), influenced if BIOCLIM data were available or missing. Occurrence records not associated with all BIOCLIM data layers were excluded from further analyses.

A total of 187 occurrence records associated with all BIOCLIM data were used as starting, parental locations in all simulations of *C. cactorum* dispersal. Explicit consideration was given to the effect of temperature on immature *C. cactorum* survival, and thus the number of *C. cactorum* dispersing from a parental location, in all simulations. A spline curve was fit to life table data on immature *C. cactorum* survival as a function of rearing temperature (Legaspi and Legaspi 2007). The mean annual temperature (BIOCLIM layer 1) at each parental location was extracted and used in conjunction with the fitted spline curve to determine percent survival of immature *C. cactorum* to adulthood at each parental location. One eggstick, containing *n* number of eggs randomly drawn from a Poisson distribution with a mean of 63 eggs (Stephens et al. 2012), was generated at each parental *C. cactorum* location. The number of individuals dispersing from each parental *C. cactorum* location was calculated as the product of the number of eggs in each eggstick and percent survival to adulthood at each parental location. All dispersing individuals were assumed to be female.

Each individual dispersed from its parental location at a randomly chosen angle θ and randomly selected distance (up to 24km) drawn from an exponential distribution. The total number of individuals dispersing was used to randomly draw the same number of values from a standard uniform distribution (minimum and maximum values of 0 and 1, respectively). Both habitat suitability values and the randomly chosen values from a uniform distribution were used to determine if dispersing individuals established a new parental location. Individuals dispersing to new raster cells successfully established if the habitat suitability value of the raster cell was greater than the corresponding randomly selected value.

A dispersal event entailed the entire process of simulating eggsticks, extracting mean annual temperature, determining the number of *C. cactorum* surviving to adulthood from each eggstick, and dispersal of individuals at all parental *C. cactorum* locations. Simulations of *C. cactorum* dispersal included between 1 to 4 dispersal events. A total of 50 replicate simulations were generated for each combination of habitat suitability map and number of dispersal events.

Evaluation of trait-based SDM, PAB, and Maxent predictions

There were two goals associated with this research. The first goal was to compare the degree of similarity (or overlap) among trait-based SDM, PAB, and Maxent predictions. The second goal was to evaluate the relative contribution of habitat suitability and number of *C. cactorum* dispersal events in generating overlap among trait-based SDM predictions. Comparing the degree of overlap among all predictions necessitated that all predictions were on the same value scale. Trait-based SDM predictions were re-scaled by classifying all raster cells in each replicate prediction as occupied (value of 1) or unoccupied (value of 0) by *C. cactorum*. Replicates generated for each model scenario (i.e., habitat suitability values and number of dispersal events) were then stacked and averaged. Raster cell values in final trait-based SDM predictions of the North American distribution of *C. cactorum* ranged from 0 to 1 and represented the proportion of replicates in which any particular cell as occupied by *C. cactorum* at the end of 1 to 4 dispersal events.

Schoener's D (Schoener 1968) was used to compare the degree of similarity (or overlap) among trait-based SDM, PAB, and Maxent predictions of the North American distribution of *C. cactorum* (package ENMeval; Muscarella et al. 2014). Values of Schoener's D range from 0, indicating predictions that are greatly divergent, to 1, which indicates predictions that are exactly similar. A Kruskal-Wallis analysis of variance was used to evaluate variation in Schoener's D values among all predictions relative to modeling constraints of habitat suitability, whether *Opuntia* traits were included in the modeling process, host *Opuntia* availability, methods of modeling *C. cactorum* dispersal, and the number of dispersal steps or events included in the simulation process. A Dunn's test (Dunn 1964) for multiple comparisons (package dunn.test; Dinno 2014) with a Bonferroni correction was also used to compare pairs of predictions relative to specific modeling constraints of interest.

Schoener's D values were also used to evaluate the relative contribution of habitat suitability and *C. cactorum* dispersal in generating variation in overlap among trait-based SDM predictions. Discrete probability distributions were fit to all Schoener's D values for trait-based SDM predictions. Akaike's Information Criterion (AIC; Burnham and Anderson 2004) was used to assess distribution fit. The Schoener's D values of the trait-based SDM predictions were best fit by a beta distribution. Nested beta regression models (package betareg; Cribari-Neto and Zeileis 2010) were then fit to assess variation in Schoener's D relative to the modeling constraints of habitat suitability and number of dispersal events. The full beta regression model included both habitat suitability and number of dispersal events as predictors, but models including each factor individually were also generated. Schoener's D values were used as the response variable in all beta regression models. Likelihood ratio tests were used to assess the fit of each beta regression model, and a coefficient test was used to evaluate predictor significance in each beta regression model (package lmtest; Zeileis and Hothorn 2002).

Results

Tissue macronutrient traits were more influential than morphological traits in assigning cladode samples to clusters (Table 4.3), but the best model for explaining cladode assignment to cluster included all morphological and tissue macronutrient traits. Across all of the possible number of clusters, models including solely *Opuntia* tissue macronutrient traits had higher log-likelihood values and generated less residual deviance in model fit than did models including solely *Opuntia* morphological traits.

Further analysis focused on the 75 cladodes associated with measurements of tissue macronutrient content because *Opuntia* morphological traits were poor predictors of *C. cactorum* infestation (Chapter Two) and were less influential than tissue macronutrient traits in determining the number of trait-based clusters generated. Clusters generated using cladodes associated with *Opuntia* tissue macronutrient traits are hereafter referred to as tissue macronutrient-based clusters. The within-cluster sum of squares was lower when cladodes were partitioned into scenarios of 5 to 8 clusters (Figure 4.2). The disorder within clusters, as measured by entropy, increased monotonically with the number of clusters (Figure 4.2). Taken together, a scenario of 5 clusters minimized both sums of squares within clusters and disorder among samples, as well as best explained *C. cactorum* infestation (Table 4.4). Infestation by *C. cactorum* was weakly correlated with Cluster 4 (Z= -1.975, P = 0.048; Table 4.4) in a 5-cluster scenario. Occurrence records associated with cladodes in each of the 5 tissue macronutrient-based *Opuntia* clusters (Figure 4.3) were used to generate Maxent predictions.

Maxent predictions were generated for each tissue macronutrient-based *Opuntia* cluster, yielding a total of 5 predictions (Table 4.5, Figures C.3-C.7). The AUC values for Maxent predictions of tissue macronutrient-based *Opuntia* clusters ranged from 0.864 to 0.978. The threshold that maximized Maxent sensitivity and specificity in generating predictions ranged from 0.19 to 0.61. Model omission rates did not exceed 17%. Cohen's kappa ranged from 0.017 to 0.370 whereas values of the TSS ranged from 0.729 to 0.955. Both the threshold-independent and threshold-dependent metrics indicate Maxent performed well in generating predictions for each tissue macronutrient-based *Opuntia* cluster.

The Maxent predictions for each tissue macronutrient-based *Opuntia* cluster were weighted and used to generate maps of habitat suitability for *C. cactorum.* Habitat suitability for *C. cactorum* varied considerably relative to its method of calculation. Suitability values ranged from 5.66×10^{-9} to 0.79, 0.0012 to 0.682, and 0.005 to 0.989 for scenarios considering the median, mean, and maximum values in each raster cell across all Maxent predictions of tissue macronutrient-based *Opuntia* cluster ranges (Figure 4.4).

Considerable variation in Schoener's D values was observed when evaluating predictions relative to the modeling constraints imposed during the modeling process (Figure 4.5). A Kruskal-Wallis analysis of variance revealed substantial variation in Schoener's D values among predictions (Figure 4.5) relative to constraints of habitat suitability ($\chi^2 = 10645$, df = 6,p-value = <0.001), the use of *Opuntia* traits in the modeling process ($\chi^2 = 5821.9$, df = 2,p-value = <0.001), host *Opuntia* availability ($\chi^2 = 7483.3$, df = 3,p-value = <0.001), method of modeling *C. cactorum* dispersal ($\chi^2 = 5929$, df = 4,p-value = <0.001), and number of dispersal steps or events included in the modeling process ($\chi^2 = 6145.4$, df = 5,p-value = <0.001). In general, trait-based SDM predictions had less median overlap than PAB or Maxent predictions. A Dunn's test for multiple pairwise comparisons revealed substantial differences in Schoener's D values for nearly every pairwise comparison of predictions (Table 4.6). Schoener's D values for trait-based SDM predictions generated using 3 or 4 dispersal events were similar, and these trait-based SDM predictions were also similar to Schoener's D values of the Maxent-generated prediction (Table 4.7, Figure 4.6).

Likelihood ratio tests of nested beta regression models were used to explore the relative contribution of habitat suitability and number of dispersal events in generating variation in Schoener's D values among trait-based SDM predictions (Table 4.8). Variation in Schoener's D values was best explained by a full beta regression model that contained both abiotic suitability and number of dispersal events as predictor variables. Number of dispersal events explained more variation in Schoener's D values than did habitat suitability (Table 4.8).

Discussion

Invaders often form novel biotic interactions with taxa present in invaded communities (Sakai et al. 2000, Mitchell et al. 2006), but these associations are an obligate requirement for the successful establishment of some invaders. Identifying the traits that mediate local-scale invader interactions in exotic ranges may inform risk analyses (Stohlgren and Schnase 2006, Kumschick and Richardson 2013), efforts to mitigate invader ecological impacts (Yokomizo et al. 2009, Baxter and Possingham 2011, Giljohann et al. 2011, Guisan et al. 2013, Guillera-Arroita et al. 2015), and predictions of invader exotic distributions. In this research, trait-based understanding of patterns of host use by an invasive consumer was used to develop a trait-based SDM method. The hypothesis of this work was generally supported in that trait-based SDM predictions differed from PAB and Maxent predictions (Chapter Three). The expectation that trait-based SDM predictions would be more constrained than both PAB and Maxent predictions was partially upheld. Some trait-based SDM predictions were considerably more constrained than PAB and Maxent predictions, but some trait-based SDM predictions were similar in overlap with a Maxent-generated prediction. The number of dispersal events included in the modeling process was the most important constraint influencing overlap among trait-based SDM predictions. The results presented here indicate that a trait-based SDM method can generate predictions of invader exotic distributions that are generally more constrained than predictions generated by other SDM methods, but similar in overlap with predictions generated by methods that consider invader occurrence relative to just environmental conditions. The importance of the number of dispersal events in generating overlap among trait-based SDM predictions also highlights the need for increased understanding of invader dispersal abilities in exotic ranges.

The trait-based SDM method described in this dissertation chapter generated estimates of habitat suitability for *C. cactorum* by integrating trait-based understanding of *C. cactorum-Opuntia* associations with Maxent predictions of abiotic suitability for tissue macronutrient-based *Opuntia* clusters. This method of estimating habitat suitability for *C. cactorum* is likely reasonable because biotic interactions can exclude species from environmentally suitable areas, and thus determine habitat suitability, at scales consistent with broad-scale environmental conditions (Peterson et al. 2011, Fraterrigo et al. 2014, Godsoe et al. 2015, Lee-Yaw et al. 2016). A potential concern was that the quality of the Maxent predictions for each tissue macronutrient-based *Opuntia* cluster may have been affected by either the small number of cladodes in each cluster or geographical bias in the location of these cladodes. In general, Maxent can perform well with small sample sizes (< 7; Pearson et al. 2007), but AUC values can become artificially inflated if a small presence-only data set is used relative to a large number of background locations (Merow et al. 2013). All AUC values, as well as all threshold-dependent performance metrics, indicated Maxent performed well in generating predictions of the ranges of the tissue macronutrient-based *Opuntia* clusters. Yet, geographical bias in cladode locations may have affected estimates of abiotic suitability for each tissue macronutrient-based *Opuntia* cluster. More specifically, the lack of cladode occurrence records in areas where *C. cactorum* is known to occur (Florida and the Gulf Coast) likely led to underestimates of habitat suitability for *C. cactorum* in these areas. Future work could simultaneously address concerns with data set size and geographical bias by quantifying the tissue macronutrient contents of cladodes collected in areas where cladode sampling effort was small and *C. cactorum* currently occurs. Additional information on *Opuntia* tissue macronutrient content could be used to generate new tissue macronutrient-based *Opuntia* clusters and Maxent predictions of *Opuntia* cluster ranges, and thus ultimately improve estimates of habitat suitability for *C. cactorum*.

It is uncommon for SDM methods to include information about traits that may mediate invader interactions with taxa in invaded communities. Implementing trait-based understanding into SDM methods will undoubtedly affect ensuing predictions of invader distributions. Yet, the limitations of implementing trait-based understanding, as well as how and to what extent SDM predictions are affected, are not immediately clear. Trait-based understanding of *C. cactorum*-*Opuntia* associations was used to estimate habitat suitability for *C. cactorum*, but this approach was predicated on accurately depicting how *C. cactorum* perceives habitat suitability. One feasible scenario is that invaders perceive habitat suitability relative to the presence of taxa that possess specific suites of traits. Previous work suggests that particular *Opuntia* taxa are important for influencing *C. cactorum* occurrence (Dodd 1940, Marsico et al. 2011, Brooks et al. 2012, Sauby et al. 2012), so regions where these important *Opuntia* taxa occur may also be good habitat for *C. cactorum*. If the presence of *Opuntia* taxa that possess specific suites of traits dictates habitat suitability for *C. cactorum*, then the maximum value of each raster cell may best estimate habitat suitability for *C. cactorum*. An alternative scenario is that an invader may perceive habitat suitability relative to the entire assemblage of taxa, and thus the entire assemblage of traits, present in an invaded community. In this case, habitat suitability for an invader may be more appropriately estimated by aggregate measures (i.e., mean or median) of habitat suitability that are weighted based on consideration given to all taxa and traits available to the invader. If *Opuntia* community composition, and thus *Opuntia* traits present in the community, influences *C. cactorum* perception of habitat suitability, then areas of greater *Opuntia* diversity may affect the future spread and North American distribution of *C. cactorum*. This scenario may be especially true if multiple tissue macronutrient-based *Opuntia* clusters that are positively correlated with *C. cactorum* infestation co-occur. Yet, the converse is also feasible; *C. cactorum* westward spread may be impeded in regions where tissue macronutrient-based *Opuntia* clusters that are negatively correlated with *C. cactorum* infestation co-occur.

The number of dispersal events included in the modeling process was more important than habitat suitability for influencing overlap among trait-based SDM predictions. In general, invader ability to capitalize on resource availability (Davis et al. 2000) or reach unoccupied, but suitable, habitat (Lonsdale 1999, Williamson 1999) is influenced by the invader's dispersal abilities. Invader dispersal in an exotic range can be influenced by both the number and dispersal abilities of dispersing propagules, and several key assumptions were made about these aspects of *C. cactorum* dispersal. The first assumption was that the mean annual temperature across the study region fell within the range of rearing temperatures utilized by Legaspi and Legaspi (2007). However, extrapolating this trait-based SDM method beyond the range of rearing temperatures used by Legaspi and Legaspi (2007) may have generated considerable error in model predictions. Secondly, it was assumed that the maximum distance a dispersing individual could move from a parental location was 24km (Dodd 1940). While some laboratory studies have quantified distances flown by *C. cactorum* (Sarvary et al. 2008a, 2008b), overestimation of *C. cactorum* dispersal abilities in North America under field conditions likely generated overestimated predictions of the North American distribution of *C. cactorum*. Conversely, *C. cactorum* dispersal abilities may have been conservatively estimated because factors such as human-mediated dispersal (Wilson et al. 2009) or inclement weather (Andraca-Gomez et al. 2015) could increase *C. cactorum* dispersal distances and rate of spread. In either case, additional research on quantifying the long-distance dispersal abilities of *C. cactorum* under field conditions is needed. Lastly, it was assumed that all dispersing individuals were female. This assumption likely inflated the number of parental locations at the beginning of each dispersal event and may have resulted in overestimates of *C. cactorum* ranges, particularly in scenarios where habitat suitability was greatest (i.e., maximum habitat suitability). Regardless, field surveys throughout the predicted range of *C. cactorum* could be used to verify all trait-based SDM predictions.

The SDM method generated in this investigation represents an initial step towards integrating trait-based understanding of invader interactions into SDM methods. More specifically, the SDM method developed here used *Opuntia* tissue macronutrient traits that were important for influencing patterns of *C. cactorum* infestation to integrate trait-based understanding of *C. cactorum-Opuntia* associations. It is worth noting that *Opuntia* tissue macronutrient traits may not mechanistically influence *C. cactorum-Opuntia* associations. However, the explicit consideration and inclusion of *Opuntia* tissue macronutrient traits in this trait-based SDM method resulted in predictions that were considerably more constrained than predictions generated by Maxent and other SDM methods (Chapter 3). The extent to which *Opuntia* traits constrained predictions generated by this trait-based SDM method thereby illustrate the importance of identifying specific *Opuntia* traits that mechanistically affect the outcome of *C. cactorum-Opuntia* interactions. Yet, a lack of understanding about the specific *Opuntia* traits mediating *C. cactorum-Opuntia* associations may complicate identifying important traits of *C. cactorum* that mediate its associations with host *Opuntia*. Future research focused on identifying specific traits of *C. cactorum* that affect female moth detection of or oviposition on *Opuntia* hosts and larval traits that affect larval development and survival within *Opuntia* hosts may be particularly fruitful. Yet, experimental studies, preferably conducted in field-based settings, will be necessary to explicitly test hypotheses associated with these suites of *C. cactorum* traits. Regardless, increased trait-based understanding of *C. cactorum-Opuntia* associations may aid in identifying novel host *Opuntia* important to future spread of *C. cactorum*, as well as how *Opuntia* hosts influence *C. cactorum* distributions and persistence.

Table 4.1 List of *Opuntia* taxa included in this study.

Opuntia taxa	N	Morpho.	Tissue Macro.	Both	Total Infested
O. humifusa var. *ammophila*	18	12 (1)	5 (2)	1 (1)	4
O. atrispina	23	17	6		
O. basilaris	1	1			
O. camanchica	22	22			
O. diplopurpurea	2	2			
O. dulcis	5	5			
O. engelmannii var. *engelmannii*	19	13	6		
O. ficus-indica	6	6			
O. humifusa var. *humifusa*	176	159 (19)	12	5 (5)	24
O. engelmannii var. *lindheimer*i	196	180	8	8	
O. engelmannii var. *linguiformis*	16	16			
O. mackensenii	8	8			
O. macrocentra	6	6			
O. macrorhiza	78	70	6	2	
O. parva	2	2			
O. phaeacantha var. phaeacantha	11	11			
O. pottsii	30	30			
O. pusilla	23	18	5		
O. stricta	49	40 (2)	6 (1)	3 (1)	4
O. strigil	8	8			
Unknown1	2	2			
Unknown2	10	8		2	
Unknown3	4	4			
	715	640 (22)	54 (3)	21 (7)	32

Table 4.2 *Opuntia* morphological traits.

Trait	Type of Variable	Number of trait states	Trait states/units of measurement
Spines per areole	Categorical	19	0, 0 or 1, 0 to 2, 0 to 3, 0 to 4, 0 to 5, 1, 1 or 2, 1 to 3,1 to 4, 1 to 5, 1 to 6, 2, 2 or 3, 2 to 3, 2 to 4, 2 to 5, 3 to 5, 3 to 6
Spine color	Categorical	17	black to white tip, black to yellow tip, brown, brown-yellow, grey, grey and red, grey-yellow, none, purple, purple to white tip, red and white bands, red to white tip, red to yellow tip, white, yellow, yellow-white
Cladode shape	Categorical	4	elliptic, lanceolate, obovate, orbicular
Plant growth form	Categorical	2	erect, sprawling
Plant height	Categorical	2	< 1m, > 1m
Spine shape	Categorical	3	curved, none, straight
Spine persistence	Categorical	2	none, persistent
Spine pattern	Categorical	3	birds-foot, none, other
Mean spines per areole	Numerical	Discrete	# spines per areole
Mean length of up to 10 spines	Numerical	Continuous	cm
Median length of up to 10 spines	Numerical	Continuous	cm
Length of longest spine	Numerical	Continuous	cm

Table 4.3 Likelihood ratio tests and PERMANOVA of GLMs evaluating cladode assignment to clusters.

		Likelihood ratio test				ANOVA			
# cluster	Model	df	LL	Chisq	P	df	Resid. Dev.	Δ Dev.	P
5	Full	16	-7.86			6	2.60		
	Macro.	5	-22.50	29.27	0.0021	17	10.48	-7.88	0.077
	Morpho.	13	-22.86	0.72	0.99	9	10.84	-0.36	
6	Full	16	-4.54			6	1.89		
	Macro.	5	-20.39	31.71	0.00085	17	8.57	-6.68	0.032
	Morpho.	13	-25.63	10.48	0.23	9	14.12	-5.55	
7	Full	16	-12.32			6	3.98		
	Macro.	5	-22.79	20.93	0.034	17	10.77	-6.79	0.51
	Morpho.	13	-32.79	20.01	0.01	9	27.93	-17.16	
8	Full	16	-19.33			6	7.75		
	Macro.	5	-26.77	14.88	0.19	17	15.74	-7.99	0.86
	Morpho.	13	-37.87	22.21	0.0045	9	45.31	-29.58	
9	Full	16	-18.00			6	6.83		
	Macro.	5	-26.62	17.23	0.1	17	15.51	-8.68	0.75
	Morpho.	13	-38.43	23.62	0.0027	9	47.78	-32.26	
10	Full	16	-20.11			6	8.35		
	Macro.	5	-27.35	14.48	0.21	17	16.64	-8.29	0.88
	Morpho.	13	-38.78	22.84	0.0036	9	49.38	-32.74	
11	Full	16	-11.84			6	3.80		
	Macro.	5	-29.09	34.51	0.0003	17	19.63	-15.83	0.009
	Morpho.	13	-39.26	20.34	0.0091	9	51.72	-32.08	
12	Full	16	-26.63			6	15.52		
	Macro.	5	-33.31	13.38	0.27	17	29.35	-13.83	0.91
	Morpho.	13	-40.68	14.73	0.065	9	59.19	-29.84	

Table 4.4 GLMs evaluating if trait-based clusters predict *C. cactorum* infestation.

Cactus moth species	# of clusters (AIC)	Cluster Identity	Infested cladodes	Uninfested cladodes	Z	Pr (>|z|)
Cactoblastis cactorum	5 (17.71)	1	4	4	0	1
		2	0	13	0	0.99
		3	5	15	-1.26	0.21
		4	1	12	-1.975	0.048
		5	0	21	0	0.99
	6 (17.92)	1	4	4	0	1
		2	0	8	0	1
		3	6	15	-1.07	0.29
		4	0	12	0	1
		5	0	4	0	1
		6	0	22	0	1

Table 4.5 List of tissue macronutrient-based *Opuntia* clusters modeled.

