5-2014

Landscape Genetics, Phylogeography, and Demographic History of a Pollinator Longhorn Beetle (Typocerus v. Velutinus)

Hossam Eldien Mohammed Abdel Moniem
Purdue University

Graduate School ETD Form 9
(Revised 01/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Hossam Eldien Mohammed Abdel Moniem

Entitled
LANDSCAPE GENETICS, PHYLOGEOGRAPHY, AND DEMOGRAPHIC HISTORY OF A POLLINATOR LONGHORN BEETLE (TYPOCERUS V. VELUTINUS).

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Jeffrey D. Holland

Brandon J. Schemerhorn

Andrew DeWoody

Steve J. Yaninek

To the best of my knowledge and as understood by the student in the *Thesis/Dissertation Agreement. Publication Delay, and Certification/Disclaimer (Graduate School Form 32)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Jeffrey D. Holland

Approved by Major Professor(s): _____

_____

Approved by: Steve J. Yaninek 04/22/2014

Head of the □epart□ent Graduate Program Date

LANDSCAPE GENETICS, PHYLOGEOGRAPHY, AND DEMOGRAPHIC HISTORY OF A POLLINATOR LONGHORN BEETLE (TYPOCERUS V. VELUTINUS)

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Hossam Eldien Mohammed Abdel Moniem

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2014

Purdue University

West Lafayette, Indiana

"To the soul of my beloved father, I wish you were here to witness this achievement that you have always dreamt about and encouraged me for. I am sure you feel it though! To my beloved mom, this achievement would have been impossible without your prayers. To my lovely wife and precious kids, you are the power and the reason that makes me carry on…"

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my major advisor Dr. Jeffrey D. Holland for all that he has offered me throughout the course of my Ph.D. since day one to the end of it. For me, Jeff was not just a major professor, he was a great mentor, a very knowledgeable teacher, a helpful lab mate, and an invaluable friend that provided support during the ease and hardships. Jeff created a work atmosphere and spirit that amalgamated professionalism, motivation and enjoyment that would make it too hard not to excel in research. Words are just not enough to thank him.

Secondly I would like to extend my gratitude to the members of my advisory committee, Dr. Brandi J. Schemerhorn and Dr. J. Andrew DeWoody. Dr. Brandi was a great help in the population genetics aspect of the study. I learned a lot from her and gained huge experience working in her lab. She was very generous making everything available to get the work done and always was available and keen to provide support and encouragement. Dr. Andrew was a great advisor and a scientific coach for many aspects of the study especially the phylogeography and molecular evolution components of it. Despite his busy schedule, commitments and research, he was always available to consult with me and discuss my research and offer encouragement. What I admire most about Dr. Andrew is that he helped me organize my research while tracking the scientific process which is not easy in multidisciplinary research.

I would like to thank Dr. Steve Yaninek, the department head, and his administrative staff for the support in the form of the departmental teaching assistantship and for their great administrative effort that helped me join the department of Entomology at Purdue.

I greatly thank former and current members of the Holland lab for helping me collect my specimens: Carolyn Foley, Shulin Yang, Kapil Raje, John Shukle, Thomas Mager, Julie Speelman, and Ashly Kissick. They have been also great lab mates and wonderful friends. I specially thank Dr. Insu Koh, our former postdoc, who helped me perform the random walk analysis, coding in R, collecting specimens and initiating fruitful scientific discussions. I also want to thank my lab mates from Dr. Brandi's USDA-ARS lab, especially Yan Crane and Alisha Johnson for help with molecular techniques, Phillip San Miguel, director of Purdue Genomics Facility, and Allison Sorg, technician, for sequencing assistance.

Great thanks are due to Dr. Kevin McGarigal and Dr. Samuel Cushman for their feedback on the surface metrics used, Anders Kühle for the great technical support with the SPIP^{TM} program, and Jeremy VanDerWal for help in implementing the SDMtools package in R.

Above all, I offer my gratitude and thankfulness to my almighty God who helped me through both hardships and ease during the course of my degree. Without his mercy and support, this work would have been impossible. Last but definitely not least, I thank my beautiful family: my wife, my kids, my parents and my sisters for their great support and sacrifice that helped me fulfill this achievement.

This research was supported by a governmental general mission scholarship administrated by the Egyptian Cultural and Education Bureau, Washington, DC, and by the Department of Entomology at Purdue University

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

ABSTRACT

Abdel Moniem, Hossam Eldien Mohammed. Ph.D., Purdue University, May 2014. Landscape Genetics, Phylogeography, and Demographic History of a Pollinator Longhorn beetle (*Typocerus v. velutinus*). Major Professor: Jeffrey D. Holland.

One of the central problems in contemporary ecology and conservation biology is the drastic change of landscapes induced by anthropogenic activities, resulting in habitat loss and fragmentation. For many wild living species, local extinctions of fragmented populations are common and re-colonization is critical for regional survival. Thus, habitat fragmentation in the landscape is a major threat to biodiversity, of which insects are a major proportion. Understanding the link between patterns, processes and population genetic continuity in the landscape is crucial for conserving genetic diversity within species. This is important for species persistence, for ecosystem functioning, and for future evolution. Herein, I use a newly introduced landscape gradient paradigm with surface metrology metrics, phylogeography, and landscape genetics to evaluate the influence of contemporary events (e.g. habitat fragmentation in the landscape) and pre-historic events (e.g. Quaternary glaciation) on the demography and population genetic structure of a pollinator longhorn beetle [*Typocerus v. velutinus* (Olivier)] in Indiana, USA and Canada.

Landsat 7 ETM+ imagery products provide researchers in many fields with a large amount of remotely sensed data that serves many applications. However, a malfunction of the scan line corrector (SLC) onboard Landsat 7 causes substantial data

gaps and data are available only as is, in the SLC-off mode. These data gaps may form an obstacle in using Landsat 7 ETM+ in many research disciplines. Several methods have been proposed to fix data gaps in Landsat 7 ETM+ imagery. These methods yield reliable results, but require sophisticated analyses and intensive computations and are still accompanied by some caveats. In the second chapter of this dissertation I demonstrate a spatial replacement method that is based on a simple neighborhood interpolation (SNI) approach. The results suggest that SNI provides an easily applicable, relatively quick and potentially reliable correction for the missing data patterns in Landsat 7 ETM+ data. I demonstrate the efficiency of the technique for two color bands across Indiana, USA. I tested the corrected imagery in calculating the normalized difference vegetation index (NDVI).

Measuring habitat connectivity in complex landscapes is a major focus of landscape ecology and conservation research. Most studies use a binary landscape or patch mosaic model for describing spatial heterogeneity and understanding pattern-process relationships. While the value of a landscape gradient approach is recognized, applications of the newly proposed three-dimensional surface metrics remain extremely under-used. In the third chapter, I created a surface habitat quality from several GIS layers and applied surface metrics to measure connectivity between 67 locations in Indiana, USA that were surveyed for one group of ecosystem service providers, flower longhorn beetles (Cerambycidae: Lepturinae). The results demonstrated great potential of surface metrics of connectivity to explain the differences of lepturine assemblages among the 2211 studied landscapes. Surface kurtosis and its interaction with geographic distance were among the most important metrics. This approach provided unique information

about the landscape through four configuration metrics. There were some uniform trends of the responses of many species to some of surface metrics, however some species responded differently to other metrics. I suggest that surface metrics of connectivity applied to a habitat surface map created with insight into species requirements is a valuable approach for understanding the spatial dynamics of species, guilds, and ecosystem services.

Historical geological processes have shaped the contemporary distribution of genetic variation in many species. However, there have been few empirical appraisals of cerambycid phylogeography despite of their economic importance and the fact that many geological processes (e.g., glaciations) should have had pronounced impacts on these insects as well as other taxa. In chapter four, I aimed to quantify phylogeographic effects on the contemporary gene pool of *Typocerus v. velutinus*. The beetle was collected from sites that were glaciated and unglaciated during the Pleistocene to determine genetic structure within and among populations from the US and Canada, to elucidate phylogenetic relationships among demes, and to determine divergence times between populations. A total of 451 beetles were sampled from 14 sites and sequenced at a mitochondrial DNA (mtDNA) gene. Maximum likelihood and Bayesian approaches were applied to analyze the mtDNA genealogy and to reconstruct phylogenetic trees whereas Bayesian skyline analyses were used to estimate divergence time. A total of sixteen haplotypes revealed weak geographical population structuring among most populations, but statistical tests identified significant differences between the Canadian and US populations. As a result of post-glacial recolonization, the US populations appear to have experienced demographic expansion while the Canadian population was influenced by a

bottleneck. The results suggest that Canadian population diverged from more southern populations around the time of last glacial maximum (~17,500 ybp).

Understanding the underling patterns and processes in the landscape that are affecting the population genetic structure and population connectivity is a major discipline in landscape genetics research. A vast number of these studies have implemented categorical approaches in analyzing both landscape and genetic data. In chapter five, I adopted a landscape gradient model and used the surface metrics of connectivity to model the genetic continuity between populations of the beetle (*Typocerus v. velutinus*) that was collected at 17 sites across a fragmentation gradient from Indiana, USA. I tested the hypothesis that landscape structure and habitat connectivity facilitate beetle movement and thus gene flow between the beetle populations against a null model of isolation by distance (IBD). I used next-generation sequencing and developed 10 polymorphic microsatellite loci and genotyped the population. Genetic dissimilarities between sites were calculated using R_{ST} and the population genetic structure was assessed using both non-spatial and spatial explicit Bayesian techniques. The connectivity in 137 landscapes was measured using surface topology metrics. The results indicated that panmixia was not evident with the beetle population. The source of genetic variation was mainly within rather than among populations. The surface metrics were found to significantly explain the variance in genetic dissimilarities between beetle populations 30 times better than IBD. I concluded that surface metrics of connectivity is a powerful extension in landscape genetics tools and need more attention especially to understand the configuration metrics. This approach might yield insightful applications in conservation management.

CHAPTER 1. GENERAL INTRODUCTION

The emerging field of landscape genetics can provide great insight towards our understanding of the mechanisms underlying population genetic structure and genetic continuity relationships with different patterns and processes in the landscape. The field could have important applications in conservation and management planning in a continuously changing environment. In this chapter, a brief introduction will be given to support general knowledge and background on different sections subsequently included with details in research chapters. Particularly, a brief introduction to longhorn beetles and the evolution of their lineages will be given. Then more specific information will be introduced on the species under the study and its importance. Following that, a brief account on the landscape connectivity and how it is measured and why it is important to study for these beetles will be given. After that, the landscape genetics approach will be introduced to show the insight of this new emerging filed in understanding the link between population genetic processes and the landscape structure and function. The chapter is then concluded with an introduction to the sampling sites of this project and the aim of work and an outline of the research chapters of the dissertation.

1.1 Family Cerambycidae

Longhorn beetles (Coleoptera: Cerambycidae) comprise a major lineage of phytophagous beetles. The adults are commonly referred to as longhorned beetles, while the larvae are known as round-headed borers. Cerambycidae is a large cosmopolitan family with approximately 9000 species known from the western Hemisphere and more than 900 species from North America (Bezark and Monné 2013). The cerambycids' body size varies from small (3 mm) to very large beetles (150 mm) with cylindrical to flattened bodies. Antennae are as commonly as long as or sometimes longer than the body (hence their common name). The antennae are flexed backward and held over the thorax and abdomen. Adults are active and feed on leaves or bark, as well as pollen. Larvae generally mine the phloem of trees or bore into the heartwood. They seem to prefer freshly injured or felled trees, and some species girdle small branches. Because adults are active and exposed, and feed on flowers, many species are aposematic and part of mimicry complexes with wasps or toxic insects (Linsely 1959, Solomon 1995).

Larvae of Cerambycidae feed mainly upon the solid tissues of living, dead, or dying plants. The various stages of a gradually disintegrating tree have their particular species. Eggs are laid in or under bark or in cracks in the wood. The larvae bore into wood and roots. Larval tunnels are usually excavated under the bark, in the sapwood, or in the heartwood of the host plant. The life histories of most species are unknown; however, host specificity in varying degrees is characteristic of cerambycids and has been an important factor in their evolution. Generally, the generalist species are mostly associated with the wood that is been dead and actively decomposing. On the other hand, almost all species with larvae that are dietary specialists have larvae that develop within

living trees. These tend to be oligophagous or monophagous such as the sugar maple borer (*Glycobius speciosus*) (Hanks 1999, Linsely 1959).

Within cerambycids, my dissertation is focused on one subfamily: Lepturinae. Lepturine beetles, commonly known as flower longhorn beetles, are a diverse and abundant subfamily with approximately 250 species described in North America (White 1983). They are mostly diurnal, often brightly colored cerambycids, and adults are commonly encountered on flowers on which they feed and mate (Michelsen 1963, Hanks 1999). Larvae of most species feed within decaying wood (Linsley 1959, Booth et al. 1990). Lepturines are providers of multiple ecosystem services: they help decompose dead wood and thus cycle nutrients and they are potential pollinators, with many species frequenting flowers of valuable hardwood trees such as the American chestnut (Benjamin 1907). This is an especially interesting group of species to study how landscape gradients influence connectivity for species in fragmented habitats because many species use complementary habitats in different life stages. Larvae require decaying wood most reliably found in forests while adults of many species are common in more open areas with abundant plants in flower.

1.2 The study species

The banded longhorn beetle, *Typocerus v. velutinus* (Olivier) is considered to be one of the important generalist lepturines in forested ecosystems. This beetle is active from May to August (Frost 1979) as adults which known to be flower feeders. They have been recorded on some wild flowering plants such as *Spirea*, *Rosa*, *Ceanothus*, *Daucus*, *Apocynum*, *Pastinaca sativa*, *Rubus*, *Rhus*, *Asclepias*, *Solidago*, *Melilotus*, *Hydrangea*,

Oxypolis, *Cirsium*, *Cesatanea*, *Sambucus*, *Passiflora*, *Eupatorium* and *Viburnum*. (Blackman 1918, Gosling 1984, Knull 1946, Bond and Philips 1999, Linsley & Chemsak 1976). Larvae hosts are decaying hardwoods including *Quercus*, *Caray*, *Betula* and *Populus* and the beetle is thought to complete its life cycle in two years (Yanega 1996). Thus, this species relies on habitat complementarity to complete the life cycle because not all required resources are contained in breeding the habitat.

The species is easily identified by the number of distinctive morphological characteristics from about ten other species in the same genus. Body size ranges from 9 to 16 mm. The body is reddish brown with transverse yellow bands on the elytra. Antennae with characteristic lateral oval pits that distinguish this species from the morphologically closest species (*Typocerus deceptus* Knull). The elytral tips are lacking strong produced outer spines. The pronotum is densely covered with hairs and its basal and apical hair bands are complete (Lingafelter 2007, Yanega 1996).

This beetle is an ecologically important species as it is providing two very important ecosystem services. As adults, they are potential pollinators (Maeto et al. 2002). They have dense pubescence, setae and spines covering the sternites and legs which helps in carrying pollen. As larvae, they are dead wood decomposers, thus they are helping in natural recycling and controlling fire fuel load in forests (Berkov and Tavakilian 1998).

1.3 Evolution of cerambycids (Coleoptera: Cerambycidae)

Beetles (Coleoptera) are one of the most diverse orders of arthropods. They comprise approximately 25% of all species in the animal kingdom (Grimaldi 2005). Coleoptera was thought to be closely related to the hemimetabolous Megaloptera and the

Strepsiptera within the Holometabola (insects with complete metamorphosis). Evolution of beetles from Megalopteran-like ancestors was supported by the structure of the elytra in Lower Permian beetles because their wing venation resembles that of a Megalopteran forewing (Lawrence 1982). However, Coleoptera is found to be more closely related to Strepsiptera because of some morphological characters such as the presence of metathoracic flight wings, free prothorax with closely associated mesothorax and metathorax, abdomen with more heavily sclerotized sternites than tergites, and the triungulin larvae (Lawrence 1982).

Coleoptera most likely arose during the Carboniferous from a generalized holometabolus insect. The ancestral adult was thought to be active, terrestrial, short lived, with two pairs of membranous flight wings and a loosely organized body (Crowson 1981). A transition from this generalized form took over towards general increase in structural integrity of the adult that helped in pre-adapting early beetles for living in both arid and aquatic environments. During the Carboniferous period, beetles were most likely phytophagous, feeding on different kinds of decomposing plant material, such as cambial tissue, rotten wood, and leaf litter. Phytophagous beetles are considered as a monophyletic group based on the structure of the tarsi, which appeared to be four-segmented with the fourth segment concealed between two tarsomeres, in addition to the reduction of the male copulatory organ (Hammond 1979, Lawrence 1982). The feeding habits of beetle larvae necessitated various morphological modifications in the basic type, such as antennal reduction and modified mouthparts, legs, and body to enable their access to more compact substrates, such as soil and less decomposed wood. However,

specialized wood-boring larvae probably did not evolve until later (Lawrence 1982). The earliest fossils that resemble modern beetles (265 MYO) are recorded from the Lower Permian beds but were not as abundant and diverse as the Upper Permian fossils. However, the only important Triassic assemblage is found in central Asia (Lawrence 1982).

Order Coleoptera is divided into four major suborders based on the structure of the prothorax and hindwing. Archostemata, which comprises about 40 recent species and is consistently indicated as the most basal lineage in all studies on the relationships of beetles as revealed from molecular studies of 18S and 28S rDNA subunits (Marvaldi et al. 2009). Myxophaga is a small group of specialized aquatic and semi-aquatic beetles. Adephaga represents close to 10% of all beetle species. These include some recent and about 5 extinct families of ground and aquatic species, which are mainly predatory. The Polyphaga are extremely diverse in diets. This group includes 90% of all beetle species and accounts for the great diversity of the order (Grimaldi 2005, Lawrence 1982).

Family Cerambycidae (longhorn beetles) belongs to the fourth suborder (Polyphaga). The suborder Polyphaga comprised of five lineages (infra-orders) extending back at least to the early part of the Triassic and comprising: Styphyliniformia, Scarabaeiformia, Elateriformia, Bostrychiformia and Cucujiformia. The last lineage (Cucujiformia) comprises the two big super-families: Chrysomeloidea (longhorned and leaf beetles) and Curculionoidea (weevils). These two super-families are the largest two groups of phytophagan beetles. This lineage is the largest assemblage of Coleoptera, with over 90 families and the majority of the current described genera and species. The ancestors of this group were thought to be characterized by larvae and adults living in the

same habitats, feeding on decaying vegetation and fungi (Lawrence and Hlavac 1979). The subject beetle of this study [*Typocerus v. velutinus* (Olivier)] belongs to the super-family Chrysomeloidea, which includes 8 distinct lineages (Crowson 1981).

The earliest apparent cerambycid fossil seems to be *Cerambycomima* sp. from the late Jurassic, early Cretaceous (about 150 MYA). However, the absence of Cretaceous cerambycid fossil records could support the idea that their fossils might be mainly Cenozoic. There are some fossils of cerambycids recorded in Eocene amber. For example they are found in the Eocene-Oligocene records from Colorado, and in Miocene amber from the Dominican Republic (Lawrence 1982).

Climatic factors and plant resources availability are the main factors controlling the distribution of cerambycids (Hanks 1999). The historical events of global climatic change and the evolution of the host plants formed, to a large extent, the distribution and evolutionary history of the current cerambycids. For example, the early Holarctic assembly of cerambycids fauna of the Northern Hemisphere was associated with the Arcto-Tertiary flora, which moved (range shift) southward during the Tertiary period and replaced pre-existing tropical floras of the Cretaceous period. These early northern types are now represented discontinuously in Europe, Eastern Asia, Western and Eastern North America and Mexico (Linsley 1959).

The distributions of the historical geological features and of the woody plants, which are the primary cerambycid hosts, are widely discontinuous. These discontinuities clearly reflect segregation in the face of gradual climatic changes during the Tertiary and centers of survival during the extremes of the Cenozoic (Linsley 1959). As a result of post-glacial recolonization, trans-tropical distributions of cerambycids are evident in both

the old and new world, but generally the Southern Hemisphere cerambycids are isolated morphologically which suggests that the geographic relationship is an ancient one (Linsley 1959, Ashworth 2001). Currently, anthropogenic factors are the major forces that influencing the ecology and evolution of the cerambycids. The most noticeable effect is that of habitat fragmentation which increased dramatically in the recent past. This fragmentation caused by habitat loss creates a patchy environment of isolated habitat paches for the cerambycids. This isolation is the initiator for various micro-evolutionary forces to take place and become significant in shaping the genetic structure of these beetles' populations. This patchy environment characterized by spatial heterogeneity among the habitat fragments further integrates with other factors (climatic, biological, anthropogenic) and could affect dispersal and gene flow among isolate to different extent based on the species response to different spatial scales.

In this dissertation, I dissertation I studied the phylogeography and demographic history of the banded longhorn beetle [*Typocerus v. velutinus* (Olivier)] as shaped by the Quaternary. I tested the hypothesis that demographic responses to climate change differentially impacted southern refugia populations of the beetle relative to northern populations that were established after retreat of the Wisconsinan ice sheet. I predicted that as sources for recolonization, southern populations would harbor more genetic variation and exhibit more evidence of recent demographic expansions than northern populations.

1.4 Landscape habitat connectivity

Habitat connectivity is defined as the degree to which the landscape facilitates or impedes movement of species among resource patches (Taylor 1993). So, landscape connectivity measures are concerned with the interactions between the species and its habitat. The species is responding via a group of behaviors to habitat change in the landscape. Dispersal is among the most important behaviors that could be influenced by the degree of habitat fragmentation. This could vary to different extents depending on the species habits [e.g. generalists vs. specialists (Tischendorf et al. 2003)]. Dispersal is important for maintaining genetic diversity, rescuing declining populations, and aiding in re-establishing extirpated populations. Adequate rates of movement (dispersal) of individuals between isolated habitats under the extinction-recolonization equilibrium can allow an entire network of populations to persist via meta-population dynamics (Hanski 1991). The importance of landscape connectivity and its impact on populations in heterogeneous landscapes, and its implications for conservation biology, resulted in increasing interest in landscape connectivity and estimating different connectivity measures (Goodwin 2003).

There are three types of landscape connectivity that have been discussed in the literature. Structural, functional (or potential) and actual connectivity. From the landscape perspective, the last two are species-specific measures as functional connectivity incorporates information about the biology of the species in question (e.g., by dispersal models) and actual connectivity is further relying more on information about the species and its relationships to its surrounding environment and available habitat in addition to its actual movement in the landscape, which is difficult to estimate (Tischendorf and Fahrig

2001). Because connectivity is determined by the connectedness of intervening habitat areas and the dispersal ability and behavior of the species (Taylor et al. 1993), factors facilitating or impeding movement will be species specific and may not be predicted by patch edges and inter-patch distances (Cushman 2006).

Wide varieties of commonly used connectivity metrics depend in their estimations on a dichotomization of focal patch and matrix habitat (Calabrese and Fagan 2004). These metrics and associated frameworks for modeling complex landscapes include the patch mosaic model (Forman and Godron 1981), the variegation model (McIntyre and Barrett 1992), and the modified habitat gradient models (Manning et al. 2004, Fischer and Lindenmayer 2006). All of these models have contributed to our understanding of biological and ecological processes in the landscape.

The patch mosaic model (PMM; Forman 1995) has been adopted in many studies and has led to many advances in our understanding of pattern-process relationships (Turner 2005). The model has great value due to its conceptual simplicity and consistency with well-developed landscape tools such as FRAGSTATS (McGarigal et al. 2002) and quantitative analysis techniques (e.g. ANOVA) (McGarigal et al. 2009). However, for some studies it is suboptimal because it is inconsistent with basic ecological theory and bypasses the continuous nature of habitat heterogeneity (McGarigal and Cushman 2005; Cushman et al. 2007, Cushman et al. 2010, McGarigal et al. 2009). The categorical representation of heterogeneity may result in an arbitrary characterization of patch classes and boundaries. Species have environmental requirements that support their survival and reproduction (Shelford 1931). These physical, chemical, and biological

conditions are usually distributed in the landscape in a continuous rather than discreet manner (Wiens 1989, Wu 2007, McGarigal et al. 2009).

In all landscape connectivity metrics, there is a trade-off between information content and data requirements (Kindlmann and Burel 2008). For example, some metrics such as the nearest neighbor measures and spatial pattern indices do not require massive data to be calculated. However, they yield only a crude estimate of structural connectivity. On the other hand, buffer radius measures and Hanski's incidence functional model (IFM) (Hanski 1994, Hanski et al. 2000), both provide detailed estimates of potential connectivity at the patch level, but they are extremely data-intensive. Also, estimates of actual connectivity require observation methods and are only applicable to small scales and are extremely data intensive. However, the graph-theory based metrics have the greatest benefit of estimating connectivity at relatively large scales. These measures provide a reasonably detailed picture of potential connectivity and have relatively moderate data requirements (Minor and Urban 2008).

One of the greatest challenges facing landscape ecologists is integrating the niche theory with spatial ecology. This challenge crystallizes in linking non-spatial niche relationships with the spatial patterns of environmental gradients in complex heterogeneous landscapes (Austin 1985, McIntyre and Barrett 1992, Urban et al. 2002, Manning et al. 2004, Cushman et al. 2007). Thus, a new paradigm that considers a gradient approach of environmental conditions and heterogeneity in the landscape is a step forward for many studies (Abdel Moniem and Holland 2013). The landscape gradient paradigm (McGarigal and Cushman 2005) and surface topology metrics are beginning to be shown to be powerful approaches to study the influence of habitat

heterogeneity on lepturine beetle species communities (Abdel Moniem and Holland 2013). The requirements of complementary habitats for these species and the inherently continuous nature of habitat quality make it important to consider habitat as a continuous attribute to avoid oversimplification of categorical landscape approaches (McGarigal and Cushman 2005, Hoechstetter et al. 2008, Kent 2009). In my dissertation, I studied the impact of habitat connectivity as measured by the newly introduced surface metrology metrics for a group of pollinator beetles in Indiana. I hypothesized that landscape connectivity enhances the movement of Lepturines in fragmented habitats and correlates with communities' dissimilarities against the null hypothesis that there is no correlation between habitat connectivity as measured by surface metrics and lepturine communities' dissimilarity.

1.5 Landscape genetics approach

Landscape genetics is a field described as an amalgamation that brings together both molecular population genetics and landscape ecology to understand the influence of patterns and processes in the landscape on the population genetics of species (Manel et al. 2003). A more distinct definition of the field was proposed by Storfer and colleagues (2007), who indicating that landscape genetics comprises research that explicitly quantifies the effects of landscape composition, configuration and matrix quality on gene flow and spatial genetic variation. Generally, landscape genetics studies combine adaptive or neutral (or both) types of population genetic data with structural landscape ecology data (Holderegger and Wagner 2008). Thus, the incorporation of the matrix (non-habitat area) component of the landscape into landscape genetics is a characteristic

difference between landscape genetics and population genetics. Population genetics often characterizes the stretches of land between occupied habitats by a simple function of geographic distance; however, in contrast, landscape genetics further analyzes the intervening matrix as an important determinant factor of biological and ecological processes at the landscape level because different quantities and qualities of the areas that separate habitats are quite important (Holderegger and Wagner 2008).

Population genetics is concerned with the distribution and changes in allele frequency due to micro-evolutionary processes acting on populations and influencing their genetic structure. These forces could be natural selection, genetic drift, mutation and gene flow. Such micro-evolutionary forces that prevent panmixia (random mating between individuals across large regions) could include ecological factors such as mating system, social structure, dispersal and spatial distribution, genetic factors such as mutation rates, genetic drift, and natural selection, and environmental factors such as climate, landscape fragmentation, and geographic barriers of gene flow. Advances in molecular biology methods have provided powerful tools to measure the relationship between species populations and detect both intra- and inter-population levels of genetic variation. These methodologies enabled estimation of genetic distances, population structures and gene flow among populations. Different types of molecular markers have been developed, tested and used widely for this purpose (Avise 2004). Microsatellites (also known as short sequence repeats SSR and short tandem repeats STR) are some of the mostly used markers. They are repetitive sequences (1–12) of nucleotides (most commonly 2–4) that are highly and frequently distributed throughout eukaryotic genomes (Ramel 1997). Their high level of polymorphism and frequency within the genome, make

them ideal markers for different applications such as paternity analysis, evolutionary genetic analyses, and population genetics (Pai et al. 2003). Nevertheless, for many reasons, the mtDNA genome has long been considered a marker of choice for phylogeography and population genetics studies (Avise et al. 1987). For example mtDNA genes are haploids, having only one set of alleles, are almost always maternally inherited, and are non-recombinant as opposed to nuclear genes. Thus they are easy to isolate and sequence and hence ideal to compare between individuals and populations. More importantly, mtDNA genes evolve at a much more rapid pace due to reduced or lacking DNA repair machinery especially at the control region genes. These characters make this genome ideal for studying population structures and phylogeography at a shallower, more recent, evolutionary scale (Avise et al. 1987).

Landscape ecology and population genetics naturally converge in the exploration of how habitat loss and the spatial isolation or fragmentation of habitat affects the movement of species across landscapes. Holderegger and Wagner (2008) argued that landscape genetics is not a scientific discipline in itself but rather provides a perspective for examining the influence of spatial, temporal, or both processes (e.g. habitat fragmentation and climate change) on the genetic structure of populations. In chapter five, I used a landscape genetics approach to study the population genetic structure and dissimilarities between *Typocerus v. velutinus* demes in the landscape of Indiana. I tested the hypothesis that landscape structure and habitat connectivity facilitate beetle movement and thus gene flow between the beetle populations.

1.6 Study area and sampling projects

Sampling sites in this project came from one study site in Canada and three different survey projects that focused on studying the longhorn beetles (Coleoptera: Cerambycidae) in the Landscape Ecology and Biodiversity Laboratory (LEBL) in the Department of Entomology at Purdue University. Beetle surveys for these projects were carried out over a period of seven years (2005–2011), however, each project ran for a particular number of years. In the following, I describe each project, sites used in each, and describe the sampling procedure in each.

1.6.1 Canada sampling site

Individuals from Canada were hand collected near Westport, Ontario, Canada, in the western edge of the St. Lawrence Lowland Eco-region. This region contains a mixture of agriculture, mixed forest, and abundant lakes and wetlands. Mixed forests of sugar maple, yellow birch, eastern hemlock, and eastern white pine are common. Other forest tree species include beech, red pine, eastern white cedar, red oak, red maple, black ash, white spruce, tamarack, and eastern white cedar. The average monthly temperatures vary from -10°C in winter to 20°C in summer, and annual precipitation is 870 mm.

1.6.2 Indiana sampling sites

Indiana sites are represented by two Omernick eco-regions (level IV). First, the northeastern area belongs to the Loamy High Lime Till Plains. The soil in this area developed from loamy, limy, glacial deposits of Wisconsinan age. The land cover in this area is dominated by corn and soybean fields with some forests that include beech forests,

oak-sugar maple forests, and elm-ash swamp forests that grew on the nearly level terrain. The second sampling region is further south in the south-central state forest area. The area belongs to the Interior Plateau eco-region with two subdivisions; the Mitchell Plain and the Norman Upland. The north of the Mitchell Plain experienced pre-Wisconsinan glaciation. Soils are leached and largely developed from loess and limestone. It was dominated by Western mesophytic forests; karst wetland vegetation and limestone glades. The Norman Upland subdivision is characterized by its hilly topology, narrow valleys, and medium to high gradient streams. The soil is derived from loess, siltstone, shale, or sandstone. It was dominated by oak-hickory forests that grew on the uplands and beech forests in the valleys. Currently the forest contains mainly chestnut oak on the upper slopes and Virginia pine on the southern uplands. Other species such as sugar maple and ash also exist. The climate of Indiana varies from north to south of the state; the annual mean temperature is 49°F–58°F (9°C–12°C) in the north and 57°F (14°C) in the south. Maximum and minimum monthly average temperatures range from a high of 88.8°F (31.5°C) to a low of 15.8°F (-7.5°C). Precipitation is evenly distributed throughout the year and the average annual precipitation in the state is 40 in (1020 mm).

Three main sampling projects that surveyed Indiana for cerambycid fauna were used in this dissertation.

1.6.2.1 Upper Wabash Ecosystem Project (UWEP)

The UWEP is a large-scale ecosystem project that was conducted in the upper basin of the Wabash River (Swihart et al. 2006). I used 43 of these sites. Among these sites, four sites represented Purdue Research Forests (PRF) and seven Purdue Agricultural

Centers (PAC). Longhorn beetle surveys were conducted on all sites for year 2005. In year 2006 only 23 sites were surveyed and only four sites among those 43 were resampled in the period between 2009–2011.

At these sites, points were selected randomly within forest using ArcGIS (ESRI Redlands, CA). To avoid edge effects, all points were located at a distance greater than 50 m from forest edges. At each selected sampling point a trapping array was placed by hanging traps from tree branches and was composed of two Lindgren funnel traps (Pherotech, Delta, Canada), one Intercept panel trap (APTIV, Portland, USA), and one transparent window pane trap that was built in the LEBL, Purdue University. A central tree was selected using a geographical positioning system unit (GPS; Magellan Meridian color) at each site and each trap in the array was setup approximately 10 m away from that central tree and randomly placed in the four cardinal directions. As a lure, each trap had a 60 ml of absolute ethanol in a 125 ml Nalgene bottle with four holes of 1 mm each in the cap to emit the attractive scent (Holland 2006). In each trap, there was a collecting bottom that contains ethylene glycol as a non-evaporating killing solution and preservative. Traps had beetles recovered every three weeks and during each visit the volume of remaining lure was recorded and refilled to 60 ml. Sites that were sampled during the year of 2011 were only sampled using sweeping nets in order to focus on the target species [*Typocerus v. velutinus* (Olivier)]. All collected longhorn beetles were identified to species level using Lingafelter (2007) and Yanega (1997) and stored in the LEBL and Purdue Entomological research collection (PERC). I used all the *Typocerus v. velutinus* specimens, checked the species identification, and each individual was given a unique ID number and recorded into a separate database.

1.6.2.2 Hardwood Ecosystem Experiment (HEE)

Many of the beetles sampled for the current study were collected as part of the Hardwood Ecosystem Experiment (HEE) being conducted at Morgan-Monroe and Yellowwood State Forests in south-central Indiana. This long-term study (100 years planned) is examining the impact of different forestry regimes on the regeneration of native oak forest, as well as on other forest flora and fauna. The HEE study consists of nine management units (MU), which are approximately 200 acres each. Three types of forestry management are being implemented: even-aged management, uneven-aged management, and a no-harvest management or control. Details about the complete experimental design of this large project is available through a base-line study on the pre-treatment assemblages of wood-boring beetles (Coleoptera: Buprestidae, Cerambycidae) of the HEE (Holland et al. 2012).

Trap arrays for the pretreatment years (2006–2008) were set up within what is called intensive sampling units (ISU). These units were selected within the management units and they are up to approximately 4 ha each. Trap arrays were approximately centered on the bird survey tree closest to the center of the ISU. Traps were randomly placed in a cardinal direction and setup about 20 m from the central tree. Traps were hung with their bottoms approximately 2 m above the ground. Each array was composed of four traps as follows: one Lindgren multiple-funnel trap (Pherotech, Delta, Canada), one Panel Trap for Bark Beetles (Alpha Scents, Portland, USA), one intersecting pane window trap designed in LEBL and one purple sticky trap (Holland 2006). We also used 0.61 m x 0.61 m rain covers on the first three trap types. For lures with the first three trap

types, we used a 125 ml Nalgene bottle containing 60 ml of absolute ethanol with caps that had four holes of 1 mm each for lure release. The collection jars for the first three trap types contained ethylene glycol as killing and preservative solution. We also added few drops of a detergent in the collecting jars to weaken the surface tension of the ethylene glycol.

The same sampling procedure was repeated for the following years. However, in 2008 trap arrays were located outside the ISU to sample the landscape matrix outside the different harvest treatments. Trap arrays were located at bird survey points that were at least 200 m from any ISU, 50 m from any road or trail, and 100 m from any previously surveyed beetle point. Within each management unit, we randomly selected four bird survey points from those that met these criteria.

Traps were checked for beetles every three weeks. We removed all insects from the traps by filtering the ethylene glycol through a strainer. At each visit we measured the amount of unevaporated ethanol and refilled the lure container to 60 ml. In the LEBL, we separated all longhorn beetles from the catch, pinned all cerambycid specimens, and identified these using Yanega (1996), Lingafelter (2007), and Linsley and Chemsak (1972, 1976). All specimens currently reside in the insect collection of the LEBL and PERC. Specimens of the target species of the current research were isolated, had the species identification confirmed, and were given a unique ID number for each individual that was recorded in a database. I preserved some specimens from the traps individually in absolute ethanol in 1.5 ml screw cap micro-centrifuge tubes (dot scientific Inc., MI, USA) while already pinned samples were kept in the lab research collection.

1.6.2.3 Landowners' Forest Properties Survey (LO)

This survey was carried out during the summer of 2009 in the LEBL (Raje et al. 2012). In this study, 19 private forest landowners whose properties were located within a 45 km radius of West Lafayette, Indiana volunteered to participate in a longhorn beetles survey. Sampling these properties involved setting two arrays of traps at each property. Each array contained a total number of four traps as follows: one Lindgren multiple-funnel trap (Pherotech, Delta, Canada), one black panel trap (APTIV, Portland, USA) and two intersecting window traps. For each of these traps we used 60 ml of 100% ethanol as a lure in a 125 ml Nalgene bottle with a perforated cap similar to those used for the UWEP sampling. Moreover, we added another type of lure to our window traps. We used similar release mechanism with benzyl acetate in an attempt to further attract flower-visiting species (Maeto et al. 2002). All of the traps had collection cups that were one-quarter filled with ethylene glycol as a non-evaporating killing and preservative solution. Insects were collected from the traps approximately every two weeks from mid-April to mid-September. In addition, a sweep net was used during each visit in an attempt to gather additional specimens of the target study species. All longhorn beetles were identified to species and voucher specimens reside in LEBL and PERC. Again, *Typocerus v. velutinus* specimens were checked for species identification, preserved, and given a unique ID number that was recorded in a database.

1.7 Extent and spatial reference in the study

All geographic information system layers and maps used in this study were set to the extent of Indiana as follows: top (4625518.7), left (403539.1), right (692139.1) and bottom (4180918.7). The spatial reference was setup to NAD1983, UTM zone 16N, with a 1 m linear unit, an angular unit of 0.0174 degrees, false easting and false northing of 50000 and 0 respectively, central meridian of -87, and latitude of origin 0. The spatial resolution (cell size) was set to 30 m x 30 m for data extraction, spatial and statistical analysis to capture finer level of variation in the variables used, then all layers were scaled up to 300 m x 300 m spatial resolution for the large scale surface metrics analysis and mapping.

1.8 Aim of work and chapters outline

The dissertation in hand aimed to study the landscape genetics, demographic history and phylogeography of the banded longhorn beetle *Typocerus v. velutinus* (Olivier) (Coleoptera: Cerambycidae: Lepturinae) as an important generalist in the forested ecosystems of Indiana that provides many ecosystem services. The dissertation contains five major chapters and a general conclusion. Following is an outline of the dissertation research chapters along with the particular hypothesis tested, the associated predictions and a brief note on the methodology used for each chapter.