Cluster	#	Thresh	AUC	Omis.	Sens.	Spec.	Prop. correct	K	TSS
Cluster 1	8	0.41	0.916	0.167	0.833	0.998	0.998	0.37	0.831
Cluster 2	13	0.61	0.978	0	1	0.955	0.955	0.045	0.955
Cluster 3	20	0.19	0.952	0	1	0.904	0.904	0.017	0.904
Cluster 4	13	0.51	0.864	0.091	0.91	0.819	0.819	0.009	0.729
Cluster 5	21	0.56	0.942	0.071	0.929	0.956	0.956	0.054	0.885

Table 4.6 Dunn test for multiple comparisons between trait-based SDM, PAB, and Maxent predictions.

Scenario	Comparisons	Z value	P
Traits included in modeling	Maxent - No traits	-15.246	< 2.2E-16
	Maxent - Traits included	6.522	1.04E-10
	No traits - Traits included	74.931	< 2.2E-16
Biotic interactions	All hosts - Trait-based clusters	68.611	< 2.2E-16
	All hosts - Eastern hosts	-40.76	< 2.2E-16
	Trait-based clusters - Eastern hosts	-83.106	< 2.2E-16
	All hosts - Maxent	13.571	< 2.2E-16
	Trait-based clusters - Maxent	-6.522	2.07E-10
	Eastern hosts - Maxent	18.512	< 2.2E-16
Habitat suitability	Cluster max - Cluster mean	7.066	1.67E-11
	Cluster max - Cluster median	12.219	< 2.2E-16
	Cluster mean - Cluster median	5.153	2.69E-06
	Cluster max - Max	-38.504	< 2.2E-16
	Cluster mean - Max	-48.302	< 2.2E-16
	Cluster median - Max	-55.448	< 2.2E-16
	Cluster max - Maxent	-2.006	0.471
	Cluster mean - Maxent	-6.475	9.93E-10
	Cluster median - Maxent	-9.735	< 2.2E-16
	Max - Maxent	17.305	< 2.2E-16
	Cluster max - Mean	-41.053	< 2.2E-16
	Cluster mean - Mean	-50.852	< 2.2E-16
	Cluster median - Mean	-57.997	< 2.2E-16
	Max - Mean	-9.178	< 2.2E-16
	Maxent - Mean	-18.598	< 2.2E-16
	Cluster max - Median	-23.484	< 2.2E-16
	Cluster mean - Median	-33.282	< 2.2E-16
	Cluster median - Median	-40.428	< 2.2E-16
	Max - Median	54.071	< 2.2E-16
	Maxent - Median	-9.684	< 2.2E-16
	Mean - Median	63.249	< 2.2E-16
Maximum dispersal distance	12km - 24km	2.905	0.0584
	12km - Local	10.253	< 2.2E-16
	12km - Maxent	15.681	< 2.2E-16
	24km - Local	7.881	1.63E-14
	24km - Maxent	15.307	< 2.2E-16

Table 4.6 (continued)

	Local - Maxent	14.006	< 2.2E-16
	12km - Diffusion	74.421	< 2.2E-16
	24km - Diffusion	73.18	< 2.2E-16
	Local - Diffusion	-66.111	< 2.2E-16
	Maxent - Diffusion	-6.522	3.46E-10
LDD Probability	Local - 0.1%	0.412	1
	Local - 1%	-18.546	< 2.2E-16
	Local - Maxent	14.006	< 2.2E-16
	0.1% - 1%	-23.218	< 2.2E-16
	0.1% - Maxent	13.999	< 2.2E-16
	1% - Maxent	16.989	< 2.2E-16
	Local - Diffusion	66.111	< 2.2E-16
	0.1% - Diffusion	68.843	< 2.2E-16
	1% - Diffusion	78.758	< 2.2E-16
	Diffusion - Maxent	-6.522	< 2.2E-16
Dispersal events	1 to 2	-7.688	1.12E-13
	1 to 3	-14.209	< 2.2E-16
	1 to 4	-16.287	< 2.2E-16
	1-Final	-51.453	< 2.2E-16
	1-Maxent	-12.63	< 2.2E-16
	2 to 3	-6.521	5.25E-10
	2 to 4	-8.599	< 2.2E-16
	2-Final	-40.634	< 2.2E-16
	2-Maxent	-7.193	4.76E-12
	3 to 4	-2.078	0.283
	3-Final	-31.458	< 2.2E-16
	3-Maxent	-2.582	0.0737
	4-Final	-28.533	< 2.2E-16
	4-Maxent	-1.113	1
	Final-Maxent	15.246	< 2.2E-16

Reported P values are exact P values calculated by the Dunn test.

Table 4.7 Results of likelihood ratio test comparing nested beta regression models.

Model	df	LogLik	Chisq	Pr (>Chisq)
Suitability + Dispersal event	5	4312.8		
Dispersal event	3	3667.4	1620.7	< 0.001
Suitability	4	2857.1	2911.5	< 0.001

Nest beta regression models were generated to evaluate the contribution of habitat suitability and the number of dispersal events to Schoener's D values for all trait-based SDM predictions.

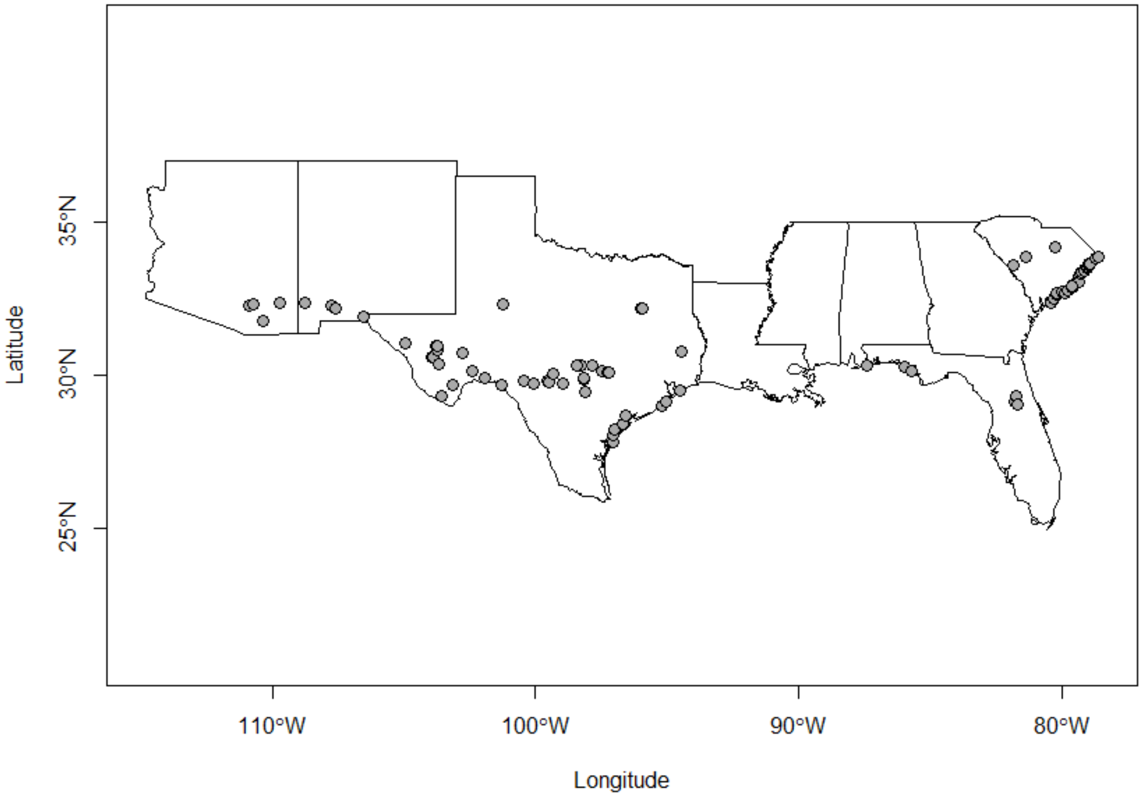

Figure 4.1 Map of the southern United States indicating where *Opuntia* cladodes collection sites.

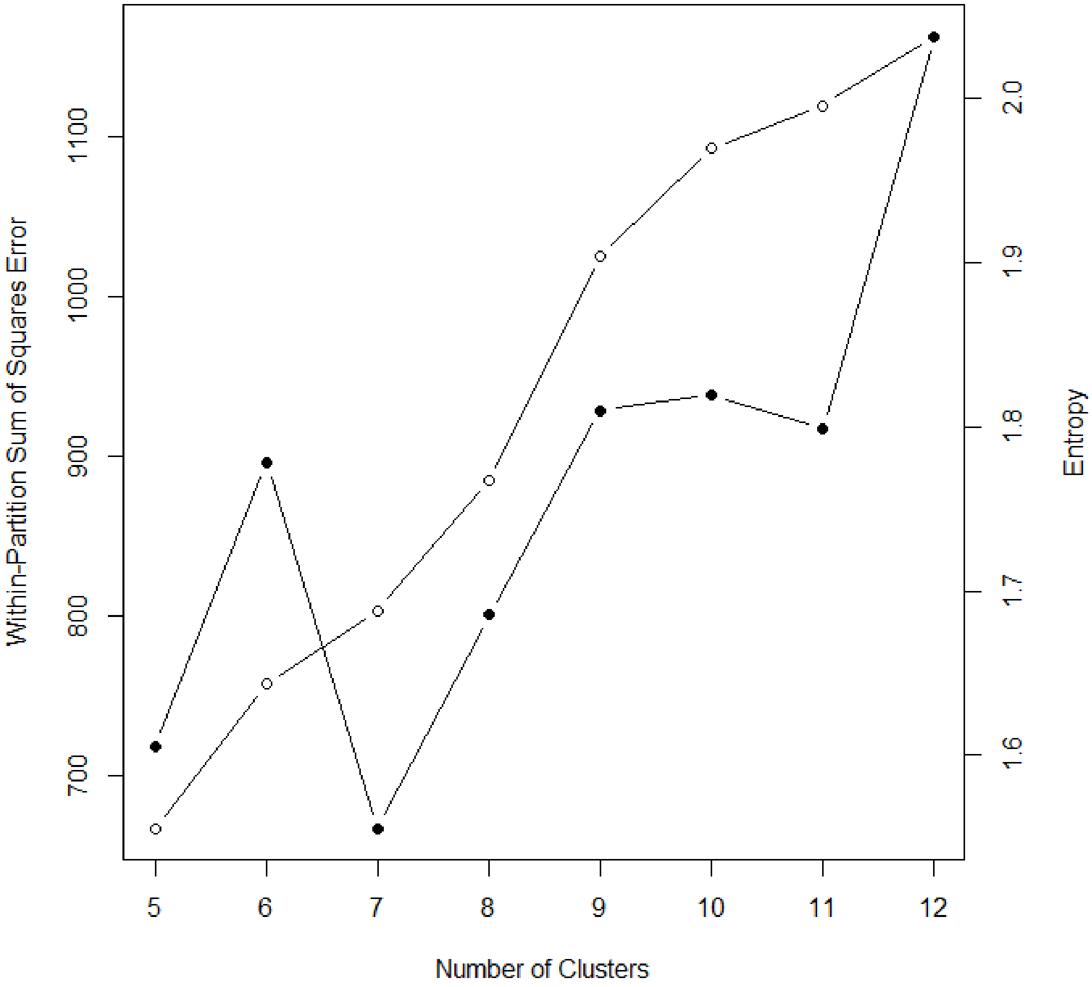

Figure 4.2 Internal (filled black circles) and external (open circles) validation criteria for all number of partition scenarios.

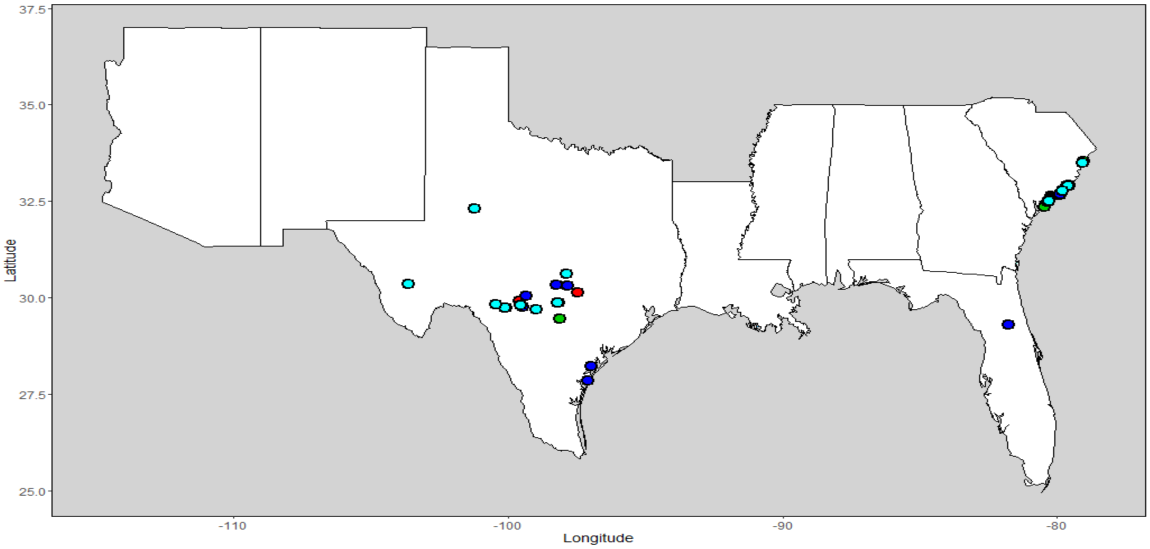

Figure 4.3 Map of the study region where 75 cladode samples were collected.

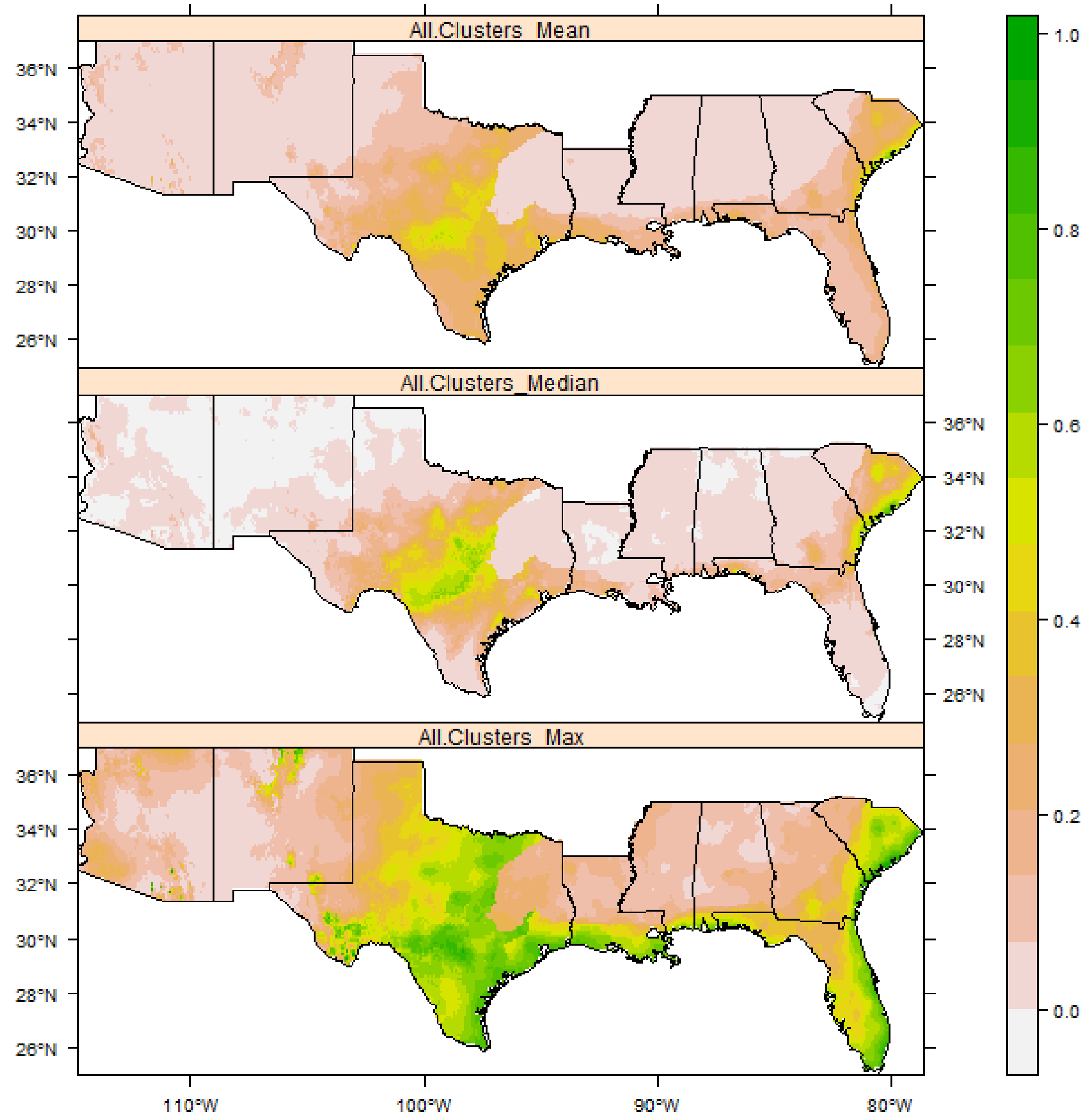

Figure 4.4 Mosaics of habitat suitability for *C. cactorum*.

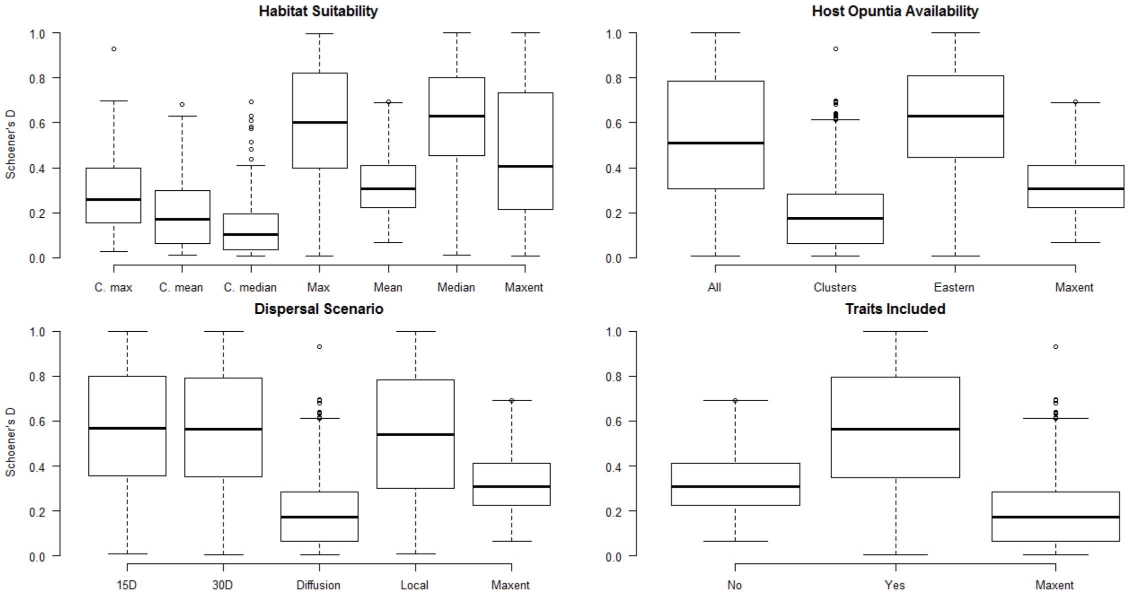

Figure 4.5 Boxplots comparing Schoener's D values associated with trait-based SDM, PAB, and Maxent predictions.

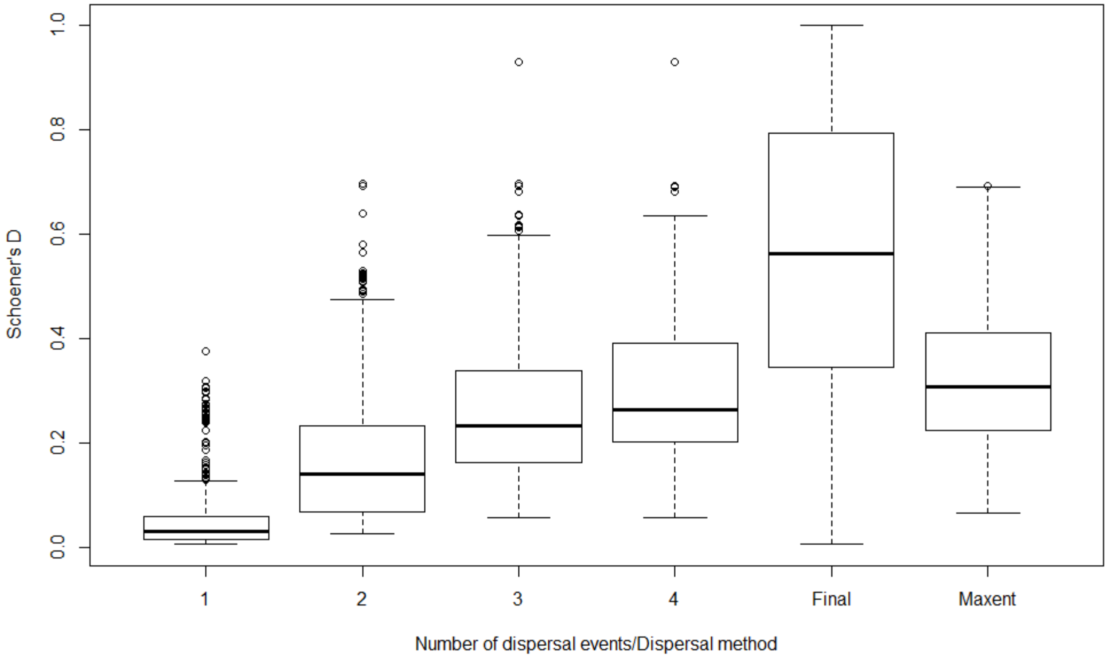

Figure 4.6 Boxplots comparing Schoener's D values associated with trait-based SDM, PAB, and Maxent predictions.

CHAPTER V

DISSERTATION SUMMARY

Understanding the sources of variation in the relationship between ecological process and pattern is crucial for extending our knowledge of ecological invasions across spatial and temporal scales. The invasion of North America by *C. cactorum* is an excellent context for exploring these relationships. The data show that tissue macronutrient concentrations in *Opuntia* tissues are predictive of patterns of *C. cactorum* host use in its exotic range (Chapter Two). Incorporating biotic interactions and dispersal into species distribution models greatly affects our predictive ability whether the model is based on host taxon (Chapter Three) or on clusters of host traits (Chapter Four).

This dissertation has primarily focused on applying trait-based approaches to a scenario of biological invasion. Results indicate predictions of an invader's exotic distribution can be constrained by utilizing trait-based understanding of the invader's biotic interactions. Similar methods could be applied when modeling the distributions of non-invasive species. Most methods for modeling species distributions consider a species' biophysical constraints relative to environmental conditions (Kearney and Porter 2009, Buckley et al. 2010, Martínez et al. 2015), but efforts to incorporate biotic interactions are uncommon and typically individual-based (Dormann et al. 2012, Meineri et al. 2015). Successfully including biotic interactions into methods for modeling species distributions is predicated on careful consideration of the biotic interactions that directly influence the occurrence of the species of interest.

Assessing the ecological fit (Janzen 1985, Agosta 2006) between the traits of interacting organisms may increase researcher ability to predict an organism's occurrence and distributions. One possible scenario is that suboptimal fit between organism traits affects how and if associations between organisms are maintained (Agosta and Klemens 2008, Gillespie and Wratten 2011). Interactions between consumers and resources whose traits are not exact ecological fits may lead to suboptimal interactions that restrict consumer dietary breadth and preferences (Forister et al. 2012). This may in turn restrict where the organism occurs. Alternatively, strong ecological fit among organism traits may play a role in the formation of novel interactions between species (Forister and Wilson 2013). For instance, larval *C. cactorum* survival is an outcome of the interaction of larval *C. cactorum* and host *Opuntia* traits. Larval *C. cactorum* survival on novel *Opuntia* hosts may thereby generate new biotic associations that influence where *C. cactorum* occurs and persists. It follows that trait-based understanding of instances of strong ecological fit between organism traits may inform predictions of an organism's distributions. In addition, understanding of the degree of ecological fit associated with biotic interactions may be useful for identifying suites of organism traits that are critical to maintaining function and structure in native communities (Wootton 1994, Brooks et al. 2006, Ackerly and Cornwell 2007, Agosta and Klemens 2008).

If species distribution modeling methods integrate trait-based understanding of the constraints and interactions that affect organism occurrence, then ensuing predictions of the organism's distribution should project both process and pattern across larger spatial scales. An important question then arises: to what extent do predictions of a species' distribution generated via trait-based modeling methods coincide with the species' niches (Sexton et al. 2009, Peterson et al. 2011, Lee-Yaw et al. 2016)? Predictions of species distributions generally appear to closely correspond to their niches, but only when these predictions are generated in the species' native range (Lee-Yaw et al. 2016). Using conditions from an invader's native range can result in a failure to predict the invader's exotic distribution (Brooks et al. 2012). It should therefore become common practice to incorporate abiotic, biotic, and dispersal constraints that are specific to the region where the prediction will be generated and projected.

In summary, the methods used in this dissertation to predict the distribution of an invasive species may be used as a guide to predict the distribution of any species of interest. Several important points should be considered prior to developing predictions. First, obligate associations that require strong ecological fit between organism traits be used to constrain model predictions. This is because instances of strong ecological fit are likely to influence the occurrence and persistence of the species of interest. Instances of strong ecological fit also present opportunities to estimate habitat suitability for the species of interest relative to what hosts, resources, or traits are available. This approach could be generally applicable to any species of interest but may be especially fruitful when the focal species is invasive. Second, multiple scenarios of propagule pressure and dispersal should be included in a modeling method because of uncertainty about organism dispersal capabilities. Multiple scenarios of dispersal also allow researchers to evaluate the sensitivity of the modeling method to the dispersal constraints being imposed. Third, modeling methods should simultaneously implement the three major factors (propagule pressure, abiotic suitability, and biotic interactions) constraining species distributions in order to assess the relative importance of each factor. Methods similar to the ones used in this dissertation could thereby facilitate additional hypothesis testing about specific mechanisms that affect or constrain the distribution of a focal species. Lastly, the quality of the information included in the modeling process should be assessed. Evaluating the information included in species distribution modeling methods will ensure that the pieces of information being assembled are of as maximum of quality as possible. Taken together, the aforementioned points will ensure that methods for predicting species distributions are carefully constructed and generally applicable to a wide variety of species. This will, in turn, provide an excellent step forward in researcher ability to develop modeling methods that address, and potentially overcome, ecology's fundamental problem.

REFERENCES

Ackerly, D.D., and W.K. Cornwell. 2007. A trait-based approach to community assembly: partitioning of species trait values into within-and among-community components. Ecol Lett. 10:135-145.

Agosta, S.J. 2006. On ecological fitting, plant-insect associations, herbivore host shifts, and host plant selection. Oikos. 114:556-565.

Agosta, S.J., and J.A. Klemens. 2008. Ecological fitting by phenotypically flexible genotypes: implications for species associations, community assembly and evolution. Ecol Lett. 11:1123-1134.

Allouche, O., A. Tsoar and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol. 43:1223-1232.

Anderson, E.F. 2001. The Cactus Family. Timber Press, Portland, Oregon, USA.

Andraca-Gómez, G., M. Ordano, K. Boege, C.A. Domínguez, D. Piñero, R. Pérez Ishiwara, J. Pérez-Camacho, M. Cañizares and J. Fornoni. 2015. A potential invasion route of *Cactoblastis cactorum* within the Caribbean region matches historical hurricane trajectories. Biol Invasions. 17:1397-1406.

Araújo, M.B., and M. Luoto. 2007. The importance of biotic interactions for modelling species distributions under climate change. Global Ecol Biogeogr, 16:743-753.

Awmack, C.S., and S.R. Leather. 2002. Host plant quality and fecundity in herbivorous insects. Ann Rev Entomol, 47:817-844.

Bartoń, K. 2013. MuMIn: multi-model inference. R package version, 1(5).

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

Boggs, C.L., and K.D. Freeman. 2005. Larval food limitation in butterflies: effects on adult resource allocation and fitness. Oecologia. 144:353-361.

Bradley, B.A., R. Early and C.J.B. Sorte. 2015. Space to invade? Comparative range infilling and potential range of invasive and native plants. Global Ecol Biogeogr. 24:348-359.

Brokensha, M.A. 1996. Incentives for cane quality. Proc S Afr Sug Technol Ass. 70:291-292.

Brooks, C.P., G.N. Ervin, L. Varone and G.A. Logarzo. 2012. Native ecotypic variation and the role of host identity in the spread of an invasive herbivore, *Cactoblastis cactorum*. Ecology. 93:402-410.

Brooks, C.P., B.H. Lambert, K.E. Sauby, G.E. Ervin, L. Varone and G.A. Logarzo. 2014. Larval morphology and host use confirms ecotypic variation in *Cactoblastis cactorum* (Berg). Biol Invasions. 16:13-22.

Buchan, L.A.J., and D.K. Padilla. 1999. Estimating the probability of long-distance overland dispersal of invading aquatic species. Ecol Appl. 9:254-265.

Buckley, L.B., M.C. Urban, M.J. Angilletta, L.G. Crozier, L.J. Rissler and M.W. Sears. 2010. Can mechanism inform species' distribution models?. Ecol Lett. 13:1041-1054.

Burnham, K.P. and D.R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York.

Burnham, K.P., and D.R. Anderson. 2004. Multimodel inference – understanding AIC and BIC in model selection. Sociol Methods Res. 33:261-304.

Byers, J.E., and E.G. Noonburg. 2003. Scale dependent effects of biotic resistance to biological invasion. Ecology. 84:1428-1433.

Calcagno, V. and C. de Mazancourt. 2010. glmulti: an R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, *34*(12), pp.1-29.

Capinha, C., L. Brotons and P. Anastácio. 2013. Geographical variability in propagule pressure and climatic suitability explain the European distribution of two highly invasive crayfish. J Biogeogr. 40:548-558.

Catford, J.A., R. Jansson and C. Nilsson. 2009. Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Divers Distrib. 15:22-40.

Chase, J.M., and M.A. Leibold, 2003. Ecological niches: linking classical and contemporary approaches. University of Chicago Press.

Chavent, M., B. Liquet, V. Kuentz and J. Saracco. 2012. ClustOfVar: An R package for the clustering of variables. J Stat Softw. 50:1-16.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educ Psychol Meas. 20:37-46.

Colasurdo, N., Y. Gélinas and E. Despland. 2009. Larval nutrition affects life history traits in a capital breeding moth. J Exp Biology. 212:1794-1800.

Colautti R.I., A. Ricciardi, I.A. Grigorovich and H.J. MacIsaac. 2004. Is invasion success explained by the enemy release hypothesis? Ecol Lett. 7:721–733.

Connell, J.H. 1961. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. Ecology. 42710-723.

Cribari-Neto F., and A. Zeileis. 2010. Beta regression in R. J Stat Softw. 34:1–24.

da Silva Doge, J., H. Valadao de Oliveria and R. Tidon. 2015. Rapid response to abiotic and biotic factors controls population growth of two invasive drosophilids (Diptera) in the Brazilian Savanna. Biol Invasions. DOI: 10.1007/s10530015-0889-0.

Daehler, C.C. 2001. Darwin's naturalization hypothesis revisited. Am Nat. 158:324-330.

Darwin, C. 1859. The Origin of Species. London: Murray.

Davis, M.A., J.P. Grime and K. Thompson. 2000. Fluctuating resources in plant communities: a general theory of invasibility. J Ecol. 88:528-534.

de Araujo, C.B., L.O. Marcondes-Machado and G.C. Costa. 2015. The importance of biotic interactions in species distribution models: a test of the Eltonian noise hypothesis using parrots. J Biogeogr. 41:513-523.

Dickel, T.S. 1991. *Cactoblastis cactorum* in Florida (Lepidoptera: Pyralidae: Phycitinae). Trop Lepidoptera. 2:117–118.

Dinno, A. 2014. dunn. test: Dunn's test of multiple comparisons using rank sums. R package version, 1(0).

Dodd, A.P. 1940. The biological campaign against prickly pear. Commonwealth Prickly Pear Board, Australia. 177 pp.

Dormann, C.F., S.J. Schymanski, J. Cabral, I. Chuine, C. Graham, F. Hartig, M. Kearney, X. Morin, C. Römermann, B. Schröder and A. Singer. 2012. Correlation and process in species distribution models: bridging a dichotomy. J Biogeogr. 39:2119-2131.

Dukes, J.S., and H.A. Mooney. 2004. Disruption of ecosystem processes in western North American by invasive species. Rev Chil Hist Nat. 77:411-437.

Dunn, O.J. 1964. Multiple comparisons using rank sums. Technometrics. 6:241-252.

Elith, J., and J.R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Ann Rev Ecol Syst. 40:677.

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

Elton, C. 1927. Animal Ecology. Macmillian Company, New York.

Engler, R., and A. Guisan. 2009. MIGCLIM: predicting plant distribution and dispersal in a changing climate. Divers Distrib. 15:590-601.

Engler, R., W. Hordijk and A. Guisan. 2012. The MIGCLIM R package seamless integration of dispersal constraints into projections of species distribution models. Ecography. 35:872- 878.

Fargione, J.E., and D. Tilman. 2005. Diversity decreases invasion via both sampling and complementarity effects. Ecol Lett. 8:604-611.

Forister, M.L., L.A. Dyer, M.S. Singer, J.O. Stireman and J.T. Lill. 2012. Revisiting the evolution of ecological specialization, with emphasis on insect–plant interactions. Ecology. 93:981-991.

Forister, M.L., and J.S. Wilson. 2013. The population ecology of novel plant-herbivore interactions. Oikos. 122:657-666.

Fraterrigo, J.M., S. Wagner and R.J. Warren. 2014. Local-scale biotic interactions embedded in macroscale climate drivers suggest Eltonian noise hypothesis distribution patterns for an invasive grass. Ecol Lett. 17:1447-1454.

Gallardo, B., M. Clavero, M.J. Sanchez and M. Vila. 2016. Global ecological impacts of invasive species in aquatic ecosystems. Glob Change Biol. 22:151-163.

Gallien, L., T. Münkemüller, C.H. Albert, I. Boulangeat and W. Thuiller. 2010. Predicting potential distributions of invasive species: where to go from here? Divers Distrib.16:331-342.

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

Gallien, L., F. Mazel, S. Lavergne, J. Renaud, R. Douzet and W. Thuiller. 2015. Contrasting the effects of environment, dispersal and biotic interactions to explain the distribution of invasive plants in alpine communities. Biol Invasions. 17:1407-1423.

Genc, H., and J.L. Nation. 2004. An artificial diet for the butterfly *Phyciodes phaon* (Lepidoptera:Nymphalidae). Fla Entomol. 87:194-198.

Gillespie, M., and S.D. Wratten. 2011. Oviposition preference of *Lycaena salustis* for, and larval performance on, a novel host plant: an example of ecological fitting. Ecol. Entomol. 36:616-624.

Godsoe, W., R. Murray and M.J. Plank. 2015. Information on biotic interactions improves transferability of distribution models. Am Nat. 185:281-290.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. Auk. 34:427-433.

Grubb, P.J. 1986. Sclerophyllys, pachyphylls and pycnophylls: the nature and significance of hard surfaces. In: Juniper, B., Southwood, T.R.E. (Eds.), Insects and the Plant Surface. Arnold, London, pp. 137–150.

Guisan, A., and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecol Model. 135:147-186.

Guisan, A., B. Petitpierre, O. Broennimann, C. Daehler and C. Kueffer. 2014. Unifying niche shift studies: insights from biological invasions. Trends Ecol Evol. 29:260-269.

Habeck, D. H., and F. D. Bennett. 1990. *Cactoblastis cactorum* Berg (Lepidoptera: Pyralidae), a phycitine new to Florida. Florida Department of Agriculture & Consumer Services, Division of Plant Industry.

Hanley, M.E., B.B. Lamont, M.M. Fairbanks and C.M. Rafferty. 2007. Plant structural traits and their role in anti-herbivore defence. Perspect Plant Ecol. 8:157-178.

Hengeveld, R. 1994. Small-step invasion research. Trends Ecol Evol 9:339-342.

Hennig, C. 2010. fpc: Flexible procedures for clustering. R package version 2(2):0-3.

Hight, S.D., J.E. Carpenter, K.A. Bloem, S. Bloem, R.W. Pemberton and P. Stiling. 2002. Expanding geographical range of *Cactoblastis cactorum* (Lepidoptera:Pyralidae) in North American. Fla Entomol. 85:527-529.

Hight, S.D., S. Bloem, K. Bloem, and J.E. Carpenter. 2003. Mating Behavior of *Cactoblastis cactorum*: Observations of courtship and mating behaviors at two locations on the Gulf Coast of Florida. Fla Entomol. 86 :400-408.

Hight, S.D. J.E. Carpenter, S. Bloem and K.A. Bloem 2005. Developing a sterile insect release program for *Cactoblastis cactorum* (Berg) (Lepidoptera: Pyralidae): effective overflooding ratios and release-recapture field studies. Environ Entomol. 34:850-856.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 25:1965-1978.

Hijmans, R. 2011. Package Raster. Version 1.6-10. http://raster.r-forge.r-project.org/

Hijmans, R.J., S. Phillips, J. Leathwick and J. Elith. 2013. dismo: Species distribution modeling. R package version 0.8-11.

Hooper, D.U., F.S. Chapin, J.J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J.H. Lawton, D.M. Lodge, M. Loreau, S. Naeem and B. Schmid. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr. 75:3-35.

Hunt, D.R. 2006. The New Cactus Lexicon. Dh books, The Manse, Chapel Lane, Milbourne Port, UK.

Hutchinson, G.E. 1957. Concluding remarks Cold Spring Harbor Symposia on Quantitative Biology 22:415–427.

Janzen, D.H. 1985. On ecological fitting. Oikos. 45:308-310.

Jarnevich, C.S., and L.V. Reynolds. 2011. Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. Biol Invasions. 13:153-163.

Jezorek, H.A., P.D. Stiling and J.E. Carpenter. 2010. Targets of an invasive species: Oviposition preference and larval performance of *Cactoblastis cactorum* (Lepidoptera: Pyralidae) on 14 North American Opuntioid cacti. Environ Entomol. 39:1884-1892.

Jiménez-Valverde, A., and A. Peterson. 2011. Use of niche models in invasive species risk assessments. Biol Invasions. 13:2785–2797.

Johnson, D.M., and P.D. Stiling. 1998. Distribution and dispersal of *Cactoblastis cactorum* (Lepidoptera:Pyralidae), an exotic *Opuntia*-feeding moth, in Florida. Fla Entomol. 1:12-22.

Jones, C.C., S. A.Acker and C.B. Halpern. 2010. Combining local-and large-scale models to predict the distributions of invasive plant species. Ecol Appl, 20:311-326.

Kearney, M. 2006. Habitat, environment and niche: what are we modelling? Oikos. 115:186-191.

Kearney, M., and W. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecol Lett. 12:334-350.

Keddy, P.A. 1992. A pragmatic approach to functional ecology. Funct Ecol, 6:621-626.

Keesing, F., Holt, R.D. and R.S. Ostfeld, 2006. Effects of species diversity on disease risk. Ecol Lett, 9:485-498.

Keller S.R., and D.R. Taylor. 2008. History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecol Lett. 11:852-866.

Kenis, M., M.A. Auger-Rozenberg, A.Rogues, L. Timms, C. Pere, M.J. Cock, I. Settele, S. Augustin and C. Lopez-Vaamonde. 2008. Ecological effects of invasive alien insects. In: Ecological Impacts of Non-Native Invertebrates and Fungi on Terrestrial Ecosystems. Spring Netherlands. pp. 21-45.

Kennedy, T.A., S. Naeem, K.M. Howe, J.M. Knops, D. Tilman and P. Reich. 2002. Biodiversity as a barrier to ecological invasion. Nature. 417:636-638.

Kissling, W.D., C.F. Dormann, J. Groeneveld, T. Hickler, I. Kühn, G.J. McInerny, J.M. Montoya, C. Römermann, K. Schiffers, F.M. Schurr and A. Singer. 2012. Towards novelapproaches to modelling biotic interactions in multispecies assemblages at large spatial extents. J Biogeogr, 39:2163-2178.

Knops, J.M., D. Tilman, N.M. Haddad, S. Naeem, C.E. Mitchell, J. Haarstad, M.E. Ritchie, K.M. Howe, P.B. Reich, E. Siemann and J. Groth. 1999. Effects of plant species richness on invasion dynamics, disease outbreaks, insect abundances and diversity. Ecol Lett. 2:286-293.

Kot, M., M.A. Lewis and P. van den Driessche. 1996. Dispersal data and the spread of invading organisms. Ecology. 77:2027-2042.

Kumschick, S. and D.M. Richardson. 2013. Species-based risk assessments for biological invasions: advances and challenges. Divers Distrib. 19:1095-1105.

Laca, E.A., L.A. Shipley and E.D. Reid, 2001. Structural anti-quality characteristics of range and pasture plants. J Range Manage. 54:413-419.

Lamouroux, N., N.L. Poff and P.L. Angermeier. 2002. Intercontinental convergence of stream fish community traits along geomorphic and hydraulic gradients. Ecology. 83:1792-1807.

Lavorel, S., and E. Garnier, 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Funct Ecol. 16:545-556.

Leach, K., W.I. Montgomery and N. Reid. 2015. Biogeography, macroecology and species' traits mediate competitive interactions in the order Lagomorpha. Mammal Rev. 45:88-102.

Lee, C.E. 2002. Evolutionary genetics of invasive species. Trends Ecol Evol. 17:386-391.

Lee-Yaw, J.A., H.M. Kharouba, M. Bontrager, C. Mahony, A.M. Csergo, A.M.E. Noreen, Q. Li, R. Schuster and A.L. Angert. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. Ecol Lett. DOI: 10.1111/ele.12604.

Legaspi, J.C. and B.C. Legaspi. 2007. Life table analysis for *Cactoblastis cactorum* immatures and female adults under five constant temperatures: implications for pest management. Ann Entomol Soc Amer. 100:497-505.

Leps, J., F. De Bello, S. Lavorel and S. Berman. 2006. Quantifying and interpreting functional diversity of natural communities: practical considerations matter. Preslia. 78:481-501.

Levin, S.A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology. 73:1943-1967.

Levine, J.M. 2000. Species diversity and biological invasions: relating local process to community pattern. Science. 288:852-854.

Liu, C., P.M. Berry, T.P. Dawson and R.G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography. 28:385-393.

Liu, C., M. White and G. Newell. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. J Biogeogr. 40:778-789.

Lockwood, J.L., P. Cassey and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. Trends Ecol Evol. 20:223-228.

Lonsdale, W.M. 1999. Global patterns of plant invasions and the concept of invasibility. Ecology. 80:1522-1536.

MacArthur R.H. 1958. Population ecology of some warblers of northeastern coniferous forests. Ecology. 39:99-619.

Majure, L.C., and G.N. Ervin. 2007. The *Opuntias* of Mississippi. Haseltonia. 14:111-126.

Majure, L.C., R. Puente, M.P. Griffith, W.S. Judd, P.S. Soltis and D.E. Soltis. 2012a. Phylogeny of *Opuntia* s.s. (Cactaceae): Clade delineation, geographic origins, and reticulate evolution. Am J Bot. 99:847-864.

Majure, L.C., W.S. Judd, P.S. Soltis and D.E. Soltis. 2012b. Cytogeography of the *Humifusa* clade of *Opuntia* ss. Mill. 1754 (Cactaceae, Opuntioideae, Opuntieae): correlations with Pleistocene refugia and morphological traits in a polyploidy complex. Comp Cytogen. 6:53-77.

Marsico, T.D., L.E. Wallace, G.N. Ervin, C.P. Brooks, J.E. McClure and M.E. Welch. 2011. Geographic patterns of genetic diversity from the native range of *Cactoblastis cactorum* (Berg) support the documented history of invasion and multiple introductions for invasive populations. Biol Invasions. 13: 857-868.

Martínez, B., F. Arenas, A. Trilla, R.M. Viejo and F. Carreño. 2015. Combining physiological threshold knowledge to species distribution models is key to improving forecasts of the future niche for macroalgae. Glob Change Biol. 21:1422-1433.

McArdle, B.H., and M.J. Anderson. 2004. Variance heterogeneity, transformations, and models of species abundance: a cautionary tale. Can J Fish Aquat Sci. 61:1294-1302.

McGill, B.J., B.J. Enquist, E. Weiher and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends Ecol Evol. 21:178-185.

Meilă, M. 2005, August. Comparing clusterings: an axiomatic view. In: Proceedings of the 22nd international conference on Machine learning. ACM. pp. 577-584.

Meilă, M. 2007. Comparing clusterings—an information based distance. J Multivariate Anal. 98:873-895.

Meineri, E., A.S. Deville, D. Grémillet, M. Gauthier-Clerc and A. Béchet. 2015. Combining correlative and mechanistic habitat suitability models to improve ecological compensation. Biol Rev. 90:314-329.