In chapter two, I introduce a spatial replacement tool that corrects for Landsat 7 ETM+ missing data patterns as this data will be used in subsequent chapters. In this chapter we hypothesized that simple neighborhood interpolation (SNI) mechanism can

fix the SLC problem and fill the imagery data gabs versus the null hypothesis that LS7 ETM+ data are available only as is with SLC substantial data gaps. I retrieved the LS7 ETM+ multispectral data for Indiana, divide the extent to 100 x 100 km polygons, and used spatial replacement with SNI algorithm to fill the gaps with nearest neighbor pixels values. Fixed polygons were then stitched to obtain the full extent of the state. To evaluate the quality of the final product, I used it to calculate the normalized difference vegetation index.

In chapter three, I studied the impact of habitat connectivity as measured by the newly introduced surface metrology metrics for a group pollinator beetles in Indiana. I hypothesized that landscape connectivity enhances the movement of lepturines in fragmented habitats and correlates with communities' dissimilarities against the null hypothesis that there is no correlation between habitat connectivity as measured by surface metrics and lepturine community dissimilarity. In this study, I sampled lepturine communities along a fragmentation gradient across Indiana. I created a habitat quality surface with insight into habitat requirements for the beetles, clipped the landscapes between sites, and measured the geographic distances between sites. Surface metrics of connectivity were measured, Bray-Curtis dissimilarity metric was calculated between sites for beetles' communities. I used a generalized additive mixed model to assess the correlation between communities' differences and surface metrics of connectivity.

In chapter four, I studied the phylogeography and demographic history of a pollinator longhorn beetle [*Typocerus v. velutinus* (Olivier)] as shaped by the Quaternary. I tested the hypothesis that demographic responses to climate change differentially impacted southern refugia populations of *Typocerus v. velutinus* relative to northern

populations that were established after retreat of the Wisconsinan ice sheet. More specifically we predicted that as sources for recolonization, southern populations would harbor more genetic variation and exhibit more evidence of recent demographic expansions than northern populations. This hypothesis was tested against the null that prehistoric climates did not affect the population structure of *Typocerus v. velutinus* in North America. In this study, I sampled the beetles across a gradient of former glacial zones between Canada and Indiana. DNA was extracted and COI was partially sequenced. The maximum likelihood and Bayesian approaches were used to analyze the COI genealogy and to construct the phylogenetic trees. A range of previously estimated mutation rates of insects' mtDNA genes were used with a strict molecular clock and Bayesian analysis was used to make an inference about the divergence date between both lineages. Bayesian Skyline plot (BSP) was used to visualize the results.

Finally in chapter five, I used a landscape genetics approach to study the population genetic structure and dissimilarities between the beetle (*Typocerus v. velutinus* Olivier) demes in the landscape of Indiana. Specifically, I tested the hypothesis that landscape structure and habitat connectivity facilitate beetle movement and thus gene flow between the beetle populations versus a null hypothesis that populations of the beetle are genetically isolated by distance alone in the landscapes. In this study, beetles were sampled across a fragmentation gradient in Indiana, DNA was extracted from samples, a number of microsatellites were developed to genotype beetles, and spatially explicit and non-explicit Bayesian techniques were used to determine population genetic structure. Genetic dissimilarities were calculated between populations in study sites. Landscape connectivity metrics were calculated between sites. A generalized additive

mixed model was used to assess the correlation between genetic distances and surface metrics of connectivity.

At the end of the dissertation, a general conclusion summarizes the major findings of this research and gives insight on the possible applications and future research suggestions of each chapter.

1.9 Bibliography

Abdel Moniem, H. E. M., Holland, J. D. 2013. Habitat connectivity for pollinator beetles using surface metrics. Landscape Ecology. 28: 1251–1267.

Austin, M. P. 1985. Continuum concept, ordination methods, and niche theory. Annual Review of Ecology and Systematics. 16(1): 39–61.

Avise, J. C. 2004. Molecular Markers: Natural History and Evolution. Sinauer & Associates, Sunderland, Massachusetts.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E. Lamb, T., Neigel, J. E., Reeb, C. A., Saunders, N. C. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics. 18: 489–522.

Benjamin, D. 1907. Annual report of the state entomologist of Indiana. Library of the Museum of Comparative Zoology, Harvard University, 216 pp.

Berkov, A., Tavakilian, G. 1998. Host utilization of the Brazil nut family (Lecythidaceae) by sympatric wood-boring species of *Palamae* (Coleoptera, Cerambycidae, Lamiinae, Acanthocinini). Biological Journal of the Linnaean Society. 67: 181–198.

Bezark, L. G., Monné, M. A. 2013. Checklist of the Oxypeltidae, Vesperidae, Disteniidae and Cerambycidae, (Coleoptera) of the Western Hemisphere, pp. 1–484.

Blackman, M. W. 1918. On the insect visitors to the blossoms of wild blackberry and wild Spirea; a study in seasonal distribution. New York City College of Technology. 10: 119–144.

Bond, W. D., Philips, T. K. 1999. Diversity, phenology, and flower hosts of anthophilous long-horned beetles (Coleoptera: Cerambycidae) in a southeastern Ohio forest. Entomological News. 110: 267–278.

Booth, R. G., Cox, M. L., Madge, R. B. 1990. IIE Guides to Insects of Importance to Man. 3. Coleoptera. CAB International, Wallingford, UK, 384 pp.

Calabrese, J., Fagan, W. 2004. A comparison shoppers' guide to connectivity metrics: trading off between data requirements and information content. Frontiers in Ecology and the Environment. 2: 529–536.

Crowson, R. A. 1981. The biology of Coleoptera. London: academic. 802 pp.

Cushman, S. A. 2006. Effects of habitat loss and fragmentation on amphibians: a review and prospectus. Biological Conservation. 128(2): 231–240.

Cushman, S. A., Gutzweiler, K., Evans, J. S., McGarigal, K. 2010. The gradient paradigm: a conceptual and analytical framework for landscape ecology. In: Cushman, S. A., and Huettmann, F. (eds). Spatial Complexity, Informatics, and Wildlife Conservation. Springer, New York, pp. 83–108.

Cushman, S. A., McKenzie, D., Peterson, D. L., Littell, J., McKelvey, K. 2007. Research agenda for integrated landscape modeling. USDA Forest Service General Technical Report RMRSGTR-194.

Fischer, J., Lindenmayer, D. B. 2006. Beyond fragmentation: the continuum model for fauna research and conservation in human–modified landscapes. Oikos. 112(2): 473–480.

Forman, R. T. T. 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, Cambridge.

Forman, R. T. T., Godron, M. 1981. Patches and structural components for a landscape ecology. Bioscience. 31(10): 733–740.

Frost, S. W. 1979. A preliminary study of North American insects associated with elderberry flowers. Florida Entomologist. 62(4): 341–355.

Goodwin, B. J. 2003. Is landscape connectivity a dependent or independent variable? Landscape Ecology. 18: 687–699.

Gosling, D. L. 1984. Flower records for anthophilous Cerambycidae (Coleoptera) in a southwestern Michigan woodland. The Great Lakes Entomologist. 17: 79–82.

Grimaldi, D., Engel, M. S. 2005. Evolution of the insects, First Eddition. Cambridge University Press.

Hammond, P. M. 1979. Wing-folding mechanisms of beetles, with special reference to investigations of Adephagan phylogeny (Coleoptera). In: Erwin, T. L., Ball, G. E., Whitehead, D. R., and Halpern, A. L. (eds). Carabid Beetles: Their Evolution, Natural History, and Classification. pp. 113–180.

Hanks, L. M. 1999. Influence of the larval host plant on reproductive strategies of cerambycid beetles. Annual Review of Entomology. 44: 483–505.

Hanski, I. 1994. Patch-occupancy dynamics in fragmented landscapes. Trends in Ecology and Evolution. 9: 131–134.

Hanski, I. 1991. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnaean Society. 42: 3–16.

Hanski, I., Alho, J., Moilanen, A. 2000. Estimating the parameters of survival and migration of individuals in metapopulations. Ecology. 81: 239–251.

Holderegger, R., Wagner, H. 2008. Landscape genetics. BioScience. 58(3): 199–207.

Hoechstetter, S., Walz, U., Dang, L. H., Thinh, N. X. 2008. Effects of topography and surface roughness in analyses of landscape structure: a proposal to modify the existing set of landscape metrics. Landscape Online. 1: 1–14.

Holland, J. D., Shukle, J. T., Abdel Moniem, H. E. M., Mager, T. W., Raje, K. R., Schnepp, K., Yang, S. 2012. Pre-treatment assemblages of wood-boring beetles (Coleoptera: Buprestidae, Cerambycidae) of the Hardwood Ecosystem Experiment. In: Swihart, R.K., Saunders, M.R., Kalb, R.A. *et al.* (eds). The Hardwood Ecosystem Experiment: a framework for studying responses to forest management. General Technical Report, NRS-P-108. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. pp. 218–236.

Holland, J. D. 2006. Cerambycidae larval host condition predicts trap efficiency. Environmental Entomology. 35(6): 1647–1653.

Kent, M. 2009. Biogeography and landscape ecology: the way forward - gradients and graph theory. Progress in Physical Geography. 33(3): 424–436.

Kindlmann, P., Burel F. 2008. Connectivity measures: a review. Landscape Ecology. 23: 879–890.

Knull, J. N. 1946. The longhorned beetles of Ohio. Ohio Biological Survey Bulletin. 39: 133–354.

Lawrence, J. F. 1982. Evolution and classification of beetles. Annual Review of Ecology and Systematics. 13: 261–290.

Lawrence, J. F., Hlavac, T. F. 1979. Review of the Derodontidae (Coleoptera: Polyphaga) with new species from North America and Chile. Coleoptrist Bulliten. 33: 369–414.

Lingafelter, S. W. 2007. Illustrated key to the longhorned woodboring beetles of the Eastern United States. Coleopterists Society Miscellaneous Publication. Special Publication No. 3. p. 206.

Linsley, E. G. 1959. Ecology of Cerambycidae. Annual Review of Entomology. 4: 99–138.

Linsley, E. G., Chemsak, J. A. 1972. The Cerambycidae of North America. Part VI (1). Taxonomy and classification of the subfamily Lepturinae. Berkeley, CA: University of California Press.

Linsley, E. G., Chemsak, J. A. 1976. The Cerambycidae of North America. Part VI (2). Taxonomy and classification of the subfamily Lepturinae. Berkeley, CA: University of California Press.

Maeto, K., Sato, S., Miyata, H. 2002. Species diversity of longicorn beetles in humid warm temperate forests: the impact of forest management practices on old-growth forest species in southwestern Japan. Biodiversity and Conservation. 11: 1919–1937.

Manel, S., Schwartz, K., Luikart, G., Taberlet, P. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18: 189–197.

Manning, A. D., Lindenmayer, D. B., Nix, H. A. 2004. Continua and umwelt: novel perspectives on viewing landscapes. Oikos 104(3): 621–628.

Marvaldi, A. E., Duckett, C. N., Kjer, K. M., Gillespie, J. J. 2009. Structural alignment of 18S and 28S rDNA sequences provides insights into phylogeny of Phytophaga (Coleoptera: Curculionoidea and Chrysomeloidea). Zoologica Scripta. 38: 63–77.

McGarigal, K., Cushman, S. A. 2005. The gradient concept of landscape structure. In: Wiens, J., Moss, M. (eds). Issues and Perspectives in Landscape Ecology. Cambridge University Press, Cambridge. pp. 112–119.

McGarigal, K., Cushman, S. A., Neel, M. C., Ene, E. 2002. FRAGSTATS: Spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available from http://www.umass.edu/landeco/research/fragstats/fragstats.html.

McGarigal, K., Tagil, S., Cushman, S. A. 2009. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landscape Ecology. 24(3): 433–450.

McGarigal, K., Cushman, S. A. 2005. The gradient concept of landscape structure. In: Wiens, J., Moss, M. (eds). Issues and perspectives in landscape ecology. Cambridge University Press, Cambridge. pp 112–119.

McIntyre, S., Barrett, G. W. 1992. Habitat variegation: an alternative to fragmentation. Conservation Biology. 6(1): 146–147.

Michelsen, A. 1963. Observations on the sexual behaviour of some longicorn beetles, Subfamily Lepturinae (Coleoptera, Cerambycidae). Behaviour 22: 152–166.

Minor, E., Urban, D. 2008. A graph-theory framework for evaluating landscape connectivity and conservation planning. Conservation Biology. 22: 297–307.

Pai, A., Sharakhov, I. V., Braginets, O., Costa, C., Yan, G. 2003. Identification of microsatellites markers in the red flour beetle, *Tribolium castaneum*. Molecular Ecology Notes. 3: 425–427.

Raje, K. R., Abdel Moniem, H. E. M., Farlee, L., Farris, V. R., Holland, J. D. 2012. Abundance of pest and benign Cerambycidae both increase with decreasing forest productivity. Agricultural and Forest Entomology. 14: 165–169.

Shelford, V. E. 1931. Some concepts of bioecology. Ecology. 12(3): 455–467.

Solomon, J. D. 1995. Guide to Insect Borers in North American Broadleaf Trees and Shrubs. Agricultural Handbook 706. Washington, DC: U.S. Department of Agriculture, Forest Service.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., Waits, L. P. 2007. Putting the landscape in landscape genetics. Heredity. 98: 128–142.

Swihart, R. K., Lusk, J. J., Duchamp, J. E., Rizkalla, C. E., Moore, J. E. 2006. The roles of landscape context, niche breadth, and range boundaries in predicting species responses to habitat alteration. Diversity and Distributions. 12: 277–287.

Taylor, P., Fahrig, L., Henein, K., Merriam, G. 1993. Connectivity is a vital element of landscape structure. Oikos. 68: 571–573.

Tischendorf, L., Bender, D., Fahrig, L. 2003. Evaluation of patch isolation metrics in mosaic landscapes for specialist vs. generalist dispersers. Landscape Ecology. 18: 41–50.

Tischendorf, L., Fahrig, L. 2001. On the use of connectivity measures in spatial ecology: a reply. Oikos. 95: 152–155.

Turner, M. G. 2005. Landscape Ecology: What is the state of the science? Annual Review of Ecology and Evolution. 36: 319–344.

Urban, D., Goslee, S., Pierce, K., Lookingbill, T. 2002. Extending community ecology to landscapes. Ecoscience. 9(2): 200–212.

White, R. E. 1983. Beetles: A Field Guide to the Beetles of North America. Houghton Mifflin, Harcourt, New York.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology. 3(4): 385–397.

Wu, J. 2007. Scale and scaling: a cross–disciplinary perspective. In: Wu, J., Hobbs, R. J. (eds). Key Topics in Landscape Ecology. Cambridge University Press, Cambridge.

Yanega, D. 1996. Field Guide to Northeastern Longhorned Beetles (Coleoptera: Cerambycidae). Illinois Natural History Survey Manual 6. 184 pp.

CHAPTER 2. SPATIAL REPLACEMENT CORRECTS FOR LANDSAT 7 ETM+ MISSING DATA PATTERNS

2.1 Abstract

Landsat 7 ETM+ imagery products provide researchers and decision makers in many fields with a large amount of remotely sensed data that serves many applications. However, a malfunction of the scan line corrector (SLC) onboard Landsat 7 causes substantial data gaps and data are available only as is with SLC-off mode. These data gaps may form an obstacle in using Landsat 7 ETM+ in many research disciplines. Several methods have been proposed to fix data gaps in Landsat 7 ETM+ imagery. These methods such as regression tree analysis, histogram-matching techniques, multi-scale segmentation approaches, and geostatistical based methods yield reliable results, but require sophisticated analyses and intensive computations and are still accompanied by some caveats. In this paper we demonstrate a spatial replacement method that is based on a simple neighborhood interpolation approach. It is implemented under Hawth's Tools and run in ArcGIS 9.2 to provide the scientific community with an easily applicable, relatively quick and potentially reliable correction for the missing data patterns in Landsat 7 ETM+ data. We demonstrate the efficiency of the technique for two color bands across Indiana, USA, and use these to calculate the normalized difference vegetation index (NDVI) that has many applications in ecological studies.

Keywords: Remote sensing; data gaps; interpolation; satellite imagery; NDVI

2.2 Introduction

The Landsat program of the National Aeronautics and Space Administration (NASA) and the U. S. Geological Survey (USGS) provides vast amounts of valuable data to researchers. Landsat 7 circles the Earth every 99 minutes at an altitude of 705 km (Arvidson et al. 2001). Landsat 7 (LS7) carries onboard the Enhanced Thematic Mapper Plus (ETM+), with 30 meter resolution visible red and near infra-red (NIR) bands, 60 meter resolution thermal band, and a 15 meter panchromatic band (USGS 2003). The imagery thus collected provides the global science community with a wealth of land-surface data that supports research in agriculture (e.g. Arvidson et al. 2000, Beltrán and Belmont 2001, Bentley et al. 2002), forestry (e.g. Rason et al. 2003, Trigg et al. 2006, Günlü *et al.* 2009), biodiversity and conservation ecology (e.g. Turner et al. 2003, Velazquez 2003, Cohen and Goward 2004), and others.

In 2003 a malfunction of the scan line corrector (SLC) began causing wedge-shaped areas of missing data ranging between a single pixel and 12 pixels in width (USGS 2003). The proposed methods to fill these data gaps (e.g. regression tree algorithm method, Quinlan 1993; linear histogram-matching method, USGS 2004) were applied in phase (I) and phase (II) gap filled product releases by earth resources and observation center (EROS) which showed great efforts on the part of United States Geological Survey (USGS) and the National Aeronautical and Space Administration (NASA) research teams in solving the SLC problem and make better use of the LS7 ETM+ data for the scientific community. However, these methods use multiple satellite scenes with different SLC modes (on and off) from different dates to build a multiple regression tree model that predicts the best closest value of the missing pixels in the data gaps. In addition to these methods, a few other proposals were introduced as potential substitutions to correct for LS7 ETM+ data gaps (e.g.

multi-scale segmentation approach (Maxwell et al. 2007), unscanned pixels' reflectance estimation via MODIS information (Roy et al. 2008) and geostatistical based methods (Pringle *et al.* 2009). Recently, a paper by Chen and colleagues (2011) introduced another approach for fixing the missing data patterns in LS7 ETM+ data. This neighborhood similar pixel interpolator method (NSPI) integrates data from different sources (e.g. LS7 with SLC-on and SLC-off mode, Landsat 5 data, Google Earth images and simulated data) to interpolate the best value for missing cells. All these techniques have contributed greatly to improve the output of the gap-free end product of LS7 ETM+ imagery. However, there are still some caveats and hurdles associated with the techniques. The methods proposed are complicated and not easily applied by non-remote sensing researchers. The methods are also quite computer-intensive and time consuming. The methods may become more challenging with issues such as cloud cover, adjacency of missing data lines in scenes, or large spatial extents. In such cases, the data that need to be manipulated will include more scenes with more overlapping areas. In turn, candidate scenes will be harder to find especially at close dates. Processing time may then also be substantial. Because the gap-filled products are no longer available on the USGS website, users may need to find an easy, applicable and reliable method to fix LS7 ETM+ data. Herein, we demonstrate one method to do so.

The technique we propose uses a simple spatial replacement method, implemented in Hawth's Tools and run in ArcGIS 9.2 (ESRI, Redlands, USA), that will provide researchers in different disciplines with an easy, relatively quick, and reliable correction for the missing data patterns in LS7 ETM+ acquired scenes with SLC-off mode. We also provide an example that illustrates the procedure and the

efficacy of its output in calculating normalized difference vegetation index (NDVI) in Indiana, USA. This layer will be used in subsequent research.

2.3 Materials and Methods

The LS7 ETM+ data for Indiana was downloaded from the USGS website (http://glovis.usgs.gov) using the USGS Global Visualization Viewer (USGS-GVV). LS7 ETM+ scenes on paths: 200, 210 and 220 and rows: 31, 32, 33 and 34, which cover the State of Indiana, USA, were acquired for the months April through October, 2008. We selected scenes of suitable dates for monitoring vegetation development in the study area (Table 2.1). For each scene band 3 and band 4, which represent the red and near Infra-red (NIR) spectra respectively, were processed. All selected scenes have a 30 m resolution, and are high quality and cloud free.

Raster data for bands 3 and 4 were processed independently in ArcGIS. For each band, rasters were stitched as mosaics of multiple input rasters into a single raster dataset that covers the extent of Indiana. The output cell value of the overlapping areas was selected to be the maximum value of the overlapped cells. The output raster was clipped to the extent of an Indiana polygon. To stay within the maximum number of pixels allowed for the spatial replacement tool to run (50 million pixels) we created a 100 x 100 km grid of polygons that were used to clip to the extent of Indiana. This yielded 16 polygons that were used to divide band 3 and band 4 composites.

A spatial replacement tool is implemented in Hawth's Tools for ESRI ArcGIS 9.x (Beyer 2004). The tool replaces cell values in a raster layer by assigning the closest acceptable permitted alternative value. It replaces unwanted categories from a raster layer by new values based on a simple deterministic neighborhood interpolator

analysis, instead of reclassification (Beyer 2004). The method works as well for linear classes like roads and rivers as it does for large patches such as agricultural fields and forests. It has an interactive interface that facilitates defining the set of acceptable replacement values. The spatial replacement tool examines the eight cells immediately surrounding the cell to be recorded for acceptable replacement values. If there are no suitable replacement values in these eight cells, the window moves out by one cell, and does the same procedure repeatedly until an acceptable replacement is found. We replaced zero values that represent missing data pixels within each LS7 ETM+ (SLC-off mode) acquired scenes with all possibilities of acceptable values from the same exact scenes composing each of band 3 (red spectrum) and band 4 (NIR spectrum) for Indiana.

The NDVI was calculated as (NIR – red) / (NIR + red) (Rouse *et al.* 1974). In ArcGIS raster calculator, the following formula (1) was used to obtain a non-truncated float NDVI raster layer with values ranging from -1 to +1.

$$NDVI = \frac{float([band4] - [band3])}{float([band4] + [band3])} \quad \ldots\ldots\ldots\ldots \text{(1)}$$

We scaled the initial values with the formula (2).

$$S_{\mathrm{NDVI}} = 100([NDVI] + 1) \quad \ldots\ldots\ldots\ldots\ldots \text{(2)}$$

This calculation will provide a range of NDVI between 0 and 200 with pixels values <100 indicating clouds and water bodies while values ≥100 indicating vegetation cover.

2.4 Results

A total number of 63 Landsat 7 ETM+ scenes were obtained. Among these scenes, 12 were classified as moderate quality scenes with up to 25 % cloud coverage. Nine scenes were of high quality and cloud free (maximum 10% cloud cover) scenes. The remaining scenes were of lesser quality and above 40% cloud cover (Table 2.1). Out of a total number of 32 separate rasters processed (16 for each spectral band) with the spatial replacement tool we produced two mosaic layers for each band (red and NIR) for Indiana as well as one mosaic raster of NDVI for the state (Figure 2.1.d).

Table 2.1. Landsat 7 ETM+ scenes downloaded from the USGS website to cover Indiana. The scenes cover the months April – October, 2008. Light shaded cells represent moderate quality scenes that were not used for calculating NDVI, while dark shaded cells are high quality scenes used for calculating NDVI. Scenes with no shading were eliminated. Blanks are outside of the extent of the Indiana polygon.

Path	Number of row															
	31				32				33				34			
200	-	-	-	-	-	-	-	-	93	125	157	205	-	-	-	-
	-	-	-	-	-	-	-	-	221	269	285	-	-	-	-	-
210	116	148	180	196	100	148	164	196	100	148	164	196	116	148	164	196
	228	260	292	-	228	260	276	-	228	260	276	-	244	260	276	-
220	107	139	171	187	107	139	171	187	107	139	171	187	107	139	171	203
	219	267	283	-	219	267	283	-	219	267	283	-	219	267	283	-

The final rasters showed a great integrity of scene features (Figure 2.1.b). There were no traces of the former patterns of missing data after the spatial replacement technique was applied to fill the gaps (Figure 2.1.a). Moreover, the resulting NDVI raster appeared to maintain feature integrity (Figure 2.1.d). To better

illustrate the efficacy of the spatial replacement technique we focused on an area of diverse land use and heterogeneous terrain adjacent to Bloomington, in Monroe County, Indiana. The patterns of missing data before correction (Figure 2.2.a) crossed forest patches, Lake Monroe, several streams, and roads. The final scene shows that these lines have completely disappeared and the replacement by the closest suitable neighbor pixels results in a gap free image (Figure 2.2.b).

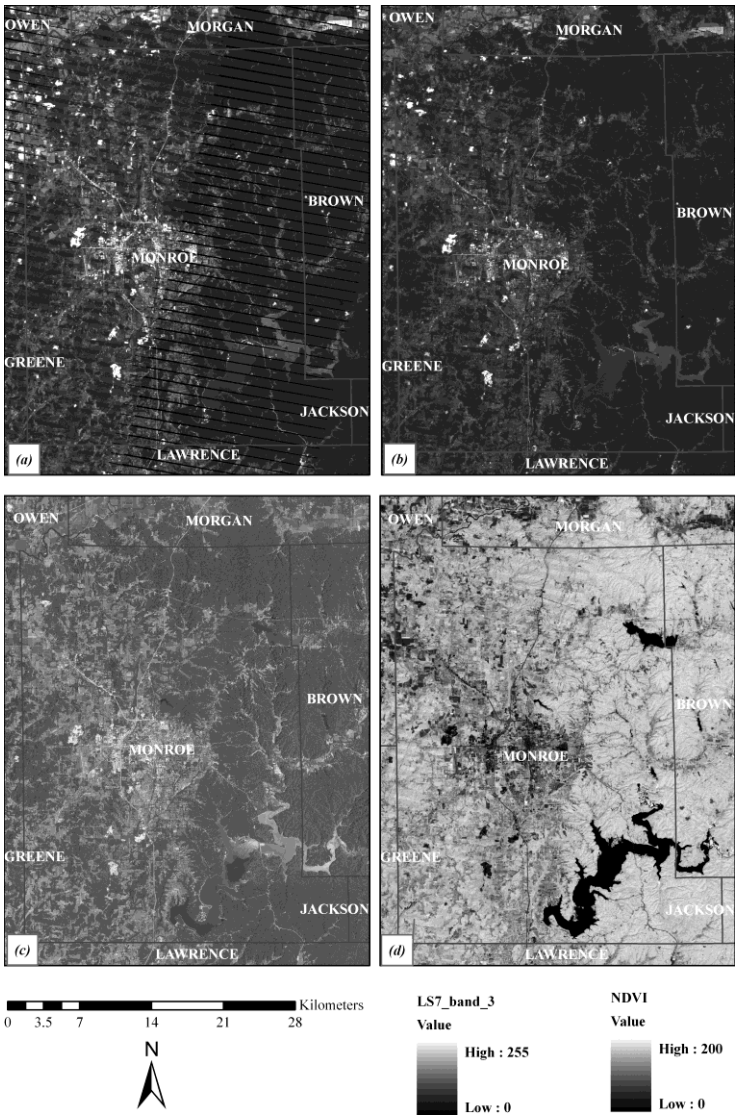


Figure 2.1. Example of the spatial replacement technique correcting for the missing data lines from LS7 ETM+ before and after spatial replacement tool is applied. (a)

Lines of missing data that occupied almost half of the clipped area to the eastern side of Monroe County. Artifact lines from the adjacent pattern are noticeable in the north western area of the scene. (b) The same area after we applied the spatial replacement tool on scene (a). Note that the lines from missing data have completely disappeared. (c) Clipped aerial photo of the same area for comparison. (d) Normalized difference vegetation index (NDVI) calculated from spectral bands 3 and 4 from LS7 ETM+ after correction.

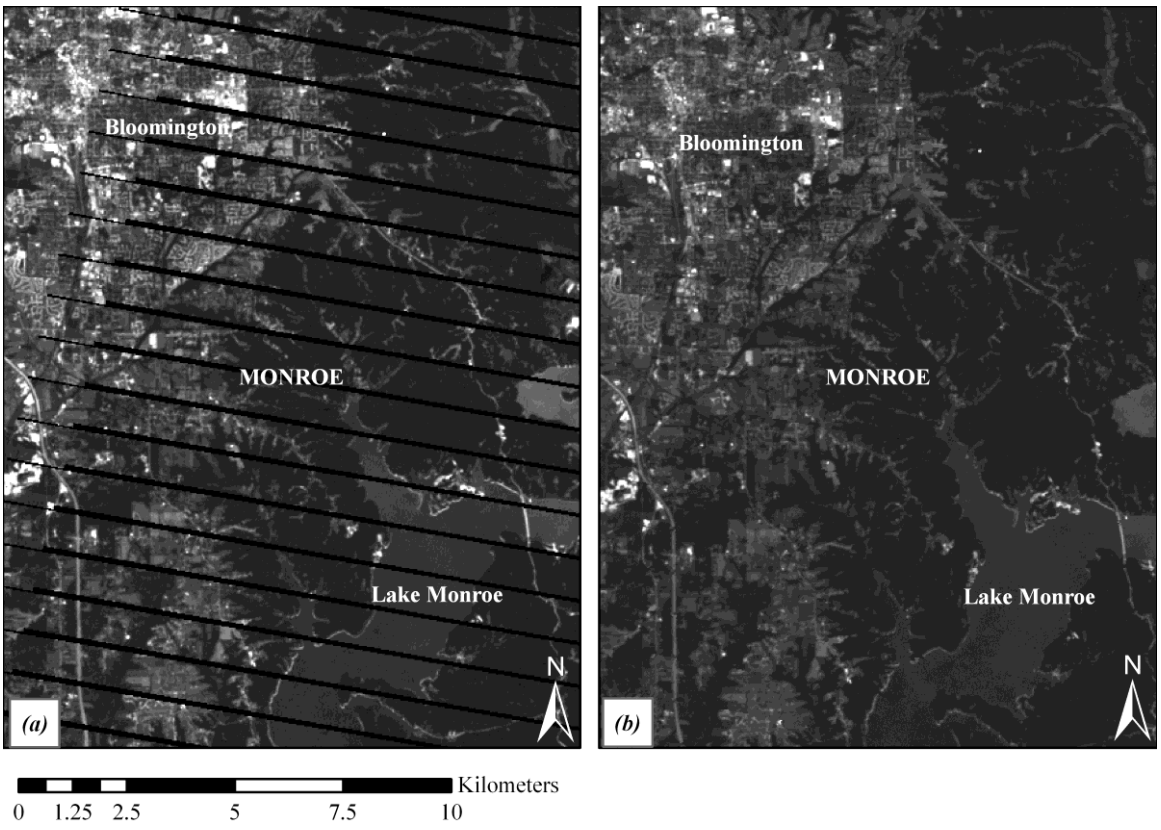


Figure 2.2. Monroe County before and after applying the spatial replacement method. Scene (a) shows the missing data patterns in a scan line corrector (SLC) off mode scene. Scene (b) shows the same County scene after correction. Note that no traces of the missing data show in (b).

2.5 Discussion

The spatial replacement method proved able to fix the patterns of missing data in LS7 ETM+ imagery products after the SLC malfunction. Although the formerly proposed methods produced some reliable results, they are still accompanied with caveats and they are not easily applicable for non-remote sensing specialist researchers. For example, the linear histogram-matching method (USGS 2004) and regression trees (USGS 2003) did not function uniformly in all scenes with missing data, especially in heterogeneous landscapes as they use scenes from different dates. As a result, a banding pattern can still occur at the site of formerly missing data, as an artifact of the differences between the remotely sensed data on different dates. This banding pattern varies from a very noticeable structure to a subtle one. Using information from other satellite systems to fill in the gaps of LS7 ETM+ (e.g. using MODIS information (Roy et al. 2008), is usually accompanied with the problem of scale. Despite the fact that MODIS has similar reflectance properties to ETM+, it has a coarser spatial resolution than LS7. Predicting the reflectance of the missing pixels in data gaps is very important and should be as accurate as possible for both small and large objects in the manipulated scenes. Both the multi-scale segmentation approach (Maxwell 2007) and geostatistical interpolation methods (e.g. Zhang et al. 2007, Pringle et al. 2009) share the disadvantage of estimating lower reflectance accuracy at the pixel level. In addition, the latter method is computationally intense and practically sophisticated. The latest proposed method, NSPI (Chen et al. 2011), showed a great advantage in accurately estimating the values of missing cells and improving results in both homogenous and heterogeneous landscapes. However, there are some hurdles associated with this method. For example, frequent cloud cover in multiple scenes and land use changes at different dates will complicate the process. In

addition, the interpolation method used cannot produce statistically uncertainty of each prediction.

In contrast to the above-mentioned approaches, the spatial replacement method provided a homogenous smooth surface at the places where missing data lines exist (Figure 2.1). The spatial replacement method produces quality corrected scenes at the lines of missing data. As with the NSPI method, the spatial replacement relies on a simple deterministic linear interpolation approach, however, it doesn't incorporate scenes from different sources at different dates. Conversely, it uses information from the same scene and hence, there is neither reflectance mismatch nor spatial scale issues with this method. The greatest advantage of the spatial replacement method lies in the fact that it is very easy to use, produces comparably accurate results and requires much less computational and processing time.

It is important to emphasize that any correction procedure has caveats that have to be dealt with carefully when using LS7 ETM+ datasets. Therefore, some limitations are also associated with the spatial replacement tool. For example, it processes only one spectral band of the raster at a time and replaces only one value at a time. On the technical side, it has been reported by Beyer (2004), the developer of Hawth's Tools, that issues have been found when using Hawth's Tools with recent versions of ArcGIS. However, the developer has provided parallel software that overcomes the incompatibility issues with recent version of ESRI ArcGIS products (Geospatial Modeling Environment, GME) (Beyer 2010). However, the spatial replacement tool is not implemented in this new software. The spatial replacement tool in Hawth's Tools (Beyer 2004) does function perfectly within ArcGIS 9.2 as we have done here.

Apparently, there is a trade-off in using LS7 ETM+ imagery products after year 2003. This trade-off lies in the fact that LS7 carries the Enhanced Thematic Mapper Plus (ETM+), providing the community with 30-meter visible and IR bands, a 60-meter spatial-resolution thermal band, and a 15-meter panchromatic band. The data support a variety of applications in areas as global change research, agriculture, forestry, geology, resource management, geography, mapping, water quality, and oceanography (USGS 2003). However, the issue of the missing data patterns might be problematic at some finer scales and high resolutions. Whether or not the users choose to fill the gaps, many users continue to find LS7 ETM+ data to be useful (Trigg et al. 2006).

Users of Geographical Information Systems (GIS) and remotely sensed data should use image processing software cautiously when attempting to repair, or minimize artifacts within, remote sensory data either for geometric or radiometric enhancement (Richards and Jia 2006). These programs may use different approaches such as Fourier transformation and Gaussian filtering. Image processing techniques may appear to yield improvements in the images; however these may or may not be conservative enough with the original dataset's values and the geospatial properties of the area being used. This is a very crucial issue that requires careful attention. In the approach we use here, the pixel values that were used to replace the missing values are quite consistent with those expected because they come from the same scene and therefore the same date and conditions. However, there may be some altering of the exact boundaries between patches of values or feature edges. The NSPI procedure will more likely preserve these edge locations at the cost of substantial processing and computational time. The user of any of these methods must first weigh these aspects of the different techniques and decide which is most suitable for their goal.

1.6 Bibliography

Arvidson, T., Gasch, J., Goward, S. 2000. Global vegetation assessing Landsat 7 ETM+ coverage of tropical rainforest and global agricultural and forest extents. Proceedings of International Geoscience and Remote Sensing Symposium. 1: 393–395.

Arvidson, T., Gasch, J., Goward, S. 2001. Landsat 7's long-term acquisition plan: an innovative approach to building a global imagery archive. Remote Sensing of Environment. 78: 13–26.

Beltrán, C. M., Belmonte, A. C. 2001. Irrigated crop area estimation using Landsat TM Imagery in La Mancha, Spain. Photogrammetric Engineering and Remote Sensing. 67: 1177–1184.

Bentley, M., Mote, T., Thebpanya, P. 2002. Using Landsat to identify thunderstorm damage in agricultural regions. American Meteorological Society. 83: 363–376.

Beyer, H. L. 2004. Hawth's analysis tools for ArcGIS. Available online at: http://www.spatialecology.com/htools (accessed February 2011).

Beyer, H. L. 2010. Geospatial modelling environment (GME). Available online at: http://www.spatialecology.com/gme (accessed February 2011).

Chen, J., Zhu, X., Vogelmann, J. E., Gao, F., Jin, S. 2011. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. Remote Sensing of Environment. 115: 1053−1064.

Cohen, W. B., Goward, S. N. 2004. Landsat's role in ecological applications of remote sensing. BioScience. 54: 535–545.

Günlü, A., Başkent, E. Z., Kadioğullar, A. I., Altun, L. 2009. Forest site classification using Landsat 7 ETM data: a case study of Maçka-Ormanüstü forest, Turkey. Environmental Monitoring and Assessment. 151: 93–104.

Maxwell, S. K., Schmidt, G. L., Storey, J. C. 2007. A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images. International Journal of Remote Sensing. 28: 5339−5356.

Pringle, M. J., Schmidt, M., Muir, J. S. 2009. Geostatistical interpolation of SLC-off Landsat ETM+ images. ISPRS Journal of Photogrammetry and Remote Sensing. 64: 654−664.

Quinlan, J. R. 1993. Combining instance-based and model-based learning. Proceedings of the 10th International Conference of Machine Learning. Morgan Kaufmann, Amherst, MA. 10: 236–243.

Ranson, K. J., Kovacs, K., Sun, G., Kharuk, V. I. 2003. Disturbance recognition in the boreal forest using radar and Landsat-7. Canadian Journal of Remote Sensing. 29: 271–285.

Richards, J. A., Jia, X. 2006. Remote Sensing Digital Image Analysis. Berlin: Springer. pp. 27−65.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W. 1974. Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the 3rd Earth Resources Technology Satellite-1 Symposium. Greenbelt, MD. pp. 301−317.

Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., Lindquist, E. 2008. Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. Remote Sensing of Environment. 112: 3112−3130.

Trigg, S. N., Lisa, M., Mcdonald, C. K., Mcdonald, A. K. 2006. Utility of Landsat 7 satellite data for continued monitoring of forest cover change in protected areas in Southeast Asia. Singapore Journal of Tropical Geography. 27: 49–66.

Turner, W. T., Spector, S., Gardiner, N. 2003. Remote sensing for biodiversity science and conservation. Trends in Ecology and Evolution. 18: 306–320.

USGS. 2003 Landsat: A global land-imaging project. Available online at: http://pubs.usgs.gov/fs/2010/3026/pdf/FS2010-3026.pdf (accessed February 2011).

USGS. 2003. Preliminary assessment of the value of Landsat 7 ETM+ data following scan line corrector malfunction. Available online at: http://landsathandbook.gsfc.nasa.gov/handbook/pdfs/SLC_off_Scientific_Usability.pdf (accessed February 2011).

USGS. 2004. SLC-off gap-filled products gap-fill algorithm methodology: Phase 2 gap-fill algorithm. Available online at: http://landsat.usgs.gov/documents/L7SLCGapFilledMethod.pdf (accessed February 2011).

Velazquez, A., Duran, E., Ramirez, I., François, J. M., Bocco, G., Ramírez, G., Palacio, J. L. 2003. Land use cover change processes in highly biodiverse areas: The case of Oaxaca, Mexico. Global and Environmental Change. 13: 175–184.