Merow, C., M.J. Smith and J.A. Silander, Jr. 2013. A practical guide to Maxent for modeling species' distributions: what it does, and why inputs and settings matter. Ecography. 36:1058-1069.

Mitchell, C.E., A.A. Agrawal, J.D. Bever, G.S. Gilbert, R.A. Hufbauer, J.N. Klironomos, J.L. Maron, W.F. Morris, I.M. Parker, A.G. Power and E.W. Seabloom. 2006. Biotic interactions and plant invasions. Ecol Lett. 9:726-740.

Mooney, H.A., and E.E. Cleland. 2001. The evolutionary impact of invasive species. P Natl Acad Sci USA. 98:5446-5451.

Moore, R.F. 1985. Artificial diet: Development and improvement. In: P. Singh and R. F. Moore [eds.], Handbook of Insect Rearing, Volume I. Elsevier Science Publishers BV, Amsterdam, The Netherlands. pp. 67-84.

Morin, X., and M.J. Lechowicz. 2008. Contemporary perspectives on the niche that can improve models of species range shifts under climate change. Biol Letters. 4:573-576.

Muscarella, R., P.J. Galante, M. Soley-Guardia, R.A. Boria, J.M. Kass, M. Uriarte and R.P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evalutations and estimating optimal model complexity for Maxent ecological niche models. Methods in Ecology and Evolution. 5:1198-1205.

Muthukrishnan, R., N.M. West, A.S. Davis, N.R. Jordan and J.D. Forester. 2015. Evaluating the role of landscape in the spread of invasive species: The case of the biomass crop *Miscanthus giganteus*. Ecol Model. 317:6-15.

Myers, J.H., J. Monro and N. Murray. 1981. Egg clumping, host plant selection and population regulation in *Cactoblastis cactorum* (Lepidoptera). Oecologia. 51:7-13.

Naeem, S., J.M. Knops, D. Tilman, K.M. Howe, T. Kennedy and S. Gale. 2000. Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. Oikos. 91:97-108.

Official Methods of Analysis of AOAC INTERNATIONAL. 2012.19th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA.

Official Methods of Analysis of AOAC INTERNATIONAL. 2012. 19th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA, Official Method 962.09.

Official Methods of Analysis of AOAC INTERNATIONAL. 2012. 19th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA, Official Method 920.39.

Official Methods of Analysis of AOAC INTERNATIONAL. 2012. 19th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA, Official Method 984.13.

Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, M.H.H. Stevens, M.J. Oksanen and M.A.S.S. Suggests. 2007. The vegan package. Community ecology package, 10.

Orestes Cerdeira, J., A.P. Duarte Silva, J. Cadima, and M. Minhoto. 2009. subselect: selecting variable subsets. R package version 0.10-1.

Pagel, J., and F.M. Schurr. 2012. Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. Global Ecol Biogeogr. 21:293-304.

Paine, R.T. 2010. Macroecology: Does it ignore or can it encourage further ecological syntheses based on spatially local experimental manipulations. Am Nat. 176:385-393.

Pearson, R.G., C.J. Raxworthy, M. Nakamura and A.T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr. 34:102-117.

Peterson, A.T. 2003. Predicting the geography of species' invasions via ecological niche modeling. The Quarterly Review of Biology. 78:419-433.

Peterson, A.T., J. Soberón, R.G. Pearson, R.P. Anderson, E. Martínez-Meyer, M. Nakamura, and M.B. Araújo. 2011. Ecological Niches and Geographic Distributions. Princeton University Press, Princeton, NJ.

Petitpierre, B., C. Kueffer, O. Broennimann, C. Randin, C. Daehler and A. Guisan. 2012. Climatic niche shifts are rare among terrestrial plant invaders. Science. 335:1344-1348.

Pettey, F.W. 1948. The biological control of prickly pears in South Africa. Dept. of Agriculture, Union of S. Africa Entomology Series No. 22.

Phillips, S.J., R.P. Anderson and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecol Model. 190:231-259.

Pigot, A.L., and J.A. Tobias. 2013. Species interactions constrain geographic range expansion over evolutionary time. Ecol Lett. 16:330-338.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. J N Am Benthol Soc. 391-409.

Pohlert, T. 2014. The pairwise multiple comparison of mean ranks package (PMCMR). R package. http://CRAN.R-project.org/package=PMCMR

Powell, A.M., and J.F. Weedin. 2004. Cacti of the Trans-Pecos and Adjacent Areas. Texas Tech University Press, Lubbock, TX.

Powell A.M., J.F. Weedin and S.A. Powell. 2008. Cacti of Texas: a Field Guide. Texas Tech University Press, Lubbock, TX.

R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBM 3900051-07-0, http://www.R-project.org.

Rand, W.M. 1971. Objective criteria for the evaluation of clustering methods. J Am Stat Assoc. 66:846-850.

Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2014.

Rebman, J.P., and D.J. Pinkava. 2001. *Opuntia* cacti of North America – An overview. Fla Entomol. 84:474-483.

Robertson, H.G. 1987. Oviposition site selection in *Cactoblastis cactorum* (Lepidoptera): constraints and compromises. Oecologia. 73:601-608.

Rose, K.E., S.M. Louda and M. Rees. 2005. Demographic and evolutionary impacts of native and invasive herbivores on *Cirsium canescens.* Ecology. 86:453-465.

Rose, R.I. 2009. Eradication of South American Cactus Moth, *Cactoblastis cactorum*, from 11 Parishes in Southeastern Louisiana. United States Department of Agriculture. http://www.aphis.usda.gov/plant_health/ea/downloads/cactoblastis-ea-louisiana.pdf.

Rose, R., R. Weeks and S. Usnick. 2011. Cactus moth, *C. cactorum* 2011 survey plan for PPQ and state cooperators. APHIS, USDA, Riverdale.

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 and D.E. McCauley. 2001. The population biology of invasive species. Ann Rev Ecol Syst. pp.305-332.

Sarvary, M.A., K.A. Bloem, S. Bloem, J.E. Carpenter, S.D. Hight and S. Dorn. 2008a. Diel flight pattern and flight performance of Cactoblastis cactorum (Lepidoptera: Pyralidae) measured on a flight mill: influence of age, gender, mating status, and body size. J Econ Entomol. 101:314-324.

Sarvary, M.A., S.D. Hight, J.E. Carpenter, S. Bloem, K.A. Bloem and S. Dorn. 2008b. Identification of factors influencing flight performance of field-collected and laboratory-reared, overwintered, and nonoverwintered cactus moths fed with field-collected host plants. Environ Entomol. 37:1291-1299.

Sauby, K.E., T.D. Marsico, G.N. Ervin and C.P. Brooks. 2012. The role of host identity in determining the distribution of the invasive moth *Cactoblastis cactorum* (Lepidoptera:Pyralidae) in Florida. Fla Entomol. 95:561-568.

Sax, D.F., J.J. Stachowicz, J.H. Brown, J.F. Bruno, M.N. Dawson, S.D. Gaines, R.K. Grosberg, A. Hastings, R.D. Holt, M.M. Mayfield and M.I. O'Connor. 2007. Ecological and evolutionary insights from species invasions. Trends Ecol Evol. 22:465-471.

Schoener, T.W. 1968. Anolis lizards of Bimini: resource partitioning in a complex fauna. Ecology. 49:704-726.

Schurr, F.M., J. Pagel, J.S. Cabral, J. Groeneveld, O. Bykova, R.B. O'Hara, F. Hartig, W.D. Kissling, H.P. Linder, G.F. Midgley and B. Schröder. 2012. How to understand species' niches and range dynamics: a demographic research agenda for biogeography. J Biogeogr. 39:2146-2162.

Sexton, J.P., P.J. McIntyre, A.L. Angert and K.J. Rice. 2009. Evolution and ecology of species range limits. Annu Rev Ecol Evol. Syst. 40:415-436.

Shea, K., and P. Chesson, 2002. Community ecology theory as a framework for biological invasions. Trends Ecol Evol. 17:170-176.

Sih, A., D.I. Bolnick, B. Luttbeg, J.L. Orrock, S.D. Paecor, L.M. Pintor, E. Preisser, J.S. Rehage and J.R. Vonesh. 2010. Predator-prey naivete, antipredator behaviour, and the ecology of predator invasions. Oikos. 119:610-621.

Silva, D.P., V.H. Gonzalez, G.A.R. Melo, M. Lucia, L.J. Alvarez and P. De Marco Jr. 2014. Seeking the flowers for the bees: Integrating biotic interactions into niche models to assess the distribution of the exotic bee species *Lithurgus huberi* in South America. Ecol Model. 273:200-209.

Simberloff, D. 2009. The role of propagule pressure in biological invasions. Annu Rev Ecol Evo Syst. 40:81-102.

Simberloff, D., J.L. Martin, P. Genovesi, V. Maris, D.A. Wardle, J. Aronson, F. Courchamp, B. Galil, E. García-Berthou, M. Pascal and P. Pyšek. 2013. Impacts of biological invasions: what's what and the way forward. Trends Ecol Evol. 28:58–66.

Simmonds, F.J., and F.D. Bennett. 1966. Biological control of *Opuntia* spp. by *Cactoblastis cactorum* in the Leeward Islands (West Indies). Entomophaga. 11:183-189.

Simonsen, T.J., R.L. Brown and F.A. Sperling. 2008. Tracing an invasion: phylogeography of *Cactoblastis cactorum* (Lepidoptera: Pyralidae) in the United States based on mitochondrial DNA. Ann Entomol Soc Amer. 101:899-905

Simonson, S.E., T.J. Stohlgren, L. Tyler, W.P. Gregg, R. Muir and L.J. Garrett. 2005. Preliminary assessment of the potential impacts and risks of the invasive cactus moth, *Cactoblastis cactorum* Berg, in the US and Mexico. International Atomic Energy Agency, Vienna, Austria.

Simpson, G.L., J. Oksanen and M.G.L. Simpson. 2015. Package 'analogue'. https://github.com/gavinsimpson/analogue

Soberon, J., J. Golubov and J. Sarukhan. 2001. The importance of *Opuntia* in Mexico and routes of invasion and impact of *Cactoblastis cactorum* (Lepidoptera: Pyralidae). Fla Entomol. 84:486-492.

Soberón, J., and A.T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform. 2:1-10.

Soberon, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecol Lett. 10:1115-1123.

Soberón, J., and M. Nakamura. 2009. Niches and distributional areas: concepts, methods, and assumptions. P Natl Acad Sci USA. 106:19644-19650.

Stang, M., P.G. Klinkhamer and E. Van Der Meijden. 2006. Size constraints and flower abundance determine the number of interactions in a plant–flower visitor web. Oikos. 112:111-121.

Stephens, F.A., A.M. Woodard and T.D. Marsico. 2012. Comparison between eggsticks of two cactophagous moths, *Cactoblastis cactorum* and *Melitara prodenialis* (Lepidoptera: Pyralidae). Fla Entomol. 95:939-943.

Stewart-Koster, B., J.D. Olden and P.T. Johnson. 2015. Integrating landscape connectivity and habitat suitability to guide offensive and defensive invasive species management. J Appl Ecol. 52:366-378.

Stigall, A.L. 2014. When and how do species achieve niche stability over long time scales? Ecography. 37:1123-1132.

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

Strubbe, D., O. Broennimann, F. Chiron and E. Matthysen. 2013. Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. Global Ecol. Biogeogr. 22:962-970.

Stuppy, W. 2002. Seed characters and the generic classification of the Opuntioideae (Cactaceas). In D.R.Hunt and N.P. Taylor [eds.], Studies in the Opuntioideae (Cactaceae), David Hunt, The Manse, Chapel Lane, Milbourne Port Sherborne, UK. pp. 25-58.

Tingley, R., M. Vallinoto, F. Sequeira and M.R. Kearney. 2014. Realized niche shift during a global biological invasion. P Natl Acad Sci USA. 111:10233-10238.

Torchin M.E., K.D. Lafferty, A.P. Dobson, V.J. McKenzie and A.M. Kuris. 2003. Introduced species and their missing parasites. Nature. 421:628–630.

Trainor, A.M, and O.J. Schmitz. 2014. Infusing considerations of trophic dependencies into species distribution modelling. Ecol Lett. 17:1507-1517.

Trainor, A.M., O.J. Schmitz, J.S. Ivan and T.M. Shenk. 2014. Enhancing species distribution modeling by characterizing predator-prey interactions. Ecol Appl. 24:204-226.

Trevino, V. and F. Falciani. 2006. GALGO: An R package for multivariate variable selection Using genetic algorithms. Bioinformatics. 22:1154-1156.

Vaclavik, T., and R.K. Meetenmeyer. 2012. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Divers Distrib. 18:73-83.

Varone, L., M.M. Acosta, G.A. Logarzo, J.A. Briano, S.D. Hight and J.E. Carpenter. 2012. Laboratory performance of *Cactoblastis cactorum* (Lepidoptera: Pyralidae) on South and North American *Opuntia* species occurring in Argentina. Fla Entomol. 95:1163-1173.

Verberk, W.C.E.P., C.G.E. Van Noordwijk and A.G. Hildrew. 2013. Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science. Freshwater Sci. 32:531-547.

Violle, C., M.L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel and E. Garnier. 2007. Let the concept of trait be functional!.Oikos. 116:882-892.

Wang, L., and D. A. Jackson. 2014. Shaping up model transferability and generality of species distribution modeling for predicting invasions: implications from a study on *Bythotrephes longimanus*. Biol Invasions. 16:2079-2103.

Warren, D.L., R.E. Glor and M. Turelli. 2008. Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. Evolution. 62:2868-2883.

Webb, C.T., J.A. Hoeting, G.M. Ames, M.I. Pyne and N. LeRoy Poff. 2010. A structured and dynamic framework to advance traits-based theory and prediction in ecology. Ecol Lett. 13:267-283.

Werner, E.E., and S.D. Peacor. 2003. A review of trait-mediated indirect interactions in ecological communities. Ecology. 84:1083-1100.

Westoby, M., and I.J. Wright. 2006. Land-plant ecology on the basis of functional traits. Trends Ecol Evol. 21:261-268.

Wiens, J.J., and C. Graham. 2005. Niche conservatism: Integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evo. Syst. 36:519-539.

Wiens J.J., D.D. Ackerly, A.P. Allen, B.L. Anacker, L.B. Buckley, H.V. Cornell, E.I. Damschen, T. Jonathan Davies, J.A. Grytnes, S.P. Harrison and B.A. Hawkins. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecol Lett. 13:1310-1324.

Williamson, M. 1999. Invasions. Ecography. 22:5-12.

Wilson, J.R., E.E. Dormontt, P.J. Prentis, A.J. Lowe and D.M. Richardson. 2009. Something in the way you move: dispersal pathways affect invasion success. Trends Ecol Evol. 24:136-144.

Wisz, M.S., J. Pottier, W.D. Kissling, L. Pellissier, J. Lenoir, C.F. Damgaard, C.F. Dormann, M.C. Forchhammer, J.A. Grytnes, A. Guisan, R.K Heikkinen, T.T. Hoye, I. Kuhn, M. Luoto, L. Maiorano, M.C. Nilsson, S. Normand, E. Ockinger, N.M. Schmidt, M. Termasen, A. Timmermann, D.A. Wardle, P. Aastrup and J.C. Svenning. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biol Rev. 88:15-30.

Wootton, J.T. 1994. The nature and consequences of indirect effects in ecological communities. Ann Rev Ecol Syst. 25:443-466.

Zeileis, A., and T. Hothorn. 2002. Diagnostic checking in regression relationships. R News. 2:7-10.

Zimmermann, H., S. Bloem and H. Klein. 2004. Biology, history, threat, surveillance and control of the cactus moth, *Cactoblastis cactorum*. Food and Agriculture Organization of the United Nations (FAO).

APPENDIX A

CHAPTER 2 SUPPLEMENTARY INFORMATION

R CODE

Evaluating if Opuntia tissue macronutrient traits predict cactus moth infestation

The following code was used to evaluate if *Opuntia* tissue macronutrient traits predict infestation by *M. prodenialis* or *C. cactorum*.

```
#LOAD APPROPRIATE PACKAGES
library(glmulti) #Version 1.0.7
library(MuMIn) #Version 1.15.6

#IMPORT DATA
setwd('working directory')
proximate=read.csv("working file",header=T)
attach(proximate)
head(proximate)

#EXAMINE TISSUE MACRONUTRIENT RELATIVE TO M. PRODENIALIS
meprall=glmulti(MePR~protein*fiber*fat,family='binomial',data=proximate,
        method='h')
meprtable=weightable(meprall)
ofinterest=meprtable[meprtable$aic <= min(meprtable$aic)+2,]
mepr1=glm(MePR ~ 1 + protein,data=proximate)
mepr2=glm(MePR ~ 1,data=proximate)
mepr3=glm(MePR ~ 1 + fiber:protein,data=proximate)
mepr4=glm(MePR ~ 1 + protein + fiber:protein,data=proximate)
mepr5=glm(MePR ~ 1 + protein + fat:protein,data=proximate)
mepr6=glm(MePR ~ 1 + protein + fiber,data=proximate)
mepr7=glm(MePR ~ 1 + fiber:protein + fat:fiber,data=proximate)
mepr8=glm(MePR ~ 1 + fiber,data=proximate)
mepr9=glm(MePR ~ 1 + fat:protein,data=proximate)
mepr10=glm(MePR ~ 1 + protein + fat,data=proximate)
mepr11=glm(MePR ~ 1 + protein + fat:fiber,data=proximate)
meprfinal=model.avg(mepr1,mepr2,mepr3,mepr4,mepr5,mepr6,mepr7,mepr8,
        mepr9,mepr10,mepr11)
meprgood=meprfinal$formula
meprgood=glm(MePR ~ fat + fiber + protein + fat:fiber + fat:protein +
          fiber:protein)
meprpredict=predict(meprgood,newdata=nutrients,type='response')
meprobserved=MePR
difference=meprobserved-meprpredict
meprMSS=sum(difference^2)
meprTSS=sum((meprobserved-mean(MePR))^2)
meprrsquared=meprMSS/meprTSS
```

```
#EXAMINE TISSUE MACRONUTRIENT TRAITS RELATIVE TO C.
        CACTORUM
ccall=glmulti(CC~protein*fiber*fat,family='binomial',data=proximate,
        method='h')
cctable=weightable(ccall)
ofinterestcc=cctable[cctable$aic <= min(cctable$aic)+2,]
cc1=glm(CC ~ 1 + fiber + fat:protein,data=proximate)
cc2=glm(CC ~ 1 + fiber:protein + fat:fiber,data=proximate)
cc3=glm(CC ~ 1 + fiber + fat,data=proximate)
cc4=glm(CC ~ 1 + fat + fiber:protein,data=proximate)
cc5=glm(CC ~ 1 + fiber + fat:fiber,data=proximate)
cc6=glm(CC ~ 1 + fiber:protein,data=proximate)
cc7=glm(CC ~ 1 + protein + fat + fiber:protein,data=proximate)
cc8=glm(CC ~ 1 + fiber + fat + fiber:protein,data=proximate)
cc9=glm(CC ~ 1 + fat:fiber,data=proximate)
cc10=glm(CC ~ 1 + fat + fiber:protein + fat:protein,data=proximate)
cc11=glm(CC ~ 1 + protein + fiber + fat,data=proximate)
cc12=glm(CC ~ 1 + protein + fiber:protein,data=proximate)
cc13=glm(CC ~ 1 + protein + fiber:protein + fat:protein,data=proximate)
cc14=glm(CC ~ 1 + protein + fiber:protein + fat:fiber,data=proximate)
cc15=glm(CC ~ 1 + fiber,data=proximate)
cc16=glm(CC ~ 1 + fiber + fat + fat:protein,data=proximate)
cc17=glm(CC ~ 1 + fiber + fiber:protein + fat:fiber,data=proximate)
cc18=glm(CC ~ 1 + fiber:protein + fat:protein,data=proximate)
cc19=glm(CC ~ 1 + fiber + fiber:protein + fat:protein,data=proximate)
cc20=glm(CC ~ 1 + fiber + fat:protein + fat:fiber,data=proximate)
cc21=glm(CC ~ 1 + protein + fiber + fat:protein,data=proximate)
ccfinal=model.avg(cc1,cc2,cc3,cc4,cc5,cc6,cc7,cc8,cc9,cc10,cc11,cc12,cc13,cc14
        ,cc15,cc16,cc17,cc18,cc19,cc20,cc21)
goodcc=ccfinal$formula
goodcc=glm(CC ~ fat + fiber + protein + fat:fiber + fat:protein + fiber:protein)
ccpredict=predict(goodcc,newdata=nutrients,type='response')
ccobserved=CC
ccdifference=ccpredict-ccobserved
ccMSS=sum((ccdifference^2))
ccTSS=sum((ccobserved-mean(CC))^2)
ccrsquared=ccMSS/ccTSS
```

Evaluating if Opuntia morphological traits predict cactus moth infestation

The following code was used to evaluate if *Opuntia* morphological traits predict infestation by *M. prodenialis* or *C. cactorum*.

```
#LOAD APPROPRIATE PACKAGES
library(glmulti) #Version 1.0.7

#IMPORT DATA
setwd('working directory')
morpho=read.csv("working file",header=T)
attach(morpho)

#EXAMINE MORPHOLOGICAL TRAITS RELATIVE TO M. PRODENIALIS
#REMEMBER TO USE “EXCLUDE” ARGUMENT TO REMOVE
        CORRELATED VARIABLES
bestmepr=glmulti("Melitara",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        data=morpho,family='binomial',method='g')
bestmepr1=glmulti("Melitara",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        ,data=morpho,family='binomial',method='g')
bestmepr2=glmulti("Melitara",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        ,data=morpho,family='binomial',method='g')
bestmepr3=glmulti("Melitara",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
         data=morpho,family='binomial',method='g')
bestmepr4=glmulti("Melitara",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
,data=morpho,family='binomial',method='g')

#ARRIVE AT CONSENSUS BEST MODEL
listmepr<-list(bestmepr,bestmepr1,bestmepr2,bestmepr3,bestmepr4)
meprconsensus=consensus(listmepr)

#EXAMINE MORPHOLOGICAL TRAITS RELATIVE TO C. CACTORUM
# REMEMBER TO USE “EXCLUDE” ARGUMENT TO REMOVE
        CORRELATED VARIABLES
```

```
bestcacto=glmulti("Cacto",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        data=morpho,family='binomial',method='g')
bestcacto1=glmulti("Cacto",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
         data=morpho,family='binomial',method='g')
bestcacto2=glmulti("Cacto",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        data=morpho,family='binomial',method='g')
bestcacto3=glmulti("Cacto",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        data=morpho,family='binomial',method='g')
bestcacto4=glmulti("Cacto",xr=c('Mean.spines.per.areole','Median.length',
        'Mean.length.of.10.spines','Longest.Spine.Length','Spine.Color',
        'Spines.per.areole','Spine.Pattern','Cladode.Color','Cladode.Shape',
        'Plant.Height','Spine.Shape','Growth.Form','Spine.Persistence'),
        data=morpho,family='binomial',method='g')

# ARRIVE AT CONSENSUS BEST MODEL
listcacto<-list(bestcacto,bestcacto1,bestcacto2,bestcacto3,bestcacto4)
5cactoconsensus=consensus(listcacto)
#EVALUATE CONSENSUS BEST MODEL
```