Zhang, C., Li, W., Travis, D. 2007. Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach. International Journal of Remote Sensing. 28: 5103−5122.

CHAPTER 3: HABITAT CONNECTIVITY FOR POLLINATOR BEETLES USING SURFACE METRICS

Hossam Eldien M. Abdel Moniem · Jeffrey D. Holland

H.M. Abdel Moniem[1,2] · J. D. Holland[1]

[1]Department of Entomology, Purdue University, West Lafayette, IN 47907, USA.

E-mail: jdhollan@purdue.edu

Phone: (765) 494-7739

[2]Department of Zoology, Faculty of Science, Suez Canal University, Ismailia 41522, Egypt.

Habitat connectivity for pollinator beetles using surface metrics. Landscape Ecology. 2013. 28: 1251–1267.

3.1 Abstract

Measuring habitat connectivity in complex landscapes is a major focus of landscape ecology and conservation research. Most studies use a binary landscape or patch mosaic model for describing spatial heterogeneity and understanding pattern-process relationships. While the value of landscape gradient approaches proposed by McGarigal and Cushman are recognized, applications of these newly proposed three dimensional surface metrics remain under-used. We created a gradient map of habitat quality from several GIS layers and applied three dimensional surface metrics to measure connectivity between 67 locations in Indiana, USA surveyed for one group of ecosystem service providers, flower longicorn beetles (Cerambycidae: Lepturinae). The three dimensional surface metrics applied to the landscape gradient model showed great potential to explain the differences of lepturine assemblages among the 2211 studied landscapes (between site pairs). Surface kurtosis and its interaction with geographic distance were among the most important metrics. This approach provided unique information about the landscape through four configuration metrics. There were some uniform trends of the responses of many species to some of surface metrics, however some species responded differently to other metrics. We suggest that three dimensional surface metrics applied to a habitat surface map created with insight into species requirements is a valuable approach to understanding the spatial dynamics of species, guilds, and ecosystem services.

Keywords:

Cerambycidae · Fragmentation · Geographical Information System · Lepturinae · Spatial Modeling · Landscape Gradient Model · Surface Metrology

3.2 Introduction

Landscape ecologists have developed different paradigms to model landscapes to understand pattern-process relationships and help make more informed management decisions. These paradigms or frameworks for modeling complex landscapes include the patch mosaic model (Forman and Godron 1981), the variegation model (Mcintyre and Barrett 1992), and the modified habitat gradient models (Manning et al. 2004; Fischer and Lindenmayer 2006). All of these models have contributed to our understanding of biological and ecological processes in the landscape. The adoption of surface metrology for describing gradients across landscapes holds great promise to increase the tools and types of metrics available to landscape ecologists.

The patch mosaic model (PMM; Forman 1995) has been adopted in many studies and has led to many advances in our understanding of pattern-process relationships (Turner 2005). The model has great value due to its conceptual simplicity and consistency with well-developed landscape tools such as FRAGSTATS (McGarigal et al. 2002) and quantitative analysis techniques (e.g. ANOVA) (McGarigal et al. 2009). However, for some studies it is suboptimal because it is inconsistent with basic ecological theory and bypasses the continuous nature of habitat heterogeneity (McGarigal and Cushman 2005; Cushman et al. 2007; McGarigal et al. 2009; Cushman et al. 2010). The categorical representation of heterogeneity may result in an arbitrary characterization of patch classes and boundaries. Species have environmental requirements that support their survival and reproduction (Shelford 1931). These physical, chemical, and biological conditions are usually distributed in the landscape in a continuous rather than discreet manner (Wiens 1989; Wu 2007; McGarigal et al. 2009). Species respond to environmental

gradients (Whittaker 1967; Austin 2002; Cushman et al. 2007) which are biologically important for determining the optimum realized niche (Hutchinson 1957). Species composition of communities shifts along these gradients according to the intersection of species' niches and the spatial structure of the environment (Hutchinson 1957; Whittaker 1967 ; Rehfeld et al. 2006; Cushman et al. 2010). The variegation and modified habitat gradient models have advanced landscape modeling by using a less simplified conceptual framework for pattern-process studies. Although they do not provide a general conceptual approach to landscape structure (McGarigal et al. 2009), they have the benefit of considering the gradient nature of habitat heterogeneity. Models based upon habitat gradients such as the variegation model (Mcintyre and Barrett 1992) and further refined versions such as the continua-umwelt model (Manning et al. 2004; Fischer and Lindenmayer 2006; Farina 2010) view environmental variables and habitat heterogeneity as continuous entities in the landscape and analyze species responses as gradient attributes that correspond to habitat requirements. Newer landscape gradient paradigms (McGarigal and Cushman 2005) may be useful by allowing a more complex model of landscapes to be analyzed. However, this is done without the insights that may come from an umwelt perspective.

The issue of characterizing three-dimensional surfaces for ecological purposes started with the efforts of geomorphologists (e.g., Strahler 1952; Schumm 1956; Melton 1957) and biologists (e.g., Beasom et al. 1983; Sanson et al. 1995) to study geomorphological processes and wildlife habitat. For example, ecological studies on communities and species richness of vascular plants showed in many cases the connection between surface characteristics and biodiversity distribution models (Bolstad et al. 1998; Burnett et al. 1998; Sebastiá 2004). Such studies also showed the impact of relief on the differentiation of ecosystems and ecological functions as soil

moisture, temperature, solar irradiation, and microclimates (Swanson et al. 1988; Bailey 2009). Despite the fact that a number of techniques were developed to quantify and analyze surface complexity via a group of surface metrics (Pike 2000; Wilson and Gallant 2000; Jenness 2004) these methods were either on a cell based scale or focused on correcting planimetric projection of slopes (topography, as opposed to topology) in patch metrics (McGarigal et al. 2009). It was not until the recent work of several researchers (McGarigal and Cushman 2005; Hoechstetter 2008; Evans and Cushman 2009; McGarigal et al. 2009; Cushman et al. 2010) that real attention was given to the application of surface metrics for quantifying surface heterogeneity at a landscapes scale.

McGarigal and colleagues (2009) introduced a number of powerful and promising surface metrics to landscape ecologists. These metrics retain the continuous nature of environmental gradients. They are classified into three categories: amplitude, configuration, and bearing metrics. Some of the metrics are unique to surface metrology; they have no analogous metric in categorical approaches to landscape description. They may therefore open a new chapter in landscape ecology and lead to novel pattern-process hypotheses.

The characterization of habitat heterogeneity is a cornerstone for understanding pattern-process relationships in the landscape (Wu and Richard 2002; Cushman et al. 2010). Any of the above models may be appropriate depending on the study. Herein, we adopt the landscape gradient paradigm (McGarigal and Cushman 2005) and use three dimensional surface metrology metrics (surface metrics hereafter) of topology (not topography) to estimate connectivity across a surface of habitat quality to investigate how landscape habitat structure and heterogeneity shape the

lepturine beetle (Coleoptera: Cerambycidae) community in the fragmented forests of Indiana, USA.

Longicorn beetles (Coleoptera: Cerambycidae) play important ecological roles in forest ecosystems. Lepturine beetles, also known as flower longicorn beetles, are a diverse and abundant subfamily of these beetles with approximately 250 species described in North America (White 1983). They are mostly diurnal, often brightly colored cerambycids, and adults are commonly encountered on flowers on which they feed and mate (Michelsen 1963; Hanks 1999). Larvae of most species feed within decaying wood (Linsley 1959; Booth et al. 1990). Lepturines are providers of multiple ecosystem services: they help decompose dead wood and thus cycle nutrients and they are potential pollinators, with many species frequenting flowers of valuable hardwood trees such as the American chestnut (Benjamin 1907). Many species in this group use complementary habitats in different life stages. Larvae require decaying wood most reliably found in forests while adults of many species are common in more open areas with flowers. We adopted a landscape gradient approach to create a map of habitat quality for lepturines in Indiana. We then analyzed this map surface using surface metrics. We predicted that lepturine community similarity would correlate more to surface metrics that describe connectivity for these beetles than to Euclidian distance between communities, as these metrics contain much information on the intervening landscape. Our assumption in this study is that connectivity between study points is more important in determining community similarity than is habitat similarity at the points or neutral processes.

3.3 Materials and Methods

Our 67 study sites spanned a gradient of forest fragmentation across the State of Indiana, USA (Figure 3.1.a). Sites were sampled for one to six summers during 2005 – 2011 using similar but not identical arrays of beetle traps at each site. There were slight differences in the specific mix of traps used at these sites, but in all cases they included at least: one Lindgren funnel trap, one window flight intercept trap, and one panel trap for bark beetles (Figure 3.1.a). Each site also included either additional window traps, a purple sticky trap, or an additional Lindgren funnel trap. We used a subset of the data from each site representing beetles caught by the former three traps common to all sites. We further corrected for sampling effort by dividing by the years sampled. Lepturine beetles caught were identified to species using Yanega (1996), Lingafelter (2007), and Linsley and Chemsak (1972, 1976). All specimens collected reside in the research collection of the Landscape Ecology and Biodiversity Laboratory at Purdue University. We applied a cube root transformation to the effort- and trap-corrected abundances of species caught. We used the package ecodist (Goslee and Urban 2012) in R (R Development Core Team 2012) to calculate the Bray-Curtis dissimilarity index (BC index) between sites for the lepturine community. We also calculated a dissimilarity matrix for each species individually because we predicted that different species would respond differently, reducing the variance explained within the overall community results. This was the simple difference in corrected abundance between site pairs.

To create a raster map of habitat quality for lepturines, we incorporated six geographical information system (GIS) layers for Indiana. These biological and geophysical layers were chosen to represent habitat quality, food resources for both larvae and adults, and structural components of habitat connectivity. All map

calculations and geoprocessing were conducted using ArcGIS 9.2 (ESRI, Redlands, California) and the R packages raster (Hijmans and van Etten 2011) and SDMTools (VanDerWal et al. 2012). To consider the habitat gradients at an appropriate scale, we applied a moving window of 2.1 km to all GIS layers. This window size was based on the scale at which a common representative lepturine species, *Typocerus v. velutinus* (Olivier), responds to habitat amount and quality in the landscape (Yang 2010). We transformed all gradient layers to a mean of zero and a unit variance to facilitate comparing coefficients from a predictive model in the next step.

Land Cover - Land Use (biological) layers. The National Land Cover Data (NLCD) for 2001 is a 16 class land cover classification scheme that has been applied consistently across all states at a spatial resolution of 30 m (Homer et al. 2004). We clipped the NLCD to Indiana and reclassified it using the level II NLCD classification scheme. We created a binary forest layer by grouping all forest classes (deciduous, evergreen and mixed forest) into one class and designated the remaining pixels as non-forest. Many lepturine species use well-decayed wood and can develop within either conifer or deciduous logs and snags. For the final habitat quality surface, we resampled this layer to 300 m x 300 m resolution and used it to generate anther two layers: (1) percentage forest (Figure 3.1.b) and (2) splitting index (Fig. 1c) (Jaeger 2000) layers that were calculated using the same moving window (2.1 km) approach on each pixel in the State of Indiana. Using this coarse grain to measure forest cover leads to the loss of some precision in the percent forest, but has the benefit of aggregating larger forest patches that are separated by short distances that most lepturines can readily cross.

Normalized Difference Vegetation Index (NDVI; biological) layer. We used the NDVI as an indicator of the condition of forested areas. This index has been used

for detecting live green plant canopies in multispectral remote sensing data (Sellers 1985; Myneni et al. 1995). We included NDVI because forest productivity influences the predominantly dead-wood feeding lepturine species (Raje et al. 2012). While the link between NDVI and productivity or dead wood availability is not direct, we assume that NDVI serves as an indirect indicator of this. We created the NDVI layer for Indiana using remote sensing imagery from the Landsat 5 TM NASA satellite. Images covered the months June through September, 2008. We selected scenes of suitable dates for monitoring vegetation development in the study area. All selected scenes had a 30 m x 30 m spatial resolution, were of high quality, and were relatively cloud free (<10%). The NDVI was calculated according to Rouse et al. (1974). We scaled the initial values to a range between 0 and 200 with pixels values <100 indicating clouds and water bodies while values ≥100 indicating vegetation cover. We clipped this layer to the forest cover of Indiana to insure that our NDVI surface values will only represent forest vegetation and not be biased by the spectral absorbance of other features in the landscape, then we resampled this layer to the coarser spatial resolution of 300 m x 300 m (Figure 3.1.d).

Geophysical properties of landscapes partially determine soil quality, availability of nutrients, forest productivity and moisture content (Schoenholtz et al. 2000; Sebastiá 2004), and thus can influence habitat quality and biodiversity of longicorn species. We used three geophysical layers (DEM, curvature index and solar insolation) to create our habitat quality surface. We created the GIS surfaces for these layers as follows:

Digital Elevation Model (DEM; geophysical) layer. We clipped the 30 m x 30 m raster DEM of Indiana from the U.S. Geological Survey (USGS) National Elevation Dataset to the extent of Indiana and scaled up its resolution to 300 m x 300

m to match that of the coarsest resolution layer used (Figure 3.1.e). This will smooth out the terrain information in a small proportion of the state with more rugged terrain, but most of Indiana consists of quite gently varying elevations. We are interested here in differences between areas further apart, at the cost of information on terrain effects in a small area of the state.

Curvature Index (geophysical) layer. We used the 300 m x 300 m DEM layer to calculate the topographical curvature index for Indiana. While this could be done with the original 30 m x 30 m data, we were interested in the coarser-grained changes between the hilly areas of Indiana and the relatively flat areas. The curvature of the DEM surface was calculated as a second derivative of the surface slope. The calculation is conducted on a cell-by-cell basis, as fitted through that cell and its eight surrounding neighbors. The output was chosen to be the plan curvature that is perpendicular to the direction of the maximum slope (Figure 3.1.f).

Solar Insolation (geophysical) layer. Insolation is important for all stages of cerambycids (Barbalat 1996; Moretti and Barbalat 2004). We calculated the solar insolation layer for Indiana using the 300 m x 300 m DEM. The insolation was calculated for a multi-day solar radiation index (14 days intervals), measured as watt hours per square meter (WH/m2) and averaged for the period that spanned the adult activity season for the common lepturine species Typocerus v. velutinus from mid-June to late August 2008 (Figure 3.1.g).

To create the final 3D surface with the value of the z axis representing habitat quality for lepturine beetles, we first determined the relative importance of our variables using a generalized linear model (GLM) with a Poisson distribution. We extracted the values of each habitat variable around each site from the GIS surfaces and used these as predictor variables to model the transformed count of the flower

longicorn beetle *Typocerus v. velutinus*. We needed to select a representative species to determine surface coefficients and a common window size for all species because we are comparing the entire community (with BC index). *Typocerus v. velutinus* was chosen because it responds at a scale close to the average for this beetle family (Yang 2010) and because we are particularly interested in the dynamics of this species. The standardized coefficients of all significant predictors were then used as weights for each layer in combining them into a single map of habitat quality using raster math. Using the smoothed surfaces from the moving window analysis allowed us to apply the coefficients from the regression analysis to the layers in constructing the final habitat quality surface. This carried the cost however, of losing information on finer scale heterogeneity. After the final habitat quality surface was created, we reclassified all 'NoData' pixels to the minimum fitted value of that surface in order to obtain a continuous (non-perforated) final surface (Figure 3.1.h).

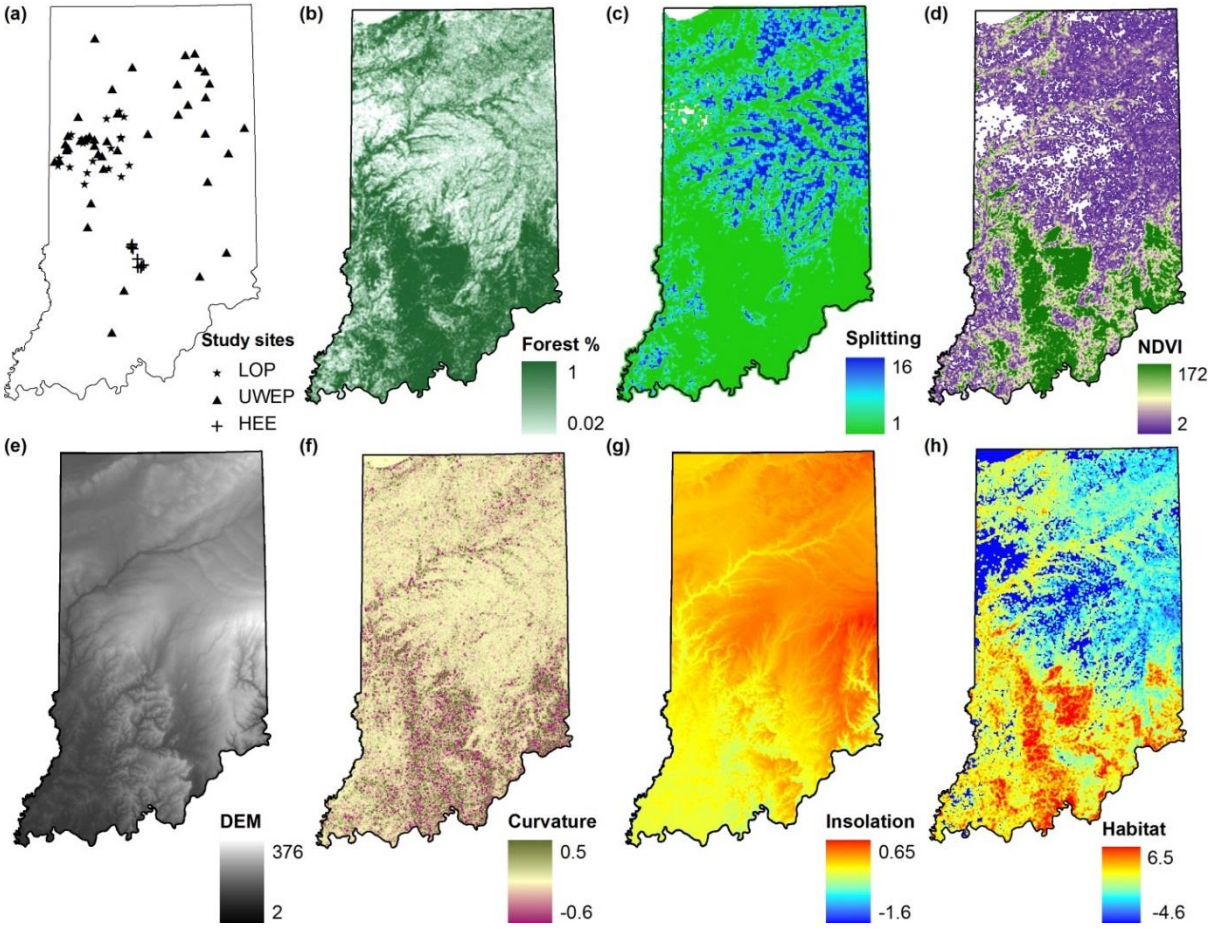


Figure 3.1 (a) Map of Indiana, USA, showing the 67 sampling sites. Triangles, Upper Wabash Ecosystem Project (UWEP) sites (3 array/yr); stars, Landowners Project (LOP) sites (1 array/yr); crosses, Hardwood Ecosystem Experiment (HEE) sites (5 array/yr). (b to g) the six GIS surfaces calculated for Indiana: (b) percent forest, (c) splitting index, (d) NDVI, (e) DEM, (f) curvature index, and (g) solar insolation. (h) Surface of habitat quality for the lepturine beetles.

We adopted a correlated random walk approach to determine the size and shape of a landscape most likely encountered by beetles dispersing between each possible pair of the 67 sampling sites (Okubo and Kareiva 2001). At a series of different distances apart, we used R to set 100,000 random walkers moving according to specific distributions of turning angle, step length, and total number of steps. For all walkers that successfully reached the other patch in the pair we averaged the minimum and maximum radii of an ellipse that entirely contained the path. The resulting relationship between distance between patches, and major and minor radii were used to determine the landscape between each pair of points depending on the distance between them (details in Koh et al. 2013). In R, we created the 2,211 elliptical landscapes and used these to clip the habitat quality surface for analysis. We measured ten surface metrics (Supplementary material) that demonstrate different characteristics of the habitat quality surface while possessing minimum possible redundancy among them (McGarigal et al. 2009). We used the Scanning Probe Image Processor (SPIP^{TM}) software to calculate the chosen surface metrics. These metrics, Euclidean distance, and the surface metric-distance interactions were predictor variables and the fixed components in multiple generalized additive mixed models (GAMM). Sampling site (one of the pair) was random effect variable to avoid pseudoreplication. Analyses were done separately for overall BC dissimilarity and for individual species. For model selection, we adopted the protocol described by Zuur et al. (2009). We started with the beyond optimal models that include all possible explanatory variables and interaction terms (fixed component) and we optimized our random component (sites) in these mixed models. We used the restricted maximum likelihood estimation (REML) and Akaike information criterion (AIC) to compare our models. In this procedure, we retained explanatory variables that passed an *F* statistic

significance test of level 0.05 in the optimal models. The selection of the best models was based on the lowest AIC.

3.4 Results

We caught 16 different species of lepturine beetles at our 67 sites (Table 3.1). The two most abundant species were *Analeptura lineola* (Say) and *Typocerus v. velutinus*. Of the six GIS surfaces we used to build the final 3D surface of habitat quality, NDVI, solar insolation and DEM values varied remarkably among study sites unlike the remaining three surfaces (splitting index, percent forest and curvature index) which varied less (Table 3.2). The six surfaces together explained 24.6% of the variance in the abundance of *Typocerus v. velutinus*. Habitat characteristics as represented by all six GIS layers significantly influenced this beetle's abundance. NDVI, curvature index and solar insolation were positively correlated with the beetle abundance while splitting index, percent forest and DEM were negatively correlated. NDVI as a measure of forest productivity, and the percent forest, were the most important explanatory surfaces and combined they explained 15.1% of the variance in abundance (Table 3.2). Splitting index and solar insolation explained much less variance.

Table 3.1 Abundance of the lepturine beetles in the 67 study sites in Indiana corrected for trap array composition and sampling effort.

Site	*Analeptura lineola* (Say)	*Bellamira scalaris* (Say)	*Brachyleptura champlaini* (Casey)	*Brachyleptura rubric* (Say)	*Gaurotes cyanipennis* (Say)	*Metacmaeops vittata* (Swederus)	*Necydalis mellita* (Say)	*Stenelytrana emarginata* (Fabricius)	*Strangalepta abbreviate* (Germar)	*Strangalia bicolor* (Swederus)	*Strangalia solitaria* (Haldeman)	*Strangalia luteicornis* (Fabricius)	*Strophiona nitens* (Forster)	*Trachysida mutabilis* (Newman)	*Typocerus deceptus* (Knull)	*Typocerus velutinus* (Olivier)
1	0.25				0.50							0.25				
2											1.00	1.00	0.50			2.00
3					1.00											
4	5.00		0.17			0.33		0.33				0.33				
5	13.33				0.33			0.17	0.33							
6	48.99				3.00							1.00				2.00
7	0.25			0.25								1.25				
8												0.50				
9	2.50		0.17		1.67											
10	11.60	0.40		0.40	1.20	2.40			8.00			5.60	0.40	0.40		10.40
11	5.20	0.40	0.80		0.80	1.20	0.80		0.80			4.80	0.40	0.40	0.40	3.60
12	8.40	0.40			2.80	0.40	0.40	0.40	0.40			1.60	0.40			0.80
13	0.80	0.40		0.40	1.20	0.40		0.40	1.60			4.40		0.40	0.40	14.40
14	0.40				0.40	0.40						0.80	0.40	0.40	0.80	11.99
15	3.20				0.80	0.80				0.40		0.40			1.20	5.60
16	3.20	0.40			1.60	0.80		0.80		0.40		0.80			1.20	38.00
17	0.80	0.80		0.40	2.40			0.40				0.40		0.40		4.00
18	0.40			0.40	0.80	0.40	0.40		1.20			0.80		0.40	0.40	6.80
19	2.00						0.25					0.75				
20												0.75				
21	16.00											1.00				
22	0.25					0.25	0.25					0.75				
23	0.50											0.25				
24												0.25				
25	0.50						0.50					1.00				
26	2.25											1.50				
27						3.00						6.00				
28	2.00											0.50				
29	7.34				0.67							0.67	0.33			
30																2.00
31	1.33				5.67							0.33				0.17
32			0.50									5.00				1.00
33											1.50	0.25				
34	8.75											0.50				
35												0.25				
36	0.50				2.50	0.50						1.00				
37					3.00											
38					1.00							1.00				1.00
39	12.17				1.17	0.33						0.67				0.17
40	16.99				0.50											
41												0.25				
42	2.00				1.50					0.33		0.17			0.67	0.33
43	3.50	0.17	0.17		0.33							1.17				0.17
44	4.00				0.50				0.33							0.67
45		0.33														
46	8.67											1.00				0.33
47	2.00				0.17			0.17				0.33				
48											2.00	1.00				
49	1.00				1.17											
50	4.50		0.33		0.33			0.33								
51	1.67				0.17							0.67				
52									0.67			1.00				1.33
53	55.18				0.17											0.33
54	1.00											1.00				
55	2.33															0.33
56						1.00										1.00
57					0.50					0.33		0.33				
58					0.50							1.00				
59	1.00											0.83				
60	2.17				0.50							0.50				
61					0.50	0.50										
62					9.00	1.00			0.50							
63	1.00					0.33						0.33				0.33
64	0.33				0.17	1.00						1.00				0.50
65	2.00											2.50				1.00
66	2.00										1.50	2.75				
67	0.25					0.25										
Total	269.50	3.30	2.14	1.85	48.52	15.29	2.60	3.00	13.83	1.46	6.00	62.18	2.43	2.40	5.07	110.25

Table 3.2 Descriptive statistics of the six GIS surfaces and results of the generalized linear regression model used to create the habitat quality surface. Coefficients represent the relative importance of predictor variables calculated from standardized surfaces.

GIS Surfaces summary statistics

Surfaces	Min	Max	Mean	SD	CV
Splitting	1.00	7.86	1.71	1.57	1.45
NDVI	3.45	170.86	60.83	50.28	41.57
Percent forest	0.08	1.00	0.65	0.26	0.11
DEM	153.04	331.25	222.98	35.35	5.61
Curvature	-0.14	0.13	-0.01	0.04	~0.00
Insolation	425954.30	437553.50	431045.10	2368.97	13.02

Generalized Linear Regression Model

Predictors	Coefficients	SE	z value	P (>|z|)	Relative importance (%)
Intercept	0.51	0.11	4.607	***	-
Splitting	-0.49	0.22	-2.243	*	7.33
NDVI	1.50	0.25	6.012	***	38.24
Percent forest	-0.96	0.27	-3.513	***	22.97
DEM	-1.26	0.25	-5.012	***	9.59
Curvature	0.20	0.07	2.975	**	15.32
Insolation	0.80	0.25	3.145	**	6.85

R^2= 0.246 AIC= 435.73 Significance at: *** P<0.001, ** P<0.01, * P<0.05

Within the ten surface metrics chosen to describe the topology and heterogeneity of the habitat quality surface, surface kurtosis (*Sku*), surface skewness (*Ssk*), and surface area ratio (*Sdr*) varied remarkably among the studied landscapes. Surface roughness (*Sa*), ten point height (*S10z*) and surface dominant texture direction

(*Std*) were the second most variant metrics among the landscapes while the remaining surface metrics varied little (Table 3).

Table 3.3 Summary of the ten surface metrics calculated from the habitat quality surface for lepturine beetles.

Surface metrics	*Sa*	*S10z*	*Ssk*	*Sku*	*Sdr*	*Sbi*	*Std*	*Stdi*	*Sfd*	*Srwi*
Min	~0.00	2.85	-365.63	1.00	10.78	0.27	0.00	0.09	2.19	0.01
Max	60.63	212.80	4.23	138612.00	592160.00	10.78	167.25	0.86	2.92	0.80
Mean	28.49	141.99	0.88	66.04	53332.40	0.54	58.39	0.23	2.40	0.01
SD	8.50	23.10	7.82	2947.12	26165.15	0.25	42.87	0.11	0.07	0.02
CV	2.54	3.76	69.26	131512.60	12836.75	0.12	31.48	0.06	~0.00	0.03

The ten surface metrics we calculated (Supplementary material, Table 3.3) depicted some important characteristics of the overall habitat quality surface of Indiana (Figure 3.1.h). For the amplitude metrics, there was an overall variability in the surface heights as demonstrated by mean value of roughness metrics *Sa* and *S10z* (28.49 and 141.99 respectively). The mean value of surface kurtosis showed that habitat quality surface was generally leptokurtic with uneven distributed height surface (*Sku*= 66.04). For the surface configuration metrics, there was a large variability in the surface slope and steepness as represented by the surface area ratio (*Sdr*= 53332.4) with a relative dominance of surface texture direction over all other texture directions (*Stdi*) of 0.23. The habitat quality surface for lepturines in Indiana generally showed very dominant radial wavelengths (*Srwi*=0.01) and a fractal surface with a dominant radial wavelength (*Sfd*=2.4). The surface had many high peaks with a mean surface bearing index (*Sbi*) of 0.54.

The 16 lepturine species had different relationships with different surface metrics. The vast majority of the species individually and the total community responded to at least one half of the 21 explanatory variables (ten surface metrics, geographical distance, and ten interaction terms) used in the generalized additive mixed models (Figure 3.2.a). Based on the values of standardized coefficients associated with our explanatory variables, among all surface metrics, surface kurtosis (*Sku*) and its interaction with geographic distance (*Sku:Geo_dist*) had the strongest relationship with beetle dissimilarities for both the total community and for individual species. Among the 16 studied species and the total community, seven individual species and the total community correlated strongly and significantly with *Sku* and *Sku:Geo_dist.* Examples of these species included *Bellamira scalaris* (Say), *Strophiona nitens* (Forster) and *Typocerus v. velutinus*. Contrary to these, nine other species did not respond to these two variables, e.g., *Analeptura lineola*, *Brachyleptura champlaini* (Casey), and *Strangalia solitaria* (Haldeman). Both *Sku* and its interaction with the geographic distance showed a significant negative correlation with the BC index. The second most important variable was the interaction between the ten point height and geographic distance between sites (*S10z:Geo_dist*). The community and almost all individual species correlated negatively with *S10z:Geo_dist* metric except for two species that correlated positively: *Gaurotes cyanipennis* (Say) and *Strangalia luteicornis* (Fabricius). Only five species (*Analeptura lineola, Brachyleptura. champlaini, Necydalis mellita* (Say)*, Strangalia bicolor* (Swederus) and *Strangalia solitaria* did not respond to this variable. Geographic distance (*Geo_dist*) between sampling sites came next in importance. The BC index values among sites for the total community and seven individual species correlated negatively with the geographic distance between sites. Examples of these species are *Brachyleptura*

rubrica (Say), *Strangalia bicolor* and *Typocerus v. velutinus*. The difference in abundance for nine remaining species, however, did not correlate with geographic distance [e.g. *Analeptura lineola*, *Stenelytrana emarginata* (Fabricius), and *Strangalia luteicornis* (Fabricius)].

The interaction of the ten surface metrics we used in our study with the geographical distance between sites is another important finding in our results. The ten metrics showed three different patterns on interacting with geographical distance between sites as explanatory variables. First, the interaction term for surface metrics like *Stdi* and *S10z* was able to explain the variance in beetle dissimilarities for about twice as many beetle species than the metrics alone. For example *Stdi* explained the variance in abundance dissimilarity for four beetle species, whereas *Stdi:Geo_dist* was able to explain this for nine species plus the total community. The second pattern is found in metrics such as *Sa* and *Std* which are found to be able to explain the variance in beetle dissimilarities for more species than their interaction terms with the geographical distance can do. For instance *Sa* was able to explain the variance in abundance of seven species while *Sa:Geo_dist* explained it for just two species. Also *Std* explained the variance in six species but *Std:Geo_dist* explained it for only two species. Finally, the remaining surface metrics and their interaction terms with geographic distance were significant for approximately the same number of species (see Figure 3.2.a,b).

The total variance in the beetles' dissimilarities as explained by the best models varied among the total community and the 16 individual species as shown by the values of the adjusted R^2 (Figure 3.2.c). The surface metrics worked very well with some lepturine species such as *Trachysida mutabilis* (Newman), *Bellamira scalaris* and *Typocerus v. velutinus* where the best models explained a moderate

amount of variance of 22.9%, 20.9% and 17.3% respectively for these three species. Surface metrics explained lower amounts of variance in beetles' dissimilarities for most other species such as *Gaurotes cyanipennis, Typocerus deceptus* (Knull) and *Metacmaeops vittata* (Swederus) where the variance explained was 11.8%, 10.8% and 9.9% respectively. On the other hand, the total variance explained was less than 5% for the remaining species. Also, surface metrics were able to explain only 5.31% of the variance in dissimilarities between sites for the total lepturine community.

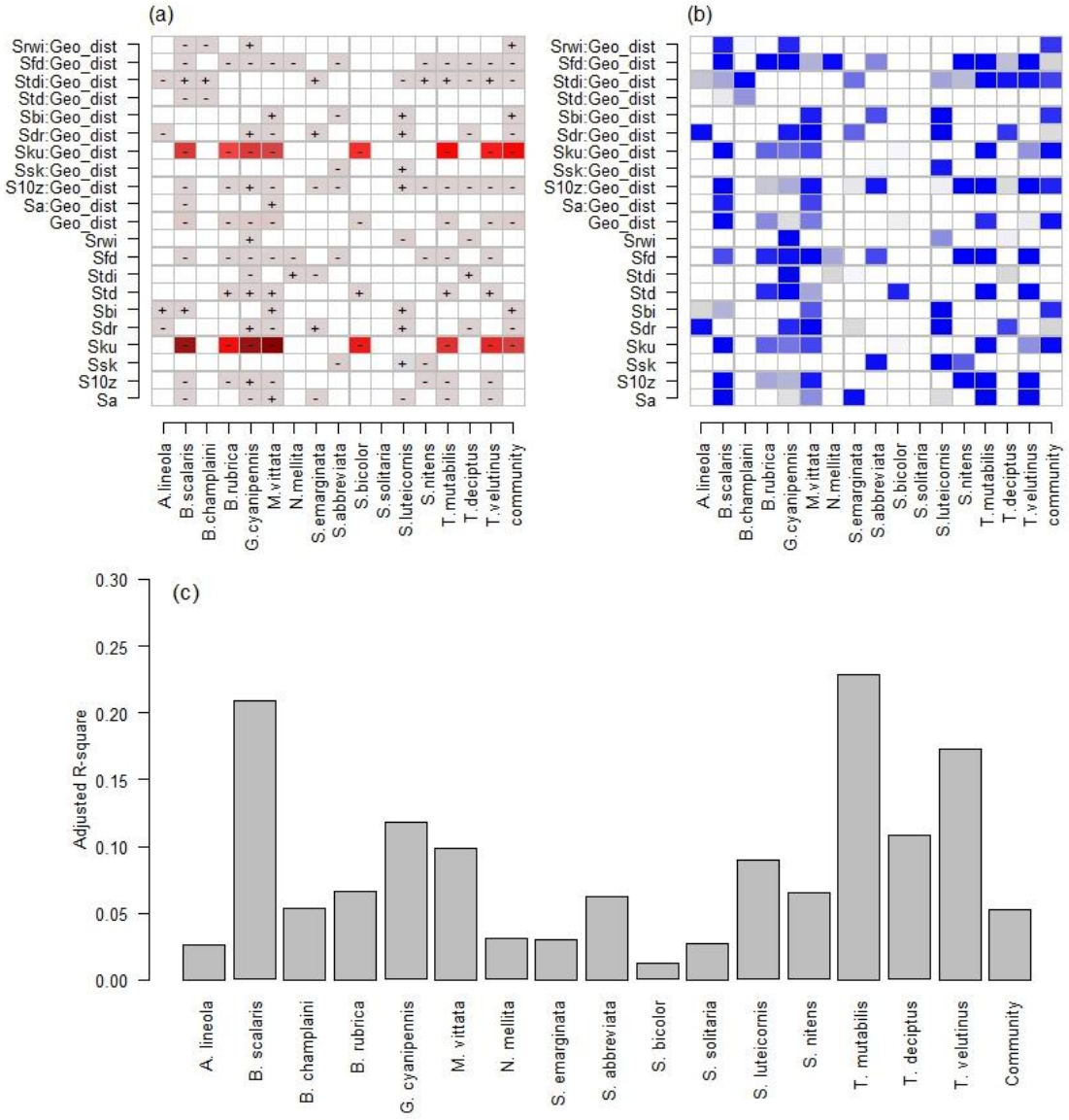


Figure 3.2 Summary of best generalized additive mixed models showing relative importance and significance of surface metrics in explaining variances in Bray-Curtis dissimilarity between sites for the beetle community and the simple difference in abundance of the 16 species. In a and b the Y-axis represents the 10 surface metrics measured, the geographical distance between sites, and the interaction term between each surface metric and geographic distance. Boxes in figure (a) are colored in a gradient from grey to dark red showing the relative importance of each explanatory variable based on the value of the standardized coefficients resulted from the models (values: -179.95 – 2.52). Signs in boxes are direction of the relationship. Blank boxes are variables that were eliminated during model selection to obtain the optimal models based on lowest AIC values. Boxes in figure (b) are the corresponding significance levels of each explanatory variable in (a). Significance is represented as a grey to blue gradient ($P<0.05$ to $P<0.001$, $n = 2211$, $df = 22$). Figure (c) represents the values of adjusted R^2 associated with best models for each species and the community.

3.5 Discussion

Topology metrics of the habitat quality surface explained differences in lepturine beetles species. Surface metrics thus seem able to serve as landscape analysis tools. A powerful characteristic of these metrics lies in their capability to describe both spatial and non-spatial aspects of a surface and to describe the continuous nature of gradients. Surface kurtosis (*Sku*) is an example of a non-spatial metric. This metric describes the peakedness of the surface height distribution and provides information on the heterogeneity of the surface. Higher values of kurtosis indicate high contrast landscapes dominated by high and low values (e.g., of habitat quality). Higher values of kurtosis thus indicate landscapes with a greater contrast, for

example, between habitat and matrix (Supplementary material). Seven of our 16 species and the overall community were more similar with higher kurtosis in the intervening landscape (negative values in Figure 3.2.a). This counters an expectation that a higher contrast landscape would be less conducive to movement. This raises the possibility that dissimilarity in beetle abundances are much a result of habitat similarity as they are of movement. Another possibility is that the higher contrast landscapes contain more high quality habitat and this is important for movement, while the lower contrast landscapes contain more area of intermediate-value 'habitat' that is less used and difficult to traverse.

Kurtosis in combination with skewness (*Ssk*) could be informative as these describe the degree and nature of land cover dominance in the landscape (McGarigal et al. 2009). Skewness is a measure of whether high or low values dominate the landscape (Supplementary material). Our results did not allow us to examine the effect of these two phenomena together because we had no beetle responses to both metrics. This may have been caused by little variation in skewness across our landscapes (Table 3.3).