APPENDIX B

CHAPTER 3 SUPPLEMENTARY INFORMATION

R CODE

Generating Maxent predictions for each taxa

The following code was used to generate Maxent predictions for each taxa modeled in this dissertation chapter. This example code was used to predict the distribution of *O. humifusa* var. *ammophila*.

```
#LOAD LIBRARIES
library(dismo) #Version 1.1.1
library(rgdal) # for readOGR(...), Version 1.1-10
library(SDMTools) #Version 1.1-221

# IMPORT REGION SHAPEFILE
setwd("working directory")
states=readOGR(dsn='states.shp',layer='states')
region=states[c(37,41:46,48,49),]
region <- spTransform(region,CRS('+proj=longlat +datum=WGS84'))

# IMPORT OCCURRENCE DATA FOR TAXA OF INTEREST
setwd("working directory")
stuff=read.csv('working file',header=T)
attach(stuff)
coords=stuff[,2:3]

#IMPORT ENVIRONMENTAL PREDICTORS (BIOCLIM) AND EXTRACT
        RELEVANT INFORMATION
files<list.files(path='working directory',pattern='grd',
        full.names=TRUE)
bio=stack(files)
predictors<-extract(bio,coords)
head(predictors)
ofinterest=predictors
everything=cbind(stuff[1:6],ofinterest)
#PREP OCCURRENCE AND ENVIRONMENTAL DATA FOR MODELING
goodtogo=cbind(coords,ofinterest)
set.seed(2000)
backgr=spsample(region,10000,type='random')
backgrcoords=coordinates(backgr)

#GENERATE MAXENT PREDICTION
modelaccuracy=c(rep(1,nrow(coords)),rep(0,nrow(backgrcoords)))
```

```
colnames(backgrcoords)<-c('ln','lat')
allpoints=data.frame(cbind(modelaccuracy,rbind(coords,backgrcoords)))
group_p = kfold(coords, 5)
group_a = kfold(backgrcoords, 5)
test = 3
train_p = coords[group_p!=test, c("ln","lat")]
train_a = backgrcoords[group_a!=test, c("ln","lat")]
test_p = coords[group_p==test, c("ln","lat")]
test_a = backgrcoords[group_a==test, c("ln","lat")]
xm <-maxent(bio,p=train_p,a=train_a)
plot(xm)

#EVALUATE MAXENT PREDICTION AND SAVE
e1<-evaluate(test_p,test_a,xm,bio)
e1
tr1=threshold(e1,'spec_sens')
px<-predict(bio,xm,ext=region,progress=")
px=mask(px,region)
plot(px,main='Maxent,raw values')
plot(region,add=TRUE,border='dark grey')
plot(px > tr1, main='presence/absence')
plot(region,add=TRUE,border='dark grey')
points(coords,pch='+')
points(backgrcoords,pch="x")
things=allpoints[,2:3]
modelvalues1=extract(px,things)
try2=cbind(modelaccuracy,modelvalues1)
try2=na.omit(try2)
modelaccuracy=try2[,1]
modelvalues1=try2[,2]
measure=accuracy(modelaccuracy,modelvalues1,threshold=tr1)
thresh=optim.thresh(modelaccuracy,modelvalues1,threshold=101)
measure1=accuracy(modelaccuracy,modelvalues1,
        threshold=thresh$`max.sensitivity+specificity`)
setwd("working directory")
writeRaster(px,"file name",bandorder='BIL',overwrite=T)
```

Generating mosaics of habitat suitability for *C. cactorum*

The following code was used to generate mosaics of habitat suitability for *C. cactorum* by integrating abiotic suitability and biotic interactions.

```
#LOAD LIBRARIES
library(dismo) #Version 1.1.1
library(rgdal) # for readOGR(...),Version 1.1-10
require(SDMTools) #Version 1.1-221

#IMPORT ALL RELEVANT DATA (MAXENT PREDICTIONS TO BE
        COMIBINED BY HOST AVAILABILITY SCENARIO)
setwd("working data")
files<-list.files(path='working directory',pattern='grd',full.names=TRUE)
preferred=stack(files[1],files[3],files[7],files[8],files[9],files[10],files[13],files[17]
        ,files[19])
all=stack(files[1],files[2],files[3],files[4],files[5],files[6],files[7],files[8],files[9],
        files[10],files[11],files[12],files[13],files[14],files[15],files[16],files[17],
        files[18],files[19],files[20],files[21])

#GENERATE MOSAICS AND WRITE TO DATA STORAGE
prefmean=mosaic(preferred$layer.1,preferred$layer.2,preferred$layer.3,
        preferred$layer.4,preferred$layer.5,preferred$layer.6,preferred$layer.7,
        preferred$layer.8,preferred$layer.9,fun=mean)
writeRaster(prefmean,"MaxEnt_Preferred_Mean.grd",bandorder='BIL',
        overwrite=T)
prefmaximum=mosaic(preferred$layer.1,preferred$layer.2,preferred$layer.3,
        preferred$layer.4,preferred$layer.5,preferred$layer.6,preferred$layer.7,
        preferred$layer.8,preferred$layer.9,fun=max)
writeRaster(prefmaximum,"MaxEnt_Preferred_Max.grd",bandorder='BIL',
        overwrite=T)
predmedian=mosaic(preferred$layer.1,preferred$layer.2,preferred$layer.3,
        preferred$layer.4,preferred$layer.5,preferred$layer.6,preferred$layer.7,
        preferred$layer.8,preferred$layer.9,fun=median)
writeRaster(prefmedian,'MaxEnt_Preferred_Median.grd',bandorder='BIL',
        overwrite=T)

allmean=mosaic(all$layer.1,all$layer.2,all$layer.3,all$layer.4,all$layer.5,
        all$layer.6,all$layer.7,all$layer.8,all$layer.9,all$layer.10,all$layer.11,
        all$layer.12,all$layer.13,all$layer.14,all$layer.15,all$layer.16,all$layer.17,
        all$layer.18,all$layer.19,all$layer.20,all$layer.21,fun=mean)
writeRaster(allmean,'MaxEnt_All_Mean.grd',bandorder='BIL',overwrite=T)
allmaximum=mosaic(all$layer.1,all$layer.2,all$layer.3,all$layer.4,all$layer.5,
        all$layer.6,all$layer.7,all$layer.8,all$layer.9,all$layer.10,all$layer.11,
        all$layer.12,all$layer.13,all$layer.14,all$layer.15,all$layer.16,all$layer.17,
        all$layer.18,all$layer.19,all$layer.20,all$layer.21,fun=max)
writeRaster(allmaximum,'MaxEnt_All_Max.grd',bandorder='BIL',overwrite=T)
allmedian=mosaic(all$layer.1,all$layer.2,all$layer.3,all$layer.4,all$layer.5,
        all$layer.6,all$layer.7,all$layer.8,all$layer.9,all$layer.10,all$layer.11,
```

```
        all$layer.12,all$layer.13,all$layer.14,all$layer.15,all$layer.16,all$layer.17,
        all$layer.18,all$layer.19,all$layer.20,all$layer.21,fun=median)
writeRaster(allmedian,'MaxEnt_All_Median.grd',bandorder='BIL',overwrite=T)
```

Generating PAB predictions

The following code was used to generate PAB predictions. Parameters can be varied relative to the constraints being imposed.

```
#LOAD LIBRARIES
library(dismo) #Version 1.1.1
library(rgdal) # for readOGR(...),Version 1.1-10
require(SDMTools) #Version 1.1-221

#IMPORT C. CACTORUM POINTS
setwd("working directory")
cacto=read.csv('working file',header=T)
attach(cacto)
coordinates=cacto[,2:3]

#IMPORT REGION SHAPEFILE
setwd("working directory")
states=readOGR(dsn='states.shp',layer='states')
region=states[c(37,41:46,48,49),]
region <- spTransform(region,CRS('+proj=longlat +datum=WGS84'))

#PREP C. CACTORUM OCCURRENCE POINTS
prespoints=rep.int(1,nrow(coordinates))
preslocation=cbind(coordinates,prespoints)
colnames(preslocation)<-c('ln','lat','presabs')
pseudos=spsample(region,500,type='random')
pseudocoords=coordinates(pseudos)
abpoints=rep.int(0,nrow(pseudocoords))
ablocation=cbind(pseudocoords,abpoints)
colnames(ablocation)<-c('ln','lat','presabs')
allcactopoints=rbind(preslocation,ablocation)
colnames(coordinates)<-c('ln','lat')
colnames(pseudocoords)<-c('ln','lat')
cactocoords=allcactopoints[,1:2]
cactocoords=na.omit(cactocoords)
values=allcactopoints[,3]
readycacto=SpatialPoints(list(cactocoords[,1],cactocoords[,2]))
```

```
#EXTRACT HABITAT SUITABILITY VALUES FOR ALL C. CACTORUM
        OCCURRENCE RECORDS
setwd("working directory")
files<-list.files(path='working directory',pattern='grd',full.names=TRUE)
preferredmean=raster(files[5])
preferredmax=raster(files[4])
preferredmedian=raster(files[6])
allmean=raster(files[2])
allmax=raster(files[1])
allmedian=raster(files[3])
cacto_pref_mean=rasterize(readycacto, preferredmean, values, fun='first',
        background=0, mask=F,update=F,updateValue='all',na.rm=T)
cacto_pref_max=rasterize(readycacto, preferredmax, values, fun='first',
        background=0, mask=F,update=F,updateValue='all',na.rm=T)
cacto_pref_median=rasterize(readycacto,preferredmedian,values,fun='first',
        background=0, mask=F,update=F,updateValue='all',na.rm=T)
cacto_all_mean=rasterize(readycacto, allmean, values, fun='first', background=0,
        mask=F,update=F,updateValue='all',na.rm=T)
cacto_all_max=rasterize(readycacto, allmax, values, fun='first', background=0,
        mask=F,update=F,updateValue='all',na.rm=T)
cacto_all_median=rasterize(readycacto,allmedian,values,fun='first',
        background=0, mask=F,update=F,updateValue='all',na.rm=T)
cacto_pref_mean=writeRaster(cacto_pref_mean,'cacto_pref_mean1',
        format='GTiff',overwrite=T)
cacto_pref_max=writeRaster(cacto_pref_max,'cacto_pref_max1',
        format='GTiff',overwrite=T)
cacto_pref_median=writeRaster(cacto_pref_median,'cacto_pref_median1',
        format='GTiff',overwrite=T)
cacto_all_mean=writeRaster(cacto_all_mean,'cacto_all_mean1',
        format='GTiff',overwrite=T)
cacto_all_max=writeRaster(cacto_all_max,'cacto_all_max1',
        format='GTiff',overwrite=T)
cacto_all_median=writeRaster(cacto_all_median,'cacto_all_median1',
        format='GTiff',overwrite=T)

#GENERATE DISPERSAL KERNELS
setwd("working directory")
rates=seq(0.01,1,0.1)
one=NULL
two=NULL
three=NULL
four=NULL
five=NULL
for (i in 1:10){
  one[i]=sum(dexp(1:5,rate=rates[i]))
```

```
  two[i]=sum(dexp(2:5,rate=rates[i]))
  three[i]=sum(dexp(3:5,rate=rates[i]))
  four[i]=sum(dexp(4:5,rate=rates[i]))
  five[i]=sum(dexp(5:5,rate=rates[i]))
}
all=cbind(one,two,three,four,five)
firstcorr=1/sum(all[1,])
secondcorr=1/sum(all[2,])
thirdcorr=1/sum(all[3,])
fourthcorr=1/sum(all[4,])
fifthcorr=1/sum(all[5,])
sixthcorr=1/sum(all[6,])
seventhcorr=1/sum(all[7,])
eigthcorr=1/sum(all[8,])
ninthcorr=1/sum(all[9,])
tenthcorr=1/sum(all[10,])
first=all[1,]*firstcorr
second=all[2,]*secondcorr
third=all[3,]*thirdcorr
fourth=all[4,]*fourthcorr
fifth=all[5,]*fifthcorr
sixth=all[6,]*sixthcorr
seventh=all[7,]*seventhcorr
eigth=all[8,]*eigthcorr
ninth=all[9,]*ninthcorr
tenth=all[10,]*tenthcorr
rescaled=rbind(first,second,third,fourth,fifth,sixth,seventh,eigth,ninth,tenth)

#RUNNING MIGCLIM DISPERSAL SIMULATIONS
#THIS IS WHERE DISPERSAL CONSTRAINT PARAMETERS CAN BE
        CHANGED
require(MigClim) #Version 1.6
setwd("working directory")
nsim=10
prefix<-'ScaledKernels_AllMedian_NoLDD'
suffix<-seq(1:nsim)
sim_name=paste(prefix,suffix,sep='')
for (i in 1:10){
MigClim.migrate(iniDist='cacto_all_median1',hsMap='AllMedian',rcThreshold=0
        ,envChgSteps=1,dispSteps=300,dispKernel=rescaled[i,],iniMatAge=1,
        overWrite=T,simulName=sim_name[i],replicateNb=25)
        }
```

Rescaling PAB predictions

The following code was used to rescale PAB predictions for comparison to Maxent predictions.

```
#RE-WRITE ALL .ASC FILES TO .TIF
library(raster) #Version 2.5-8
setwd("working directory")
folderslist=list.files(path='working directory',full.names=TRUE)
allfiles<-list(NULL)
for (i in 1:length(folderslist)){
  allfiles=list.files(folderslist[i],pattern='.tif',full.names=TRUE)
           }
  for (i in 1:length(allfiles)){
    layer=raster(allfiles[i])
    layer[values(layer) %in% -101:0]<-0
    layer[values(layer) %in% 30000]<-0
    layer[values(layer) %in% 1:29999]<-1
    prefix=allfiles[i]
    middle<-"Occupied"
    filetype<-'.grd'
    sim_name=paste(prefix,middle,filetype)
            }

#TAKE.TIF FILES, STACK THEM, CALCULATE MEAN OF SIMULATIONS
library(raster) #Version 2.5-8
setwd("working directory")
folderslist=list.files(path='working directory',full.names=TRUE)
for (i in 1:length(folderslist)){
  everything=list.files(folderslist[i],pattern='.tif',full.names=TRUE)
  layers<-list(NULL)
  for (i in 1:length(everything)){
    layers[i]=raster(everything[i])
    mess=stack(layers)
    messing=calc(mess,fun=mean)
    prefix=folderslist[i]
    middle='Occupied'
    sim_name=paste(prefix,middle)
    writeRaster(messing,filename=sim_name,format='GTiff',overwrite=TRUE)
  }
```

Analyzing Schoener’s D overlap among predictions

The following code was used to analyze Schoener’s D overlap among all predictions generated in this dissertation chapter.

```
#LOAD APPROPRIATE PACKAGES
library(dunn.test) #Version 1.3.2

#IMPORT APPROPRIATE DATA
setwd('working data')
comparisons=read.csv('working file',header=T)
attach(comparisons)
scores=Score

###DUNN’S TEST FOR MULTIPLE COMPARISONS
host=dunn.test(Score,Host,method='bonferroni')
hostinterest=cbind(host$comparisons,host$Z,host$P.adjusted)
write.csv(hostinterest,"hostComparison.csv")

suit=dunn.test(Score,Suit,method='bonferroni')
suitinterest=cbind(suit$comparisons,suit$Z,suit$P.adjusted)
write.csv(suitinterest,"SuitComparison.csv")

disp=dunn.test(Score,Disp,method='bonferroni')
dispinterest=cbind(disp$comparisons,disp$Z,disp$P.adjusted)
write.csv(dispinterest,"dispComparison.csv")

prob=dunn.test(Score,LDD.Prob,method='bonferroni')
probinterest=cbind(prob$comparisons,prob$Z,prob$P.adjusted)
write.csv(probinterest,"probComparison.csv")
```

Analyzing relative importance of modeling constraints

The following code was used to analyze the relative importance of the modeling constraints used to generate predictions in this dissertation chapter.

```
#IMPORT SCHOENER’S COMPARISONS
setwd('working directory')
newstuff=read.csv('working file',header=T)
attach(newstuff)
```

```
#LOAD LIBRARY
library(betareg) #Version 3.0-5

#FIT BETA REGRESSIONS, RUN LIKELIHOOD TEST
try1=betareg(Score~Host+Suit+Kernel+LDD.Prob+Disp,newstuff)
try2=betareg(Score~Host)
try3=betareg(Score~Suit)
try4=betareg(Score~Kernel)
try5=betareg(Score~LDD.Prob)
try6=betareg(Score~Disp)
lrtest(try1,try2,try3,try4,try5,try6)
```

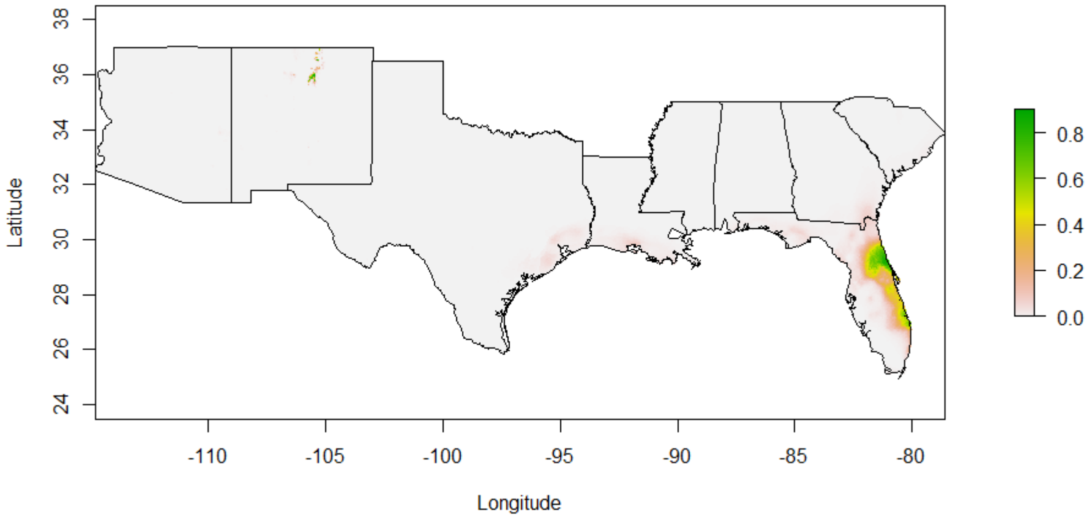

Figure B.1 Maxent prediction for *O. humifusa* var. *ammophila*.

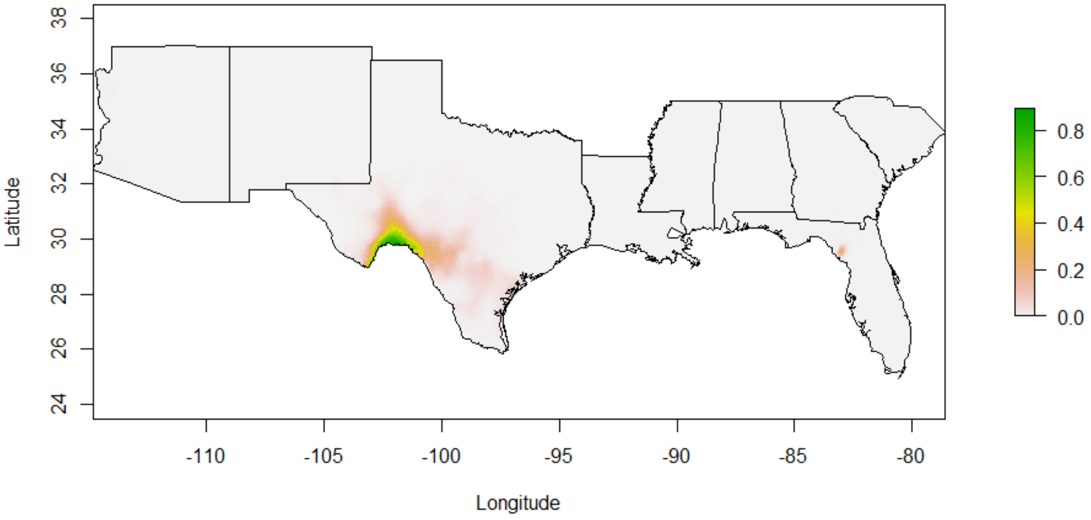

Figure B.2 Maxent prediction for *O. atrispina*.

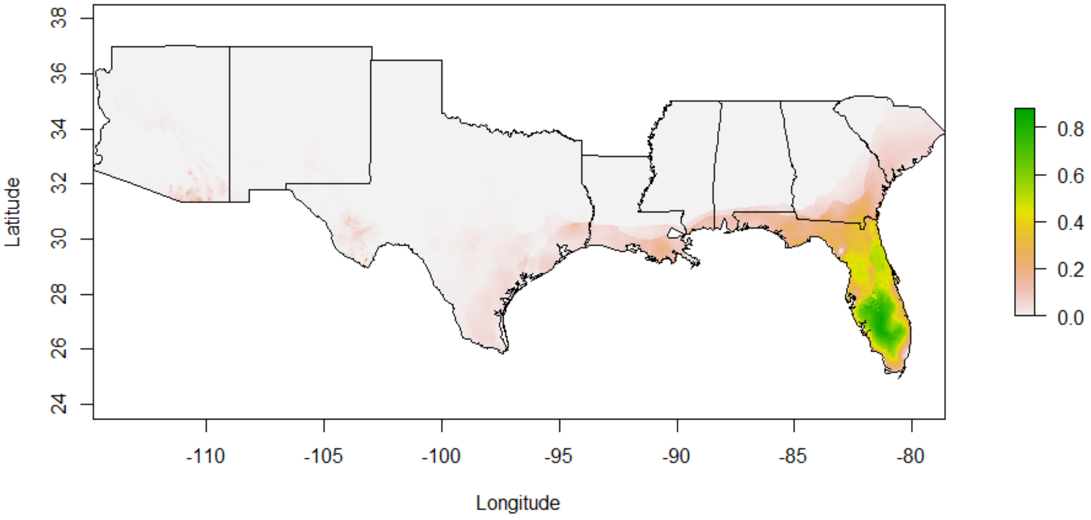

Figure B.3 Maxent prediction for *O. austrina*.

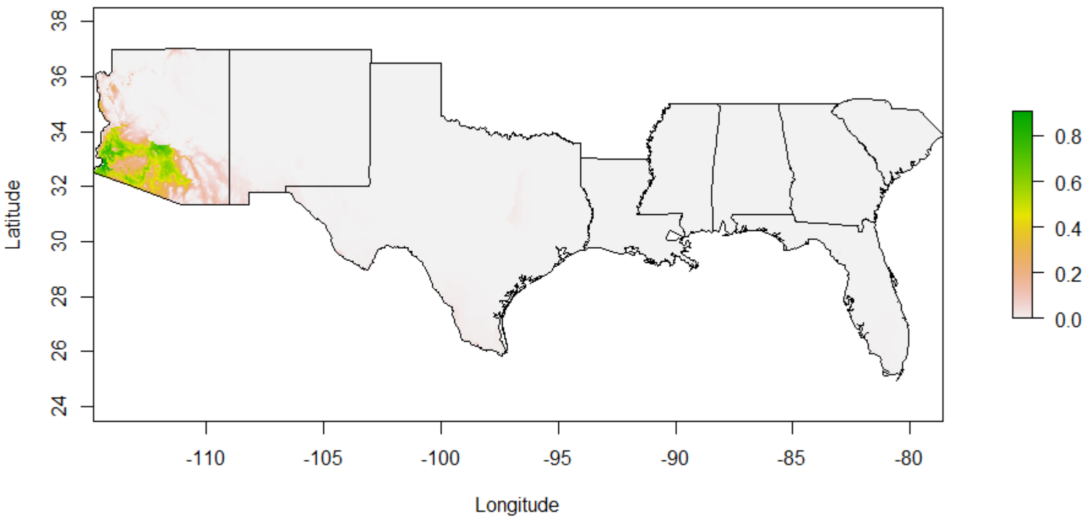

Figure B.4 Maxent prediction for *O. basilaris*.

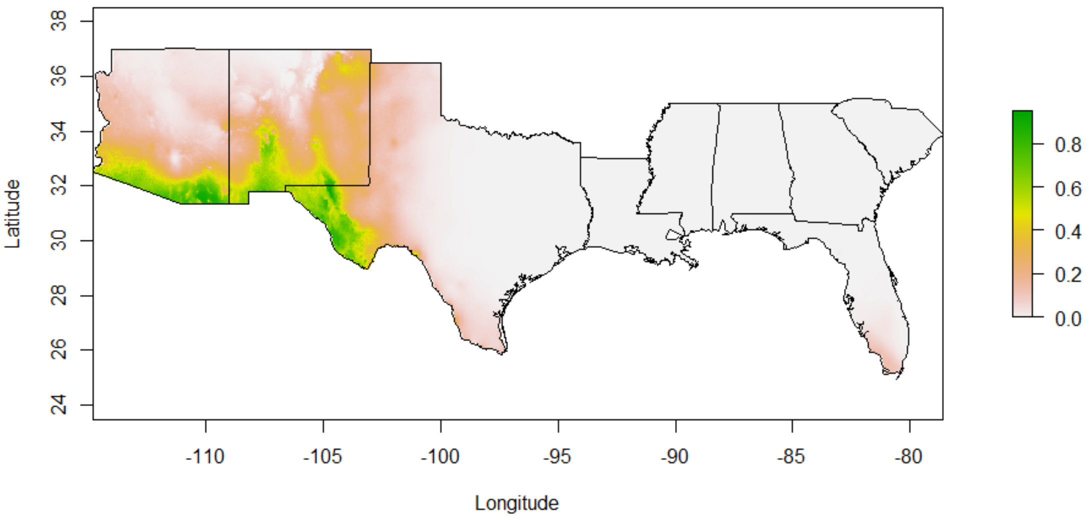

Figure B.5 Maxent prediction for *O. camanchica*.

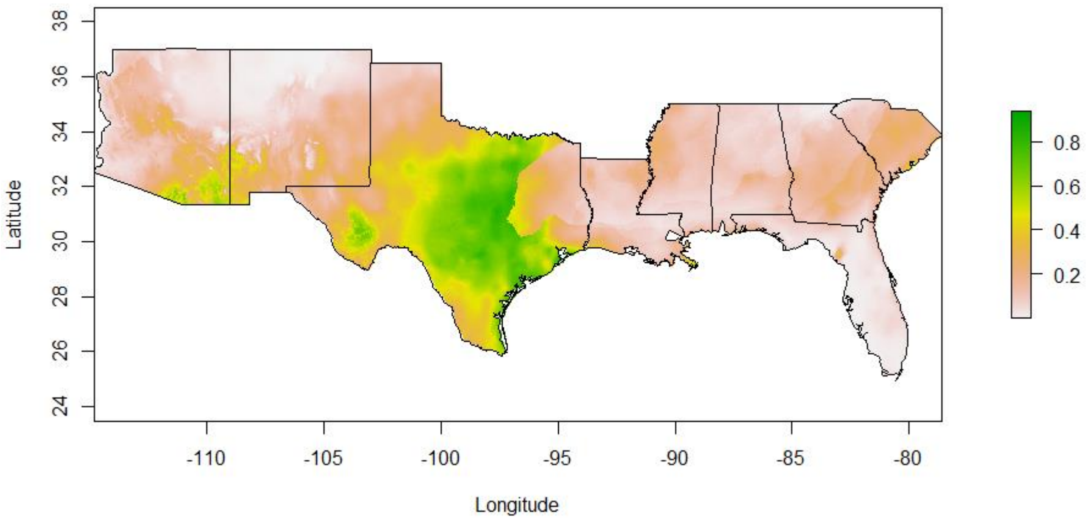

Figure B.6 Maxent prediction for *O. ellisiana*.

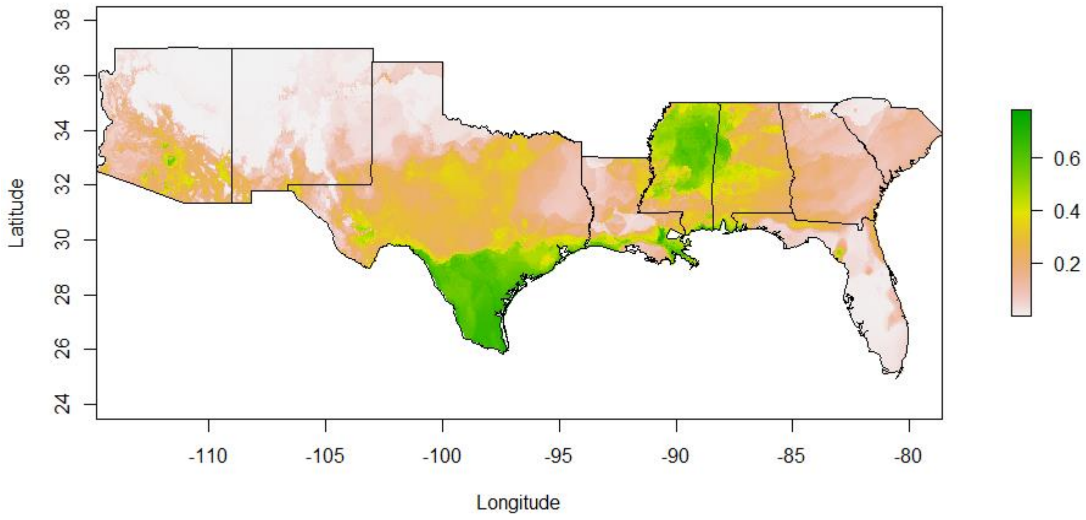

Figure B.7 Maxent prediction for *O. engelmannii* var. *engelmannii*.

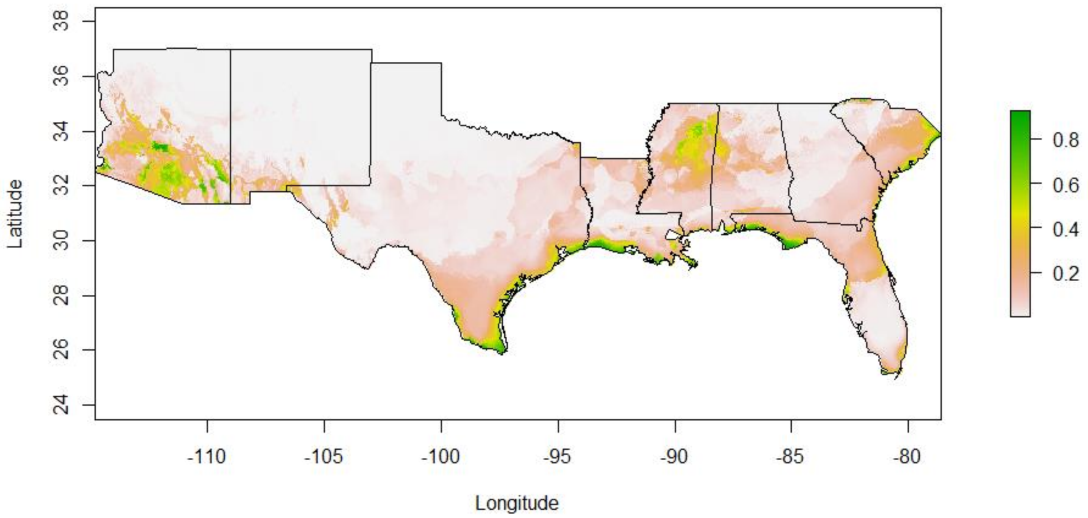

Figure B.8 Maxent prediction for *O. ficus-indica*.

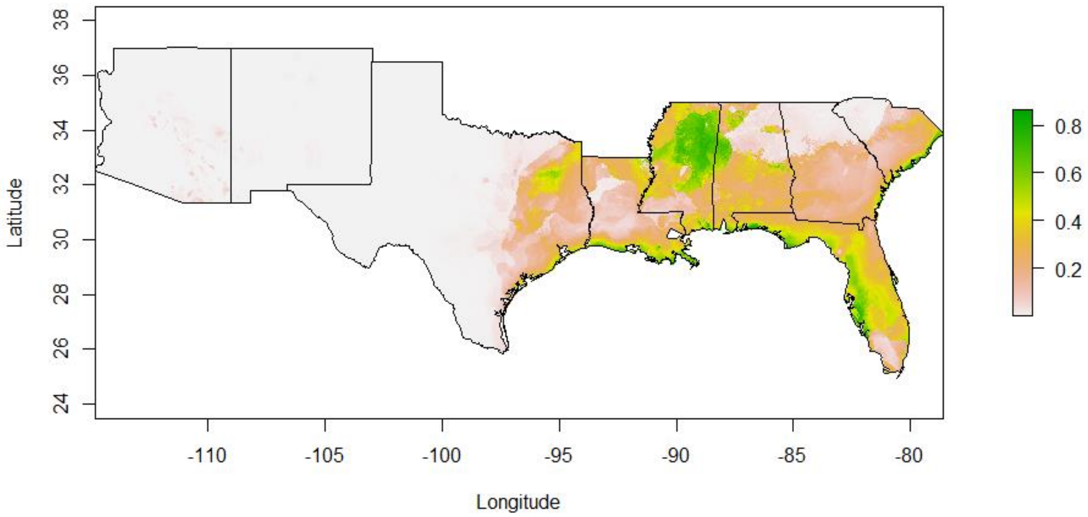

Figure B.9 Maxent prediction for *O. humifusa* var. *humifusa*.

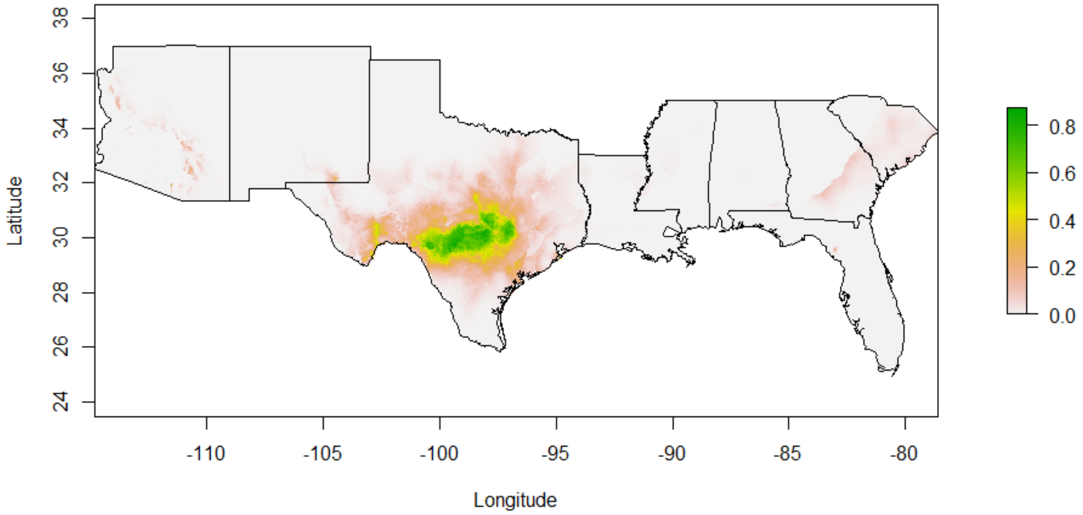

Figure B.10 Maxent prediction for *O. engelmannii* var. *lindheimeri*.

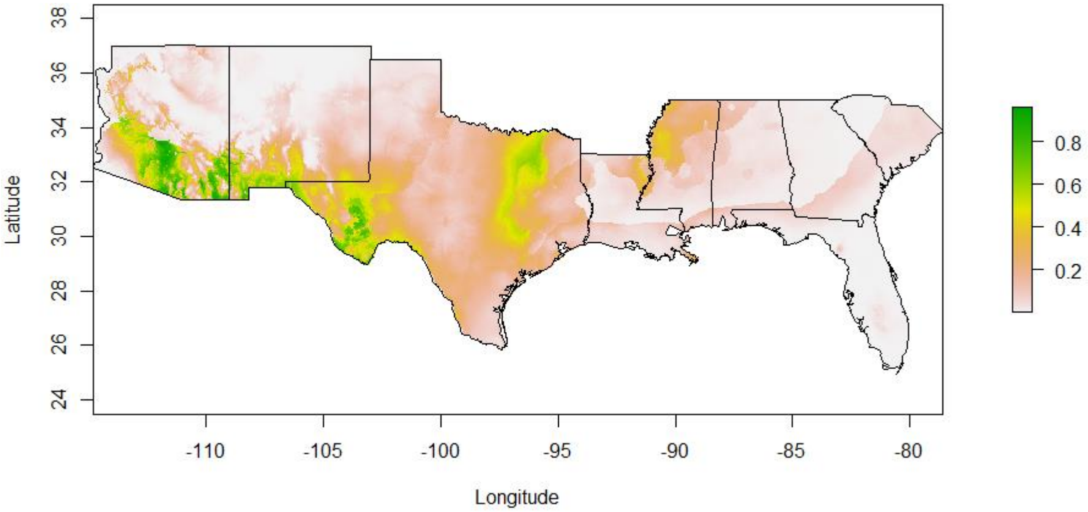

Figure B.11 Maxent prediction for *O. engelmannii* var. *linguiformis*.

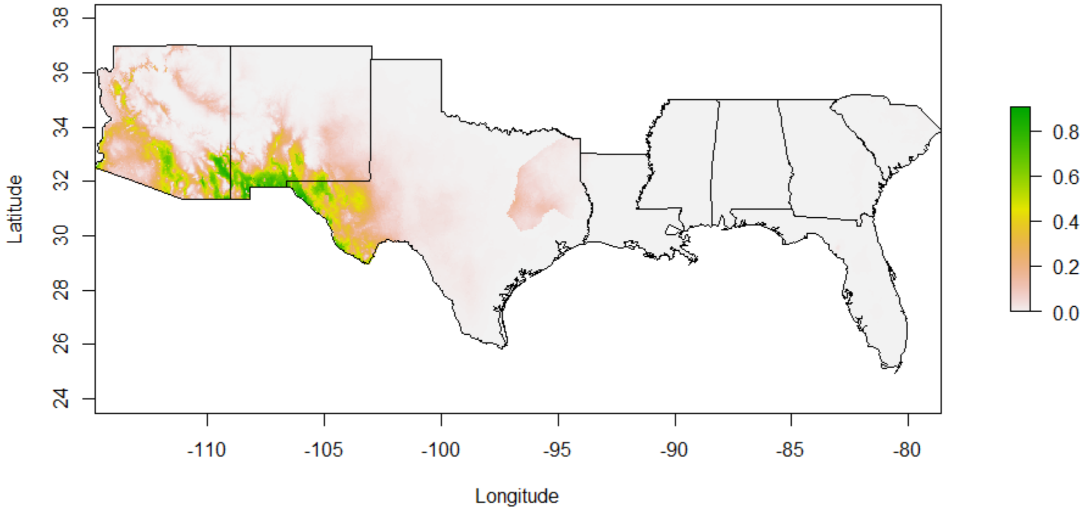

Figure B.12 Maxent prediction for *O. macrocentra*.

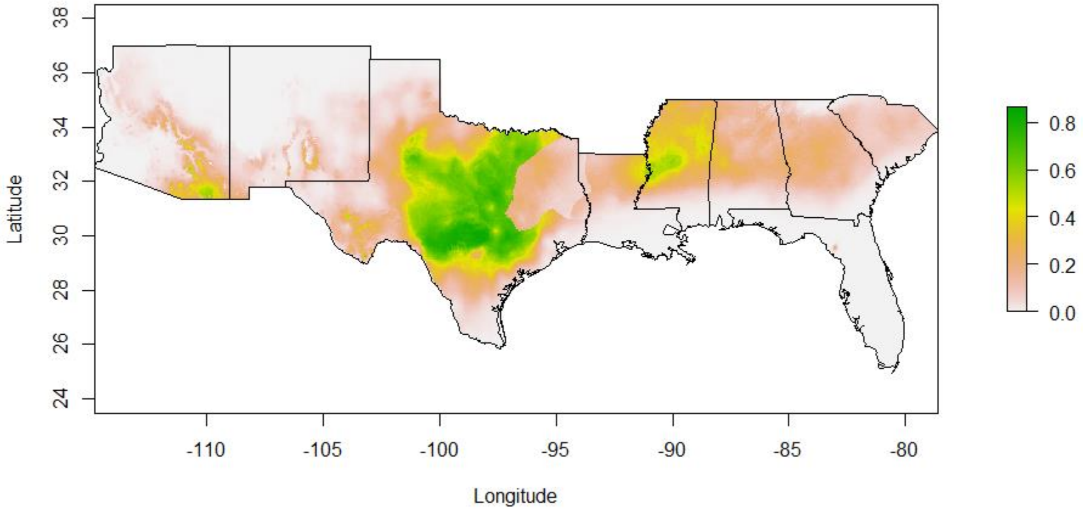

Figure B.13 Maxent prediction for *O. macrorhiza*.

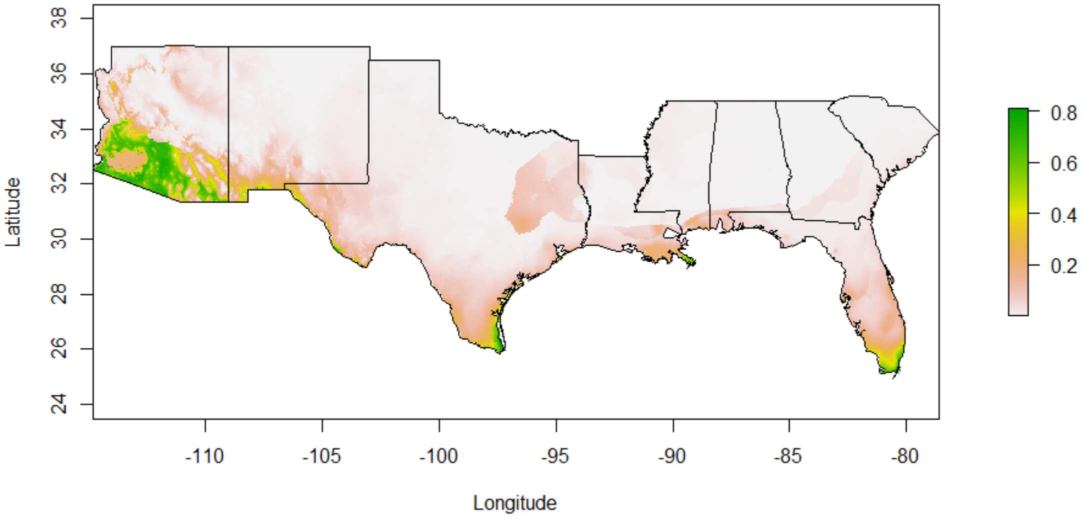

Figure B.14 Maxent prediction for *O. microdasys*.

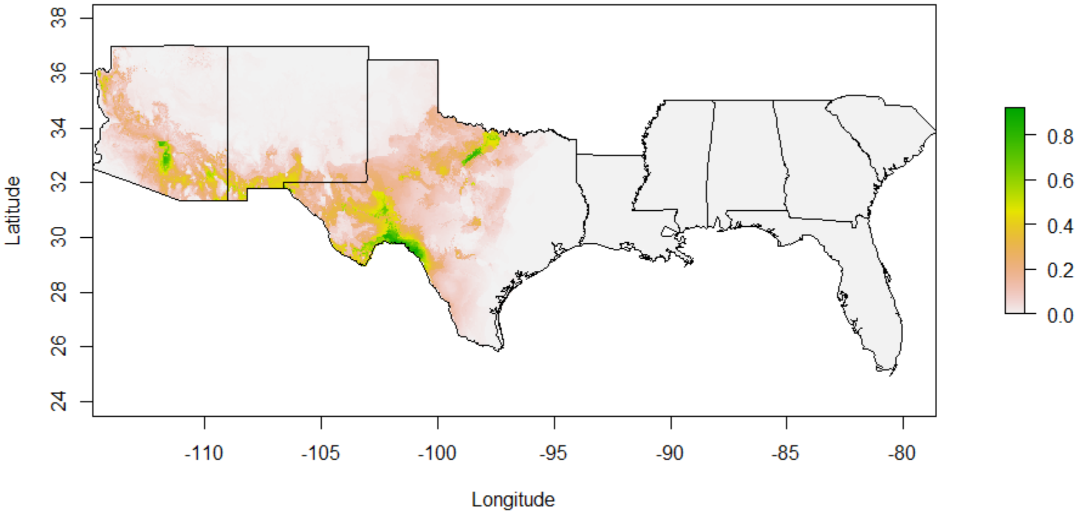

Figure B.15 Maxent prediction for *O. phaeacantha*.