Other amplitude-based metrics such as surface roughness (*Sa*) and ten point height (*S10z*) are potential measures of overall heterogeneity of the habitat quality surface. These two metrics are analogous to the patch-based diversity index from the PMM (McGarigal et al. 2009). We revisit these metrics below when we discuss the interaction between surface metrics and geographical distance.

In addition to amplitude-based surface metrics, our results emphasized four landscape configuration metrics that provided unique information about the landscape structure that are unavailable with categorical approaches of landscape analysis (Supplementary material; McGarigal et al. 2009; Cushman et al. 2010). For example,

the dominant texture direction (*Std*) measures the orientation of the dominant undulations of habitat quality in the landscape. Thus, this metric becomes valuable only if repeated changes in habitat quality occur in a particular direction and these repeated changes are of greater amplitude than those in other directions. An example of this is illustrated in the *Std* values of the landscapes in figure 3 which show a dominant direction to the changes of habitat quality in one landscape but not the other. This information could have application in determining the orientation of repeated high contrast areas that could constrain movement. This could provide warning of cumulative effects of repeated barriers that would individually not have a large effect on movement. This must be interpreted with care however, as the *Stdi* is calculated in comparison to other directions rather than in an absolute sense.

The Abbott curve calculated from the cumulative height distribution (Figure 3.3) may be another useful new landscape analysis tool. It can be used to graphically show the relative amounts of high, medium, and low values. This must be interpreted with caution because as with the original height distribution histogram or cumulative histogram, there is no indication of spatial location of these values. A large proportion of medium values therefore, does not necessarily come from a landscape with gradual transitions.

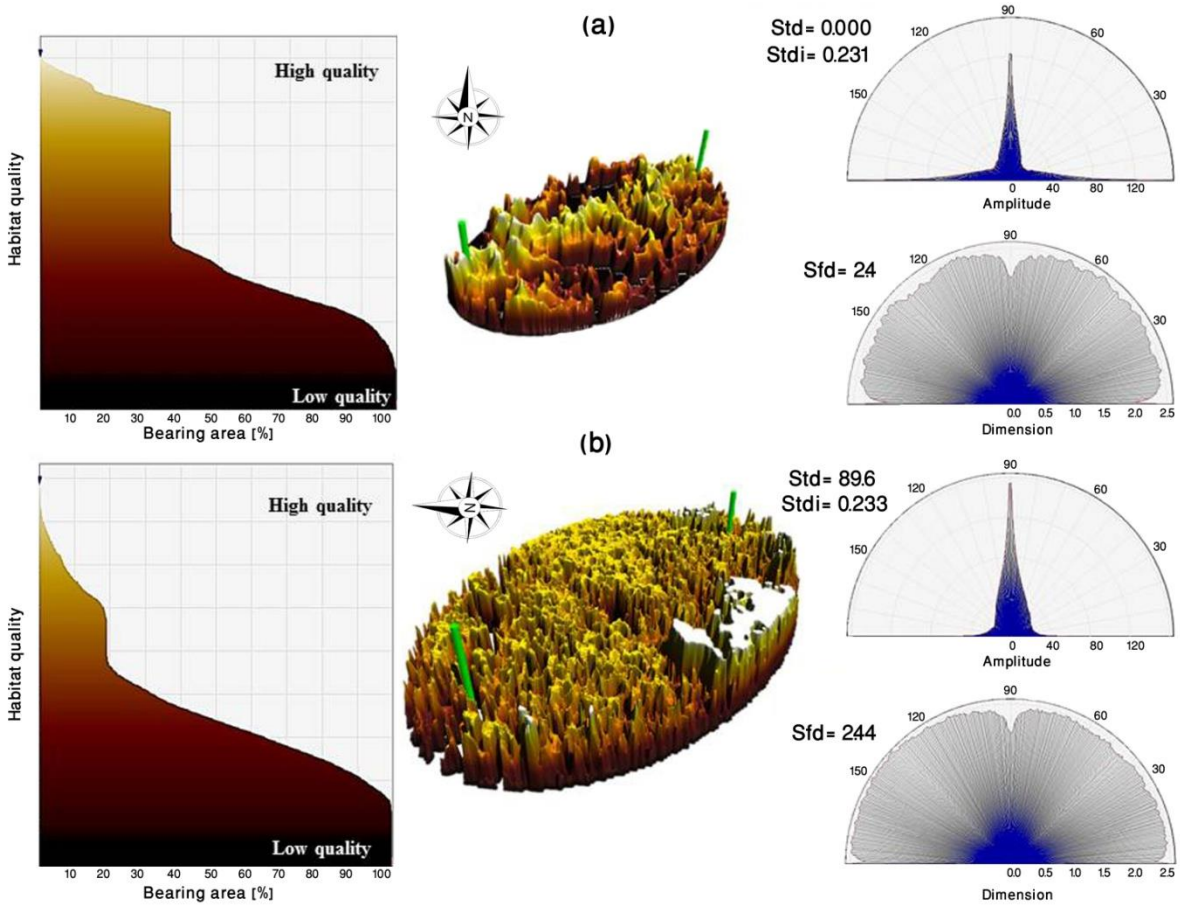


Figure 3.3 Example of two landscapes used in the study. (a) A landscape that is dominated by high quality habitat (peaks) as shown in the corresponding Abbott curve, and the *Std* figure of the same landscape showing a central direction of the high peaks. (b) A landscape dominated by medium quality habitat (see Abbott curve) but with a dominant direction of quality undulation. Both landscapes were similar in their (*Sfd*). Green bars represent the pair of sampling locations around which landscape ellipses were drawn.

Because some measures of habitat composition and configuration may influence animals differently depending on the distance traversed, we explored the interaction of surface metrics and geographic distances between sites. Our findings show that distance between habitats should be considered when using surface metrics. Measures most likely to predict movement between habitat areas are those focused on

quantifying high quality habitat. Several lepturine species and the overall community responded to some habitat quality surface metrics differently depending on the distance between sites, most obviously surface kurtosis (Figure 3.2). It is not only the amount of habitat and less hospital area that matters, but also the distance between points with this profile.

Two other examples of important interaction terms that explain differences between sites are the ten point height (*S10z*) and dominant texture direction index (*Stdi*) (Figure 3.2). While *S10z* and *Stdi* alone did not explain dissimilarities in the overall lepturine community, they did explain differences in seven and four species, respectively. However, the *S10z:Geo_dist* and *Stdi:Geo_dist* terms were significant for the overall community and eleven and nine individual species, respectively. The negative relationship between difference in abundance and both *S10z* and *S10z:Geo_dist* for almost all species is not surprising because *s10z* measures the mean amplitude of the most extreme values, both high and low, in the landscape. Higher values result from both better quality in the best habitats available and more hostile environments in the non-habitat areas. It seems reasonable to presume that it is the mean value from the highest parts of the surface that is responsible for the more similar numbers (Figure 3.2) in most species. Ten point height may be an indicator of how beneficial and how hostile different areas are, but this should likely be used in parallel with other metrics to ensure that this is not being caused by relatively rare peaks and valleys. Metrics that characterize the relative size of the upper and lower tails would do this. Interpretation of the *Stdi:Geo_dist* metric is difficult because the complementary dominant texture direction that gives the direction of the strongest amplitude pattern was not a predictor for most species.

Another amplitude metric that we found interesting in its behavior is average roughness (*Sa*). It is a measure of overall heterogeneity of the habitat quality surface calculated as the average height difference from the mean height. The metric is analogous to a patch-based diversity index from the PMM (McGarigal et al. 2009), especially if the latter is area-weighted. This metric is interesting in that it is a predictor of abundance difference for only two species as an interaction term with *Geo_dist*, but on its own it has much better explanatory power, predicting differences in seven species. This suggests that variation in the landscape has an influence, but that the distance over which this operates does not. An alternative explanation that we cannot discount is that species differences were related to *Geo_dist* interactions because species assemblages are less similar with distance (e.g., beta diversity) due to factors other than frequent dispersal.

Connectivity will be determined by the connectedness of intervening habitat areas and the dispersal ability and behavior of the species (Taylor et al. 1993). Factors facilitating or impeding movement are species specific and may not be predicted by patch edges and inter-patch distances (Cushman 2006). Given the species specific nature of connectivity, it is not surprising that there is variation in the responses among species to the metrics and to the distance interaction terms (Figure 3.2). Although all species included in this study have similar complementary habitat requirements and thus may be expected to respond similarly to the integrated habitat quality surface, the species possessed some interspecific variation in their responses. Almost all species showed a uniform response to several surface metrics (e.g., *Sku*, *Sku_Geo*:*dist*, *Sbi*, *Sbi_Geo*:*dist* and *Std*). In contrast, some other metrics exhibited variation in the nature of the lepturines' responses. Longicorn beetles have different larval host plant requirements (Hanks 1999) and even generalist species may have

some preference depending on the availability and the condition of the larval host wood. Ulyshen et al. (2004) and Makino et al. (2007), attributed the diversity of saproxylic beetles (dependent upon dead wood) and the differences in the assemblages among sites to the amount of coarse woody debris and forest open areas with flowers. As these two requirements are crucial for the lepturines to complete their life cycle, differences in availability and distribution of both requirements in our landscapes may have contributed to shaping the lepturine community in our study. Species may also have different preferences for environmental conditions such as solar insolation, which varied among the study landscapes. With regard to the overall community response, apparently, the low amount of variance explained by the best surface metrics model (5.3%) is a reflection of the different individual responses of species. The species that did not show strong correlations (e.g. *Analepturalineola, Strangalia bicolori* and *Strangalia solitaria*) with the surface metrics may have resulted in an overall drop of the community trend by diluting a portion of the higher correlations with the same metrics in other species (e.g. *Trachysida mutabilis*, *Typocerus v. velutinus* and *Bellamira scalaris*).

One of the greatest challenges facing landscape ecologists is the integration of niche theory with spatial ecology. This challenge crystallizes in linking non-spatial niche relationships with the spatial patterns of environmental gradients in complex heterogeneous landscapes (Austin 1985; Mcintyre and Barrett 1992; Urban et al. 2002; Manning et al. 2004; Cushman et al. 2007). Consequently, a new paradigm that considers a gradient approach of environmental conditions and heterogeneity is a step forward for many studies. As shown in our study, the landscape gradient paradigm (McGarigal and Cushman 2005) and surface topology metrics are powerful approaches to study the influence of habitat heterogeneity on lepturine beetle species

communities. The requirements of these species for complementary habitats and habitat quality determinants that have an inherently continuous range make it important to consider habitat as a continuous attribute to avoid oversimplification (McGarigal and Cushman 2005; Hoechstetter 2008; Kent 2009).

Before applying a landscape gradient approach and using surface metrics, it is important to consider environmental gradients relevant for the species of interest. Because habitat suitability is largely determined by availability of resources and conditions that support survival and reproduction of organisms (Hutchinson 1957), we incorporated six habitat requirement gradients into a final surface of habitat quality. These gradients were sampled at the same spatial resolution with insight into an appropriate scale for the beetles (Yang 2010). By integrating the biologically-important landscape gradients into a final surface of habitat quality we analyzed the responses of many lepturine species simultaneously with each responding individually to multiple landscape gradients (Cushman et al. 2010). It remains a possibility that some of the lepturine beetles in this study respond to the gradients used at a spatial scale different from that which we settled upon, weakening the perceived relationships.

Our study shows that 3D surface metrology metrics are a valuable extension of the existing set of landscape metrics. More effort and attention should be directed towards this new landscape gradient paradigm. Future studies should examine how to interpret multiple metrics in concert (e.g., skewness + kurtosis) to better resolve response trends.

Acknowledgements:

We thank Dr. Kevin McGarigal and Dr. Samuel Cushman for their feedback on the metrics used, Anders Kühle for the great technical support with the SPIP™ program, and Jeremy VanDerWal for help in implementing the SDMtools package in R. Three very helpful reviews helped us improve the analysis and the text. Insu Koh performed the random walk analysis to determine landscape shape. Kapil Raje, John Shukle and Tommy Mager helped with field collections. This research was supported by a governmental general mission scholarship administrated by the Egyptian Cultural and Education Bureau, Washington, DC, and by the Department of Entomology at Purdue University.

Supplement 3. 1 Explanation, biological interpretation, and PMM analog of the ten surface metrics calculated for 2211 landscapes. This table is based largely on the supplementary material from McGarigal et al (2009).

Metric		Characteristics	Meaning	Interpretation	PMM
Amplitude Metrics					
Sa	Average surface roughness	–Measures aspects of landscape composition but not configuration. –Sensitive to overall height distribution.	Average deviation of height from mean height.	–Non-spatial measures of landscape diversity. –The larger the values of *Sa* and *S10z* the larger the landscape richness.	Patch-based diversity metrics
S10z	Ten- point height		Average difference between surface mean and most extreme heights and depths.		
Ssk	Skewness	Measures symmetry of the surface height distribution.	Skewness of height distribution towards high or low values.	High (positive or negative) skewness indicates a landscape dominated by high or low values.	Patch-based evenness metrics.
Sku	Kurtosis	A measure of the shape of the surface height distribution. Sensitive to deep valleys or high peaks.	–Peakedness of the height distribution. –More constant height = Platykurtic (*Sku* < 3). –Large-tailed height distribution =Leptokurtic (*Sku* > 3).	–High kurtosis: landscape with greater abundance of high and low values –Low kurtosis: landscape with greater abundance of mean values. –In combination with surface skewness it explains the degree and nature of landscape dominance.	Contrast between habitat and matrix.
Configuration Metrics					
Sdr	Surface area ratio	Configuration metrics common shared characteristics: – Measure compositional and configurational aspects of the landscape. – Measure horizontal and vertical aspects of surface deviation. – Sensitive to variability in distribution and spatial arrangement of heights.	–Ratio between the surface area to the area of a flat plane with the same x-y dimensions. –Totally flat surface: *Sdr* = 0 %. –*Sdr* increases with the local slope variability.	Increasing variability and steepness of local slopes: increasing density of edges and the magnitude contrast between abutting high quality habitat areas along those edges.	Contrast-weighted edge density metric.
Std	Dominant texture direction		Direction of the dominant amplitude calculated from the Fourier spectrum.	–Ranges between 0-180 –Only meaningful if there is a dominate direction and is =0 otherwise. –Direction that crosses repeated higher contrasts.	NA
Stdi	Texture direction index		Relative dominance of amplitude in direction *Std* over other directions.	–Ranges from 0 - 1. –Surfaces with very dominant directions: *Stdi* ~ 0. –If all directions are similar: *Stdi* ~ 1.	NA
Srwi	Radial wavelength index		Relative dominance wavelengths over all other radial distances.	–Ranges from 0 – 1. –Surfaces with very dominant radial wavelengths: *Srwi* ~ 0. –If there is no dominating wavelength: *Srwi* ~ 1.	NA
Sfd	Fractal dimension		Calculated for the different angles of the angular spectrum by analyzing the Fourier amplitude spectrum.	–Ranges from 2 – 4. –Larger values indicate a fractal surface with an increasing dominant radial wavelength.	NA
Bearing Metric					
Sbi	Surface bearing index	–Cumulative measure of vertical aspects of surface deviation based on Abbott curve. –Landscape composition only metrics. –Measure of the surface height shape profile. –Sensitive to occasional high peaks and not deep valleys.	–Ratio of the root mean square roughness to the height from the top of the surface to the height at 5% bearing area. –Normal height distribution: *Sbi* = 0.608. –Relatively few high peaks: *Sbi* < 0.608. –Relatively many high peaks or no high peaks: *Sbi* > 0.608.	Measure of landscape dominance and nature of the surface composition.	Matrix and patch distribution in the landscape.

3.6 Bibliography

Austin MP (1985) Continuum concept, ordination methods, and niche theory. Annu Rev Ecol Syst 16(1):39–61

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol Model 157(2–3):101–118

Bailey R (2009) Mesoscale: landform differentiation (landscape mosaics). In: Bailey R (ed) Ecosystem geography, from ecoregions to sites, 2nd edn. Springer, New York, pp 127–144

Barbalat S (1996) Influence of forest management on three wood-eating beetle families in the Areuse gorges (Canton of Neuchatel, Switzerland). Rev Suisse Zool 103:553–564

Beasom SL, Wiggers EP, Giardino JR (1983) A technique for assessing land surface ruggedness. J Wildlife Manage 47(4):1163–1166

Benjamin D (1907) Annual report of the State Entomologist of Indiana. Library of the museum of comparative zoology, Harvard University, 216 pp

Bolstad PV, Swank W, Vose J (1998) Predicting southern Appalachian overstory vegetation with digital terraindata. Landscape Ecol 13:271–283

Booth RG, Cox M L, Madge RB (1990) IIE guides to insects of importance to man. 3. Coleoptera. CAB International, Wallingford, UK, 384 pp

Burnett MR, August PV, Brown JH et al. (1998) The influence of geomorphological heterogeneity on biodiversity: a patch–scale perspective. Conserv Biol 12(2):363–370

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: a review and prospectus. Biol Conserv 128(2):231–240

Cushman SA, Gutzweiler K, Evans JS et al. (2010) The Gradient Paradigm: a conceptual and analytical framework for landscape ecology. In: Cushman, SA and Huettmann F (eds) Spatial Complexity, Informatics, and Wildlife Conservation. Springer, New York, pp 83–108

Cushman SA, McKenzie D, Peterson DL et al. (2007) Research agenda for integrated landscape modelling. USDA For. Serv. General Technical Report RMRS-GTR-194

Evans JS, Cushman SA (2009) Gradient modeling of conifer species using random forests. Landscape Ecol 24(5):673–683

Farina A (2010) Ecology, Cognition and Landscape. Linking Natural and Social Systems. Springer, New York

Fischer J, Lindenmayer DB (2006) Beyond fragmentation: the continuum model for fauna research and conservation in human–modified landscapes. Oikos 112(2):473–480

Forman RTT (1995) Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, Cambridge

Forman RTT, Godron M (1981) Patches and structural components for a landscape ecology. Bioscience 31(10):733–740

Goslee S, Urban D (2012) ecodist: Dissimilarity-based functions for ecological analysis. R package version 1.2.7. http:// cran.r-project.org/web/packages/ecodist. Accessed 15 Feb 2012

Hanks LM (1999) Influence of the larval host plant on reproductive strategies of cerambycid beetles. Annu Rev Entomol 44:483–505

Hijmans RJ, van Etten J (2011) raster: Geographic analysis and modeling with raster data. R package version 1.8. http:// cran.r-project.org/web/packages/raster. Accessed 15 Feb 2012

Hoechstetter S, Walz U, Dang LH, et al. (2008) Effects of topography and surface roughness in analyses of landscape structure: a proposal to modify the existing set of landscape metrics. Landsc Online 1:1–14

Homer C, Huang C, Yang L, et al. (2004) Development of a 2001 national land-cover database for the United States. Photogramm Eng Rem S 70(7):829–840

Hutchinson GE (1957) Concluding remarks. Cold Spring Harb Symp Quant Biol 22:415–427

Jaeger JAG (2000) Landscape division, splitting index, and effective mesh size: new measures of landscape fragmentation. Landscape Ecol 15(2):115–130

Jenness J (2004) Calculating landscape surface area from digital elevation models. Wildl Soc Bull 30:829–839

Kent M (2009) Biogeography and landscape ecology: the way forward – gradients and graph theory. Prog Phys Geog 33(3):424–436

Koh I, Rowe H, Holland JD (2013) Graph and circuit theory connectivity models of conservation biological control. Ecol Appl 23:1554–1573.

Lingafelter SW (2007) Illustrated Key to the Longhorned Woodboring Beetles of the Eastern United States. Terry NS (ed) Coleopterists Society Special Publication No.3, Maryland

Linsley EG (1959) Ecology of Cerambycidae. Ann Rev Entomol 4:99–138

Linsley EG, Chemsak JA (1972) Taxonomy and classification of the subfamily Lepturinae, Part VI, No. 1. University of California Press, Berkely

Linsley EG, Chemsak JA (1976) Taxonomy and classification of the subfamily Lepturinae, Part VI, No. 2. University of California Press, Berkely

Makino Si, Goto H, Hasegawa M et al. (2007) Degradation of longicorn beetle (Coleoptera, Cerambycidae, Disteniidae) fauna caused by conversion from broad-leaved to man-made conifer stands of Cryptomeria japonica (Taxodiaceae) in central Japan. Ecol Res 22(3):372–381

Manning AD, Lindenmayer DB, Nix HA (2004) Continua and umwelt: novel perspectives on viewing landscapes. Oikos 104(3):621–628

McGarigal K, Cushman SA (2005) The gradient concept of landscape structure. In: Wiens J, Moss M (eds) Issues and Perspectives in Landscape Ecology. Cambridge University Press, Cambridge, pp 112–119

McGarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS: spatial pattern analysis program for categorical maps. Available from http://www.umass.edu/landeco/research/fragstats/fragstats.html

McGarigal K, Tagil S, Cushman SA (2009) Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landscape Ecol 24(3):433–450

Mcintyre S, Barrett GW (1992) Habitat variegation: an alternative to fragmentation. Conserv Biol 6(1):146–147

Melton MA (1957) An analysis of the relations among elements of climate, surface properties, and geomorphology. ONR Report 11 NR 389-042 Department of Geology, Columbia University, New York

Michelsen A (1963) Observations on the sexual behaviour of some longicorn beetles, Subfamily Lepturinae (Coleoptera, Cerambycidae). Behaviour 22(1–2):152–166

Moretti M, Barbalat S (2004) The effects of wildfires on wood-eating beetles in deciduous forests on the southern slope of the Swiss Alps. Forest Ecol Manag 187(1):85–103

Myneni RB, Hall FG, Sellers PJ, et al. (1995) The interpretation of spectral vegetation indexes. IEEE T Geosci Remote 33(2):481–486

Okubo A, Kareiva P (2001) Some examples of animal diffusion. In: Okubo A and Levin SA (eds), Diffusion and Ecological Problems: Modern Perspectives. Springer, New York, pp 170–196

Pike RJ (2000) Geomorphometry diversity in quantitative surface analysis. Prog Phys Geogr 24:1–20

R Development Core Team (2012) R: A language and environment for statistical computing. R foundation for statistical computing, version 2.14.1. Vienna, Austria. ISBN 3–900051–07–9. Available from http://www.R–project.org

Raje KR, Abdel-Moniem HEM, Farlee L, et al. (2012) Abundance of pest and benign Cerambycidae both increase with decreasing forest productivity. Agr Forest Entomol 14(2):165–169

Rehfeld GE, Nicholas LC, Marcus VW, et al. (2006) Empirical analyses of plant-climate relationships for the western United States. Int J Plant Sci 167(6):1123–1150

Rouse JW, Haas RH, Schell JA, et al. (1974) Monitoring vegetation systems in the Great Plains with ERTS Third ERTS–1 Symposium, Section A. NASA/GSFC United States, pp 309–317

Sanson GD, Stolk R, Downes BJ (1995) A new method for characterizing surface roughness and available space in biological systems. Funct Ecol 9:127–135

Schumm SA (1956) Evolution of drainage basins and slopes in badlands at Perth Amboy, New Jersey. Bull Geol Soc Am 67:597–646

Sebastiá MT (2004) Role of topography and soils in grassland structuring at the landscape and community scales. Basic Appl Ecol 5(4):331–346

Sellers PJ (1985) Canopy reflectance, photosynthesis and transpiration. Int J Remote Sens 6(8):1335–1372

Shelford VE (1931) Some concepts of bioecology. Ecology 12(3):455–467

Schoenholtz SH, Miegroet HV, Burger JA (2000) A review of chemical and physical properties as indicators of forest soil quality: challenges and opportunities. Forest Ecol Manag 138(1–3):335-356

SPIP™ The scanning probe image processor. Image metrology. APS, Lyngby. Available from http://www.imagemet.com/

Strahler AN (1952) Hypsometric (area-altitude) analysis of erosional topography. Bull Geol Soc Am 63:1117–1142

Swanson FJ, Kratz TK, Caine N, et al. (1988) Landform effects on ecosystem patterns and Processes. Bioscience 38(2):92–98

Taylor P, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. Oikos 68:571–573

Turner MG (2005) Landscape ecology: what is the state of the science? Annu Rev Ecol Evol S 36:319–344

Ulyshen MD, Hanula JL, Horn S, et al. (2004) Spatial and temporal patterns of beetles associated with coarse woody debris in managed bottomland hardwood forests. Forest Ecol Manag 199(2–3):259–272

Urban D, Goslee S, Pierce K, et al. (2002) Extending community ecology to landscapes. Ecoscience 9(2):200–212

VanDerWal J, Falconi L, Januchowski S, et al. (2012) SDMtools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises. R package version 1.1-5. http://cran.r-project.org/web/packages/SDMTools. Accessed 15 Feb

White RE (1983) Beetles: A Field Guide to the Beetles of North America. Houghton Mifflin, Harcourt, New York

Whittaker RH (1967) Gradient analysis of vegetation. Biol Rev 49:207–264

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3(4):385–397

Wilson JP, Gallant JC (2000) Terrain Analysis: Principles and Applications. Wiley, New York

Wu J (2007) Scale and scaling: a cross-disciplinary perspective. In: Wu J, Hobbs RJ (eds) Key Topics in Landscape Ecology. Cambridge University Press, Cambridge

Wu J, Richard H (2002) Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. Landscape Ecol 17:355–365

Yang S (2010) Landscape Scaling and Occupancy Modelling with Indiana Longhorrned Beetles (Coleoptera: Cerambycidae). Dissertation, Purdue University

Yanega D (1996) Field Guide to Northeastern Longhorned Beetles (Coleoptera: Cerambycidae). Illinois Natural History Survey, Champaign, Illinois

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed effects modelling for nested data. In: Gail M, et al (eds) Mixed Effects Models and Extensions in Ecology with R. Springer, New York pp 101–142

CHAPTER 4: PHYLOGEOGRAPHY AND DEMOGRAPHIC HISTORY OF A POLLINATOR LONGHORN BEETLE (*TYPOCERUS V. VELUTINUS*) SHAPED BY THE QUATERNARY

Abdel Moniem, H. M.[1], Schemerhorn, B. J.[2], DeWoody, J. A.[3] and Holland, J. D.[4]

1. Department of Entomology, Purdue University, 901 W. State St., West Lafayette, IN 47907, USA & Department of Zoology, Faculty of Science, Suez Canal University, Ismailia 41522, Egypt

2. USDA-ARS & Department of Entomology, Purdue University, 901 W. State St., West Lafayette, IN 47907, USA

3. Department of Forestry and Natural Resources & Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

4. Department of Entomology, Purdue University, 901 W. State St., West Lafayette, IN 47907, USA

Phylogeography and demographic history of a pollinator longhorn beetle (*Typocerus v. velutinus*) shaped by the Quaternary. Insect Molecular Biology (*In review*).

4.1 Abstract

Historical geological processes have shaped the contemporary distribution of genetic variation in many species such as flowering plants and mammals. However, there have been few empirical appraisals of insect phylogeography despite the fact that many geological processes (e.g., glaciations) should have had more pronounced impacts on insects than on mammals or other taxonomic groups. Our aim herein was to quantify phylogeographic effects on the contemporary gene pool of an ecologically important insect, the longhorned beetle *Typocerus v. velutinus*. We collected *T. v. velutinus* from sites that were glaciated and unglaciated during the Pleistocene to determine genetic structure within and among populations from the US and Canada, to elucidate phylogenetic relationships among demes, and to determine divergence times between populations. A total of 451 beetles were sampled from 14 sites and sequenced at a mitochondrial DNA (mtDNA) gene. Maximum likelihood and Bayesian approaches were applied to analyze the mtDNA genealogy and to reconstruct phylogenetic trees whereas Bayesian skyline analyses were used to estimate divergence time. A total of sixteen haplotypes revealed weak geographical population structuring among most populations, but statistical tests identified significant differences between the Canadian and US populations. Allelic and nucleotide diversities were lower in the Canadian populations, consistent with a recent population expansion in southern US populations and a recent bottleneck for the Canadian population. As a result of post-glacial recolonization, the US populations appear to have experienced demographic expansion while the Canadian population was influenced by a bottleneck. The Canadian population diverged from more southern populations around the time of last glacial maximum (~17,500 ybp).

Keywords:

Bayesian skyline plots; Cerambycidae; Mitochondrial DNA; pollinator; post-glacial recolonization

4.2 Introduction

Global climatic change since the Quaternary era has shaped the demographic history of many taxa in the northern hemisphere. For example, phylogeographic studies in North America have shown patterns of population expansion and contraction with the advance and retreat of ice sheets (e.g. Avise, 2000; Lessa et al., 2003; Rowe et al., 2004). Other studies have illustrated the influence of glaciation on genetic diversity, divergence due to glacial vicariance, and post-glacial recolonization in different species (Hewitt, 2004; Harris & Taylor, 2010; Breen et al., 2012; Duennes et al., 2012). Although these studies have contributed greatly to our understanding of the contemporary distribution of biodiversity in North America, we still know very little about the phylogeography of insect species (DeChain & Martin, 2005). Herein, we studied the phylogeography of an important ecosystem services provider, the banded flower longhorn beetle *Typocerus v. velutinus* (Olivier), in response to the last glaciation.

Typocerus v. velutinus belongs to the subfamily Lepturinae, within the Cerambycidae (Yanega, 1996). Anthophilous adults are active from May to August (Frost, 1979), are fairly active flyers, and feed and mate on flowers. They have been widely recorded on various flowering plants within and around forests (Linsley & Chemsak, 1976; Golsing, 1984; Bond & Philips, 1999). The larvae tunnel within decomposed hardwoods (including *Quercus*, *Caryae*, *Betula* and *Populus*) and thus they

help recycle nutrients and reduce fire fuel loads in forest ecosystems (Berkov & Tavakilian, 1998). This and many other lepturines are thought to be important pollinators in contemporary and historic forests (Benjamin, 1907; Maeto et al., 2002).

Climate and host plant availability are the main factors controlling the distribution of cerambycids (Linsley, 1959; Hanks, 1999). In the late Pleistocene, the Wisconsinan ice sheet covered the Midwestern United States and its northern borders with Canada and arctic beetle fauna existed only in a narrow zone at south of that ice sheet (Schwert & Ashworth, 1988). The Wisconsinan ice sheet extended to the central part of Indiana (Figure 4.1) and began retreating approximately 20,000 ybp (Wilson, 2008). The dispersal and recolonization of populations from southern refugia to previously glaciated landscapes occurred very rapidly after the last glacial maxima (LGM) of 24,000 – 16,000 ybp and was largely completed approximately 7000 ybp (Downes & Kavanaugh 1988). Thus, the latitudinal range of beetles has shifted repeatedly in response to historic climate change and contemporary northern populations were likely colonized from southern refugia as ice sheets retreated (Soltis et al., 2006). Thus, northern biotas may now be discontinuous (Downes & Kavanaugh, 1988).

Many species surely underwent substantial demographic changes as their distributions shifted in the Quaternary. Such demographic changes may affect population genetic structure (Pamilo & Savolainen, 1999; Hewitt, 2004). Recent evolutionary changes in genetic diversity can be examined with modern molecular techniques, such as those focusing on the mtDNA genome (Avise, 2009). We hypothesized that demographic responses to climate change differentially impacted southern refugia populations of *T. v. velutinus* relative to northern populations that were established after retreat of the

Wisconsinan ice sheet. More specifically we predicted that as sources for recolonization, southern populations would harbor more genetic variation and exhibit more evidence of recent demographic expansions than northern populations. We tested these ideas by analyzing the phylogeographic and demographic history of *T. v. velutinus* using mitochondrial DNA sequences from sites across the LGM. Our primary aim was to examine the effect of Pleistocene glaciations on the genetic diversity within and among contemporary beetle populations. Our secondary aim was to estimate phylogeographic relationships and divergence times among populations to quantify the genetic impacts of historic environmental change on beetle populations. Collectively, these data add to a growing appreciation of how insect populations evolve in response to climate change.

4.3 Materials and Methods

4.3.1 Study sites and sampling:

Individual beetles were sampled from fourteen sites in the U.S. and Canada. These sites represent three distinct zones of glaciation with respect to the late Pleistocene (Figure 4.1; Table 4.1). The site near Westport, Ontario, Canada, in the western edge of the St. Lawrence lowland eco-region, represents an area that was fully submerged under an ice sheet during much of the glacial period. The U.S. sites represent two zones, each in a different Omernick eco-region. The LO zone, which now includes agricultural landscapes with small fragmented forest patches, was just within the region covered with ice at the LGM and was likely within the tundra zone during the last glacial retreat. Sites within the HEE zone were located beyond the LGM (i.e., they represent unglaciated regions). HEE zone sites now occur in more continuous forested habitat (Figure 4.1). The

HEE zone belongs to the interior plateau eco-region with two subdivisions; the Mitchell Plain which exhibited pre-Wisconsinan glaciation and the Norman Upland (Raje et al., 2012; Holland et al., 2013; Abdel Moniem et al., 2013).

Beetles were collected during 2005 – 2011, although individual sites were sampled over a shorter period (Table 1). Individuals were confirmed as *T. v. velutinus* using criteria detailed in Lingafelter (2007) and Yanega (1996). Voucher specimens were deposited in the Purdue Entomological Research Collection (PERC).

Table 4.1 Fourteen study sites across Indiana and Ontario. GPS coordinates (UTM NAD83 zone 16N), sampling years and different sampling projects are recorded. CAN: Ontario, Canada population; LO: Land-owners sites (Raje et al., 2011); HEE: Hardwood Ecosystem Experiment (Holland et al., 2013).

Site	Name	N	E	Year	Project
1	Canada	4984900	1361850	2010	CAN
2	HEE_1	4356090	548512	2005–2010	HEE
3	HEE_2	4354720	548023	2005–2010	HEE
4	HEE_3	4352430	547793	2005–2010	HEE
5	HEE_4	4350830	549554	2005–2010	HEE
6	HEE_5	4339480	554443	2005–2010	HEE
7	HEE_6	4330210	554904	2005–2010	HEE
8	HEE_7	4331690	558954	2005–2010	HEE
9	HEE_8	4329640	558471	2005–2010	HEE
10	HEE_9	4332500	561220	2005–2010	HEE
11	LO_1	4505645	534188	2009	LO
12	LO_10	4475457	496660	2005–2006 2009–2011	LO
13	LO_20	4474143	505748	2005–2006 2009–2011	LO
14	LO_21	4442803	478158	2009	LO

Figure 4.1 Ice coverage of North America during late Wisconsinan glaciation. The three sampling regions are shown, CAN (red), LO (blue) and HEE (green). The continental map modified from the United States Geological Survey glaciers map (Lambert cylindrical equal-area projection). Smaller extent map created in ArcGIS 9.2 (ESRI 2009).

4.3.2 DNA extraction, PCR, and sequencing:

DNA was extracted from three legs from each beetle with Qiagen DNeasy blood and tissue kits (Qiagen, Inc. Valencia, CA) following company protocol. DNA was eluted in 200 µl of elution buffer and stored at -20°C. To amplify a partial fragment (~648bp) of the cytochrome oxidase I (COI) gene, we used primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer et al., 1994). PCRs were performed in 96 well plates with a total volume of 50 µl per well. The reaction cocktail contained 25 µl MyFi™ high fidelity mix (Bioline Inc. USA), 3 µl of template DNA, 1 µl of LCO1490 primer, 1 µl of HCO2198 primer and 20 µl water. The thermal profile consisted of: 95°C for 1 min, 95°C for 1.5 min, 45°C for 1.5 min, and 72°C for 1.5 min. The cycle repeated at step 2 for 5 times, then changed to 95°C for 30 sec, 50°C for 1 min, 72°C for 1 min and repeated 35 times. Finally, samples were incubated at 72°C for 5 min then kept at 4°C. All PCR products were tested by electrophoresis on a 2% agarose gel. The resulting PCR product was purified using the ZR-96 DNA Clean-up Kit™ as directed by the manufacturer.

Sequencing was performed through the Purdue University Genomics Center (PUGC). A sequencing reaction containing 100 µl of BigDye® Terminator v3.1 (Applied Biosystems, cat# 4336913), 12 µl of appropriate of primer R4 (5'-CTCACTAAAGGGACTAGTCCTG-3') or primer F5 (5' CTCACTATAGGGCGAATTGA-3'), 500 µl of BigDye® Terminator V1.1, V3.1 5X sequencing buffer (Applied Biosystems, cat# 4336701) and 1388 µl of double distilled

H_2O water per reaction was added. A Prism 3730XL genetic analyzer (Applied Biosystems, Foster City, CA, USA) was used for sequencing in both directions.

Fifteen individual samples representing three subsets of the populations were cloned and compared to the direct sequences of the same individuals to check for accuracy. Amplicons were cloned with a TOPO® TA cloning kit with pCR2.1®-TOPO® TA vector (Invitrogen, Carlsbad, CA) according to the manufacturer's protocols. The Wizard® Plus SV Minipreps DNA Purification System kit (Promega) was used for purification before sequencing following manufacturer's directions.

4.3.3 Sequences alignment and editing:

A Perl script was used to build the consensus sequence and merge the forward and reverse sequences for all the COI sequences. MEGA v.5.05 (Tamura *et al*., 2011) and BioEdit v.7.0.5 (Hall, 1999) were used to align and edit the sequences using the ClustalW alignment algorithm implemented in the software. A final COI sequence fragment of in-frame 462 bp with no gaps or ambiguous bases in all 451 individual beetles was considered for analysis. A multiple BLAST for the sequences was performed to verify species. All COI haplotype sequences were submitted to GenBank, and the accession numbers are reported.

GenBank Accession numbers for sequences of the 16 recorded haplotypes:

H01 KF768080, H02 KF768081, H03 KF768082, H04 KF768083, H05 KF768084, H06 KF768085, H07 KF768086, H08 KF768087, H09 KF768088, H10 KF768089, H11 KF768090, H12 KF768091, H13 KF768092, H14 KF768093, H15 KF768094, H16 KF768095.

4.3.4 Phylogenetic and population genetics analysis

The best DNA substitution model was calculated in MEGA v.5.05. A neighbor joining tree was used, and model selection was obtained by maximum likelihood (ML) estimation. The nucleotide type substitutions for this test included all 1^{st}, 2^{nd}, 3^{rd} and non-coding sites. The maximum likelihood tree method was used for phylogenetic reconstruction of the COI haplotypes sequences from 451 *T. v. velutinus* samples. Tamura's 3 parameter model (+I) was implemented with 1000 permutations for bootstrapping. The nearest neighbor interchange (NNI) ML heuristic method was adopted with all nucleotide type substitutions.

The analysis of molecular variance (AMOVA) was conducted with Arlequin v3.5.1.3 (Excoffier *et al*., 2005) for haplotypic data using 10,000 permutations and Tamura's three parameter model. The hierarchical island model (Slatkin & Voelm, 1991) was used to perform coalescent simulations leading to the joint null distributions of hierarchical F-statistics (F_{SC}, F_{CT}, and F_{ST}) and heterozygosities, from which locus-specific p-values were estimated. The pairwise F_{ST} were computed with 1000 permutations at α=0.05. The exact test of population differentiation was done with 100,000 Markov Chain Monte Carlo (MCMC) steps and 10,000 steps for burning and reaching conversion at α=0.05. Haplotypes frequencies were estimated by counting. The linkage disequilibrium (LD) test between all pairs of loci was done with 10,000 MCMC steps and 1000 steps for burning and reaching convergence. To calculate the LD coefficients between pairs of alleles at different loci, D, D', and r^2 were computed. Mismatch distributions of pair-wise molecular distance were calculated with 100 bootstrap replicates to estimate demographic parameters. The standard diversity indices and molecular diversity indices were calculated using Tamura's molecular distance and

estimated Theta (Homozygosity), Theta (k) and Theta (pi). The pegas package (Paradise & Potts, 2013) was used in R (R Development Core Team, 2012) to perform Ramos-Onsins and Rozas test of neutrality (Ramos-Onsins & Rozas, 2002). The default setup of Theta=1 was used with 1000 permutations. The same package was used to create the haplotype network (Bandelt *et al*., 1999) and the minimum spanning tree using Kruskal's algorithm (Kruskal, 1956).