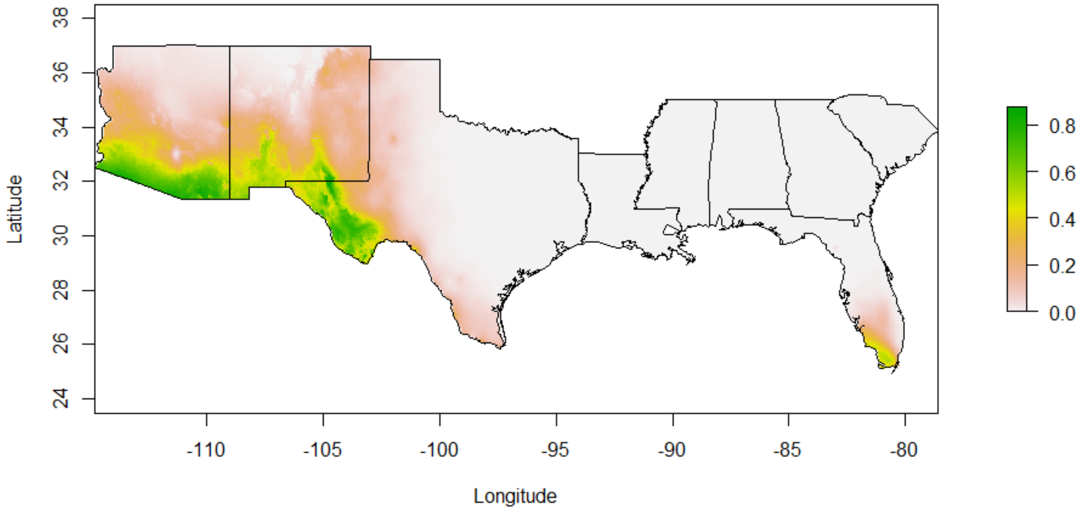

Figure B.16 Maxent prediction for *O. pottsii*.

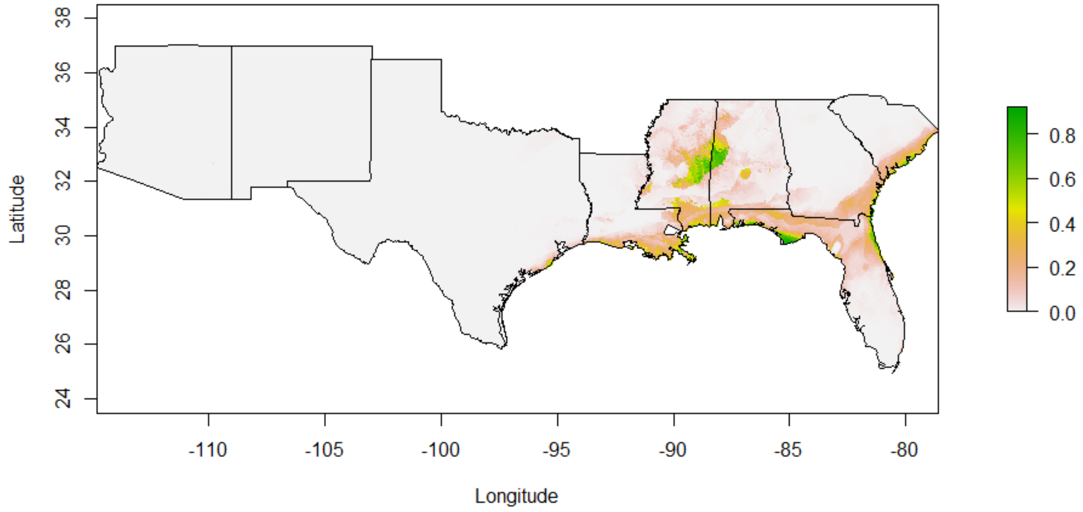

Figure B.17 Maxent prediction for *O. pusilla*.

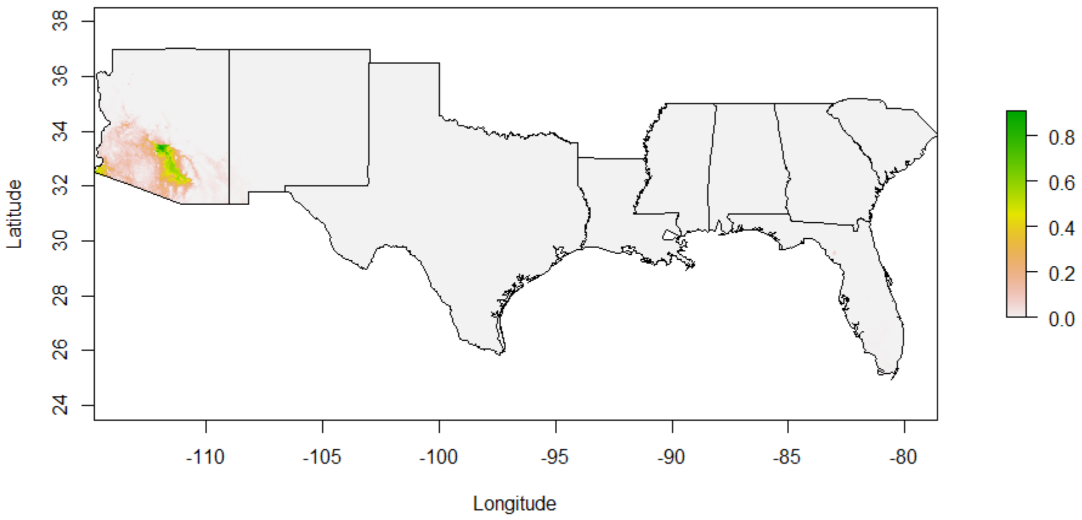

Figure B.18 Maxent prediction for *O. santa-rita*.

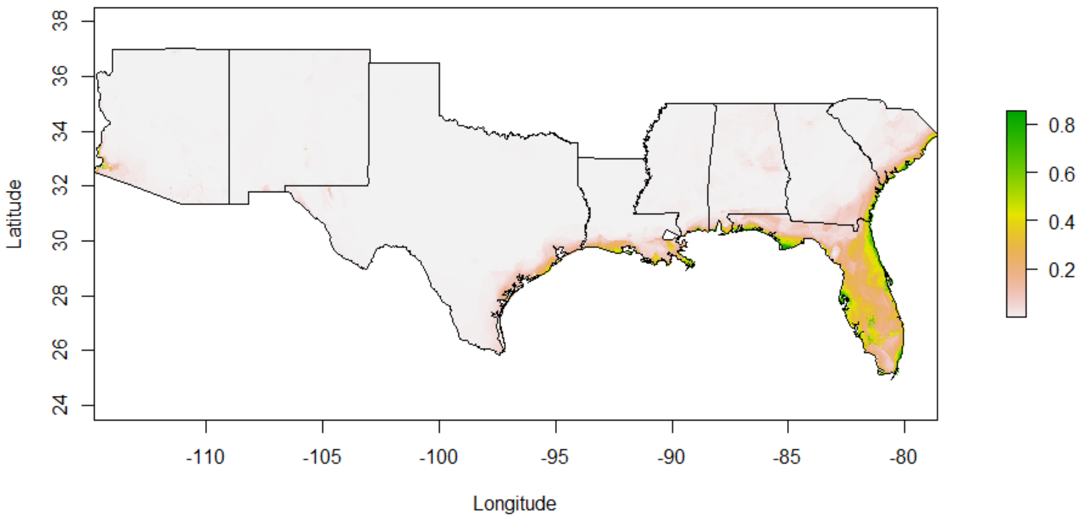

Figure B.19 Maxent prediction for *O. stricta*.

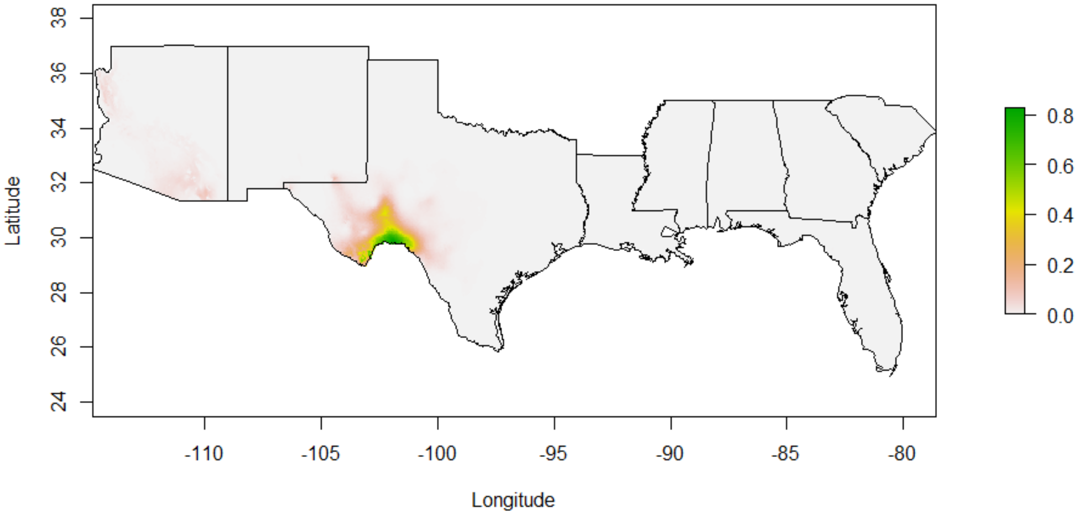

Figure B.20 Maxent prediction for *O. strigil*.

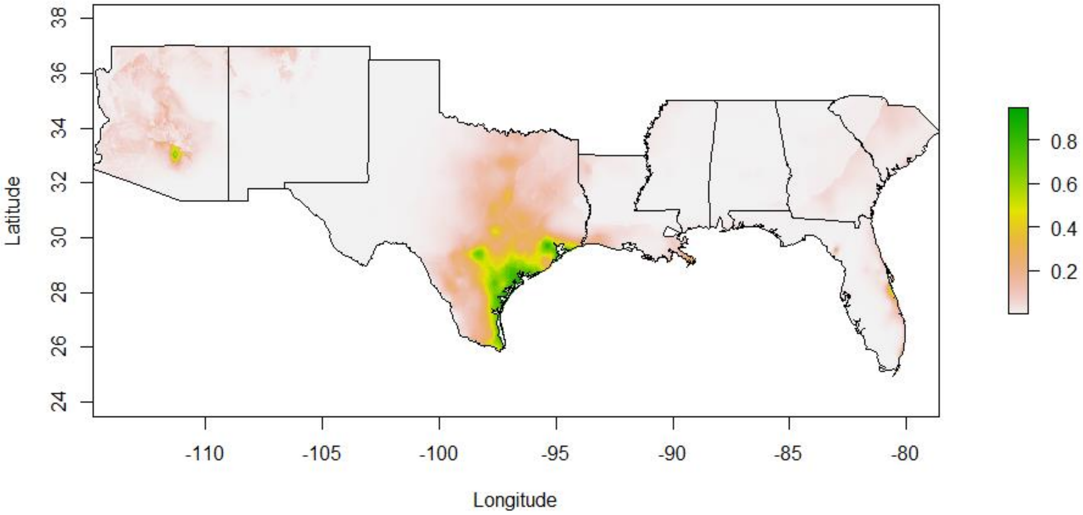

Figure B.21 Maxent prediction for *O. undulata*.

APPENDIX C

CHAPTER 4 SUPPLEMENTARY INFORMATION

R CODE

Determining the number of trait-based clusters

The following code was used to determine the centroid location for each *Opuntia* morphological and tissue macronutrient trait relative to the numer of clusters being considered.

```
#LOAD APPROPRIATE PACKAGES
library(ClustOfVar) #Version 0.8

#IMPORT DATA
setwd('working directory')
everything=read.csv('working file’,header=T)
attach(everything)

#GENERATE DENDROGRAM
traits=everything[,c(11:26)]
goodtraits=traits[c(1:9,11:16)]
quantitative=traits[,c(1:4,14:16)]
qualitative=traits[,c(5:9,11:13)]
tree <- hclustvar(quantitative,qualitative)

#EVALUATE DENDROGRAM STABILITY RELATIVE TO NUMBER OF
        CLUSTERS
stable=stability(tree,B=25,graph=TRUE)
meanscores=stable$meanCR
interest=stable$matCR
means=colMeans(interest)

#EVALUATE DIFFERENCES BETWEEN RAND SCORES
library(PMCMR) #Version 4.1
posthoc.kruskal.dunn.test(means,cluster,p.adjust.method="none")
```

Determining centroid location for *Opuntia* morphological and tissue macronutrient traits

The following code was used to determine the centroid location for each *Opuntia* morphological and tissue macronutrient trait relative to the numer of clusters being considered.

```
#LOAD PACKAGES
library(ClustOfVar) #Version 0.8

#IMPORT DATA ON OPUNTIA TRAITS
setwd("working directory")
messing=read.csv('working file',header=T)
attach(messing)

#START CLUSTERING PROCESS
traits=messing[,c(10:25)]
goodtraits=traits[c(1:9,11:16)]
quantitative=traits[,c(1:4,14:16)]
qualitative=traits[,c(5:13)]
qualitative=qualitative[,c(1:5,7:9)]
tree <- hclustvar(quantitative,qualitative)
morpho=goodtraits[,c(1:12)]
quanti_morpho=morpho[,c(1:4)]
quali_morpho=morpho[,c(5:12)]
nutrients=goodtraits[,c(13:15)]

#GENERATE CENTROIDS FOR TRAITS FROM 21 CLADODES
clust.centroid = function(i, dat, clusters) {
  ind = (clusters == i)
  colMeans(dat[ind,])
}

#GENERATE CENTROIDS FOR TISSUE MACRONUTRIENT TRAITS
          RELATIVE TO SCENARIOS OF 5  TO 12 CLUSTERS
nut_clusters=cutree(hclust(dist(nutrients)), k=5)
nutrient_centroids=sapply(unique(nut_clusters), clust.centroid,
          nutrients, nut_clusters)

#GENERATE CENTROIDS FOR QUANTITATIVE MORPHOLOGICAL
          TRAITS
quanti_clusters=cutree(hclust(dist(quanti_morpho)),k=5)
quantitative_centroids=sapply(unique(quanti_clusters),clust.centroid,
          quanti_morpho,quanti_clusters)

#GENERATE CENTROIDS FOR QUALITATIVE MORPHOLOGICAL
          TRAITS
chr2num=function(x){
  key=cbind(unique(x),as.numeric(as.factor(unique(x))))
  x=as.numeric(as.factor(x))
  return(list(data=x,key=key))}
```

```
spines_per=chr2num(quali_morpho[,1])
spine_shape=chr2num(quali_morpho[,2])
spine_color=chr2num(quali_morpho[,3])
spine_persist=chr2num(quali_morpho[,4])
spine_pattern=chr2num(quali_morpho[,5])
cladode_shape=chr2num(quali_morpho[,6])
plant_height=chr2num(quali_morpho[,7])
growth_form=chr2num(quali_morpho[,8])
new_quali=cbind(spines_per$data,spine_shape$data,spine_color$data,
        spine_persist$data,spine_pattern$data,cladode_shape$data,
        plant_height$data,growth_form$data)
quali_clusters=cutree(hclust(dist(new_quali)),k=5)
qualitative_centroids=as.data.frame(sapply(unique(quali_clusters),clust.centroid,
        new_quali,quali_clusters))
row.names(qualitative_centroids)<-c('spines_per','spine_shape','spine_color',
        'spine_persist','spine_pattern','cladode_shape','plant_height','growth_form')
final_centroids=rbind(nutrient_centroids,quantitative_centroids,
        qualitative_centroids)
```

Determine contribution of *Opuntia* morphological and tissue macronutrient traits in assigning cladode to cluster

The following code was used to evaluate the importance of *Opuntia* traits in assigning cladodes to cluster.

```
#LOAD APPROPRIATE PACKAGES
library(lmtest) #Version 0.9-34

#IMPORT DATA
setwd('working directory')
five=read.csv('StandarizedFiveClusters.csv',header=T)
six=read.csv('StandardizedSixClusters.csv',header=T)
seven=read.csv('StandardizedSevenClusters.csv',header=T)
eight=read.csv('StandardizedEightClusters.csv',header=T)
nine=read.csv('StandardizedNineClusters.csv',header=T)
ten=read.csv('StandardizedTenClusters.csv',header=T)
eleven=read.csv('StandardizedElevenClusters.csv',header=T)
twelve=read.csv('StandardizedTwelveClusters.csv',header=T)

#EVALUATE CONTRIBUTION OF TRAITS FOR EACH NUMBER OF
        CLUSTERS SCENARIO
#5 cluster scenario
attach(five)
five=as.matrix(t(five))
```

```
five=five[2:22,]
clustermembership5=as.vector(apply(five,1,which.min))
attempt1=glm(clustermembership5~Protein+Fiber+Fat+Mean.spines.per.areole+
          Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
          spines_per+spine_shape+spine_color+spine_persist+spine_pattern+
          cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership5~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership5~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern+cladode_shape+
          plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#6 cluster scenario
attach(six)
six=as.matrix(t(six))
six=six[2:22,]
clustermembership6=as.vector(apply(six,1,which.min))
attempt1=glm(clustermembership6~Protein+Fiber+Fat+Mean.spines.per.areole+
          Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
          spines_per+spine_shape+spine_color+spine_persist+spine_pattern+
          cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership6~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership6~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#7 cluster scenario
attach(seven)
seven=as.matrix(t(seven))
seven=seven[2:22,]
clustermembership7=as.vector(apply(seven,1,which.min))
attempt1=glm(clustermembership7~Protein+Fiber+Fat+Mean.spines.per.areole+
          Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
          spines_per+spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership7~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership7~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
```

```
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#8 cluster scenario
attach(eight)
eight=as.matrix(t(eight))
eight=eight[2:22,]
clustermembership8=as.vector(apply(eight,1,which.min))
attempt1=glm(clustermembership8~Protein+Fiber+Fat+Mean.spines.per.areole+
        Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
        spines_per+spine_shape+spine_color+spine_persist+spine_pattern
         +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership8~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership8~Mean.spines.per.areole+Median.length+
        Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
        spine_shape+spine_color+spine_persist+spine_pattern
         +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#9 cluster scenario
attach(nine)
nine=as.matrix(t(nine))
nine=nine[2:22,]
clustermembership9=as.vector(apply(nine,1,which.min))
attempt1=glm(clustermembership9~Protein+Fiber+Fat+Mean.spines.per.areole+
        Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
        spines_per+spine_shape+spine_color+spine_persist+spine_pattern
         +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership9~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership9~Mean.spines.per.areole+Median.length+
        Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
        spine_shape+spine_color+spine_persist+spine_pattern
         +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#10 cluster scenario
attach(ten)
ten=as.matrix(t(ten))
ten=ten[2:22,]
clustermembership10=as.vector(apply(ten,1,which.min))
attempt1=glm(clustermembership10~Protein+Fiber+Fat+Mean.spines.per.areole+
        Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
        spines_per+spine_shape+spine_color+spine_persist+spine_pattern
```

```
           +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership10~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership10~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#11 cluster scenario
attach(eleven)
eleven=as.matrix(t(eleven))
eleven=eleven[2:22,]
clustermembership11=as.vector(apply(eleven,1,which.min))
attempt1=glm(clustermembership11~Protein+Fiber+Fat+Mean.spines.per.areole+
          Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
          spines_per+spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership11~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership11~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)

#12 cluster scenario
attach(twelve)
twelve=as.matrix(t(twelve))
twelve=twelve[2:22,]
clustermembership12=as.vector(apply(twelve,1,which.min))
attempt1=glm(clustermembership12~Protein+Fiber+Fat+Mean.spines.per.areole+
          Median.length+Mean.length.of.10.spines+Longest.Spine.Length+
          spines_per+spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
attempt2=glm(clustermembership12~Protein+Fiber+Fat,family=gaussian)
attempt3=glm(clustermembership12~Mean.spines.per.areole+Median.length+
          Mean.length.of.10.spines+Longest.Spine.Length+spines_per+
          spine_shape+spine_color+spine_persist+spine_pattern
           +cladode_shape+plant_height+growth_form,family=gaussian)
try1=anova(attempt1,attempt2,attempt3,test='Chisq')
lrtest(attempt1,attempt2,attempt3)
```

Generating Maxent predictions for each tissue macronutrient-based cluster

Maxent predictions for each tissue macronutrient-based cluster were generated using code that is very similar to the code used to generate Maxent predictions for each taxa modeled in Chapter 3 (Appendix B).

Generating mosaics of habitat suitability for *C. cactorum*

Mosaics of habitat suitability for *C. cactorum* were generated using the Maxent predictions for each trait-based *Opuntia* cluster. These mosaics of habitat suitability were generated using code that is very similar to the code used to generate mosaics of habitat suitability in Chapter 3 (Appendix B).