4.3.5 Demographic history and divergence time estimation

The 14 sampled populations were tested for population expansion or contraction using both neutrality tests and mismatch distributions (reviewed in Fahey *et al*. 2014). For neutrality tests, Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1985) we used in Arlequin v.3.5.1.3 assuming a stepwise expansion model for mtDNA sequences (Schneider & Excoffier, 1999). We calculated three demographic parameters, $\theta_0 = 2\ \mu N0$, $\theta_1 = 2\ \mu N1$, and $\tau = 2\mu t$, where μ is the mutation rate for the COI gene. The sum of squared deviation (SSD) was calculated between mismatch distributions as test statistics for the estimated stepwise expansion models (Harpending *et al*., 1998) along with the raggedness index of Harpending (1994).

The molecular clock test was performed by comparing the ML value for the topology tree with and without the molecular clock constraints under T92+I model. A proportion of sites (44%) were allowed to remain invariant (I) in the evolutionary rate model. The analysis involved 16 haplotype sequences and all codon positions. To estimate the time to most recent common ancestor (TMRCA), the 451 COI sequences were analyzed in BEAST v.1.7.4 (Drummond & Rambaut, 2007). Using the T92+ I model, the best fit model to our sequences, the MCMC simulations were conducted with

the coalescent Bayesian skyline tree model and the strict clock model (Drummond et al., 2005). The relaxed uncorrelated lognormal molecular clock model was used to assess the clock-like nature of the data. We specified a range of possible substitution rates for insect mtDNA genes using two publications (Brower, 1994; Farrell, 2001) and by using a flat prior ranging from $1.7x10^{-9}$ to $1.7x10^{-7}$ substitutions per site per year with a median initial value of $1.7x10^{-8}$. Four independent simulations each for 20 million generations, sampling every 2000 generations were conducted to confirm convergence. Log files were compiled from the independent runs using LogCombiner v.1.7.4 (Drummond & Rambaut, 2007) and the Bayesian skyline plot (BSP) and analysis were done in Tracer v1.5 (Drummond & Rambaut, 2007). This method uses a MCMC procedure to sample the distribution of generalized skyline plots, given the data and according to their posterior probabilities, then combines these plots to generate estimates and credibility intervals for the effective population size at every point backward in time until the most common recent ancestor (MCRA) of the sampled sequences is reached.

4.4 Results

Within the 451 individual beetle COI sequences (after trimming, 462 bp each), 16 haplotypes were characterized. Haplotype 3 (H3) was the most frequently recorded in the 14 populations (16.2% of total individuals), while haplotype 10 (H10) was the least prevalent (1.7% of total individuals). Five haplotypes (H1, H2, H3, H5 and H16) were shared between the three major sampling sites (CAN, HEE and LO). Within the 451 sequences, there was an average of 10 ± 2 polymorphic sites per sequence. Sequences from the sampling site HEE7 encountered the highest number of polymorphic sites (14

sites), while sequences from Canada showed the lowest (6 sites). Transitions were more frequent than transversions in general with mean calculated transition/transversion of 9.07±1.9 / 2.79 ± 0.58 with average substitutions number of 11.86 ± 2.28 (Table 4.2).

The best substitution model found to fit the COI sequences for *T. v. velutinus* is the Tamura's 3 parameter model with evolutionarily invariable sites (T92+I). The frequencies of the base pairs A, T, C and G for the best model were 0.332, 0.332, 0.168 and 0.168 respectively.

Molecular diversity indices (Table 4.2) showed a lower level of both gene and nucleotide diversities (1.39 ± 1.25, 2.95 ± 0.57 respectively) in the Canadian population than the U.S. populations (3.16 ± 1.25, 3.76 ± 0.57 respectively). The same trend was found with other measures of molecular diversity (Table 4.2).

The neutrality tests, especially Fu's Fs (Fu, 1995), showed a pattern of negative values for populations collected from the two southern zones (LO and HEE), reflecting a population expansion. Contrary to this, the positive value for the Canadian population indicated a population bottleneck (Table 4.3). Fu's simulations suggest that Fs is a more sensitive indicator of population expansion and genetic hitchhiking than Tajima's D (Fu, 1997). Both neutrality tests and raggedness index were not significant at $\alpha = 0.05$. The bimodal pattern of the mismatch distribution for the Canadian population is an indicator of a population contraction or bottleneck while the unimodal pattern of the mismatch distribution for the southern populations indicates a demographic expansion (Figure 4.2; Table 4.4). The Ramos-Onsins and Rozas test of neutrality showed that there was no selection force acting on the any of the 462 COI loci ($R^2 = 0.134$, d.f. = 461, $P = 0.957$, $\alpha = 0.05$).

Table 4.2 Molecular diversity indices from14 sampled populations. A total number of 462 loci per gene were considered. Molecular diversity estimators: θ_k obtained from the observed number of alleles k; θ_H: obtained from the observed homozygosity H; θ_S: obtained from the observed number of segregating site S and θ_Π: obtained from the mean number of pairwise differences $\hat{\pi}$.

Statistics	CAN	HEE 1	HEE 2	HEE 3	HEE 4	HEE 5	HEE 6	HEE 7	HEE 8	HEE 9	LO 1	LO 10	LO 20	LO 21	Mean	s.d.
Gene copies	33	45	19	16	9	38	63	70	40	36	26	13	24	19	32.21	18.133
Poly. sites	6	12	12	10	9	13	13	14	12	12	11	12	12	13	10	2.065
Transitions	5	9	9	7	6	10	10	12	10	10	9	9	11	10	9.07	1.90
Transversions	1	3	3	3	3	3	3	3	3	3	2	3	3	3	2.79	0.58
Substitutions	6	12	12	10	9	13	13	15	13	13	11	12	14	13	11.86	2.28
Indels	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Subst. Sites	6	12	12	10	9	13	13	14	12	12	11	12	13	13	11.57	2.07
Π	2.95	3.85	4.11	2.88	2.94	3.44	3.63	3.59	3.97	3.49	3.52	4.77	4.52	4.26	3.71	0.58
θ_k	1.39	5.01	7.82	3.02	6.69	5.65	4.13	6.18	6.29	5.89	4.45	11.65	5.91	7.82	5.85	2.42
θ_k_lower	0.52	2.53	3.33	1.12	2.08	2.80	2.13	3.43	3.19	2.90	1.97	4.21	2.65	3.33	2.58	0.97
θ_k_upper	3.44	9.61	18.30	7.77	22.18	11.05	7.69	10.81	12.06	11.61	9.68	33.50	12.89	18.30	13.49	7.54
θ_H	2.69	9.82	6.35	3.36	9.61	4.55	5.35	8.49	9.61	7.28	8.17	10.57	8.69	11.77	7.59	2.77
s.d. θ_H	0.62	1.97	3.81	1.45	9.98	1.47	1.26	1.62	2.52	2.16	2.62	9.30	3.17	6.59	3.47	2.99
θ_S	1.48	2.74	3.43	3.01	3.31	3.09	2.76	2.91	2.82	2.89	2.88	3.87	3.48	3.72	3.03	0.57
s.d. θ_S	0.73	1.08	1.49	1.39	1.70	1.22	1.03	1.06	1.12	1.17	1.23	1.78	1.44	1.59	1.29	0.29
θ_Π	2.95	3.85	4.11	2.88	2.94	3.44	3.63	3.59	3.97	3.49	3.52	4.77	4.52	4.26	3.71	0.58
s.d. θ_Π	1.76	2.19	2.39	1.79	1.93	2.00	2.07	2.05	2.25	2.03	2.06	2.80	2.57	2.47	2.17	0.30
Mean Exp. H	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
s.d. Exp. H	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.06	0.07	0.06	0.06	0.06	0.05
Mean #	1.01	1.03	1.03	1.02	1.02	1.03	1.03	1.03	1.03	1.03	1.02	1.03	1.03	1.03	1.03	1.01
s.d.	0.11	0.16	0.16	0.15	0.14	0.17	0.17	0.19	0.18	0.18	0.15	0.16	0.18	0.17	0.16	0.11

Table 4.3 Tajima's D and Fu's Fs neutrality tests indicating the demographic history for the 14 sampled populations.

Tajima's D test

Statistics	CAN	HEE 1	HEE 2	HEE 3	HEE 4	HEE 5	HEE 6	HEE 7	HEE 8	HEE 9	LO 1	LO 10	LO 20	LO 21	Mean	s.d.
Sample size	33	45	19	16	9	38	63	70	40	36	26	13	24	19	32.21	18.13
Substitutions	6	12	12	10	9	13	13	14	12	12	11	12	13	13	11.57	2.07
θ_Π	2.95	3.85	4.11	2.88	2.94	3.44	3.63	3.59	3.97	3.49	3.52	4.77	4.52	4.26	3.71	0.58
Tajima's D	2.77	1.21	0.71	-0.16	-0.51	0.35	0.91	0.68	1.25	0.65	0.73	0.95	1.03	0.53	0.79	0.75
D, p-value	0.99	0.89	0.81	0.46	0.33	0.68	0.83	0.79	0.91	0.78	0.79	0.85	0.88	0.74	0.77	0.18

Fu's Fs test

Statistics	CAN	HEE 1	HEE 2	HEE 3	HEE 4	HEE 5	HEE 6	HEE 7	HEE 8	HEE 9	LO 1	LO 10	LO 20	LO 21	Mean	s.d.
Real no. of alleles	5	12	10	6	6	12	12	16	13	12	9	9	10	10	10.14	3.03
Exp. No. of alleles	7.86	10.26	7.52	5.87	4.53	9.04	11.06	11.35	10.02	8.95	7.95	6.65	8.76	7.66	8.39	1.94
FS	2.95	-0.85	-1.74	0.30	-1.21	-1.79	-0.32	-2.62	-1.72	-1.88	-0.47	-1.91	-0.60	-1.61	-0.96	1.37
FS, p-value	0.90	0.40	0.20	0.57	0.18	0.23	0.51	0.16	0.25	0.21	0.42	0.14	0.41	0.22	0.34	0.21

Table 4. 4. Mismatch distribution analysis and estimates of demographic expansion parameters for the 14 sampled populations.

Statistics	CAN	HEE 1	HEE 2	HEE 3	HEE 4	HEE 5	HEE 6	HEE 7	HEE 8	HEE 9	LO 1	LO 10	LO 20	LO 21	Mean	s.d.
Tau (τ)	5.72	5.28	5.54	4.79	1.64	5.37	5.07	4.39	5.18	4.88	4.80	6.09	5.63	5.78	5.01	1.07
(τ) 95% qt	9.41	7.19	7.40	7.59	5.66	8.39	7.19	6.40	6.94	8.19	6.99	8.42	7.02	7.40	7.44	0.93
Theta 0 (Θ 0)	0.00	0.00	0.00	0.00	1.74	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.13	0.46
(Θ 0) 95% qt	2.78	1.44	1.25	0.55	4.40	0.97	0.24	1.13	0.75	1.21	1.49	2.35	1.46	1.85	1.56	1.05
Theta 1 (Θ 1)	5.66	11.46	15.31	4.89	109.53	7.49	9.31	14.62	14.32	8.59	10.47	30.00	34.18	19.10	21.07	26.87
SSD	0.04	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.01	0.01
Raggedness index	0.12	0.02	0.03	0.04	0.05	0.02	0.02	0.01	0.01	0.02	0.02	0.05	0.05	0.04	0.04	0.03

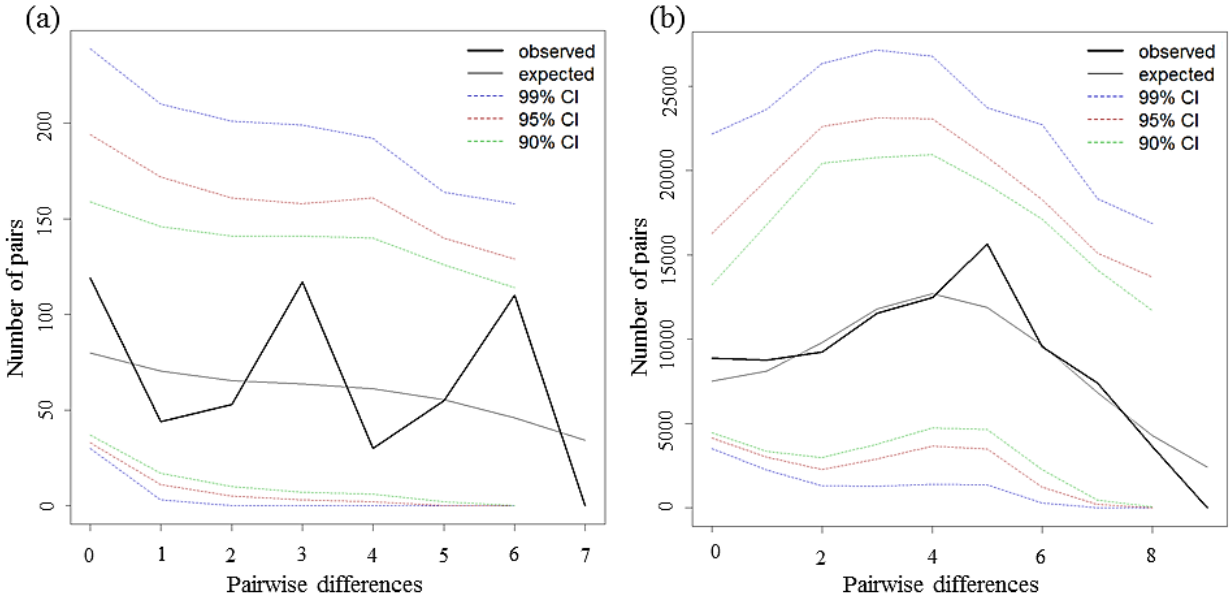


Figure 4.2 The mismatch distribution plot for (a) group 1 (Canada population) and (b) group 2 (HEE and LO populations) with 90%, 95% and 99% confidence intervals.

The AMOVA revealed a weak, but significant level of genetic structuring. In total, 5.7% of the genetic variance was partitioned between two major mtDNA haplogroups (Canada and U.S.). Only 0.89% of the molecular variance resulted from differences among populations within these groups, and the majority of variance (93.4%) was within the populations (Table 4.5). The mean pairwise value of F_{ST} was 0.07 which indicates a high level of gene flow among populations within these two groups. However, the pairwise F_{ST} values between populations and exact test of populations differentiation based on haplotype frequencies (Table 4.6) showed significant levels of differentiation (α=0.05) between the population from Canada and those from the U.S.

Table 4.5 Analysis of molecular variance (AMOVA) design and results. Significance for test statistics was calculated with 10,000 permutations.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	9.452	Va = 0.11435	5.74
Among populations within groups	12	28.977	Vb = 0.01772	0.89
Within populations	437	812.797	Vc = 1.85995	93.37
Total	450	851.226	1.99202	

Fixation Indices

Source	Index	P-value
F_{SC}	0.00994	0.00366±0.00065
F_{ST}	0.0663	0.09257±0.00244
F_{CT}	0.0574	0.06762±0.00242

Table 4.6 Pairwise F_{ST} values between populations as calculated by Tajima & Nei distance. Significance was tested at α=0.05 with 1000 permutations. Values in bold were significant. Grey boxes show the significant pairwise exact test of populations differentiation based.

	CAN	HEE_1	HEE_2	HEE_3	HEE_4	HEE_5	HEE_6	HEE_7	HEE_8	HEE_9	LO_1	LO_10	LO_20	LO_21
CAN	0													
HEE_1	**0.078**	0												
HEE_2	**0.169**	0.025	0											
HEE_3	**0.079**	0.032	**0.152**	0										
HEE_4	0.084	-0.034	0.028	0.034	0									
HEE_5	**0.086**	0.015	0.008	**0.068**	0.010	0								
HEE_6	**0.044**	-0.002	**0.052**	0.003	-0.011	0.005	0							
HEE_7	**0.066**	-0.009	0.013	**0.059**	-0.027	0.000	0.001	0						
HEE_8	**0.066**	-0.014	0.006	0.038	-0.023	0.001	-0.009	-0.014	0					
HEE_9	**0.077**	-0.012	0.033	0.014	-0.035	0.002	-0.010	-0.006	-0.013	0				
LO_1	**0.087**	0.009	**0.105**	0.054	-0.023	**0.093**	0.034	**0.035**	0.027	0.042	0			
LO_10	**0.128**	-0.017	-0.033	**0.108**	-0.007	0.030	0.028	-0.005	-0.017	0.016	0.019	0		
LO_20	0.044	0.010	0.031	**0.076**	0.039	0.023	0.018	0.012	0.004	0.025	**0.060**	-0.012	0	
LO_21	**0.066**	-0.034	0.020	0.031	-0.041	0.018	-0.007	-0.014	-0.021	-0.011	-0.017	-0.040	-0.012	0

The tree showing the most likely evolutionary history of the 16 haplotypes (LL= -790.23; Figure 4.3) allowed us to reject the null hypothesis of equal evolutionary rate throughout the tree (d.f.= 15, $P < 0.045$). The haplotype network map (Figure 4.4) illustrates the relationships between the 16 characterized haplotypes along with the relative frequencies of each sampled geographical location (CAN, HEE and LO) represented by these haplotypes. The spatiotemporal haplotype network (Figure 4.5) represents the relation between three haplotype networks generated in parallel to compare the differences in haplotype network structure between geographically different samples. The network showed a lower presentation of genetic diversity from Canadian population as depicted by white ellipses representing lost haplotypes.

Bayesian analysis and Bayesian tree topology as inferred from MCMC simulation confirmed the monophyly of the Canadian beetles' clade with a posterior probability value of 100% (Figure 4.6). The sister clade of beetles from the two clusters of sites shared the five aforementioned haplotypes with the Canadian group. This sister clade had a posterior probability of 35%. The remaining clades were represented only by beetles from the U.S. (Indiana), and there was a pattern of polyphyly of HEE and LO beetles due to inadequate resolution in the tree.

The Bayesian skyline plot indicated that the Indiana and Canadian populations diverged approximately 17,500 ybp (Figure 4.7). As predicted, this divergence time coincides with the beginning of the latest deglaciation following the LGM. This also demonstrates a recent bottleneck of the Canadian population as inferred from the decline in the effective population size.

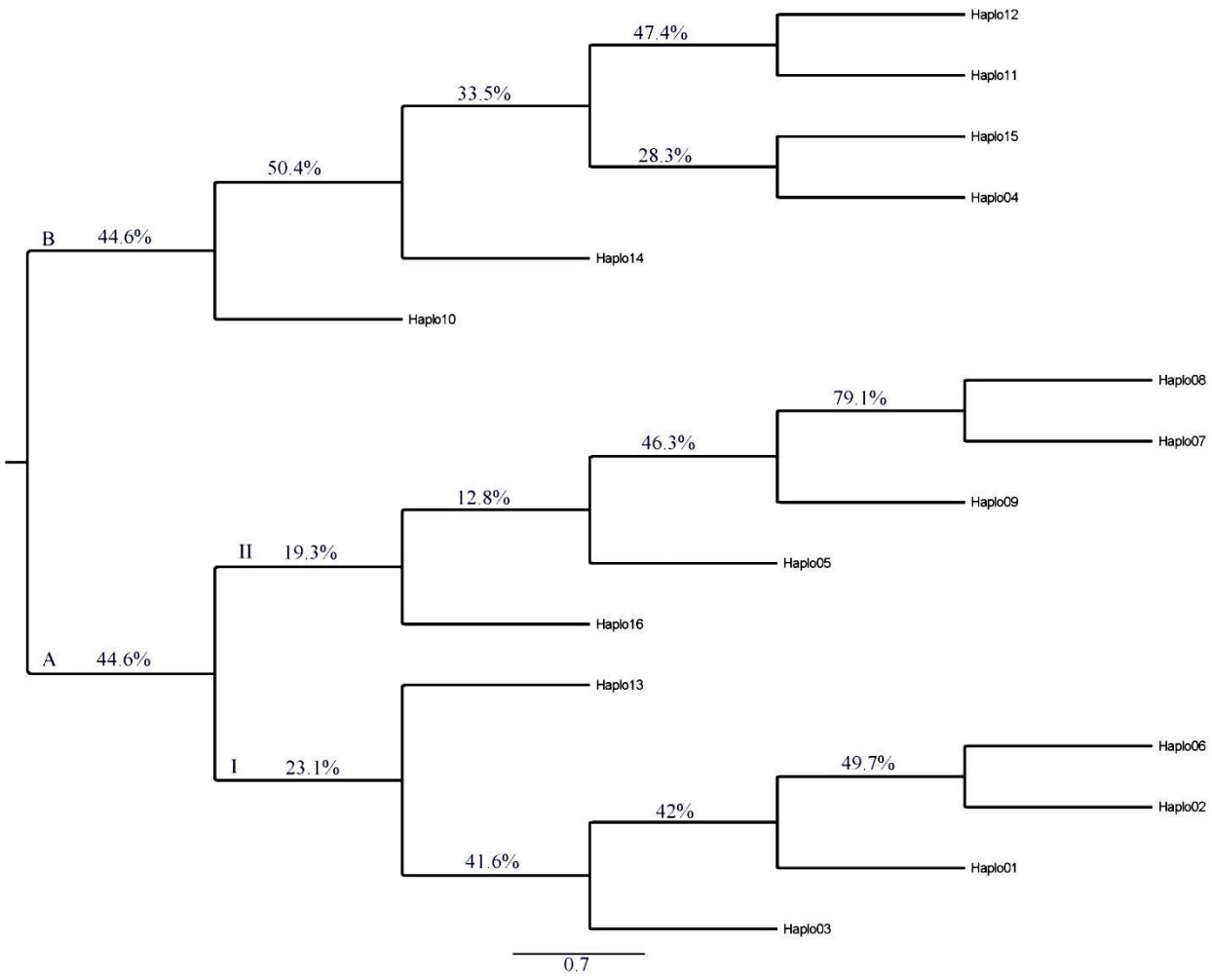


Figure 4.3 Maximum likelihood tree and bootstrap support as inferred from the molecular phylogenetic analysis for the 16 observed haplotypes. The rate variation model allowed for some sites to be evolutionarily invariable (+I, 0.2314% sites) as inferred by the best substitution model.

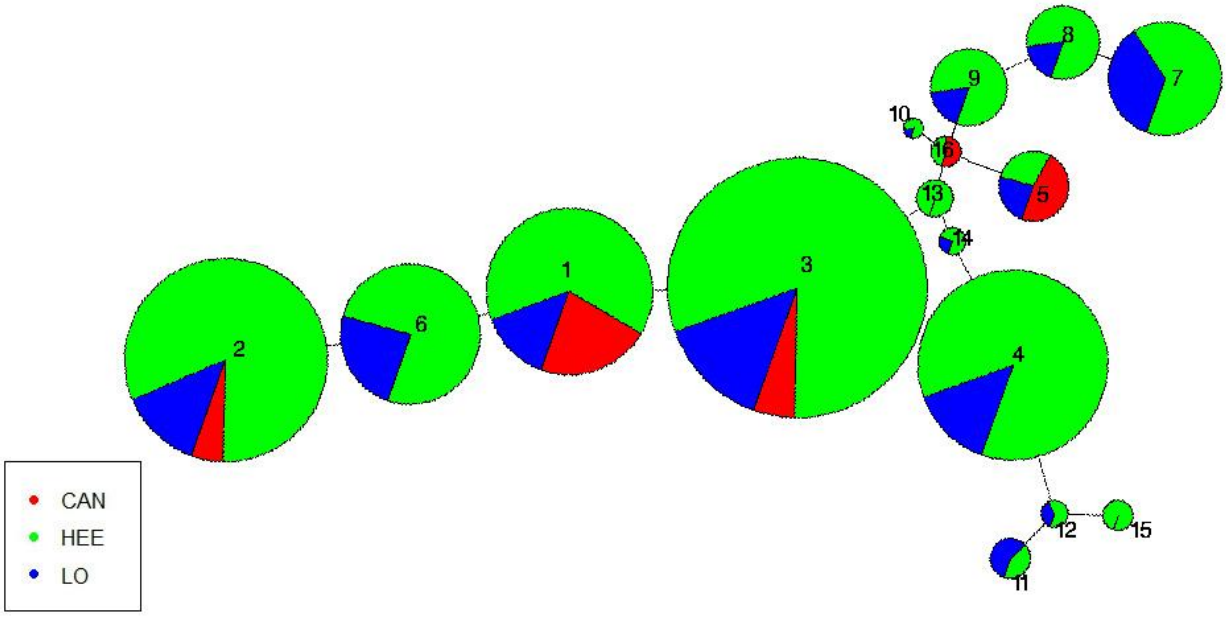


Figure 4.4 Haplotype network map for 16 recorded haplotypes. Size of circles proportional to haplotypes frequencies. Colors indicate frequencies of each population in each haplotype. Length of connection lines proportional to mutation steps between haplotypes.

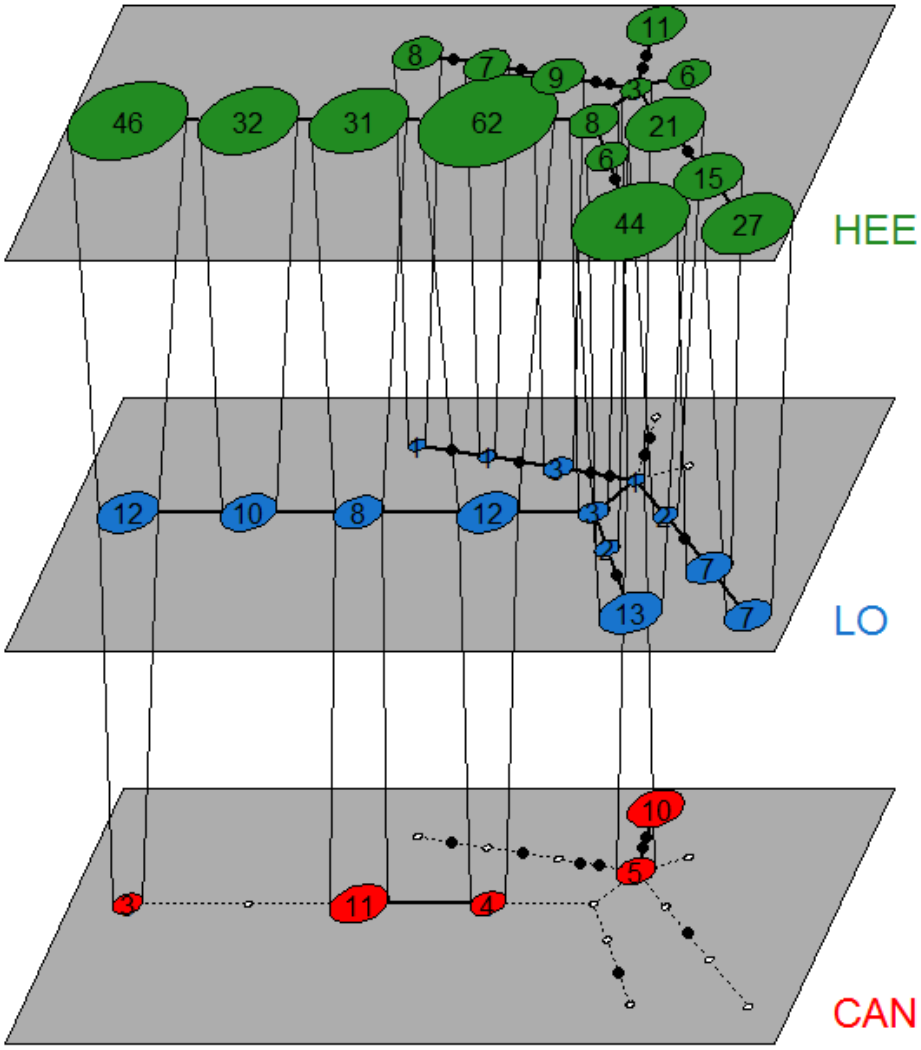


Figure 4.5 Spatiotemporal haplotype network summarizing relationship between the three major populations. Ellipse size represents the frequency of each haplotype. Black nodes show mutational steps. Empty ellipses are haplotypes lost from the population. Connections between populations illustrate shared haplotypes.

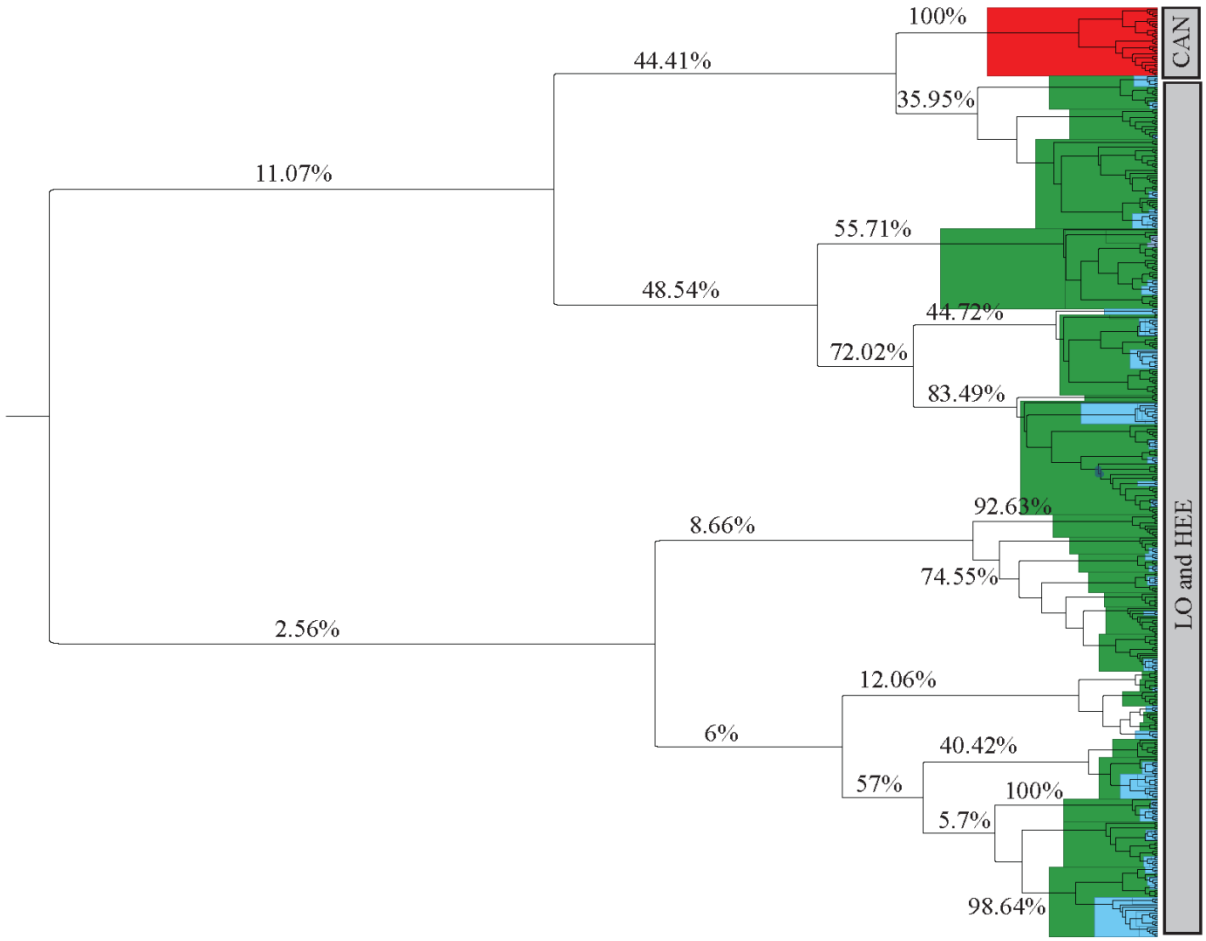


Figure 4.6 Bayesian tree topology as inferred from MCMC simulation. The posterior probabilities are recorded as percentage at the major branches. Populations are color coded as Canada (red), HEE (green) and LO (blue).

4.5 Discussion

We hypothesized that demographic responses to climate change differentially impacted southern refugia *T. v. velutinus* populations relative to northern populations established after retreat of the Wisconsinan ice sheet. Our empirical results, which document higher levels of genetic diversity in southern populations, are consistent with theoretical predictions associated with source populations (Hewitt, 2004). The geologic record suggests the most northern (Canadian) population must be the youngest, and our molecular data are consistent with this idea. The five major haplotypes shared between northern and southern populations, along with higher numbers of beetles representing these haplotypes in the south, and higher nucleotide diversity suggest both a common gene pool and southern refugia for *T. v. velutinus*. The spatiotemporal haplotype network (Figure 4.4) illustrates that the genetic diversity in the Canadian populations is a subset of the diversity in the more southern populations. Haplotypes missing in Canada might not have survived the glacial climates or were eliminated from the original gene pool due to the stochastic dynamics of early post-glacial recolonization and population expansion. Similar latitudinal gradients in genetic diversity have been reported with other beetles in North America (e.g., Stauffer et al., 1999; Reiss et al., 2004; Ruiz, 2010); these phylogeographic patterns are consistent with known geological processes.

Neutrality tests (especially Fu's Fs) showed a general trend of negative values for southern populations, and a positive value for the Canadian population, supporting our predictions regarding the beetle's historical demography. The mismatch distribution analysis for southern populations showed a pattern that does not deviate from a unimodal distribution of pairwise differences among haplotypes which suggests a recent population

expansion (Slatkin & Hudson, 1991; Rogers & Harpending, 1992). The non-significant mismatch distribution means that we cannot reject the null hypothesis of population expansion, and the non-significant raggedness index indicates a good fit of our data to a model of population expansion. An exact opposite scenario was clear with the Canadian population that showed a multimodal distribution curve. This was further supported by the steep decline in the effective population size. Furthermore, the recent evolutionary time frame of the LGM might be too recent for either speciation or drift-migration equilibrium to exist in this population (Varvio et al., 1986; Avise, 2000). Because the Quaternary period covers approximately 2.4 Myr, most DNA sequences will diverge little over the ice ages, and few new mutations will characterize post-glacial haplotypes (Hewitt, 2000). Consequently, cautious interpretations of results are required to differentiate between recent evolutionary mutations characterizing post-glacial populations versus those ancient mutations that are more likely attributed to the more distant past (Templeton, 1998; Hewitt, 1999; Fahey et al. 2014).

The pairwise F_{ST} and the exact test of differentiation between populations showed significant levels of differences between the Canadian and the U.S. (Indiana) populations. These results were further supported by the ML and Bayesian phylogenetic analysis of haplotypes. These findings revealed a phylogeographic pattern of haplotype distribution into northern and southern groups. This pattern was confirmed with the Bayesian tree which supported the monophyly of the Canadian clade. In addition, the high level of gene flow between the southern populations that is apparent in the panmixed structure of the clades representing these demes could be attributed to the habitat connectivity in the southern part of the state. From a LGM time perspective, the differentiation between the

two clusters of sites is unlikely to lead to a speciation event, especially with species that have large geographic distributions and high dispersal potential (Ashworth, 2001; Stewart et al., 2010).

The coalescence of the Canadian population to its most recent common ancestor (MRCA) and estimating its divergence has enabled us to test whether this coincided with the beginning of deglaciation. As predicted, the Bayesian analysis and BSP showed two important characteristics of the Canadian population. First, the population could have passed through a recent severe bottleneck represented by the steep decline in its effective population size. Second, the divergence time of this population goes back to about 17,500 ybp which coincides with the last glacial retreat after LGM (24,000 – 16,000 ybp). This finding further supports the south to north post-glacial recolonization pattern and the validity of the southern refugium theory for this beetle.

Anthropogenic activity and its influence on natural habitats is another substantial factor that is shaping contemporary structure of species populations (Lande, 1987; Winchester, 1997). Habitat loss and fragmentation form potential barriers to gene flow, thus resulting in patchy, isolated sub-populations. Depending on the beetles' dispersal ability and the spatial scale at which they respond to their environment, their genetic diversity and population structure will be affected (Fahrig, 2001). One of the surprises of the Quaternary record of beetle fossils is that species extinction was rarely associated with glacial climate change (Ashworth, 2001). Habitat loss and fragmentation due to human activities appear to be more responsible for range contractions and extinctions than climate change during the Quaternary period (Ashworth, 2001).

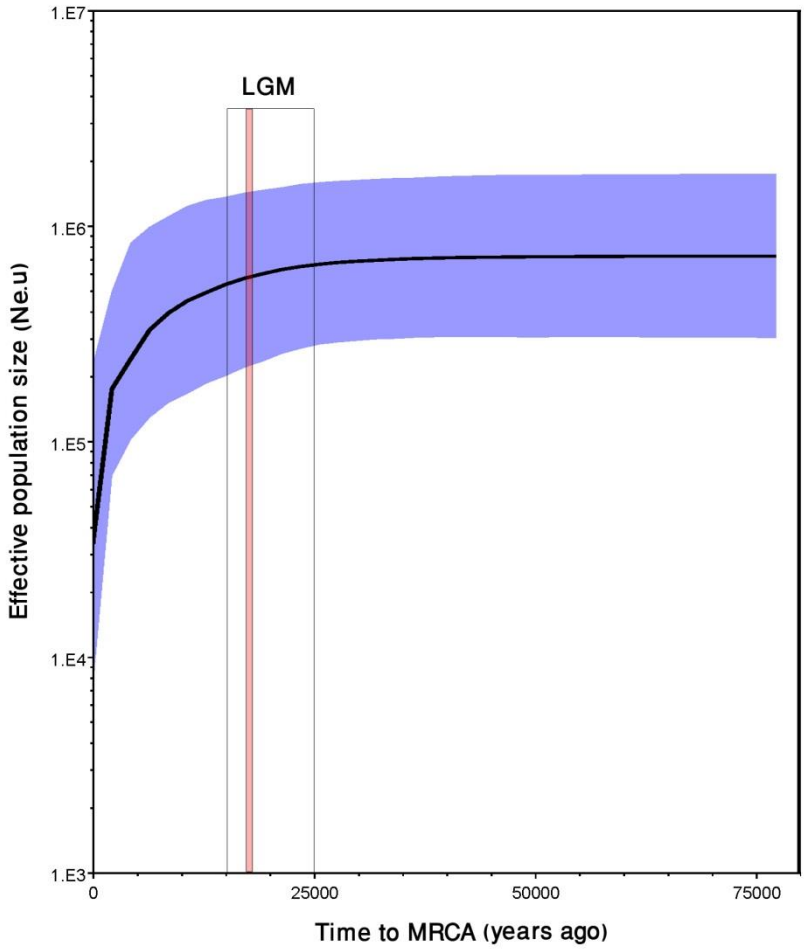


Figure 4.7 Bayesian skyline plot showing the historical demography of Canada population as inferred from COI sequences. Along the y-axis, the effective population size estimated as Ne.μ (Ne: effective population size, μ: mutation rate per haplotype per generation). The x-axis represents years before present time since divergence. Solid line is the median estimate and shaded area is the 95% confidence intervals. Period of last glacial maxima LGM is shown in box. Red line marks divergence time approximately 17,500 ybp.

In conclusion, the banded flower longhorn beetle has survived periods of climatic change in the past mainly because of their population demography dynamics and their dispersal ability. These traits helped the more southern populations to survive the Quaternary in warmer refugia, then recolonize the north as new habitat became available. Patterns of phylogeographic distribution, differences in genetic diversity, and molecular evidence for demographic population expansion and contraction support this scenario. Our results pertain not only to how species and populations responded to historic climatic changes (Mikkola, 1991; Pamilo & Savolainen, 1999), but may provide valuable context for predicted range and demographic shifts due to future climate change.

Acknowledgments**:**

We thank Yan Crane and Alisha Johnson from the USDA-ARS laboratory at Purdue University for help with molecular techniques; Phillip San Miguel, director of Purdue Genomics Facility, and Allison Sorg, technician, for sequencing assistance. Kapil Raje, John Shukle and Tommy Mager helped with field collections. This research was supported by a governmental general mission scholarship administrated by the Egyptian Cultural and Education Bureau, Washington, DC, and by the Department of Entomology at Purdue University.

4.6 Bibliography

Abdel Moniem, H. E. M. & Holland, J. D. (2013) Habitat connectivity for pollinator beetles using surface metrics. *Landscape Ecology*, **28,** 1251–1267.

Ashworth, A. C. (2001) Perspectives on Quaternary beetles and climate change. *Geological Perspectives of Global Climate Change.* (ed. by L. C. Gerhard, W. E. Harrison and B. M. Hanson). American Association of Petroleum Geologists. Studies in Geology, 47. Tulsa, Oklahoma. pp. 153–168.

Avise, J. C. (2000) *Phylogeography: The history and formation of species*, 3rd edn. Harvard University Press, Cambridge, MA.

Avise, J. C. (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36,** 3–15.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, E., Reed, C. A. & Saunders, N. C. (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18,** 489–522.

Bandelt, H. J., Forster, P. & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16,** 37–48.

Benjamin, W. D. (1907) *The first Annual Report of the state Entomologist of Indiana*. Harvard University Press, Cambridge, MA.

Berkov, A. & Tavakilian, G. (1998) Host utilization of the Brazil nut family (Lecythidaceae) by sympatric wood-boring species of *Palamae* (Coleoptera, Cerambycidae, Lamiinae, Acanthocinini). *Biological Journal of the Linnaean Society*, **67,** 181–198.

Bond, W. D. & Philips, T. K. (1999) Diversity, phenology, and flower hosts of anthophilous long-horned beetles (Coleoptera: Cerambycidae) in a southeastern Ohio forest. *Entomological News*, **110,** 267–278.

Breen, A. L., Murray, D. F. & Olson, M. S. (2012) Genetic consequences of glacial survival: The late Quaternary history of balsam poplar (*Populus balsamifera* L.) in North America. *Journal of Biogeography*, **39,** 918–928.

Brower, A. V. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of National Academy of Science*, *USA,* **91,** 6491–6495.

DeChaine, E. G. & Martin, A. P. (2005) Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *Journal of Biogeography*, **32,** 1943–1956.

Downes, J. A. & Kavanaugh, D. H. (1988) Origins of the North American insect fauna. *Memoirs of the Entomological Society of Canada*, **144,** 1–11.

Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22,** 1185–1192.

Drummond, A. J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC. *Evolutionary Biology*, **7,** 214–221.

Duennes, M. A., Lozier, J. D., Hines, H. M. & Cameron, S. A. (2012) Geographical patterns of genetic divergence in the widespread Mesoamerican bumble bee *Bombus ephippiatus* (Hymenoptera: Apidae). *Molecular Phylogenetics and Evolution*, **64,** 219–231.

ESRI, Environmental Systems Resource Institute (2009). ArcMap, *version 9.2*. ESRI, Redlands, California.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin, *version 3.0*: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Fahey, A. L., Ricklefs, R. E. & DeWoody, J. A. (2014) DNA-based approaches for evaluating historical demography in terrestrial vertebrates. *Biological Journal of the Linnaean Society*, in press.

Farrell, B. D. (2001) Evolutionary assembly of the milkweed fauna: cytochrome oxidase I and the age of Tetraopes beetles. *Molecular Phylogenetics and Evolution*, **18,** 467–478.

Fahrig, L. (2001) How much habitat is enough? *Biological Conservation*, **100,** 65–74.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3,** 294–299.

Frost, S. W. (1979) A preliminary study of North American insects associated with elderberry flowers. *Florida Entomologist*, **62**, 341–355.

Fu, Y. X. (1995) Statistical properties of segregating sites. *Theoretical Population Biology*, **48,** 172–197.

Fu, Y. X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics*, **147,** 915–925.

Gosling, D. L. (1984) Flower records for anthophilous Cerambycidae (Coleoptera) in a southwestern Michigan woodland. *The Great Lakes Entomologist*, **17,** 79–82.

Hall, T. A. (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41,** 95–98.

Hanks, L. M. (1999) Influence of the larval host plant on reproductive strategies of cerambycid beetles. *Annual Review of Entomology*, **44,** 483–505.

Harpending, H. C. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66,** 591–600.

Harpending, H. C., Batzer, M. A., Gurven, M., Jorde, L. B., Rogers, A. R. & Sherry, S. T. (1998) Genetic traces of ancient demography. *Proceedings of National Academy of Science, USA*, **95,** 1961–1967.

Harris, L. N. & Taylor, E. B. (2010) Pleistocene glaciations and contemporary genetic diversity in a Beringian fish, the broad whitefish, *Coregonus nasus* (Pallas): inferences from microsatellite DNA variation. *Journal of Evolutionary Biology*, **23,** 72–86.

Hewitt, G. M. (1999) Post-glacial recolonization of European biota. *Biological Journal of Linnaean Society*, **68,** 87–112.

Hewitt, G. M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405,** 907–913.

Hewitt, G. M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B*, **359,** 183–195.

Holland, J. D., Shukle, J. T., Abdel Moniem, H. E. M., Mager, T. W., Raje, K. R., Schnepp, K. E. & Yang, S. (2013) Pre-treatment assemblages of wood-boring beetles (Coleoptera: Buprestidae, Cerambycidae) of the hardwood ecosystem experiment. *The Hardwood Ecosystem Experiment: a Framework for Studying Responses to Forest Management, Northern Forestry Research Station* (ed. by R. Swihart, M. Saunders, R. Kalb, S. Haulton and C. Michler). USDA Division of Forestry, General Technical Report NRS-P-108. pp. 218–236.

Kruskal, J. B. Jr. (1956) On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, **7,** 48–50.

Lande, R. (1987) Extinction thresholds in demographic models of territorial populations. *The American Naturalist*, **130,** 624–635.

Lessa, E. P., Cook, J. A. & Patton, J. L. (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceedings of National Academy of Science USA*, **100,** 10331–10334.

Lingafelter, S.W. (2007) *Illustrated Key to the Longhorned Woodboring Beetles of the Eastern United States* (ed. by N. S. Terry) Coleopterists Society Special Publication No.3, Maryland.

Linsley, E. G. (1959) Ecology of Cerambycidae. *Annual Review of Entomology*, **4,** 99–138.

Linsley, E. G. & Chemsak, J. A. (1976) *Cerambycidae of North America. Taxonomy and classification of the subfamily Lepturinae*, Part VI, No. 2. University of California Publications in Entomology, **80,** 1–186.

Maeto, K., Sato, S. & Miyata, H. (2002) Species diversity of longicorn beetles in humid warm temperate forests: the impact of forest management practices on old-growth forest species in southwestern Japan. *Biodiversity and Conservation*, **11,** 1919–1937.

Mikkola, K. (1991) The conservation of insects and their habitats in northern and eastern Europe. *The Conservation of Insects and their Habitats* (ed. by N. M. Collins and J. A. Thomas). Academic Press, London. pp. 109–119.

Pamilo, P. & Savolainen, O. (1999) Post-glacial colonization, drift, local selection and conservation value of populations: a northern perspective. *Hereditas*, **130,** 229–238.

Paradis, E. & Potts, A. (2013) pegas: Population and evolutionary genetics analysis system. *Version 0.4.4*. http:// cran.r-project.org/web/packages/pegas.

Raje, K., Abdel Moniem, H. M., Farlee, L., Ferris, V. R. & Holland, J. D. (2011) Abundance of pest and benign Cerambycidae both increase with decreasing forest productivity. *Agricultural and Forest Entomology*, **14,** 165–169.

R Development Core Team (2012) R: *A language and environment for statistical computing*. R foundation for statistical computing, *version 2.14.1*. Vienna, Austria.

Ramos-Onsins, S. E. & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19,** 2092–2100.

Reiss, R. A., Ashworth, A. C. & Schwert, D. P. (2004) Molecular genetic evidence for the post-Pleistocene divergence of populations of the arctic-alpine ground beetle *Amara alpina* (Paykull) (Coleoptera: Carabidae). *Journal of Biogeography*, **26,** 785–794.

Rogers, A. R. & Harpending, H. C. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9,** 552–69.

Rowe, K. C., Heske, E. J., Brown, P. W. & Paige, K. N. (2004) Surviving the ice: Northern refugia and post-glacial colonization. *Proceedings of National Academy of Science USA*, **101,** 10355–10359.

Ruiz, E. A., Rinehart, J. E., Hayes, J. L. & Zuniga, G. (2010) Historical Demography and Phylogeography of a specialist Bark Beetle, *Dendroctonus pseudotsugae* Hopkins (Curculionidae: Scolytinae). *Molecular Ecology and Evolution*, **39,** 1685–1697.

Schneider, S. & Excoffier, L. (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152,** 1079–1089.

Schwert, D. P. & Ashworth, A. C. (1988) Late Quaternary history of the northern beetle fauna of North America: A synthesis of fossil and distributional evidence. *Origins of the North American Insect Fauna.* (ed. by J. A. Downes and D. H. Kavanaugh). *Memoirs of the Entomological Society of Canada*, **144**, 93–107.

Slatkin, M. & Hudson, R. R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129,** 555–562.

Slatkin, M. & Voelm, L. (1991) F_{ST} in a hierarchical island model. *Genetics*, **127,** 627–629.

Soltis, D. E., Morris, A. B., Mclachlan, J. S., Manos, P. A. & Soltis, P. S. (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15,** 4261–4293.

Stauffer, C., Lakatos, F. & Hewitt, G. M. (1999) Phylogeography and postglacial colonization routes of *Ips typographus* (Coleoptera: Scolytidae). *Molecular Ecology*, **8,** 763–774.

Stewart, J. R., Lister, A. M., Barnes, I. & Dalén, L. (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London B*, **277,** 661–671.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123,** 585–595.

Tamura, K. (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. *Molecular Biology and Evolution*, **9,** 678–687.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, **28,** 2731–2739.

Templeton, A. R. (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7,** 381–397.

Varvio, S. L., Chakraborty, R. & Nei, M. (1986) Genetic variation in subdivided populations and conservation genetics. *Heredity*, **57,** 189–198.

Wilson, J. (2008) *How the Ice Age shaped Indiana*. Wilstar Media Press. Indianapolis, Indiana, USA, 30 pp.

Winchester, N. N. (1997) The arboreal superhighway: arthropods and landscape dynamics. *The Canadian Entomologist*, **129,** 595–599.

Yanega, D. (1996) *Field Guide to Northeastern Longhorned Beetles (Coleoptera: Cerambycidae).*Natural History Survey. Champaign, Illinois.

CHAPTER 5: LANDSCAPE GENETICS OF A POLLINATOR LONGHORN BEETLE [*TYPOCERUS V. VELUTINUS* (OLIVIER)]: A SURFACE METRICS APPROACH.

Abdel Moniem, H. M.[1], Schemerhorn, B. J.[2], DeWoody, J. A.[3] and Holland, J. D.[2]

1. Department of Entomology, Purdue University, 901 W. State St., West Lafayette, Indiana, USA 47907 & Department of Zoology, Faculty of Science, Suez Canal University, Ismailia, Egypt 41522

2. Department of Entomology, Purdue University, 901 W. State St., West Lafayette, Indiana, USA 47907

3. Department of Forestry and Natural Resources and Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Landscape genetics of a pollinator longhorn beetle [*Typocerus v. velutinus* (Olivier)]: a surface metrics approach. Molecular Ecology (*In review*).

5.1 Abstract

Understanding the underlying patterns and processes in the landscape that are affecting the population genetic structure and population connectivity is a major goal of landscape genetics research. A vast number of these researches have implemented categorical approaches in analyzing both landscape and genetic data. The landscape gradient paradigm and surface topology metrics were shown as powerful alternative approach that simultaneously maintains the continuous nature of landscape heterogeneity and hold true to niche theory. Herein, we adopted a landscape gradient model and used surface metrics of connectivity to model the genetic continuity between populations of the banded flower longhorn beetle [*Typocerus v. velutinus* (Olivier)] collected at 17 sites across a fragmentation gradient in Indiana, USA. We tested the hypothesis that landscape structure and habitat connectivity facilitate gene flow between beetle populations against a null model of isolation by distance (IBD). We used next-generation sequencing to develop 10 polymorphic microsatellite loci and genotype the populations to assess genetic structure. Panmixia was not evident in the beetle populations, although there was greater genetic variation within populations than among populations. The surface metrics were found to significantly explain the variance in genetic dissimilarities between the beetle populations, and did so 30 times better than the IBD model. We conclude that surface metrology of habitat maps is a powerful extension of landscape genetics tools that needs more attention.

Keywords:

Gene flow; isolation by distance; habitat connectivity; landscape configuration; microsatellites; R_{ST}

5.2 Introduction

A major focus of landscape genetics is to understand how population genetic processes are affected by the complexity and heterogeneity of spatial and temporal environmental patterns in the landscape (Bolliger *et al.* 2014; Manel *et al.* 2003; Storfer *et al.* 2007). Research questions in this discipline mostly focus on studying effects of geographic barriers and landscape variables or both on the genetic continuity and structure of different taxa at the landscape scale (Storfer *et al.* 2007; Storfer *et al.* 2010). The vast majority of these studies adopt discrete landscape ecology paradigms such as the patch mosaic model (Forman& Godron 1981), the variegation model (Mcintyre& Barrett 1992), or modified habitat gradient models (Fischer& Lindenmayer 2006; Manning *et al.* 2004) to quantify the effects of landscape composition, configuration, habitat quality, and connectivity on gene flow and spatial variation of population structure (examples reviewed in Storfer *et al.* 2007). However, natural populations usually exhibit continuous gradients of continuity, divergence, and structure across the landscapes in response to the continuous nature of habitat heterogeneity and landscape features. Thus, it is more appropriate to represent population structure and their patterns of genetic connectivity as a gradient rather than a categorical or patch-based phenomenon in complex landscapes (Cushman *et al.* 2010).

Adopting appropriate quantitative approaches to estimate the underlying landscape processes of interest and to assess habitat connectivity is crucial prior to linking genetic data and making inferences about population genetic structure. Landscape connectivity has been identified and refined to reflect the degree to which landscape structure facilitates or impedes movement of organisms and thus gene flow (Merriam

1984; Taylor *et al.* 1993; Tischendorf & Fahrig 2001). The importance of landscape connectivity to conservation and land management has increased interest in developing connectivity measures (Goodwin 2003). In many studies that use these measures, they are used for assessing patterns of adaptive traits (selection) (Holderegger *et al.* 2006) or using gene flow and drift (neutral variation) (Holderegger& Wagner 2008) to investigate ecological processes. They rely primarily on binary or categorical views of landscapes to explain the variance in genetic structure (examples reviewed in Storfer *et al.* 2007). Categorical landscape models have contributed much to our understanding of pattern-process interactions between species and environments. Yet, these models have caveats that are oversimplifying the multivariate spatial aspect and continuous nature of many environmental and genetic processes in the landscape (Cushman *et al.* 2010; McGarigal& Cushman 2005). McGarigal *et al.* (2009) introduced surface metrics that retain the continuous nature of environmental gradients to be implemented with the landscape gradient model. These metrics are classified into three categories: amplitude, configuration, and bearing metrics. Some of them are unique to surface metrology; they have no analogous metrics in categorical approaches to landscape description, especially the configuration metrics. The ability to detect drivers of genetic variation in the landscape is very sensitive to the composition (matrix and habitat quality) and the configuration (spatial arrangement) aspects of the landscapes (Jaquiéry *et al.* 2011). Although many landscape genetic studies have explored the effects of landscape composition and habitat quality on the population genetic structure (e.g. Angelone *et al.* 2011; Keller *et al.* 2013a; Keller *et al.* 2004) configuration of landscapes has been largely neglected (Bolliger *et al.* 2014). The new metrics of connectivity have shown promising

results when applied to a large-scale habitat quality surface (Abdel Moniem& Holland 2013). Therefore, the landscape gradient model and the surface metrology metrics of connectivity could offer a great advance to landscape genetics.

The banded flower longhorn beetle (*Typocerus v. velutinus* Olivier) belongs to family Cerambycidae (Coleoptera: Cerambycidae) which is a large cosmopolitan family of beetles comprising more than 9000 species known from the western Hemisphere and more than 900 species from North America (Bezark& Monné 2013). They are an important component of the biodiversity in almost any forested ecosystem. Several members of the family are serious pests, with the larvae boring into wood where they can cause extensive damage to either living trees or untreated lumber. However, many other species are important mediators of ecosystem services (Hanks 1999; Michelsen 1963). *Typocerus. v. velutinus,* a member of the subfamily Lepturinae (flower visiting cerambycids), is an important generalist that helps decompose decaying wood and cycle nutrients and acts as a pollinator of valuable hardwood trees such as the American chestnut (Benjamin 1907). They depend on landscape complementarity to complete their life cycles and hence, care should be taken in modeling their habitat requirements. Herein, we developed a set of polymorphic microsatellite markers, and adopted a landscape gradient model (McGarigal & Cushman 2005) and surface metrology (Abdel Moniem& Holland 2013; McGarigal *et al.* 2009) to model the genetic dissimilarities between populations and evaluate the predictive power of surface connectivity metrics. More explicitly, we tested the hypothesis that landscape structure and habitat connectivity facilitate beetle movement and thus gene flow between the beetle populations against a null model of isolation by distance (IBD). We predicted that habitat connectivity as

measured by surface metrics in the studied landscapes would better explain the genetic dissimilarities between the beetle populations than would geographic distance alone.

5.3 Materials and Methods

5.3.1 Beetle sampling and assessment of habitat connectivity using surface metrics:

Typocerus v. velutinus was sampled from 17 sites across Indiana, USA. Sites were scattered across a fragmentation gradient that varies from an area of connected forests in south-central Indiana to an area of highly fragmented forests to the north (Figure 5.1). Names and coordinates of these sites, sampling project and numbers of individuals caught are reported in Table 5.1 and see previous studies for sampling details (Abdel Moniem& Holland 2013; Holland *et al.* 2012; Raje *et al.* 2012). A habitat quality surface was created using six geographical information system (GIS) layers for Indiana that represent biological and geophysical requirements for both larvae and adult beetles. These layers were smoothed from a spatial resolution of 300 x 300 m to a spatial scale of 2.1 km representing the appropriate response scale of the beetle to the surrounding landscape (Yang 2010). A multiple Poisson regression was done to model the abundance of these beetles across the whole state and generate the state-wide habitat quality surface (Abdel Moniem& Holland 2013). To assess the habitat connectivity between all sites, a correlated random walk approach (Koh *et al.* 2013; Okubo& Kareiva 2001) was adopted to delineate 136 spatial polygons (landscapes between sites) that encompass likely routes of beetle dispersal between sites. Within these elliptical landscapes we measured ten surface metrics of connectivity (Abdel Moniem& Holland 2013; supplementary material). These metrics demonstrate different characteristics of the habitat quality surface while

possessing minimum possible redundancy among them (McGarigal *et al.* 2009). We used the Scanning Probe Image Processor (SPIP)^{TM} software to calculate the chosen surface metrics. These metrics, Euclidean distance between sites, and the surface metric-distance interactions were standardized and used as predictor variables and the fixed components in multiple generalized additive mixed models (GAMM). One of the pair of sampling sites was a random effect variable in the model to avoid pseudoreplication and possible type I errors from techniques such as Mantel tests (Legendre& Fortin 2010). Our response variable was the genetic dissimilarity matrix of R_{ST} measured between populations at each of the sampling sites after a Box-Cox transformation (Box & Cox 1964) to meet the model's assumptions. We started with the beyond optimal models that include all possible explanatory variables and interaction terms (fixed component) and we optimized our random component (sites) in these mixed models. We used the restricted maximum likelihood estimation (REML) and adjusted R^2 to compare our models. In this procedure, we retained explanatory variables that passed a *t* statistical significance test of level 0.05 in the optimal model. The surface metrics model was compared to a model that contains Euclidian distance only between sites as a null model representing isolation by distance only (IBD). All geoprocessing and statistical analysis were done in R (R Development Core Team 2013) and ArcGIS 9.2 (ESRI 2006).

5.3.2 Isolation of DNA and developing the Microsatellites

For each beetle, DNA was extracted from three legs using Qiagen DNeasy blood and tissue kits (Qiagen, Inc. Valencia, CA) following company protocol. Each individuals' DNA was eluted in 200 µl of elution buffer and stored in -20°C and voucher

specimens were stored in the Purdue Entomological research collection (PERC). High quality DNA was purified prior to sequencing using DNA Clean & Concentrator™-5 (Zymo Research, Orange, CA) following manufactures' procedure. To isolate microsatellites, a 1 µg of sample DNA was converted to a TruSeq library using methodology supplied by the manufacturer (Illumina, San Diego, CA) with the following modifications. DNA was sonicated using the Covaris 800 machine and subjected to size selection using a 1:0.6X sample:Agencourt AMPure XP (Beckman Coulter, Inc. Pasadena, CA) beads purification. This resulted in fragments largely of 500–1500 bp in length. Only 4 cycles of enrichment PCR were undertaken, rather than 10. The library ranged in size from 400–2000 bp, with an average length of just less than 1 kb. The library was titrated with a qPCR kit (KAPA Biosystems, Wilmington, MA) on a StepOne™ instrument (Applied Biosystems, Inc. Foster City, CA) then clustered at 15 pM on a MiSeq 500 cycle cassette for 250 cycles in both directions. A subsequent MiSeq v2 300 cycle run was undertaken subsequently and the sequences from both runs were assembled using ABySS (Simpson *et al.* 2009). To scan the resulted sequences for SSRs, scaffolds of length between 501 to 5498 bp (no missing bases, and GC% 32) were considered. Using a Perl code, SSRs were defined as any tandem repeat of 6 or greater of a sequence at least 2 bases long. The output was further filtered by discarding any SSRs within the first or last 100 bp of the scaffold sequence. A total of 24 candidate di- and tri-nucleotides SSRs were selected for screening in the beetle populations. Forward and reverse primers for these SSRs were designed in Primer3 (Rozen & Skaletsky 2000) to amplify three different fragment sizes of DNA (90–130, 180–220, and 270–320 bp) and were tagged with three florescent dyes (black, green and blue). Polymerase chain

reactions (PCR) were done independently per locus in 25 µl reactions. The reactions contained 2.5 µl of 10X PCR buffer (Promega, Madison, WI), 1.5 µl of 25 nM MgCl2 (Promega), 5 µM dNTPs (Promega), 0.75 U of Taq DNA polymerase (Promega), 1.0 µl of reverse non-fluorescent primer (5 µM) and 0.5 µl of a fluorescently labeled forward primer (10 µM), labeled with one of three Beckman-Coulter florescent dyes (Invitrogen, Carlsbad, CA), and ~25 ng template DNA. The cycling program was set to 95°C for 4 min, 6 cycles of 95°C for 1 min, 50°C for 1 min, 72°C for 1 min, 31 cycles of 95°C for 30 sec, 50°C for 30 sec, 72°C for 55 sec, and a final extension of 72°C for 30 min performed in a DNA Engine Dyad thermo-cycler (MJ Research, Watertown, MA). Post amplification, PCR products were multiplexed for genotyping. The multiplexed mixture composed of four groups of six loci, and adding 1 µl of blue, 4 µl of green, and 10 µl of black florescent tagged products. Genotyping reactions were prepared by mixing 1 µl aliquot of multiplexes, 0.5 µl of a 600 bp size standard (Beckman-Coulter), and 40 µl SLS buffer (Beckman-Coulter). Genotyping was performed on a Beckman-Coulter CEQ8000, following the manufacturer's instructions, and sized with CEQ8000 software. Recorded genotypes were checked for null alleles, stuttering, and scoring errors at individual populations level using Micro-Checker (Van Oosterhout *et al.* 2004). After the loci were assessed in Micro-Checker, ten final microsatellites were chosen for the analysis of the beetle population structure (Table 5.2).

5.3.3 Analysis of population genetics structure

To assess genetic diversity, allele frequencies and observed and expected heterozygosities for each locus in each sample were estimated. Departures from Hardy-Weinberg equilibrium for multiple alleles was examined using a test analogous to an

extension of Fisher's exact test of differentiation with 1,000,000 steps (Guo& Thompson 1992) of Markov chain Monte Carlo simulation (MCMC) and 100,000 dememorization steps done as implemented in Arlequin 3.5.1.3 (Excoffier *et al.* 2005). Both standard and molecular diversity indices were calculated for each locus. Linkage equilibrium between all pairs of the 10 loci were tested using the likelihood-ratio test (Slatkin & Excoffier 1996) with 10,000 MCMC and 1000 steps for burn-in. The pairwise R_{ST} (Slatkin 1995) was computed as a genetic differentiation index between populations. R_{ST} is a more appropriate measure for multi-allelic microsatellite data because it accounts for a suitable mutation model (stepwise mutation model, SMM) for SSRs (Hardy *et al.* 2003; Slatkin 1995). Analysis of population subdivision under the AMOVA framework (Excoffier 2003; Excoffier *et al.* 1992), was conducted to detect variation between individuals within populations, and among populations groups along with computation of F-statistics (inbreeding coefficient, F_{IS} and index of population differentiation F_{ST}). AMOVA was done after inferring the most likely number of true populations (K). To assess the subdivision of the beetle population structure, two approaches were followed. First, we used a non-spatially explicit Bayesian technique in STRUCTURE v2.3 (Pritchard *et al.* 2000). All 453 genotypes were entered as unique individuals and assigned putative populations based on the 17 sampling sites. The run length contained 5000 runs as burn-in period and 50,000 MCMC steps. The admixture model with correlated allele frequencies was used and assumed a range of K = 1–7 each to run for 10 iterations to calculate statistics. Output used Evanno's ΔK method based on the rate of change in the log probability of data between successive K values (Evanno *et al.* 2005) in STRUCTURE HARVESTER v0.6.7 (Earl& Vonholdt 2012). DISTRUCT (Rosenberg

2004) was used to graphically display population structure. Second, a spatially explicit Bayesian technique that is sensitive to a priori information given about sampling coordinates thus isolation by distance (IBD). The package Geneland 'R' (Guillot *et al.* 2005) was used with similar MCMC settings to infer the number of true populations (best K) and of the spatial location of genetic discontinuities.

5.4 Results

The total number of *T. v. velutinus* collected was 453. Individuals per site used varied between 5 and 69 (Table 5.1). All 10 microsatellite loci were polymorphic in all populations except for one locus (M34_3) in a single population (FPAC) (Supplement 5.1). Total number of alleles ranged between seven and 65. Observed heterozygosities varied from 0.225 to 0.834. Genetic diversity as inferred from average number of alleles ($\mu.A$) and mean observed heterozygosities ($\mu.H_o$) varied among populations. The lowest of these measures were scored in FPAC ($\mu.A$=3.4 and $\mu.H_o$=0.42) and LO04 ($\mu.A$=3.3 and $\mu.H_o$=0.48) while the highest were scored in HEE6 ($\mu.A$=12.7 and $\mu.H_o$=0.55) and HEE7 ($\mu.A$=11.6 and $\mu.H_o$=0.58).

Across populations, all loci were significantly different from what expected under Hardy–Weinberg equilibrium (Table 5.2). However, the number of loci that deviated from Hardy–Weinberg equilibrium varied among populations, ranging from one to eight (Supplement 5.1). Among the 765 tests of linkage equilibrium between all loci pairs at all populations, 11.7% of these tests were significant at $P<0.05$. The population genetic

dissimilarity between sites as calculated by R_{ST} (Figure 5.2) showed a pattern of higher dissimilarities between the populations FPAC, HEE9, LO10, LO11, and LO04 and the rest of the populations in the study.

Population structure analysis using both non-spatial and spatial explicit Bayesian clustering techniques showed that the best K for the beetle population is five. The non-spatial explicit Evanno's method indicated a highest ΔK value of 19.03 and a mean LnP(K) = -15818.52 ± 14.09 associated with K = 5 (Figure 5.3.a). The Geneland analysis indicated a number of discontinued populations of five, associated with a maximum probability density along the MCMC after burn-in of ~ 0.6 (Figure 5.3.b). The map of estimated cluster membership delineated the five spatial population clusters in a similar clustering pattern comparable to results from non-spatial clustering (Figure 5.4). The map of posterior probabilities associated with population cluster five showed possible genetic discontinuities between populations (e.g. Pop 1 and 2, and Pop 4 and 5) but it also indicated a possibility of genetic connectivity between Pop 4 and the site LO03 from Pop 3 (Figure 5.5).

The AMOVA conducted on these five populations revealed that most of the variance (83.95%) was explained by within-populations variation as compared to 16.04% of the explained variance by among populations differences (Table 5.4). The fixation index $F_{ST} = 0.16$ was significant at $P<0.0001$.

The GAMM containing the surface metrics of connectivity and their interaction terms with geographic distance between sites explained 30.5% of the variance in R_{ST} measured between beetle populations at these sites (overall model significance: $P<0.01$,

n=137, df=16). Among the surface metrics used in the model, two amplitude metrics, surface skewness (*Ssk*), and surface kurtosis (*Sku*) and one bearing metric, surface bearing index (*Sbi*), were the most significant (P<0.0001) and important predictors with coefficients ± standard deviations of 0.434 ± 0.102, -0.427 ± 0.084, and 0.176 ± 0.047 successively. Surface fractal dimension (*Sfd*) was a second important predictor (coeff. = -0.558 ± 0.166, P<0.001). The interaction term with geographic distance of these predictors was less important and significant (P<0.01) in explaining the variance in R_{ST} with coefficients: 0.232 ± 0.082, 0.128 ± 0.042, and -0.343 ± 0.129 respectively, for *Sku:Geo_dist*, *Sbi:Geo_dist*, and *Sfd:Geo_dist*. However, the interaction term was meaningful only with one configuration surface metric (surface radial wavelength index; *Srwi*) with a coefficient 0.144 ± 0.07 at P<0.05 (Table 5.4). The IBD model explained only 1.1% of the variance in R_{ST} with overall significance at P=0.02 (Table 5.4).

5.5 Discussion

Habitat connectivity in the landscape as measured by surface metrology metrics explained the population genetic dissimilarities between the banded longhorn beetle [*Typocerus v. velutinus* (Olivier)] populations. Thus, surface metrics of connectivity appear to have the potential to be powerful analysis tools in landscape genetics. Applying a suite of surface metrics to a habitat quality surface can provide information on both non-spatial and spatial characteristics of the habitat while maintaining the continuous nature of heterogeneity in the landscape. This can give insights into specific biological and geophysical requirements of species. Also, because habitat heterogeneity and landscape composition and configuration may influence the genetic continuity of

populations depending on the distance traversed by these animals, the interaction of surface metrics and geographic distances between sites should be considered. Considering the continuous nature of habitat heterogeneity in complex landscapes and the population genetic structure of species could provide much more insight in our ability to understand the link between patterns and processes in a landscape genetics context over categorical approaches (Cushman *et al.* 2010; Murphy *et al.* 2008).

Surface kurtosis (*Sku*) and surface skewness (*Ssk*) are non-spatial metrics (amplitude metrics) that can provide information on habitat heterogeneity. Surface kurtosis explains the peakedness of the surface height distribution, while surface skewness describes whether high or low values dominate the landscape. Thus, coupling these complementary metrics can yield inference on the degree and nature of land cover dominance in the landscape (McGarigal *et al.* 2009). Higher values of surface kurtosis indicate a high contrast between the high and low habitat quality values (peaks and valleys) in the landscape. The genetic dissimilarities between the beetle populations varied significantly and inversely with kurtosis (Table 5.4). Thus, the higher the contrast between habitat quality in the landscape, the more similar the populations are. This finding counters the expectation that a higher contrast landscape would hinder individuals' movement and thus the gene flow through generations of the beetle populations. However, a possible interpretation could be that higher contrast landscapes contain more high quality habitat which is more conductive for movement while low contrast landscapes will be dominated more with intermediate and low quality habitat that are less used and difficult to traverse by the beetles. This interpretation becomes more likely upon considering the relationship between surface skewness and genetic dissimilarity. The

high positive and significant relationship indicates that beetle populations will be more similar with the dominance of higher values of habitat in the landscape. Thus it is important to interpret these two complementary amplitude metrics together to get a better picture of the response to the landscape. A similar pattern of response was reported for seven of 16 lepturine species, and for the overall community in a study that looked at abundance dissimilarities in the same landscapes (Abdel Moniem & Holland 2013), however, skewness was less informative than for the genetic similarities reported here.

These two amplitude metrics (*Ssk* and *Sku*) were less important predictors of the genetic dissimilarities when including their interaction with geographic distance between sites. A possible explanation for kurtosis is that when the distance increases between sites the numbers of both beaks and valleys will increase dramatically. Because this metric is very sensitive to deep valleys and high peaks (McGarigal *et al.* 2009) the sensitivity of the metric to calculate informative and interpretable contrast measure may drop. Similarly, with skewness, a larger area will comprise a mixture of smaller areas that are dominated differently by high or low habitat values in the landscape, consequently, an overall trend of dominance between farther sites will be harder to define.

A great advantage of surface metrics of habitat quality is the availability of landscape configuration metrics. These metrics are unique to surface metrology metrics and have no analogues in categorical measures (e.g. measures implemented in FRAGSTAT based on the PMM) of the landscape (Cushman *et al.* 2010; McGarigal & Marks 1995; McGarigal *et al.* 2009). Among these new metrics, our results indicated that surface fractal dimension is an important significant predictor of the genetic dissimilarities between sites in the studied landscapes. High values of this configuration

metric (the metric values range between 2–4) indicate a dominant direction of high peaks (high quality habitat) in the landscape as inferred from the Fourier transformation of their radial wavelengths (McGarigal *et al.* 2009). Interpreting this metric ecologically could be difficult without coupling it with the surface dominant texture direction (*Std*), anther configuration metric. Surface dominant texture direction measures the orientation of the dominant undulations of habitat quality across the landscape. It is only meaningful if there is a dominating direction of the peaks of habitat quality. Thus, in our case, since *Std* was not related to the genetic dissimilarities we suggest that landscapes with high habitat quality areas aggregated in a certain direction (as indicated from high *Sfd* values) will limit the movement of individuals to this area leaving beyond the surroundings low quality habitats. Therefore, genetic similarity will be driven mainly by the gene flow between the populations localized within these high habitat quality spots. Interestingly, this configuration metric (*Sfd*) also predicted genetic dissimilarities when considering its interaction with geographic distance. This pattern supports our interpretation mentioned above, because it would be predicted that genetic dissimilarities will start to increase as other high quality habitat spots will appear in the landscape, but then there will be no dominant texture direction unless the new habitat are also aggregated. The surface radial wavelength index (*Srwi*) is another configuration metric that was found only meaningful and significant when considered in combination with geographic distance. The metric could be interpreted ecologically similarly to *Std* and *Sfd*. However, it could be conceptually related to the coefficient of variation in nearest neighbor distance from PMM because this index indicates the change in spacing of surface height deviations (McGarigal *et al.* 2009). These configuration metrics could have important applications

in determining the orientation of repeated high contrast areas that could impede animal movement (Abdel Moniem & Holland 2013). Thus they could be used as warning of cumulative effects of repeated barriers of gene flow that would have a large effect on genetic diversity and population structure.

The surface bearing metric (*Sbi*) is a landscape composition metric that is found to be an important and significant predictor to the genetic dissimilarities with and without its interaction with geographic distance. The metric describes the cumulative distribution of the habitat quality in the landscape. It is more sensitive to occasional peaks of high quality habitat than the valleys of low quality ones (McGarigal *et al.* 2009). Graphically, this metric is represented as Abbott curves which can be used to make a visual inference about the relative amounts of high, medium, and low habitat values (Abdel Moniem & Holland 2013; McGarigal *et al.* 2009). Because high *Sbi* values (>0.608) could reflect either the presence of many high peaks or their absence (McGarigal *et al.* 2009) it must be interpreted with caution. This metric does not have a spatial component to indicate locations of these habitat values, thus proportions of habitat quality measures on Abbott curves will not necessarily reflect gradual transitions in the landscape. We suggest further investigations into the behavior of this metric with model landscapes to compare areas with different degrees of connectedness of the high peaks.

Opposite to our findings from the habitat connectivity model explained above, isolation by distance (IBD) was not supported as an underlying mechanism that explained differences in genetic dissimilarities. Euclidian distance between sites is often used as a null model in landscape genetics, because population genetic theory predicts that genetic distances between populations will increase with increasing geographical distance

(Allendorf & Luikart 2007). Other studies on ground beetles, and grasshoppers, have rejected the null model of IBD in explaining the population genetic structure (Keller *et al.* 2013b; Keller & Largiader 2003; Keller *et al.* 2004). The connectivity model expressed by conducting surface metrology on a biologically-informed continuous habitat quality surface has much more information about the biological and geophysical requirements of the species, and on the composition and configuration of the landscape. This was shown to be much more informative when studying drivers of genetic variations in a landscape genetics context.

A concordance was found between both non-spatially and spatially explicit Bayesian techniques in estimating the true number of the beetle populations in Indiana. The five populations inferred from these two approaches seem to belong, spatially, to different forested areas in the landscape across a certain fragmentation gradient. For example, one large population was characterized in the state forests of southern Indiana (Pop 2 in Fig. 4). The landscape in this area encloses large connected forests with high quality habitat that can be highly permeable for movement and gene flow. However, the noticed differences between beetles from this population and an adjacent one (Pop1 in Fig. 4) might be attributed to the presence of a deep valley of poor habitat quality represented by some physical barriers such as agriculture and urbanized areas between the two populations. This pattern was illustrated also by the steep contours in the heat map of posterior probabilities associated with the five clusters detected by Geneland (Figure 5.5). Physical features due to anthropogenic activities have been shown in many studies as potential disruptors of the genetic continuity of populations in the landscape (Cushman 2006; Keller *et al.* 2013b; Keller *et al.* 2004; Paetkau *et al.* 1995). Our results

further support the idea that similarities between beetle communities in the studied landscapes is more related to the intervening habitat quality and connectivity between sampled sites than just due to localized habitat similarity at these sites.