Simulating *C. cactorum* dispersal across mosaic of habitat suitability

The following code was used to simulate *C. cactorum* dispersal across mosaics of habitat suitability.

```
#LOAD PACKAGES
library(dismo) #Version 1.1.1
library(rgdal) # for readOGR(...),Version 1.1-10
require(SDMTools) #Version 1.1-221

#IMPORT C. CACTORUM OCCURRENCE POINTS
setwd("working directory")
cacto=read.csv('working file',header=T)
attach(cacto)
coordinates=cacto[,2:3]

#IMPORT ENVIRONMENTAL PREDICTORS
files<list.files(path='working directory',pattern='grd',
        full.names=TRUE)
bio=stack(files)
predictors<-extract(bio,coordinates)

#CONVERT TEMPERATURE PREDICTORS TO CELSIUS
temp=(predictors[,1])/10
ofinterest=cbind(coordinates,temp)
cactopointstouse=na.omit(ofinterest)
```

```
#IMPORT REGION SHAPEFILE
setwd("working directory")
states=readOGR(dsn='states.shp',layer='states')
region=states[c(37,41:46,48,49),]
region <- spTransform(region,CRS('+proj=longlat +datum=WGS84'))

#USE LEGASPI AND LEGASPI (2007) TO GENERATE NUMBER
          INDIVIDUALS HATCHING FROM EGGSTICKS
setwd("working directory")
legaspi=read.csv('working file',header=T)
attach(legaspi)
temperature=c(18,22,26,30,34)
egg_female=c(33.96,41.73,31.80,30.39,10.84)

#USE STEPHENS ET AL. (2012) TO SIMULATE MEAN NUMBER OF EGGS
          PER EGGSTICK
sticksize=abs(round(rpois(nrow(cactopointstouse),63)))
step1=cbind(cactopointstouse,sticksize)

#FIT CURVES TO TEMPERATURE DATA, CALCULATE FEMALE
          SURVIVAL
plot(temperature,egg_female)
femalesurvival1=smooth.spline(temperature,egg_female)
first_femalesurv=(predict(femalesurvival1,step1$temp)$y)/100

#COMPILE INITIAL DATA
perstick_female=round(step1$sticksize*first_femalesurv)
everything=cbind(step1,perstick_female)

#IMPORT TRAIT-BASED CLUSTER SUITABILITIES AND CROP
setwd("working directory")
files<-list.files(path='working directory',pattern='grd',full.names=TRUE)
clusterfiles<-list(NULL)
for (i in 1:length(files)){
  clusterfiles[i]=raster(files[i])
}
names(clusterfiles)<-c('All Cluster Max', 'All Cluster Mean', 'All Cluster Median')
cluster_stack=stack(clusterfiles)
clusterbio=crop(bio,cluster_stack)
final_clusterstack=stack(cluster_stack,clusterbio)

#FUNCTION FOR GENERATING OFFSPRING LOCATIONS
offspring.xy<-function(parent_x,parent_y,x_females,y_females,dispdist){
  parent_xloc=rep(parent_x,x_females)
```

```
  parent_yloc=rep(parent_y,y_females)
  angle=runif(sum(x_females),min=0,max=2*pi)
  dist=rexp(sum(x_females),1/dispdist)
  offspring_lat=parent_xloc++cos(angle)*dist
  offspring_ln=parent_yloc++sin(angle)*dist
  return(cbind(offspring_ln,offspring_lat))
}

#GENERATE DISPERSING INDIVIDUALS, EXTRACT ENVIRONMENTAL
          INFO AT NEW LOCATION
dispersers<function(parent_x,parent_y,x_females,y_females,dispdist
          ,final_clusterstack){
  pointstostart=offspring.xy(parent_x,parent_y,x_females,y_females,dispdist)
  Vgood=extract(final_clusterstack,pointstostart)
  return(cbind(par_y,par_x,pointstostart,Vgood))
}

#ENSURE ALL DISPERSERS LAND IN LOCATION WITH
          ENVIRONMENTAL DATA
dispcheck<-function(offspring){
  while(sum(is.na(offspring))){
    good=offspring[complete.cases(offspring),]
    bad=offspring[!complete.cases(offspring),]
    parent_coords=bad[,1:2]
    repeats=1
    newlocs<offspring.xy(parent_coords[,2],parent_coords[,1],repeats,
          repeats,dispdist)
    preds2=extract(final_clusterstack,newlocs)
    badpos=cbind(bad[,1:2],newlocs,preds2)
    offspring=rbind(good,badpos)
  }
  return(offspring)
}

#DETERMINE DISPERSER PROBABILITY OF ESTABLISHMENT
probestab=function(good.offspring){
  loc_ofinterest=good.offspring[,3:4]
  scenario_probs=good.offspring[,5:16]
  prob_estabs=runif(nrow(scenario_probs),0,1)
  scenario_probs[scenario_probs <= prob_estabs] <-NA
  return(scenario_probs)
}

#REMOVE DISPERSERS IF PROB. ESTAB > SUITABILITY
#MODIFY CODE RELATIVE TO SUITABILITY SCENARIO
```

```
removedispersers=function(good.offspring,disp.estab) {
  offspringloc=good.offspring[,3:4]
  allmean=cbind(offspringloc,disp.estab[,2])
  all_mean=allmean[complete.cases(allmean),]
  return(all_mean)
}

#REMOVE DISPERSERS WITH "NA"
all_max=allmax[complete.cases(allmax),]
all_mean=allmean[complete.cases(allmean),]
all_median=allmedian[complete.cases(allmedian),]

#SET UP DISPERSAL OF SUBSEQUENT GENERATIONS
nextgen_info=function(next.gen){
  newcoords=next.gen[,1:2]
  newinfo=extract(bio,newcoords)
  temp_good=((newinfo[,1])/10)
  sticksize=abs(round(rpois(nrow(newinfo),63)))
  step1=cbind(newcoords,temp_good,sticksize)
  femalesurv=(predict(femalesurvival1,step1[,3])$y)/100
  female_per_stick=round(step1[,4]*femalesurv)
  nextgen_good=cbind(step1,female_per_stick)
  return(nextgen_good)
}

#FUNCTION TO RASTERIZE DISPERSER POINTS, CONVERT TO
          OCCUPIED/UNOCCUPIED
rasterize.points=function(everything,final_clusterstack){
  setwd('working directory')
  new_coords=everything[,1:2]
  new_info=rasterize(new_coords,bio,background=0)
  new_info[values(new_info) %in% 1:50000000]<-1
  writeRaster(new_info,'nsim[i]',overwrite=T)
}

#USE ALL FUNCTIONS TO RUN SIMULATION
#VARY INITIAL PARAMETERS RELATIVE TO NUMBER OF DISPERSAL
          EVENTS DESIRED
nsim=2
steps=4
for (i in 1:steps){
  parent_x=everything[,2]
  parent_y=everything[,1]
  x_females=everything[,5]
  y_females=everything[,5]
```

```
  par_x=rep(parent_x,x_females)
  par_y=rep(parent_y,y_females)
  dispdist=.3
offspring=dispersers(parent_x,parent_y,x_females,y_females,dispdist,final_clusterstack)
  good.offspring=dispcheck(offspring)
  disp.estab=probestab(good.offspring)
  next.gen=removedispersers(good.offspring,disp.estab)
  everything=nextgen_info(next.gen)
}
```

Analyzing Schoener's D overlap among predictions

The following code was used to analyze Schoener's D overlap among all predictions generated in this dissertation chapter.

```
#LOAD PACKAGES
library(dunn.test) #Version 1.3.2

#IMPORT APPROPRIATE DATA
setwd('working data')
comparisons=read.csv('working file',header=T)
attach(comparisons)

###Kruskal-Wallis tests for multiple comparisons
traits=dunn.test(Score,Traits,method='bonferroni')
traitsinterest=cbind(traits$comparisons,traits$Z,traits$P.adjusted)
write.csv(traitsinterest,"TraitComparison.csv")

step=dunn.test(Score,Step,method='bonferroni')
stepinterest=cbind(step$comparisons,step$Z,step$P.adjusted)
write.csv(stepinterest,"StepComparison.csv")

host=dunn.test(Score,Host,method='bonferroni')
hostinterest=cbind(host$comparisons,host$Z,host$P.adjusted)
write.csv(hostinterest,"hostComparison.csv")

suit=dunn.test(Score,Suit,method='bonferroni')
suitinterest=cbind(suit$comparisons,suit$Z,suit$P.adjusted)
write.csv(suitinterest,"SuitComparison.csv")

disp=dunn.test(Score,Disp,method='bonferroni')
dispinterest=cbind(disp$comparisons,disp$Z,disp$P.adjusted)
write.csv(dispinterest,"dispComparison.csv")
```

```
prob=dunn.test(Score,LDD.Prob,method='bonferroni')
probinterest=cbind(prob$comparisons,prob$Z,prob$P.adjusted)
write.csv(probinterest,"probComparison.csv")
```

Analyzing relative importance of modeling constraints

The following code was used to analyze the relative importance of the modeling constraints used to generate predictions in this dissertation chapter.

```
#LOAD LIBRARY
library(betareg) #Version 3.0-5
library(lmtest) #Version 0.9-34

#IMPORT SCHOENER'S COMPARISONS
setwd('working directory')
newstuff=read.csv('working file',header=T)
attach(newstuff)

#FIT BETA REGRESSIONS, RUN LIKELIHOOD TEST
try1=betareg(Score~Suitability+Dispersal event,newstuff)
try2=betareg(Score~Suitabilty)
try3=betareg(Score~Dispersal event)
lrtest(try1,try2,try3)
```

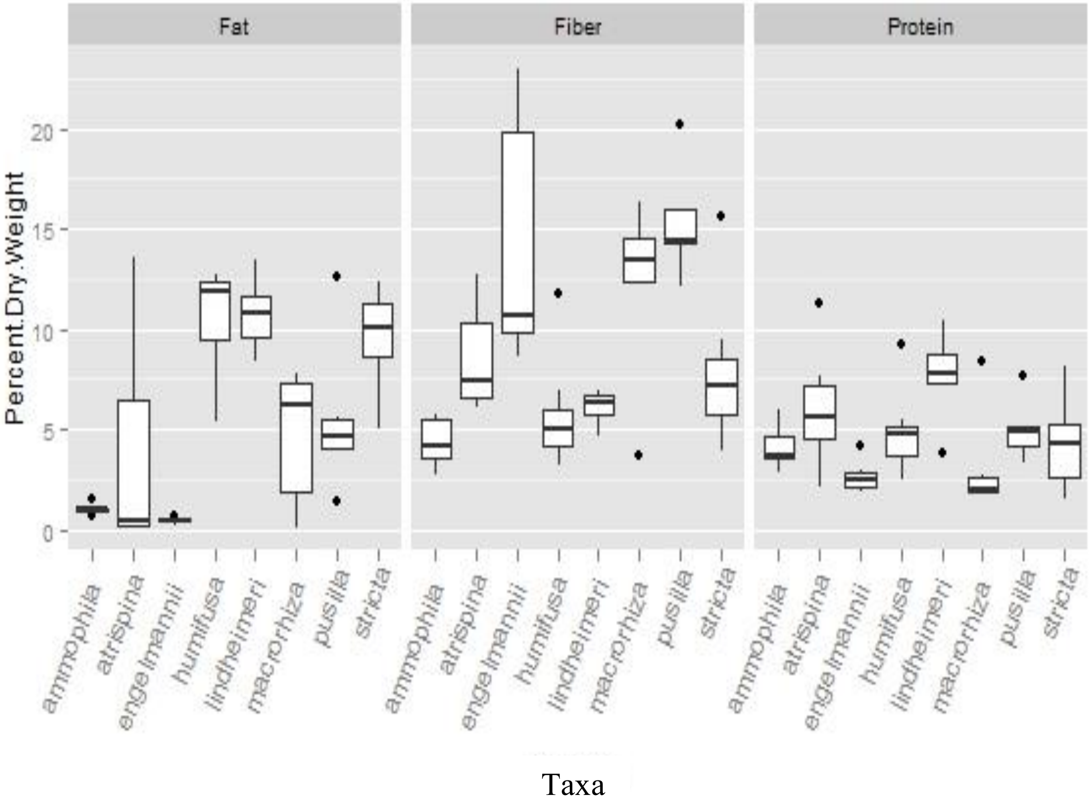

Figure C.1 Boxplot comparing tissue macronutrient content of *Opuntia* tissues across 8 *Opuntia* taxa.

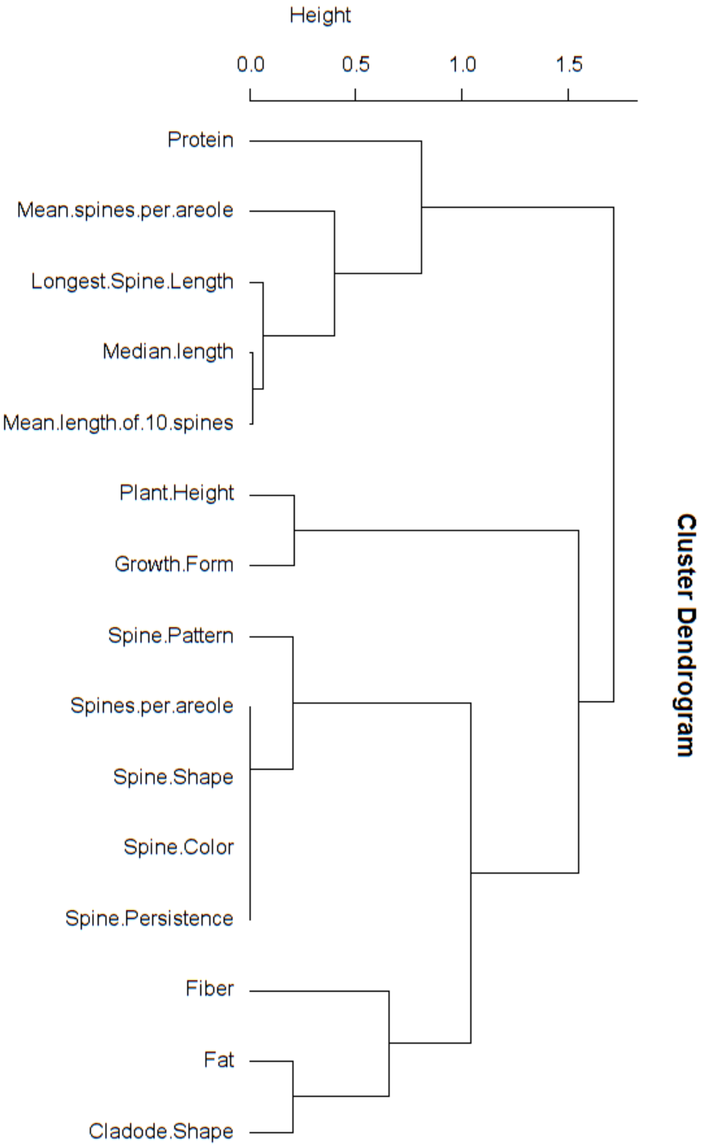

Figure C.2 Dendogram depicting the relationships between *Opuntia* morphological and tissue macronutrient traits.

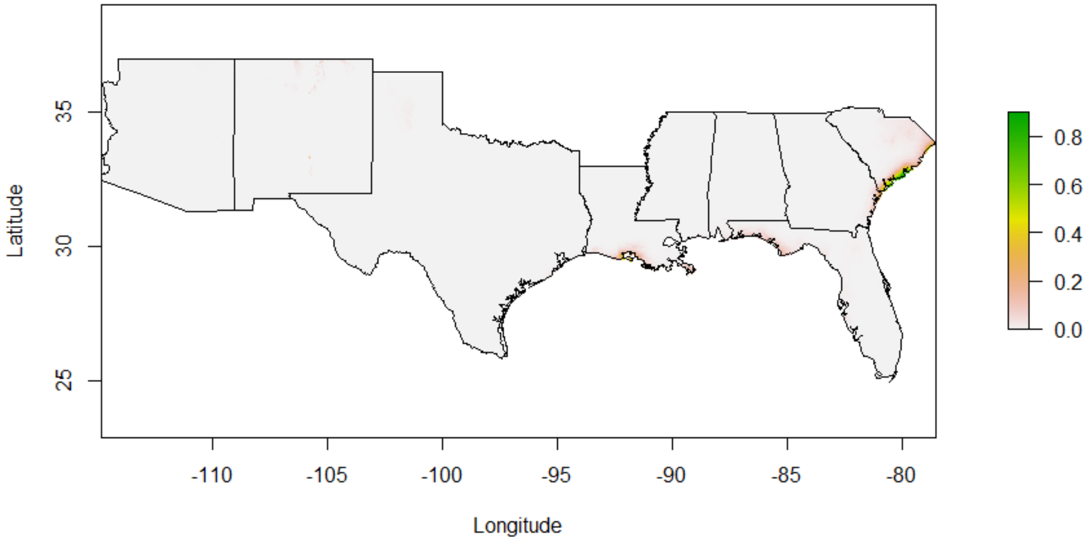

Figure C.3 Maxent prediction for the first tissue macronutrient-based *Opuntia* cluster.

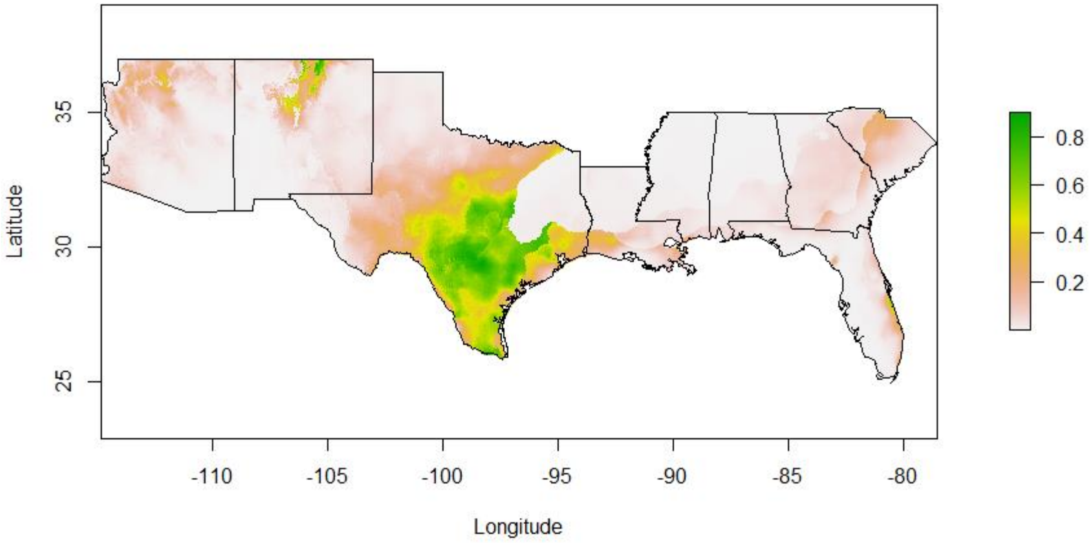

Figure C.4 Maxent prediction for the second tissue macronutrient-based *Opuntia* cluster.

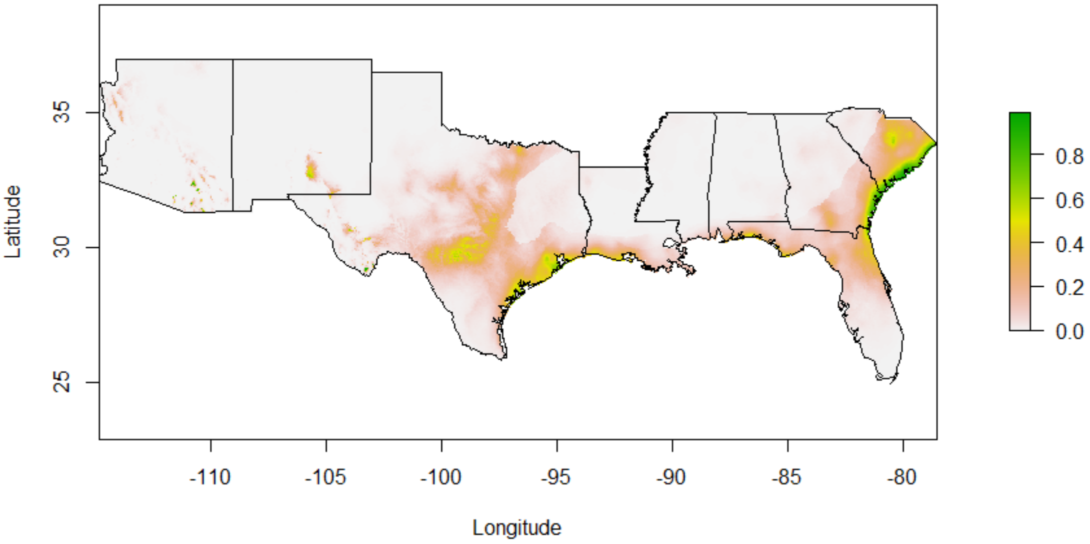

Figure C.5 Maxent prediction for the third tissue macronutrient-based *Opuntia* cluster.

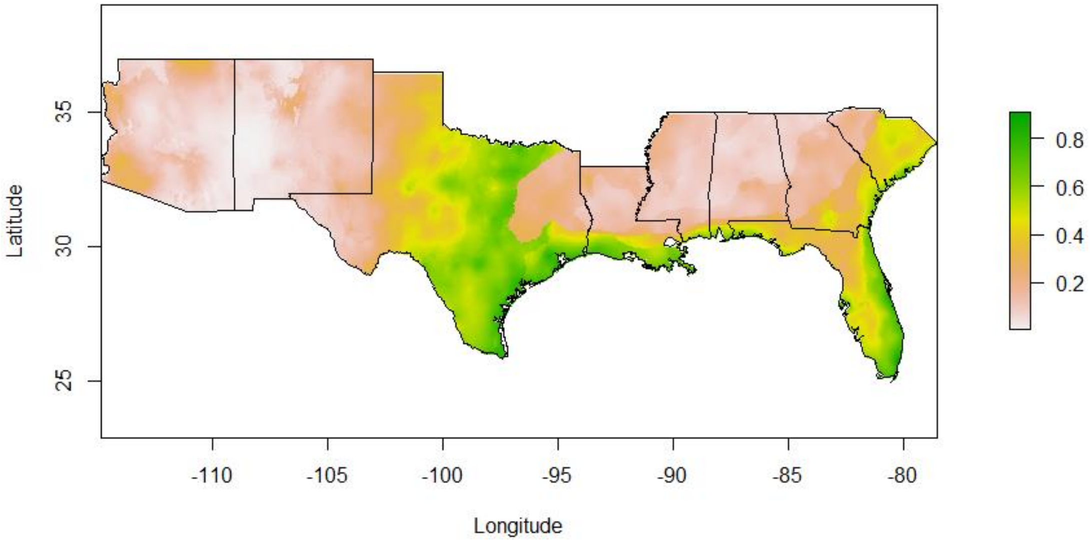

Figure C.6 Maxent prediction for the fourth tissue macronutrient-based *Opuntia* cluster.

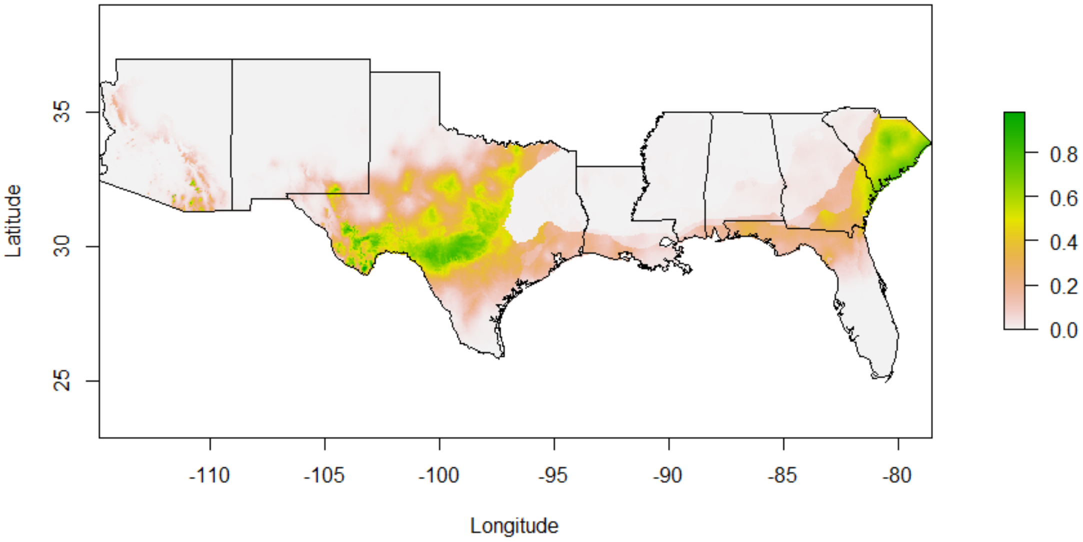

Figure C.7 Maxent prediction for the fifth tissue macronutrient-based *Opuntia* cluster.