The beetle population across all sites was not found to be a single panmictic one as inferred from the population structure analysis and the multilocus Hardy-Weinberg (HW) tests. Thus, there is some level of population structure found which could be attributed primarily to the within population variations and partially due to among populations as inferred from the AMOVA results. At the subpopulation level, there were more loci found to be at both HW and linkage equilibrium (Supplement 5.1). These results in combination with the population genetic structure results support our hypothesis because within high habitat quality areas the populations seem to be more connected and move freely with less barriers of gene flow and thus with less genetic dissimilarity than populations that coexist at a fragmented area of high and low habitat quality. However, within populations in few sites we found few loci with significant departure from HW equilibrium and linkage disequilibrium (LD). A possible explanation of that could be the presence of some finer level structure within individual populations at sites that was not detected at large landscape scale we used, in addition to the possibility that null alleles cannot be completely avoided at a finer sampling scale. With regard to the LD some of the significant tests we reported (Supplement 5.1) involved different pairs of loci in different samples. We could conclude that they were more likely to be a result of type I errors or due to within-population structuring as a result of limited sample size at few populations and not due to an actual physical linkage between loci. Actual linkage between loci would be more likely to be predicted if a significant linkage was

found between same pairs of loci in several samples (Avise 1994; Hartel & Clark 1997).

In a continuously changing landscape and increasingly fragmented habitat that caused by anthropogenic activities and changing environmental conditions, understanding the factors affecting population connectivity is essential for conservation and management of biological diversity (Cushman *et al.* 2006; Murphy *et al.* 2010) especially for ecologically important species. Landscape genetics approaches seem to provide insightful conclusions that help us understanding these dynamics. However, a challenge for landscape geneticists and ecologists is to integrate three components: non-spatial niche relationships, spatial patterns of environmental gradients and continuity of genetic structure in complex heterogeneous landscapes (Austin 2007; Austin 1985; Cushman *et al.* 2007; Manning *et al.* 2004). In an attempt to tackle this challenge, in our study we accounted for non-spatial niche component by creating a continuous surface of habitat quality with insight into biological and geophysical requirements of the beetle species at an optimum spatial response scale. We considered the gradient model of the heterogeneity in the landscape and modeled the genetic continuity of the population in a landscape genetics context. We conclude that surface metrology of habitat quality is a powerful tool that considers both composition and configuration of the landscape and can potentially explain the variation in genetic dissimilarities and population structure. We suggest that more effort should be applied to understand the behavior of these metrics, especially the ones concerned with the configuration of the landscape. We also suggest that direct methods of estimating gene flow should be tested with these metrics at various spatial scales, as this could be indirect promising extension in studying dispersal of organisms when traditional techniques are hard to implement.

Acknowledgments

We thank Dr. Samuel Cushman and Dr. Kevin McGarigarl for encouragement and feedback on the surface metrics used. We thank Anders Kühle for the technical support with the SPIP™ program, and Yan Crane and Alisha Johnson from the USDA-ARS laboratory at Purdue University for help with molecular techniques. We also thank Phillip San Miguel, Rick Westerman, and Viktoria Krasnyanskaya; Purdue Genomics Facility, for next-generation sequencing assistance. Kapil Raje, John Shukle and Tommy Mager helped with field collections. This research was supported by a governmental general mission scholarship administrated by the Egyptian Cultural and Education Bureau, Washington, DC, and by the Department of Entomology at Purdue University.

5.6 Bibliography

Abdel Moniem HEM, Holland JD (2013) Habitat connectivity for pollinator beetles using surface metrics. *Landscape Ecology*, **28**, 1251–1267.

Allendorf FW, Luikart G (2007) *Conservation and the Genetics of Populations*. Blackwell, Malden, MA.

Angelone S, Kienast F, Holderegger R (2011) Where movement happens: scale–dependent landscape effects on genetic differentiation in the European tree frog. *Ecography*, **34**, 714–722.

Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.

Austin MP (1985) Continuum concept, ordination methods, and niche theory. *Annual Review of Ecology and Systematics*, **16**, 39–61.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York, USA.

Benjamin D (1907) Annual report of the state entomologist of Indiana. Library of the museum of comparative zoology, Harvard University, 216 pp.

Bezark LG, Monné MA (2013) Checklist of the Oxypeltidae, Vesperidae, Disteniidae and Cerambycidae, (Coleoptera) of the Western Hemisphere, pp. 1–484.

Bolliger J, Lander T, Balkenhol N (2014) Landscape genetics since 2003: status, challenges and future directions. *Landscape Ecology*, **29**, 361–366.

Box GEP, Cox DR (1964) An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, **26**, 211–252.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation*, **128**, 231–240.

Cushman SA, Gutzweiler K, Evans JS *et al.* (2010) The gradient paradigm: A conceptual and analytical framework for landscape ecology. In: Cushman SA, Huettmann F (eds) *Spatial Complexity, Informatics and Wildlife Conservation*. Springer, Tokyo, pp. 83–110.

Cushman SA, McKelvey KS, Hayden J *et al.* (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist*, **168**, 486–499.

Cushman SA, McKenzie D, Peterson DL *et al.* (2007) Research agenda for integrated landscape modelling. USDA Forest Service General Technical Report RMRS–GTR–194.

Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L (2003) Analysis of Population Subdivision. In: Balding D, Bishop M, Cannings C (eds). *Handbook of Statistical Genetics*. John Wiley & Sons, Ltd, New York, pp. 713–750.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes – Application to human mitochondrial–DNA restriction data. *Genetics*, **131**, 479–491.

Fischer J, Lindenmayer DB (2006) Beyond fragmentation: the continuum model for fauna research and conservation in human–modified landscapes. *Oikos*, **112**, 473–480.

Forman RTT, Godron M (1981) Patches and structural components for a landscape ecology. *BioScience*, **31**, 733–740.

Goodwin BJ (2003) Is landscape connectivity a dependent or independent variable? *Landscape Ecology*, **18**, 687–699.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hanks LM (1999) Influence of the larval host plant on reproductive strategies of cerambycid beetles. *Annual Review of Entomology*, **44**, 483–505.

Hardy OJ, Charbonnel N, Freville H, Heuertz M (2003) Microsatellite allele sizes: A simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Hartel DL, Clark AG (1997) *Principles of Population Genetics*, Third edn. Sinauer Associates, Inc, Sunderland, MA.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21**, 797–807.

Holderegger R, Wagner HH (2008) Landscape genetics. *BioScience*, **58**, 199–207.

Holland JD, Shukle JT, Abdel Moniem HEM*, et al.* (2013) Pre–treatment assemblages of wood–boring beetles (Coleoptera: Buprestidae, Cerambycidae) of the Hardwood Ecosystem Experiment. In: Swihart RK, Saunders MR, Kalb RA *et al.* (eds). *The Hardwood Ecosystem Experiment: a framework for studying responses to forest management*. General Technical Report, NRS-P-108. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station, pp. 218–236.

Jaquiéry J, Broquet T, Hirzel AH *et al.* (2011) Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology*, **20**, 692–705.

Keller D, van Strien MJ, Herrmann M*, et al.* (2013a) Is functional connectivity in common grasshopper species affected by fragmentation in an agricultural landscape? *Agriculture, Ecosystems & Environment*, **175**, 39–46.

Keller D, Holderegger R, van Strien MJ (2013b) Spatial scale affects landscape genetic analysis of a wetland grasshopper. *Molecular Ecology*, **22**, 2467–2482.

Keller I, Largiader CR (2003) Recent habitat fragmentation caused by major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. *Proceedings of the Royal Society*, **270**, 417–423.

Keller I, Nentwig W, Largiader CR (2004) Recent habitat fragmentation due to roads can lead to significant genetic differentiation in an abundant flightless ground beetle. *Molecular Ecology*, **13**, 2983–2994.

Koh I, Rowe HI, Holland JD (2013) Graph and circuit theory connectivity models of conservation biological control agents. *Ecological Applications*, **23**, 1554–1573.

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, **10**, 831–844.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Manning AD, Lindenmayer DB, Nix HA (2004) Continua and Umwelt: novel perspectives on viewing landscapes. *Oikos*, **104**, 621–628.

McGarigal K, Cushman SA (2005) The gradient concept of landscape structure. In: Wiens J, Moss M (eds) *Issues and Perspectives in Landscape Ecology*. Cambridge University Press, Cambridge, pp. 112–119.

McGarigal K, Marks BJ (1995) FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR. p. 122.

McGarigal K, Tagil S, Cushman SA (2009) Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology*, **24**, 433–450.

Mcintyre S, Barrett GW (1992) Habitat variegation, an alternative to fragmentation. *Conservation Biology*, **6**, 146–147.

Merriam G (1984) Connectivity: a fundamental ecological characteristic of landscape pattern. In: Brandt J, Agger P (eds) *Proceedings of First International Seminar on Methodology in Landscape Ecology Research and Planning*, *vol I*. Roskilde Universitessforlag GeoRue, Roskilde, Denmark, pp. 5–15.

Michelsen A (1963) Observations on the sexual behaviour of some longicorn beetles, subfamily Lepturinae (Coleoptera, Cerambycidae). *Behaviour*, **22**, 152–166.

Murphy MA, Evans JS, Cushman SA, Storfer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography*, **31**, 685–697.

Murphy MA, Evans JS, Storfer A (2010) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology*, **91**, 252–261.

Okubo A, Kareiva P (2001) Some examples of animal diffusion. In: Okubo A, Levin SA (eds) *Diffusion and Ecological Problems: Modern Perspectives*. Springer, New York, USA. pp. 170–196.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population–structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austia.

Raje KR, Abdel Moniem HEM, Farlee L *et al.* (2012) Abundance of pest and benign Cerambycidae both increase with decreasing forest productivity. *Agricultural and Forest Entomology*, **14**, 165–169.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rozen S, Skaletsky H (2000) Primer3 on the www for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.

Simpson JT, Wong K, Jackman SD*, et al.* (2009) ABySS: a parallel assembler for short read sequence data. *Genome Research*, **19**, 1117–1123.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the expectation–maximization algorithm. *Heredity*, **76**, 377–383.

Storfer A, Murphy MA, Evans JS*, et al.* (2007) Putting the "landscape" in landscape genetics. *Heredity*, **98**, 128–142.

Storfer A, Murphy MA, Spear SF *et al.* (2010) Landscape genetics: where are we now? *Molecular Ecology*, **19**, 3496–3514.

Taylor PD, Fahrig L, Henein K *et al.* (1993) Connectivity is a vital element of landscape structure. *Oikos*, **68**, 571–573.

Tischendorf L, Fahrig L (2001) On the use of connectivity measures in spatial ecology. A reply. *Oikos*, **95**, 152–155.

Van Oosterhout C, Hutchinson WF, Wills DPM *et al.* (2004) Micro–checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Yang S (2010) *Landscape scaling and occupancy modelling with Indiana longhorned beetles (Coleoptera: Cerambycidae)*. Dissertation, Purdue University.

Table 5.1 Seventeen study sites across Indiana. GPS coordinates (WGS84 UTM NAD83 zone 16N), number of individuals sampled, and sampling projects are recorded. LO: Land-owners sites (Raje *et al*., 2011); HEE: Hardwood Ecosystem Experiment (Holland *et al*., 2013), and UWEP: Upper Wabash Ecosystem Project.

Site	Name	N	E	# individuals	Project
1	FPAC	4304340	539052	5	UWEP
2	HEE1	4356090	548512	49	HEE
3	HEE2	4354720	548023	19	HEE
4	HEE3	4352430	547793	16	HEE
5	HEE4	4350830	549554	11	HEE
6	HEE5	4339480	554443	38	HEE
7	HEE6	4330210	554904	68	HEE
8	HEE7	4331690	558954	69	HEE
9	HEE8	4329640	558471	41	HEE
10	HEE9	4332500	561220	35	HEE
11	LO01	4442803	478158	26	LO
12	LO02	4505645	534188	13	LO
13	LO03	4464372	524447	5	LO
14	LO04	4423882	494783	5	LO
15	LO10	4475457	496660	25	LO
16	LO11	4474143	505748	19	LO
17	LO13	4499491	487722	9	UWEP

Table 5.2 Characteristics of 10 microsatellite loci isolated from the *Typocerus v. velutinus*. Reported are: locus name, GenBank Accession no., sequences of forward (F) and reverse (R) primers, repeat motif, allelic range in (bp), PCR annealing temperature, total number of alleles and observed and expected Heterozygosities (*Ho* & *He*). Asterisks marks significance of Hardy-Weinberg statistics.

Locus and GenBank Accession No.	Forward and Reverse primer sequences (5'-3') (Promega)	Repeat motif	Ta (°C)	Allelic range (bp)	Total num. alleles	*Ho*	*He*
M29_2 KJ415366	F:AAACGTACAGCGGTAAGAAA R:ACGTTGACTAACAGAAAATGCT	(GA)6	55.25	225–240	15	0.487	0.683*
M14_3 KJ415367	F:GTGGAGAATTTGGAGCAGTA R:TGTAAATGTGGTTGGGAGAC	(AGT)6	55.37	100–115	8	0.319	0.383*
M34_3 KJ415368	F:ACAGCGTACTTTTTCTAGGGTA R:GTTGAGGCTTGTATGGAAGA	(AAT)6	55.14	200–250	7	0.225	0.250*
M26_2 KJ415369	F:GCAATATTAAATCGCAATGG R:ATCGCCCCTAAGGTAAAATA	(TC)11	55.85	200–215	17	0.825	0.800*
M31_2 KJ415370	F:GAAGCGTCAGACAAAGAGAG R:CGGGTTTCGAGCTTTATATT	(GA)8	55.31	210–300	18	0.602	0.732*
M21_2 KJ415371	F:TACAATGCTCATGTTCACCA R:GAAACAACGACCATATCGAG	(AC)10	55.97	210–300	65	0.674	0.916*
M17_3 KJ415372	F:AATTTTGTTGCAAAGCTACTG R:AAAAAGGTTTAGTTTGGATTCAT	(TAG)6	55	100–110	24	0.372	0.627*
M8_2 KJ415373	F:CAGGCAGCAACTACTTTGAG R:TGTTACTGTTTTCGCCTTCT	(GA)12	56.34	100–120	27	0.569	0.799*
M37_3 KJ415374	F:TGCTTTGCTGATTATGTTGA R:GTTCATTTTCCATTTGTGCT	(TAA)9	55.47	230–245	18	0.834	0.794*
M12_3 KJ415375	F:CGTTTAAATCTGGGACACC R:GCTCTAAGCTAAACTTCACTTTGT	(ATA)9	55.49	90–100	23	0.662	0.803*

Table 5.3 Analysis of molecular variance (AMOVA) design and results conducted on the five populations identified from Indiana. Significance for test statistics was calculated with a MCMC chain length of 100000 steps with 10000 dememorization steps.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	4	162578.08	Va= 425.63	16.05
Within populations	901	2005905.39	Vb= 2226.31	83.95
Total	905	2168483.45	2651.94	
Fixation Indices Fst: 0.16	P-value < 0.0001			

Table 5.4 Summary of Generalized Linear Mixed Models (GLMM) with R_{ST} between sites as response variable and surface metrics of connectivity and their interaction with geographic distance between sites (Model 1) and the model explaining the variance in R_{ST} values under the isolation by distance only (Model 2). Significance codes: P= 0 '***' P<0.001 '**' P<0.01 '*' P<0.05 '.' 0.1

Model 1 Surface metrics of connectivity and their interaction with geographic distance. Adjusted R^2= 0.305 Overall model P-value = 0.01				
	Coefficient ± s.d.	t value	Pr(>|t|)	Sig.
(Intercept)	-3.213 ± 0.117	-27.409	<2e-16	***
Sa	0.14 ± 0.159	0.876	0.384	
S10z	0.152 ± 0.109	1.398	0.166	
Ssk	0.434 ± 0.102	4.258	0.0001	***
Sku	-0.427 ± 0.084	-5.095	0.0005	***
Sdr	-0.809 ± 0.535	-1.511	0.135	
Sbi	0.176 ± 0.047	3.717	0.0003	***
Std	0.148 ± 0.094	1.576	0.119	
Stdi	0.026 ± 0.043	0.619	0.537	
Sfd	-0.558 ± 0.166	-3.352	0.0012	**
Srwi	0.007 ± 0.071	0.104	0.917	
Geo_dist	-0.008 ± 0.156	-0.053	0.958	
Sa:Geo_dist	0.074 ± 0.15	0.494	0.622	
S10z:Geo_dist	0.146 ± 0.107	1.361	0.177	
Ssk:Geo_dist	0.195 ± 0.1	1.96	0.0534	.
Sku:Geo_dist	-0.232 ± 0.082	-2.808	0.0063	**
Sdr:Geo_dist	-0.624 ± 0.398	-1.567	0.1212	
Sbi:Geo_dist	0.128 ± 0.042	3.031	0.0033	**
Std:Geo_dist	0.149 ± 0.094	1.587	0.1166	
Stdi:Geo_dist	-0.035 ± 0.041	-0.86	0.3923	
Sfd:Geo_dist	-0.343 ± 0.129	-2.659	0.0095	**
Srwi:Geo_dist	0.144 ± 0.07	2.059	0.0429	*
Model 2 Isolation by distance only (IBD). Adjusted R^2= 0.011 Overall model P-value = 0.02				
	Coefficient ± s.d.	t value	Pr(>|t|)	Sig.
(Intercept)	-2.961 ± 0.019	-153.89	<2e-16	***
Distance	0.025 ± 0.017	1.47	0.145	

Figure 5.1 Habitat quality surface of the banded longhorn beetle across the State of Indiana along with the 17 study sites.

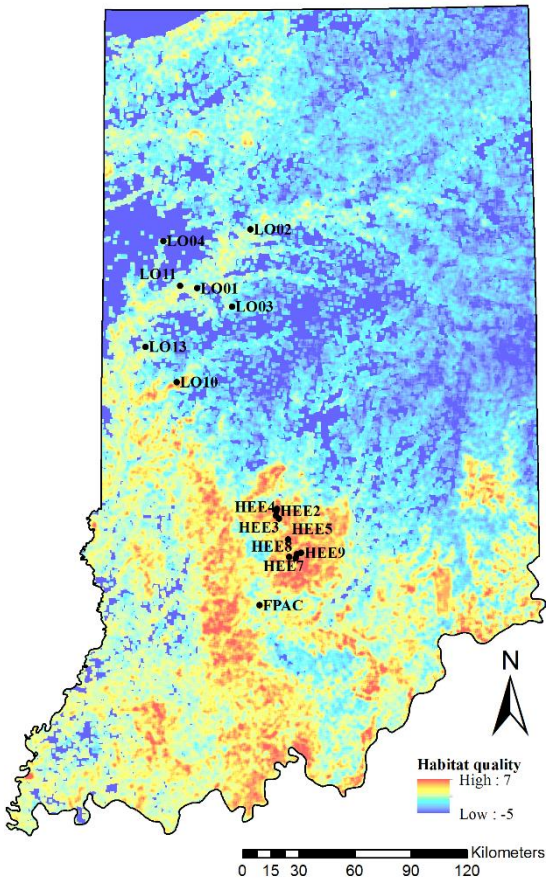


Figure 5.2 Population pairwise dissimilarity matrix as measured by the sum of squared differences in allelic size (R_{ST}) as implemented in Arlequin. Dots in boxes of pairwise comparisons indicate significance with $P<0.05$.

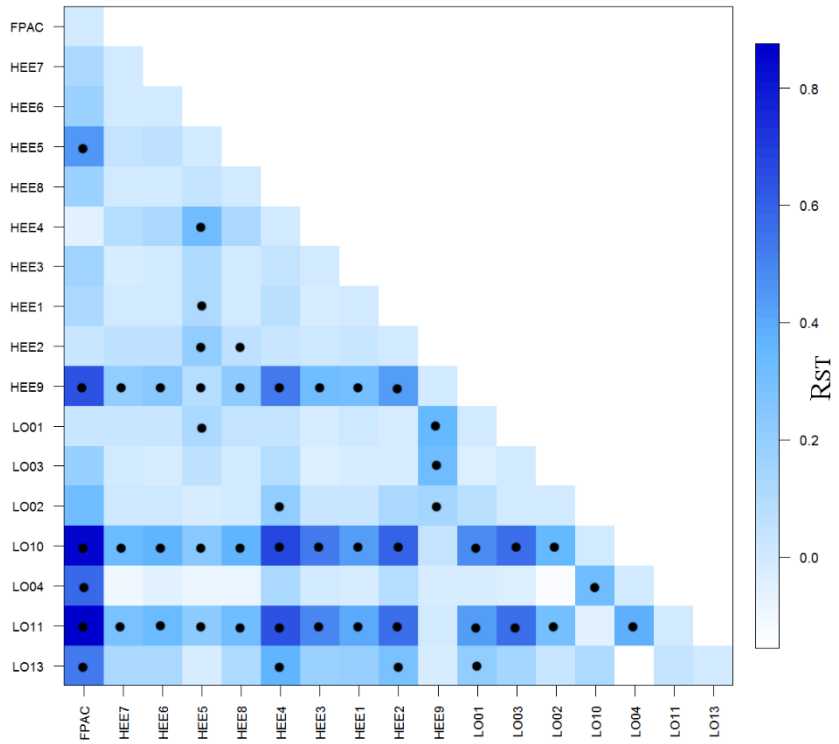


Figure 5.3 Inferring the beetle population structure. (a) Number of population classes investigated across the whole MCMC and the number of spatial population clusters and their probability density as inferred from Genelend. (b) Bar plots of admixture assignments for the beetle population across the state based on Bayesian clustering implemented in STRUCTURE, showing K = 5. Individual bars represent individual beetles with the colors indicating the likelihood assignment of each individual to an inferred genetic cluster. Population names abbreviated as in Table 1.

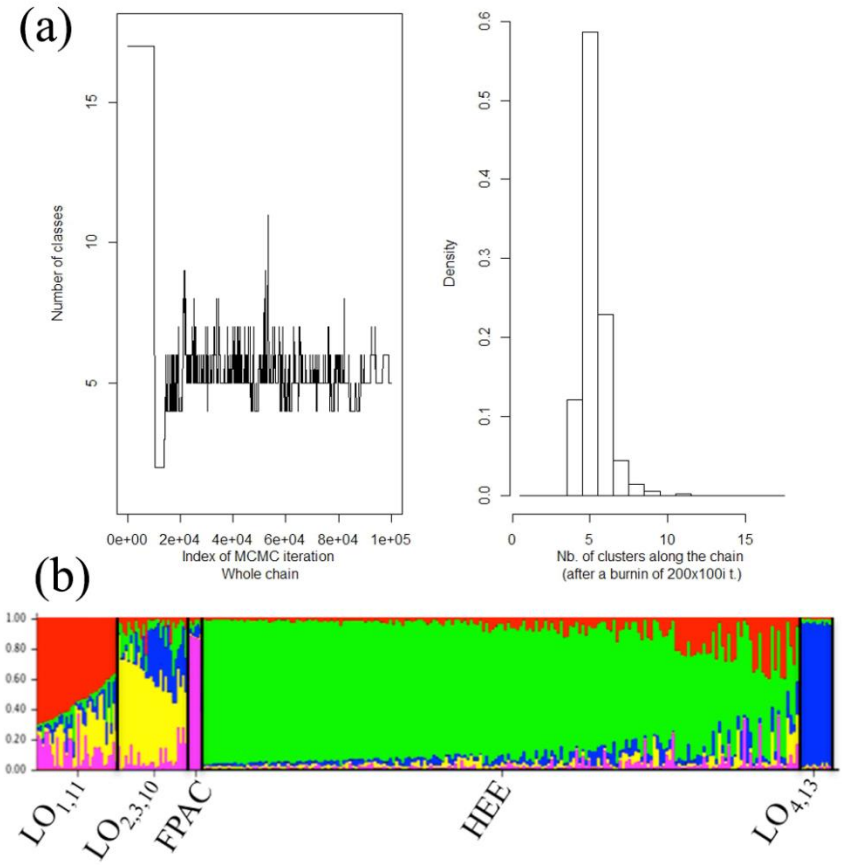


Figure 5.4 Thematic map of population membership clusters with coordinate axis as inferred from Geneland to the left and the corresponding clipped area from the habitat quality map with delineation of these population clusters to the right. Population clusters at sites are: Pop1= FPAC; Pop 2= $HEE_{1–9}$; Pop 3= LO $_{2,3,10}$; Pop 4= LO $_{4,13}$; and Pop 5= LO $_{1,11}$.

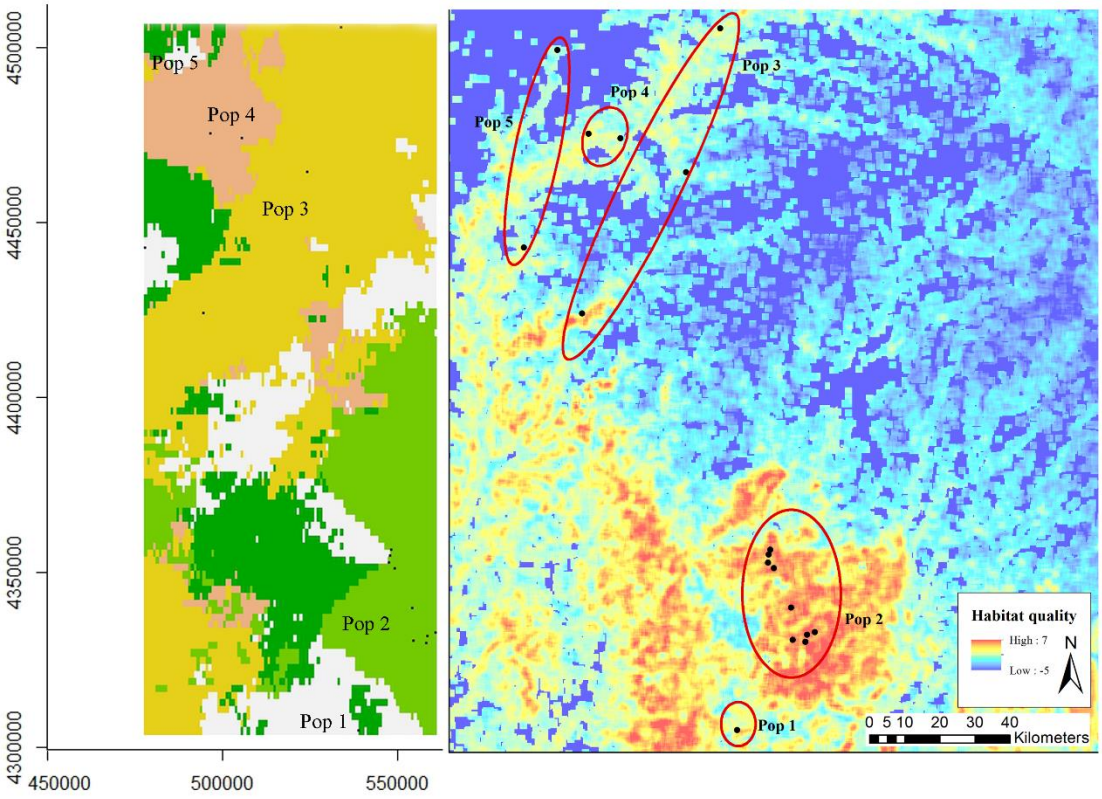
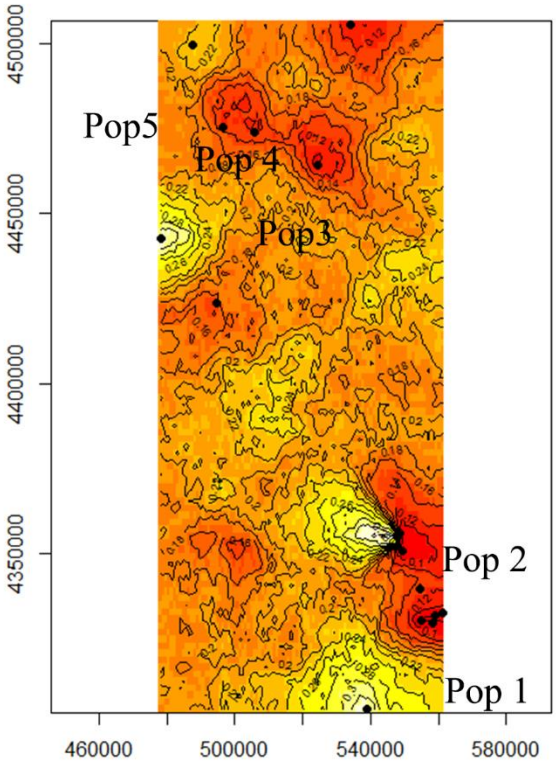


Figure 5.5 Map of posterior probabilities associated with population cluster five resulted from the spatial explicit Bayesian clustering performed in Geneland. The x and y axes represent easting and northing geographic coordinates consecutively. The heat map and the contours depict the spatial location of genetic discontinuities (i.e. possible barriers of gene flow between populations).



Supplement 5.1 Molecular diversity indices and Hardy Weinberg statistics for the 10 microsatellite markers in the 17 studied populations. L: locus number; *#G*: number of gene copies; *#A*: number of alleles; H_o: observed heterozygosities; H_e: expected heterozygosities; and *A.r*: allelic range.

			LO01					LO02					LO03		
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	52	5	0.308	0.649*	8	26	3	0.308	0.335	4	10	4	0.800	0.711	8
M14_3	52	5	0.385	0.402	12	26	2	0.154	0.148	6	-	-	-	-	-
M34_3	52	4	0.308	0.280	9	26	3	0.077	0.218*	6	10	2	0.000	0.356	3
M26_2	52	9	0.846	0.833	22	26	8	0.923	0.815	20	10	4	0.800	0.644	8
M31_2	49	13	0.625	0.664	42	26	6	0.615	0.606	12	10	4	0.600	0.533	18
M21_2	52	25	0.846	0.944*	224	26	12	0.692	0.868*	122	10	8	0.800	0.933	120
M17_3	52	8	0.462	0.637*	93	26	7	0.769	0.735	63	10	4	0.800	0.711	42
M8_2	52	13	0.731	0.850*	112	26	5	0.385	0.757*	8	10	3	0.200	0.644	10
M37_3	52	8	0.923	0.801*	30	26	4	0.769	0.686	162	8	3	1.000	0.750	6
M12_3	52	8	0.769	0.812	24	24	6	0.333	0.641*	66	10	4	0.200	0.733*	36
Mean	51.7	9.8	0.620	0.687	57.6	25.8	5.6	0.503	0.581	46.9	9.778	4	0.578	0.669	27.889
s.d.	0.949	6.161	0.237	0.209	68.479	0.632	2.951	0.289	0.255	55.929	0.667	1.658	0.353	0.159	37.177

			LO04					LO10					LO11		
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	10	2	0.600	0.467	2	50	5	0.560	0.666	8	36	5	0.500	0.651	8
M14_3	10	3	0.400	0.511	6	50	3	0.240	0.222	6	38	3	0.263	0.351	18
M34_3	10	2	0.200	0.200	3	50	3	0.240	0.223	6	38	3	0.263	0.243	6
M26_2	10	4	1.000	0.733*	8	49	8	0.917	0.762	18	38	10	0.895	0.772	30
M31_2	6	4	0.667	0.800	18	50	8	0.560	0.722	24	32	6	0.375	0.692*	24
M21_2	6	5	0.667	0.933	102	50	10	0.560	0.744	112	32	6	0.375	0.692*	24
M17_3	10	3	0.400	0.378	42	46	4	0.261	0.488*	15	38	6	0.421	0.565	33
M8_2	8	4	0.500	0.786	8	48	11	0.750	0.714	40	38	9	0.632	0.799*	38
M37_3	10	3	0.200	0.689*	6	50	8	0.960	0.830	27	38	8	0.947	0.788*	27
M12_3	10	3	0.200	0.733*	33	50	8	0.440	0.780*	87	38	7	0.579	0.787*	93
Mean	9	3.3	0.483	0.623	22.8	49.3	6.8	0.549	0.615	34.3	36.6	6.3	0.525	0.634	30.1
s.d.	1.7	0.949	0.259	0.226	30.904	1.337	2.86	0.265	0.226	36.421	2.503	2.312	0.241	0.194	24.329

	LO13					FPAC					HEE1				
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	14	5	0.429	0.813*	14	6	2	0.000	0.533	2	96	5	0.458	0.687*	8
M14_3	16	2	0.500	0.400	6	10	2	0.400	0.356	6	96	4	0.292	0.351	9
M34_3	18	3	0.111	0.307	9	-	-	-	-	-	96	3	0.167	0.245*	6
M26_2	18	8	1.000	0.824	22	10	5	1.000	0.822	14	96	8	0.813	0.800	22
M31_2	18	5	0.333	0.680*	12	10	4	0.400	0.711	30	96	9	0.542	0.752*	42
M21_2	18	8	0.444	0.797*	124	10	5	0.400	0.822	64	96	28	0.667	0.925*	146
M17_3	18	4	0.444	0.739	15	10	3	0.400	0.622	15	86	9	0.256	0.498*	96
M8_2	16	7	0.500	0.833*	38	10	2	0.000	0.356	4	82	12	0.390	0.664*	42
M37_3	18	6	0.667	0.843	30	10	5	0.800	0.867	24	78	7	0.744	0.777	24
M12_3	18	4	0.222	0.627*	57	10	3	0.400	0.511	33	86	11	0.674	0.799*	57
Mean	17.2	5.2	0.465	0.686	32.7	9.556	3.444	0.422	0.622	21.333	90.8	9.6	0.500	0.650	45.2
s.d.	1.398	2.044	0.243	0.190	35.656	1.333	1.333	0.323	0.197	19.5	7.068	7.09	0.222	0.217	44.87

	HEE2					HEE3					HEE4				
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	38	4	0.684	0.622	6	32	5	0.375	0.613*	10	22	7	0.364	0.693*	32
M14_3	36	3	0.389	0.624*	6	30	2	0.200	0.186	6	22	4	0.545	0.519	12
M34_3	36	3	0.278	0.256	6	32	4	0.313	0.286	9	22	5	0.364	0.632*	18
M26_2	38	6	0.684	0.802	22	32	8	0.688	0.821	14	22	9	0.727	0.883	28
M31_2	36	8	0.722	0.741	42	28	6	0.786	0.730	32	22	8	0.909	0.831	24
M21_2	38	19	0.895	0.942	202	28	16	0.857	0.939	132	22	19	1.000	0.987	198
M17_3	36	4	0.333	0.654*	90	30	5	0.400	0.602*	39	20	5	0.200	0.758*	42
M8_2	36	12	0.556	0.829*	42	28	9	0.286	0.709*	42	22	7	0.545	0.597	34
M37_3	38	6	0.842	0.811	24	28	4	0.643	0.706	21	16	5	0.750	0.842*	30
M12_3	38	6	0.789	0.760	27	30	10	0.733	0.731	54	22	10	0.545	0.853*	66
Mean	37	7.1	0.617	0.704	46.7	29.8	6.9	0.528	0.632	35.9	21.2	7.9	0.595	0.760	48.4
s.d.	1.054	4.977	0.218	0.187	60.111	1.751	4.04	0.238	0.231	37.439	1.932	4.358	0.253	0.146	54.576

	HEE5					HEE6					HEE7				
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	74	5	0.514	0.683	8	132	8	0.455	0.661*	80	138	7	0.565	0.691	14
M14_3	76	4	0.263	0.355	9	132	5	0.318	0.387*	18	138	4	0.261	0.259	9
M34_3	76	4	0.316	0.3607*	9	132	4	0.212	0.234	9	138	3	0.290	0.282	6
M26_2	75	11	0.946	0.807	24	132	11	0.773	0.797*	20	135	12	0.881	0.807*	26
M31_2	64	8	0.344	0.7057*	40	134	17	0.642	0.738*	42	136	12	0.676	0.775*	42
M21_2	64	20	0.375	0.8427*	142	134	37	0.701	0.925*	150	136	33	0.721	0.932*	192
M17_3	74	9	0.270	0.5717*	141	110	8	0.291	0.494*	126	98	11	0.388	0.716*	111
M8_2	76	15	0.579	0.757	42	124	17	0.581	0.805*	44	134	13	0.537	0.798*	46
M37_3	76	12	0.921	0.8267*	99	127	8	0.825	0.771	162	112	7	0.750	0.772	27
M12_3	76	12	0.526	0.7627*	75	136	12	0.691	0.777*	54	136	14	0.824	0.854*	72
Mean	73.1	10	0.505	0.667	58.9	129.3	12.7	0.549	0.659	70.5	130.1	11.6	0.589	0.688	54.5
s.d.	4.864	5.121	0.251	0.181	52.647	7.631	9.615	0.216	0.217	56.508	13.699	8.435	0.219	0.230	57.989

	HEE8					HEE9				
L	*#G*	*#A*	H_o	H_e	*A.r*	*#G*	*#A*	H_o	H_e	*A.r*
M29_2	82	6	0.463	0.718*	34	68	5	0.441	0.710*	8
M14_3	82	5	0.488	0.450	12	68	4	0.500	0.462*	9
M34_3	82	3	0.195	0.182	6	68	3	0.206	0.190	6
M26_2	82	10	0.829	0.795	18	65	9	0.781	0.801	28
M31_2	82	10	0.537	0.583	42	70	9	0.600	0.746	42
M21_2	82	26	0.585	0.905*	168	70	18	0.657	0.809*	124
M17_3	68	6	0.588	0.691*	33	58	7	0.207	0.491*	54
M8_2	78	14	0.692	0.840*	52	66	10	0.455	0.760*	58
M37_3	72	7	0.944	0.785	33	56	10	0.857	0.820	129
M12_3	82	14	0.683	0.801*	57	70	10	0.657	0.807	48
Mean	79.2	10.1	0.601	0.675	45.5	65.9	8.5	0.536	0.660	50.6
s.d.	5.095	6.691	0.207	0.218	46.039	4.999	4.249	0.219	0.210	44.38

CHAPTER 6. GENERAL CONCLUSION

Landscape genetics as a multidisciplinary research approach can yield generous information about the study system. This information can be used in multiple different applications either individually in each discipline or with even more insight if used as integrative suit as it meant to be. In the research chapters of this dissertation there were some overarching conclusions, lessons learned and implications stated.

The simple neighborhood interpolation approach we use to correct for the data gaps in Landsat 7 ETM+ seems effective and more applicable for non-GIS specialist researchers than more specialized solutions. The pixel values that were used to replace the missing values are quite consistent with those expected because they come from the same scene and therefore the same date and conditions. However, there may be some altering of the exact boundaries between patches of values or feature edges. Users of Geographical Information Systems (GIS) and remotely sensed data should use image processing software cautiously when attempting to repair, or minimize artifacts within, remote sensory data either for geometric or radiometric corrections. Image processing techniques may appear to yield improvements in the images; however these may or may not be conservative enough with the original dataset's values and the geospatial properties of the area being used.

More sophisticated correction methods such as neighborhood similar pixel interpolator (NSPI) and regression trees procedures will more likely preserve these edge locations at the cost of substantial processing and computational time. The user of any of these methods must first weigh these aspects of the different techniques and decide which is most suitable for their goal.

As shown in chapter three, the landscape gradient paradigm and surface topology metrics are powerful approaches to study the influence of habitat heterogeneity on lepturine beetle species communities. The requirements of these species for complementary habitats and habitat quality determinants that have an inherently continuous range make it important to consider habitat as a continuous attribute to avoid oversimplification.

Before applying a landscape gradient approach and using surface metrics, it is important to consider environmental gradients relevant for the species of interest because habitat suitability is largely determined by availability of resources and conditions that support survival and reproduction of organisms. Considering the spatial scale at which the organism responds to different gradients in the landscape is very important prior to generating these gradients to be able to correctly interpret their biological and ecological roles. By integrating the biologically-important landscape gradients into a final surface of habitat quality, we were able to analyze the responses of many lepturine species simultaneously with each responding individually to multiple landscape gradients. It remains a possibility that some of the lepturine beetles in this study respond to the gradients used at a spatial scale different from that which we settled upon, weakening the perceived relationships.

This study shows that 3D surface metrology metrics are a valuable extension of the existing set of landscape metrics. More effort and attention should be directed towards this new landscape gradient paradigm. Future studies should examine how to interpret multiple metrics in concert (e.g., skewness + kurtosis) to better resolve different response trends.

In chapter four, the coalescence of the Canadian population to its most recent common ancestor (MRCA) and estimating its divergence has enabled us to test whether this coincided with the beginning of deglaciation. The Canadian population could have passed through a recent severe bottleneck represented by the steep decline in its effective population size. The divergence time of this population goes back to about 17,500 ybp which coincides with the last glacial retreat after LGM (24,000 – 16,000 ybp). This finding further supports the south to north post-glacial recolonization pattern and the validity of the southern refugium theory for this beetle. The banded longhorn beetle has survived periods of climatic change in the past mainly because of their population demography dynamics and their dispersal ability. These traits helped the more southern populations to survive the Quaternary in warmer refugia, then recolonize the north as new habitat became available. Patterns of phylogeographic distribution, differences in genetic diversity, and molecular evidence for demographic population expansion and contraction support this scenario. Our results pertain not only to how species and populations responded to pre-historic climatic changes, but may provide valuable context for predicted range and demographic shifts due to future climate change.

Landscape genetics approaches in chapter five seem to have a potential insight towards understanding the population genetic processes in the landscape scale. With the

integration of spatial, biological, and genetic data on the beetle under the study, we were able to tackle one of the big challenges in the field which is coupling non-spatial niche relationships, spatial patterns of environmental gradients and continuity of genetic structure in complex heterogeneous landscapes.

Surface metrics of connectivity is a valid powerful tool that considers both composition and configuration of the landscape and can potentially explain the variation in genetic dissimilarities and population structure. However, more effort should be applied to understand more about these metrics especially those measuring the configuration of the landscape, as this remains a challenge in this field. We suggest that direct methods of estimating gene flow should be tested with these metrics at various spatial scales, as this could be an indirect promising extension in studying dispersal of organisms when traditional techniques are hard to implement.

This landscape genetics study could have some important implications in different fields. It might have a potential towards explaining patterns of genetic variation between demes at finer spatial scales with insight to habitat requirements. This might enable testing different hypothesis about latent processes that could shape genetic structure such as population density, local dispersal and migration with overlapping generations in natural populations. The study could also help in initiating investigations on the temporal aspect of connectivity in the landscapes. The surface metrics approach might have more to offer in assessing the past and the future fragmentation predicted scenarios in the landscapes and relating the contemporary and historic genetic responses to landscape changes in time. This could also be further useful if direct methods of estimating gene

flow are coupled with this approach which could potentially give insight into identifying habitat source-sink dynamics in the landscape.

The power and insight of landscape genetics and surface metrics approach can also provide important lamina for applied conservation management. For example it can provide information on species movement in a spatial context, assessments of the spatial need for management measures, and evaluate the efficacy of existing management measures.

APPENDICES

Appendix A Landscape calculations in R and ArcGIS

Batch file for calculating insolation at 30m x 30m resolution for the 71 landscapes

```
##############################################################################
AreaSolarRadiation C:\centerspace\radiation\andij C:\centerspace\solars\sol_andij 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\cunni C:\centerspace\solars\sol_cunni 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\cups C:\centerspace\solars\sol_cups 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\dargton C:\centerspace\solars\sol_dargton 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\dpac C:\centerspace\solars\sol_dpac 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\finmemo C:\centerspace\solars\sol_finmemo 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\fpac2 C:\centerspace\solars\sol_fpac2 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\geyer C:\centerspace\solars\sol_geyer 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\harrold C:\centerspace\solars\sol_harrold 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\hee_1 C:\centerspace\solars\sol_hee_1 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_2 C:\centerspace\solars\sol_hee_2 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_3 C:\centerspace\solars\sol_hee_3 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_4 C:\centerspace\solars\sol_hee_4 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_5 C:\centerspace\solars\sol_hee_5 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_6 C:\centerspace\solars\sol_hee_6 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_7 C:\centerspace\solars\sol_hee_7 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_8 C:\centerspace\solars\sol_hee_8 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hee_9 C:\centerspace\solars\sol_hee_9 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\hughp C:\centerspace\solars\sol_hughp 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\jeffr C:\centerspace\solars\sol_jeffr 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\jimbrown C:\centerspace\solars\sol_jimbrown 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY
0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\jimdicks C:\centerspace\solars\sol_jimdicks 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY
0.3 0.5
```

```
AreaSolarRadiation C:\centerspace\radiation\jimdroste C:\centerspace\solars\sol_jimdroste 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY
0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\jimspence C:\centerspace\solars\sol_jimspence 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY
0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\kenny C:\centerspace\solars\sol_kenny 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\landon C:\centerspace\solars\sol_landon 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\lewal C:\centerspace\solars\sol_lewal 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\marklaf C:\centerspace\solars\sol_marklaf 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\martal C:\centerspace\solars\sol_martal 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\mccormic C:\centerspace\solars\sol_mccormic 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY
0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\miked C:\centerspace\solars\sol_miked 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\miller C:\centerspace\solars\sol_miller 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\nelson C:\centerspace\solars\sol_nelson 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\nepac C:\centerspace\solars\sol_nepac 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\ppac C:\centerspace\solars\sol_ppac 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\ricks C:\centerspace\solars\sol_ricks 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\ritab C:\centerspace\solars\sol_ritab 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\rossb C:\centerspace\solars\sol_rossb 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\royw1 C:\centerspace\solars\sol_royw1 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\sepac C:\centerspace\solars\sol_sepac 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\sipac C:\centerspace\solars\sol_sipac 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\stevens C:\centerspace\solars\sol_stevens 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
AreaSolarRadiation C:\centerspace\radiation\stout C:\centerspace\solars\sol_stout 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\stuntz C:\centerspace\solars\sol_stuntz 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3
0.5
```

```
AreaSolarRadiation C:\centerspace\radiation\tpac1 C:\centerspace\solars\sol_tpac1 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw295 C:\centerspace\solars\sol_uw295 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw365 C:\centerspace\solars\sol_uw365 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw366 C:\centerspace\solars\sol_uw366 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw456 C:\centerspace\solars\sol_uw456 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw459 C:\centerspace\solars\sol_uw459 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw464 C:\centerspace\solars\sol_uw464 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw561 C:\centerspace\solars\sol_uw561 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw580 C:\centerspace\solars\sol_uw580 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw654 C:\centerspace\solars\sol_uw654 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw691 C:\centerspace\solars\sol_uw691 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw720 C:\centerspace\solars\sol_uw720 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw763 C:\centerspace\solars\sol_uw763 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw790 C:\centerspace\solars\sol_uw790 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw793 C:\centerspace\solars\sol_uw793 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw821 C:\centerspace\solars\sol_uw821 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw831 C:\centerspace\solars\sol_uw831 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw844 C:\centerspace\solars\sol_uw844 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw845 C:\centerspace\solars\sol_uw845 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw856 C:\centerspace\solars\sol_uw856 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw865 C:\centerspace\solars\sol_uw865 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw869 C:\centerspace\solars\sol_uw869 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw896 C:\centerspace\solars\sol_uw896 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw920 C:\centerspace\solars\sol_uw920 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\uw960 C:\centerspace\solars\sol_uw960 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\vermj C:\centerspace\solars\sol_vermj 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
AreaSolarRadiation C:\centerspace\radiation\wabab C:\centerspace\solars\sol_wabab 45 200 'MultiDays 2008 166 243' 14 0.5 FALSE 1 FROM_DEM 32 8 8 UNIFORM_SKY 0.3 0.5
```

Batch file for converting insolation raters to TIF format.

```
################################################

CopyRaster C:\centerspace\solars\sol_andij C:\centerspace\solars\sol_andij.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_cunni C:\centerspace\solars\sol_cunni.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_cups C:\centerspace\solars\sol_cups.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_dargton C:\centerspace\solars\sol_dargton.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_dpac C:\centerspace\solars\sol_dpac.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_fpac2 C:\centerspace\solars\sol_fpac2.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_geyer C:\centerspace\solars\sol_geyer.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_harrold C:\centerspace\solars\sol_harrold.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_1 C:\centerspace\solars\sol_hee_1.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_2 C:\centerspace\solars\sol_hee_2.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_3 C:\centerspace\solars\sol_hee_3.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_4 C:\centerspace\solars\sol_hee_4.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_5 C:\centerspace\solars\sol_hee_5.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_6 C:\centerspace\solars\sol_hee_6.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_7 C:\centerspace\solars\sol_hee_7.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_8 C:\centerspace\solars\sol_hee_8.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hee_9 C:\centerspace\solars\sol_hee_9.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_hughp C:\centerspace\solars\sol_hughp.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_jeffr C:\centerspace\solars\sol_jeffr.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_jimbrown C:\centerspace\solars\sol_jimbrown.tif # 0 0 NONE
NONE 32_BIT_FLOAT
```

```
CopyRaster C:\centerspace\solars\sol_jimdicks C:\centerspace\solars\sol_jimdicks.tif # 0 0 NONE
NONE 32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_jimdroste C:\centerspace\solars\sol_jimdroste.tif # 0 0 NONE
NONE 32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_jimspence C:\centerspace\solars\sol_jimspence.tif # 0 0 NONE
NONE 32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_kenny C:\centerspace\solars\sol_kenny.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_landon C:\centerspace\solars\sol_landon.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_lewal C:\centerspace\solars\sol_lewal.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_marklaf C:\centerspace\solars\sol_marklaf.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_martal C:\centerspace\solars\sol_martal.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_mccormic C:\centerspace\solars\sol_mccormic.tif # 0 0 NONE
NONE 32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_miked C:\centerspace\solars\sol_miked.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_miller C:\centerspace\solars\sol_miller.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_nelson C:\centerspace\solars\sol_nelson.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_nepac C:\centerspace\solars\sol_nepac.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_ppac C:\centerspace\solars\sol_ppac.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_ricks C:\centerspace\solars\sol_ricks.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_ritab C:\centerspace\solars\sol_ritab.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_rossb C:\centerspace\solars\sol_rossb.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_royw1 C:\centerspace\solars\sol_royw1.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_sepac C:\centerspace\solars\sol_sepac.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_sipac C:\centerspace\solars\sol_sipac.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_stevens C:\centerspace\solars\sol_stevens.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_stout C:\centerspace\solars\sol_stout.tif # 0 0 NONE NONE
32_BIT_FLOAT
```

```
CopyRaster C:\centerspace\solars\sol_stuntz C:\centerspace\solars\sol_stuntz.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_tpac1 C:\centerspace\solars\sol_tpac1.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw295 C:\centerspace\solars\sol_uw295.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw365 C:\centerspace\solars\sol_uw365.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw366 C:\centerspace\solars\sol_uw366.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw456 C:\centerspace\solars\sol_uw456.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw459 C:\centerspace\solars\sol_uw459.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw464 C:\centerspace\solars\sol_uw464.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw561 C:\centerspace\solars\sol_uw561.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw580 C:\centerspace\solars\sol_uw580.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw654 C:\centerspace\solars\sol_uw654.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw691 C:\centerspace\solars\sol_uw691.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw720 C:\centerspace\solars\sol_uw720.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw763 C:\centerspace\solars\sol_uw763.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw790 C:\centerspace\solars\sol_uw790.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw793 C:\centerspace\solars\sol_uw793.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw821 C:\centerspace\solars\sol_uw821.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw831 C:\centerspace\solars\sol_uw831.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw844 C:\centerspace\solars\sol_uw844.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw845 C:\centerspace\solars\sol_uw845.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw856 C:\centerspace\solars\sol_uw856.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw865 C:\centerspace\solars\sol_uw865.tif # 0 0 NONE NONE
32_BIT_FLOAT
```

```
CopyRaster C:\centerspace\solars\sol_uw869 C:\centerspace\solars\sol_uw869.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw896 C:\centerspace\solars\sol_uw896.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw920 C:\centerspace\solars\sol_uw920.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_uw960 C:\centerspace\solars\sol_uw960.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_vermj C:\centerspace\solars\sol_vermj.tif # 0 0 NONE NONE
32_BIT_FLOAT
CopyRaster C:\centerspace\solars\sol_wabab C:\centerspace\solars\sol_wabab.tif # 0 0 NONE NONE
32_BIT_FLOAT
```

R script for calculating mean and standard deviation of insolation for the 71 landscapes:

```
#############################################################################################
#### install needed packages
library(spatstat)
library(maptools)
library(raster)
library(gpclib)
library(rgdal)
library(SDMTools)

## summarizing the insolation data measured in WH/m2 units from ArcGIS
#-----------------------------------------------------------------------

##start the for loop for (cellStat)

name <- read.csv("solar_names.csv",header=1, sep=',')
ID <- name$Id
plot.name <- name$NAME
r <- list()
results_solar <- matrix(NA, nrow=length(ID), ncol=2)
colnames(results_solar)=c('mean','sd')

for (i in 1:length(ID)) {
     names<- paste(plot.name[i],".tif",sep="")
     r[[i]] <- raster(names)
     result.mean <-cellStats(r[[i]],stat='mean')
     result.sd <-cellStats(r[[i]],stat='sd')
     comb<-cbind(result.mean, result.sd)
     results_solar[i,]<-comb
        }

############################
#read the output as table #
############################

write.table(results, file = "results_solar", append = FALSE, quote = TRUE, sep = " ",
            eol = "\n", na = "NA", dec = ".", row.names = TRUE,
            col.names = TRUE, qmethod = c("escape", "double"))
```

R script for calculating percent forest in a moving window (2.1km):

```
################################################################################

#####################
# Loading libraries #
#####################
library(raster)
library(rgdal)
library(SDMTools)
library(RSAGA)

# Calculating percent forest in a moving window (2.1km):
#-------------------------------------------------------
F<-raster('forest_300.tif')
plot(F)
F
F2 <- focal(F,na.rm=T, w=matrix(1/49,nrow=7,ncol=7))
F2
plot(F2)
rf <- writeRaster(F2, filename="forest_wind2.1.tif", format="GTiff", overwrite=F)
```

R script for Calculating splitting index in a moving window (2.1km):

```
################################################################################
#####################
# Loading libraries #
#####################
library(raster)
library(rgdal)
library(SDMTools)
library(RSAGA)

# Calculating splitting index in a moving window (2.1km):
#-------------------------------------------------------
# Function to calculate splitting index (as implemented in SDMTools)
#_____________________________________________________________________

sIND<-function (mat, cellsize = 1, bkgd = NA, latlon = FALSE)
        {
    aggregation.index = function(a, g) {
        n = trunc(sqrt(a))
        m = a - n^2
        if (m == 0)
            maxg = 2 * n * (n - 1)
        if (m <= n)
            maxg = 2 * n * (n - 1) + 2 * m - 1
        if (m > n)
            maxg = 2 * n * (n - 1) + 2 * m - 2
        minp = rep(0, length(m))
        for (ii in 1:length(m)) {
            if (m[ii] == 0)
                minp[ii] = 4 * n[ii]
            if (n[ii]^2 < a[ii] & a[ii] <= n[ii] * (1 + n[ii]))
                minp[ii] = 4 * n[ii] + 2
            if (a[ii] > n[ii] * (1 + n[ii]))
                minp[ii] = 4 * n[ii] + 4
        }
        return((g/maxg) * 100)
    }
    shape.index = function(a, p) {
```

```
        n = trunc(sqrt(a))
        m = a - n^2
        minp = rep(0, length(m))
        for (ii in 1:length(m)) {
            if (m[ii] == 0)
                minp[ii] = 4 * n[ii]
            if (n[ii]^2 < a[ii] & a[ii] <= n[ii] * (1 + n[ii]))
                minp[ii] = 4 * n[ii] + 2
            if (a[ii] > n[ii] * (1 + n[ii]))
                minp[ii] = 4 * n[ii] + 4
        }
        return(p/minp)
    }
    if (any(class(mat) %in% "RasterLayer"))
        mat = asc.from.raster(mat)
    if (any(class(mat) == "SpatialGridDataFrame"))
        mat = asc.from.sp(mat)
    mat = try(as.matrix(mat))
    if (!is.matrix(mat))
        stop("objects must be a matrix")
    classes = as.numeric(na.omit(unique(as.vector(mat))))
    classes = classes[order(classes)]
    if (!is.na(bkgd))
        classes = classes[-which(classes == bkgd)]
    out = NULL
    for (cl in classes) {
        mat2 = mat
        mat2 = mat * 0
        mat2[which(mat == cl)] = 1
        out.patch = PatchStat(ConnCompLabel(mat2), cellsize = cellsize,
            latlon = latlon)
        rm(mat2)
        L.cell = sum(out.patch$n.cell)
        L.area = sum(out.patch$area)
        if (0 %in% out.patch$patchID)
            out.patch = out.patch[-which(out.patch$patchID ==
                0), ]
        tout = list(class = cl)
        tout$splitting.index = L.area/sum(out.patch$area^2)

    }
    return(tout$splitting.index)
}

splitting <- focal.function("forest_asc_300.txt",is.pixel.radius=T,
radius=7,search.mode="square",mw.to.vector=T,mw.na.rm=T, fun=sIND)

#############################################################################################
# the output is an ascii (.asc) file which is imported in ArcGIS 9.2 and exported as GRID file #
#############################################################################################




#################################################################
# Clipping landscape using CRW
# 1. CRW simulation
# 2. Extract ellipse information from success CRWs, then make equation
# 3. Clipping landscape using the ellipse
#################################################################


library(spatstat)   # for random walker simulation
library(maptools)   # read shape files
library(raster)     # read raster files and clipping
library(PBSmapping) # make polygons
```

```
################################################################
#  1. CRW simulation
#   1-1. define 3 functions for CRW simulation
#   1-2. CRW simulation (takes about 2 hours), then draw ellipse (potential success paths)
################################################################

################################################################
# 1-1.  define 3 functions for CRW simulation
################################################################

# calculate max minor axis distance for success random walker
# x0,y0 are source point, x1,y1 are target point
sminorP2L <- function (x0, y0, x1, y1, xr, yr) {

          vx0 <- rep(x0,length(xr))
          vy0 <- rep(y0,length(xr))
          vx1 <- rep(x1,length(xr))
          vy1 <- rep(y1,length(xr))

          a = (vy1-vy0)/(vx1-vx0)
          b = vy0 - a*vx0
          if (x1==x0) distance = abs(xr - vx0)
          if (y1==y0) distance = abs(yr- vy0)
          if (x1!=x0 & y1!=y0) distance = abs(a*xr -yr + b)/ sqrt(a^2 + 1)

          return (max(distance))
          }

# calculate max major axis distance for success random walker
smajorP2L <- function (x0, y0, x1, y1, xr, yr) {

          vx0 <- rep(x0,length(xr))
          vy0 <- rep(y0,length(xr))
          vx1 <- rep(x1,length(xr))
          vy1 <- rep(y1,length(xr))

          a = -(vx1-vx0)/(vy1-vy0)
          b = (vy0+vy1)/2 - a*(vx0+vx1)/2
          if (x1==x0) distance =  abs(yr - (vy0+vy1)/2)
          if (y1==y0) distance =  abs(xr - (vx0+vx1)/2)
          if (x1!=x0 & y1!=y0) distance = abs(a*xr -yr + b)/ sqrt(a^2 + 1)

          return (max(distance))
          }


# isOnlineseg function find which hitting point is the first one
#  because crossing.psp can't tell this information (just show hitting points)
#  find that random walker's successful path from start to hitting point on target area

isOnlineseg  <- function (x0, y0, x1,y1,hit.px, hit.py) {

                hit.steps <- rep(NA, length(hit.px))

                  for( hit.n in 1:length(hit.px)) {

                  hit.x   <-  rep(hit.px[hit.n], stepCount)
                  hit.y   <-  rep(hit.py[hit.n], stepCount)

                  hit.steps[hit.n] <-
                  min (which((y1- y0)/(x1-x0)*(hit.x - x1) + (y1 - hit.y) < 10^(-5)
                                    & ifelse(x1>x0, x0 < hit.x & hit.x < x1, x1 < hit.x & hit.x <
x0)
                                    & ifelse(y1>y0, y0 < hit.y & hit.y < y1, y1 < hit.y & hit.y <
y0)
                                    )
```

```
                    )
                }

                # first line segment hitting target and hitting point vector index in hit.point
                c(min(hit.steps), which(hit.steps==min(hit.steps)))

                }



##############################################################################
#  1-2. CRW simulation based on different distance between source and target
##############################################################################

xs <- 0              # source point x location
ys <- 0              # source point y location

xt <- rep(0, length(yt))                     # target x location
yt <- c(250,500,750,1000,1500,2000,3000,4000) # target y location


A <-rep(NA, length(yt))        # Eliipse major radius axis
B <-rep(NA, length(yt))        # Eliipse minor radius axis
Eangle <- rep(NA, length(yt))  # Ellipse angle

success.perc <- rep(NA,length(yt)) # store success.percentage information for each target


stepCount <- 1000    # maximum step count
n.walker  <- 100000  # number of random walker

k <- 0.85            # degree of correlation between movement directions: highly correlated


par(mfrow=c(2,4))

for ( target in 1: length(yt)) {
 print(yt[target])

 set.seed(100000)

 n.success <- 0

 frame.e <- rep(NA,n.walker)
 s.ellipse <- data.frame("Ex.c"=frame.e, "Ey.c"=frame.e, "E.angle"=frame.e,
                         "minor.r"=frame.e, "major.r"=frame.e) # success.ellipse

 plot(0,0,xlim=c(-500,500), ylim=c(-200,max(yt)+1000), asp=1,xlab="x", ylab="y")
 text(0, max(yt)+1000, paste("distance = ",yt[target], " m"), cex=1.5)

  w.extent <- c(-500000,500000,-500000,500000)

  # strart point: circle polygon
  start.C <-  as.psp(disc(2, c(xs,ys)), w.extent)
  start.P  <- runifpointOnLines(n.walker, start.C)

  # target point: circle polygon
  target.C <- as.psp(disc(2, c(xt[target],yt[target])), w.extent  )

  plot(start.C, add=T)
  plot(target.C, add=T)

  for (walker in 1:n.walker) {
```

```
    turningAngles <- round(rnorm(stepCount, mean=0, sd=(1-k)*2*pi),2) # sd means direction of
animal head
    turningAngles[1] <- runif(1, 0, 2*pi) # make sure that first step goes in random direction

    stepLength <- round(rgamma(stepCount, shape=2, scale=25),2) # each step has step length

    theta <- cumsum(turningAngles) # theta are now turning angles relative to north

    dx <- stepLength * sin(theta)
    dy <- stepLength * cos(theta)

    x <- c(start.P$x[walker], start.P$x[walker] + cumsum(dx))  # now step x,y from start and end
    y <- c(start.P$y[walker], start.P$y[walker] + cumsum(dy))

    From <- as.ppp(cbind(x[1:stepCount],y[1:stepCount]), w.extent )
    To <-   as.ppp(cbind(x[2:(stepCount+1)],y[2:(stepCount+1)]), w.extent )
    r.path <- as.psp(from=From, to=To)  # r.path is segement of line xy: psp objet

    hit.point <- crossing.psp (r.path, target.C) # to find out hitting point

    # when random walker hit the target

    if(hit.point$n > 0) {

        hit.step1 <- isOnlineseg(r.path$ends$x0, r.path$ends$y0, r.path$ends$x1, r.path$ends$y1,
                  hit.point$x, hit.point$y) # the first hitting point: hit.step1

        success.x <-c(r.path$ends$x0[1:hit.step1[1]],hit.point$x[hit.step1[2]]) # success.x means
the r.path from start to hitting step
        success.y <-c(r.path$ends$y0[1:hit.step1[1]],hit.point$y[hit.step1[2]])

        lines(success.x, success.y, col=sample(rainbow(100),1)) # success r.path
        points(success.x[1],success.y[1])                       # the first step of success r.path
        points(success.x[length(success.x)],success.y[length(success.y)]) # the end step of
success r.path

        # memorize ellipse information from the success r.path
        s.ellipse[walker,] <- c( (success.x[1]+success.x[length(success.x)])/2,

                                 (success.y[1]+success.y[length(success.y)])/2,

                                 atan((success.y[length(success.y)]-success.y[1])
                                       /(success.x[length(success.x)]-success.x[1])),

                                 sminorP2L (success.x[1], success.y[1],
                                            success.x[length(success.x)],
success.y[length(success.y)],
                                            success.x,success.y),
                                 smajorP2L (success.x[1], success.y[1],
                                            success.x[length(success.x)],
success.y[length(success.y)],
                                            success.x,success.y)
                                )

        n.success <- n.success + 1

     } # the end for success hitting random walker

  }    # the end for the each random walker simulation


#drawing ellipse for success r.paths
Xc <- mean(s.ellipse$Ex.c, na.rm=T)             # mean center point of ellipse
Yc <- mean(s.ellipse$Ey.c, na.rm=T)             # mean center point of ellipse
A[target] <- mean(s.ellipse$major.r, na.rm=T)    # A is the major radius
B[target] <- mean(s.ellipse$minor.r, na.rm=T)    # B is the major radius
```

```
Eangle[target] <- atan((yt[target]-Yc)/(xt[target]-Xc))# Eliipse angle

t <- seq(0,2*pi,0.1)
Xe <- Xc + A[target]*cos(t)*cos(Eangle[target]) - B[target]*sin(t)*sin(Eangle[target])
Ye <- Yc + A[target]*cos(t)*sin(Eangle[target]) + B[target]*sin(t)*cos(Eangle[target])

lines(Xe, Ye, lty=2, lwd=2)


success.perc[target] <- n.success / n.walker * 100

text(0, max(yt), paste("success =", round(success.perc[target],2), "%"), cex=1.5)

} # CRW simulation end.


##############################################################################
# 2. Extract ellipse information from success CRWs, then make equation
#  2-1. ploting the relationship between distance and ellipse raidus.
#  2-2. make equation from the relation
##############################################################################


##############################################################################
#  2-1. plotting the relationship between distance and ellipse raidus
##############################################################################
#distance <- dist.st[which(!is.na(A))]
distance <- yt

success.p <- success.perc[1:length(distance)]
plot(distance, success.p, ylab="Success rate of 100,000 random walker (%)")

major.r <- A/distance * 100
minor.r <- B/distance * 100

plot(distance, major.r,  ylim=c(0,max(major.r, minor.r)),
      pch=16,cex=1.5,
      ylab="Length of radius in proportion to the distance (%)",
      xlab="Distance between source and target")
points(distance, minor.r, pch=17,cex=1.5)
abline(h=50, lty=2)
legend("topright",
       legend=c("major radius (A)", "minor radius (B)"),
        pch=c(16,17))

##############################################################################
#  2-2. make equation from the relation
##############################################################################

# We assume minor and major radius converge to 40% and 60
model.minor <- lm(log(minor.r - 40) ~ distance )
y.minor<-exp(predict(model.minor,list(distance=1:10000)))+40
lines(1:10000, y.minor, lty=2)


model.major <- lm(log(major.r - 60) ~ distance )
y.minor<-exp(predict(model.major,list(distance=1:10000)))+60
lines(1:10000, y.minor, lty=3)

# store model equation

# minor radius  =  (exp(-0.001163*distance + 5.600736) + 40) * distance (between sourse and target)
# major radius  =  (exp(-0.001163*distance + 5.366365) + 60) * distance
##############################################################################
# 3. Clipping landscape with the ellipse
##############################################################################
```

```
setwd("C:/Hossam_surface/")

Indi <- raster("T_surface.tif")
plot(Indi)

spoints <- readShapeSpatial("sampling_points.shp")

## important !! id: 0 -> 70
id <- as.numeric(rownames(spoints@data))

plot(spoints, add=T, pch=1)

# creat distance matrix
dist.m <- as.matrix(dist(spoints@coords, method="euclidean", diag=T, upper=T))

for (i in 1:length(id)) {

  for (j in (i+1):length(id)) {

        Xc <- (spoints@coords[i,1] + spoints@coords[j,1])/2 #Ellipse center x
        Yc <- (spoints@coords[i,2] + spoints@coords[j,2])/2 #Ellipse center x

        Eangle  <- atan(
                  (spoints@coords[j,2]-spoints@coords[i,2])/(spoints@coords[j,1]-
spoints@coords[i,1])
                   ) # ellipse angle

        if(dist.m[i,j] < 4000) {

            A <- (exp(-0.001163*dist.m[i,j] + 5.366365) + 60)/100 * dist.m[i,j]
            B <- (exp(-0.001163*dist.m[i,j] + 5.600736) + 40)/100 * dist.m[i,j]
        }
        if(dist.m[i,j] >= 4000) {

            A <- 60/100 * dist.m[i,j]
            B <- 40/100 * dist.m[i,j]
        }

        radian <- seq(0,2*pi, length.out=360)
        Xe <- Xc + A*cos(radian)*cos(Eangle) - B*sin(radian)*sin(Eangle)
        Ye <- Yc + A*cos(radian)*sin(Eangle) + B*sin(radian)*cos(Eangle)
        Eline <- cbind(Xe, Ye)

        # make ellipse polygon from line by two steps
        Epolyset <- as.PolySet(data.frame(PID=rep(1,length(radian)),
                       SID=rep(1,length(radian)), POS=1:length(radian),
                       X= Eline[,"Xe"], Y= Eline[,"Ye"]),
                       projection=1)
        Epolygon <- PolySet2SpatialPolygons(Epolyset)
        plot(Epolygon, add=T)

        # Clipping Ellipse
        Emask   <- rasterize(Epolygon, Indi)
        Elands  <- mask(Indi, Emask)
        # Write files
        writeRaster(Elands, filename=paste("Elands_",i,"_",j,sep=""),format="GTiff",
overwrite=TRUE)

  }

}
```

R script for Calculating Landscape and Fragmentation Indices:

```
#############################################################################

#### Install required packages
#_____________________________

library(spatstat)
library(maptools)
library(raster)
library(gpclib)
library(rgdal)
library(SDMTools)


### Read names of landscape polygons
#___________________________________

name <- read.csv("I:/SMITH_GIS/landscapes.csv")
ID <- name$Id
plot.name <- name$NAME




## Generate the list for the loop and the results matrix (71*38)
#_____________________________________________________________

r <- list()
results <- matrix(NA, nrow=length(ID), ncol=38)



## Try the code for one landscape
#_______________________________

i=1
names<- paste(plot.name[i],".tif",sep="")
r[[i]] <- raster(names)
landscape <-as.matrix(r[[i]])
result <-ClassStat(landscape,cellsize=1,bkgd=NA,latlon=FALSE)## bkgd is the background value for
which statistics will not be calculated (could be NA or any other value)
values <- as.matrix(result)
if (i==1) colnames(results)<-names(result)
results[i,]<-values[2,]


## Start the for loop for Cell based metrics (ClassStat)
#_____________________________________________________

for (i in 1:length(ID)) {

        names<- paste(plot.name[i],".tif",sep="")
        r[[i]] <- raster(names)
        landscape <-as.matrix(r[[i]])
        result <-ClassStat(landscape,cellsize=1,bkgd=NA,latlon=FALSE)## bkgd is the background
value for which statistics will not be calculated (could be NA or any other value)
        values <- as.matrix(result)
        if (i==1) colnames(results)<-names(result)
        results[i,] <-values[2,]
        }

## Start the for loop for Patch based metrics (PatchStat) (71*12 matrix)
#_____________________________________________________________________
```

```
r <- list()
p_results <- matrix(NA, nrow=length(ID), ncol=12)

for (i in 1:length(ID)) {

        names<- paste(plot.name[i],".tif",sep="")
        r[[i]] <- raster(names)
        landscape <-as.matrix(r[[i]])
        P_result <-PatchStat(landscape,cellsize=30,latlon=FALSE)
        values <- as.matrix(P_result)
        if (i==1) colnames(p_results)<-names(P_result)
        p_results[i,] <-values[2,]
        }

## Write the results to a comma separated (.csv) file
#_____________________________________________________

analysis.output <- write.csv (results, file = "frag_results", append = FALSE, quote = TRUE, sep =
",",
            eol = "\r\n", na = "NA", dec = ".", row.names = TRUE,
            col.names = TRUE, qmethod = c("escape", "double"))


analysis.output <- write.csv (p_results, file = "P_frag_results", append = FALSE, quote = TRUE,
sep = ",",
            eol = "\r\n", na = "NA", dec = ".", row.names = TRUE,
            col.names = TRUE, qmethod = c("escape", "double"))
```

R script for Statistical Analysis of BC distance and surface metrics:

```
#############################################################################

library(stats)

library(vegan)

library(outliers)

full_mat<-read.table(file='full_mat.csv',header=T, sep=',')

names(full_mat)

mat_stand<- decostand(full_mat, "standardize")

names(mat_stand)

sp<-read.table(file='species.csv', header=T, sep=',')

names(sp)

all_mat<- cbind(sp,mat_stand)

names(all_mat)

attach(all_mat)

#--------------------------------------

# for color coding coefficients

#--------------------------------------

coeff_T <- matrix(data=NA, nrow=21, ncol=17)
```

```
colnames(coeff_T) <-
c("Analeptura.lineola","Bellamira.scalaris","Brachyleptura.champlaini","Brachyleptura.rubrica","Ga
urotes.cyanipennis","Metacmaeops.vittata","Necydalis.mellita","Stenelytrana.emarginata","Strangale
pta.abbreviata","Strangalia.bicolor",
                                  "Strangalia.solitaria",
"Strangalia.luteicornis","Strophiona.nitens",
"Trachysida.mutabilis","Typocerus.deceptus","Typocerus.v..velutinus","community")
rownames(coeff_T) <- c("Sa", "S10z", "Ssk", "Sku", "Sdr", "Sbi", "Std", "Stdi", "Sfd", "Srwi",
"Geo_dist", "Sa:Geo_dist", "S10z:Geo_dist", "Ssk:Geo_dist", "Sku:Geo_dist", "Sdr:Geo_dist",
"Sbi:Geo_dist", "Std:Geo_dist", "Stdi:Geo_dist", "Sfd:Geo_dist", "Srwi:Geo_dist")
#---------------------------------------
# for color coding P values
#---------------------------------------
P.table <- matrix(data=NA, nrow=21, ncol=17)
colnames(P.table) <-
c("Analeptura.lineola","Bellamira.scalaris","Brachyleptura.champlaini","Brachyleptura.rubrica","Ga
urotes.cyanipennis","Metacmaeops.vittata","Necydalis.mellita","Stenelytrana.emarginata","Strangale
pta.abbreviata","Strangalia.bicolor",
                                  "Strangalia.solitaria",
"Strangalia.luteicornis","Strophiona.nitens",
"Trachysida.mutabilis","Typocerus.deceptus","Typocerus.v..velutinus","community")
rownames(P.table) <- c("Sa", "S10z", "Ssk", "Sku", "Sdr", "Sbi", "Std", "Stdi", "Sfd", "Srwi",
"Geo_dist", "Sa:Geo_dist", "S10z:Geo_dist", "Ssk:Geo_dist", "Sku:Geo_dist", "Sdr:Geo_dist",
"Sbi:Geo_dist", "Std:Geo_dist", "Stdi:Geo_dist", "Sfd:Geo_dist", "Srwi:Geo_dist")
#---------------------
# Analysis in loop
#---------------------
modata <- list()
for (i in 1:17) {
 print (i)
 modata[[paste("sp",i,sep="")]][["mod"]] <- lm (all_mat[,i] ~ (Sa+ S10z+ Ssk+ Sku+ Sdr+ Sbi+ Std+
Stdi+ Sfd+ Srwi) * Geo_dist, data=all_mat)
 modata[[paste("sp",i,sep="")]][["modsum"]] <- summary(modata[[paste("sp",i,sep="")]][["mod"]])
```

```
 modata[[paste("sp",i,sep="")]][["bestmod"]] <-
step(modata[[paste("sp",i,sep="")]][["mod"]],trace=1, scale=0,steps = 1000, direction="backward",
k=2)
 modata[[paste("sp",i,sep="")]][["bestmodsum"]] <-
summary(modata[[paste("sp",i,sep="")]][["bestmod"]])
 stepmod <- modata[[paste("sp",i,sep="")]][["bestmodsum"]]
coefvar <- names(stepmod$coefficients[,1])[-1]
for (name in coefvar) {
coeff_T[name, i] <- stepmod$coefficients[name,1]
                        }
p.val <- names(stepmod$coefficients[,4])[-1]
for (name in p.val) {
P.table[name, i] <- stepmod$coefficients[name,4]
                                }
                        }
coef.table<- round(coeff_T,4)
write.table(coef.table, file = "coef.table", sep = ",")
write.table(P.table, file = "P.table", sep = ",")
```

Appendix B Perl script to build the consensus and merging the forward and reverse sequences for the mt.DNA COI sequences (mergeFR.perl).

```
#!/usr/bin/perl
use strict;
use warnings;
=pod
Program to attempt to merge the forward and reverse reads pair-wise for a set
of clones.
It will rip a forward and reverse fasta file set to produce paired-read fasta
files for each clone
and optionally do the same for quality value files. Then run phrap on each
fasta file. Then clean up.
=cut
use Bio::Perl;
use Bio::SeqIO;
use Getopt::Std;
use vars qw($opt_f $opt_Q $opt_r $opt_v);
if (! (getopts('vQf:r:') && $opt_f && $opt_r)) {
        die "Usage: $0 [-Qv] -f filename -r filename
         -f fasta format input filename one.
         -r fasta format input filename two.
         -Q also input qual file.
         -v verbose output mode\n";
}

#read in all seqs and quals from their respective fasta files
#First the forward reads
my %Fseqs = %{seqhash($opt_f,'fasta')};

#Then the reverse reads
```

```
my %Rseqs = %{seqhash($opt_r,'fasta')};

#Then the quality values, if any.
my (%Fquals,%Rquals);
if ($opt_Q) {
        %Fquals = %{seqhash($opt_f.".qual",'qual')};
        %Rquals = %{seqhash($opt_r.".qual",'qual')};
}

#Create a list of all uniq keys
my %keylist;
foreach (keys(%Fseqs),keys(%Rseqs)) {
        $keylist{$_}++;
}

#Create objects to do cumulative output to
my ($infilename) = $opt_f =~ m%(?:.*/)?(.*)%;
my ($basefilename) = $infilename =~ /^(.*?)\./;
$basefilename   =~s/(.*)_.*/$1/;    #strip primer field off basefilename
my $mergeoutfilename = "$basefilename.merged.fasta";
my $merged_out = Bio::SeqIO->new( -format => 'fasta' , -file =>
">$mergeoutfilename");
my $q_merged_out = Bio::SeqIO->new( -format => 'qual' , -file =>
">$mergeoutfilename.qual") if $opt_Q;
my $singlets_out = Bio::SeqIO->new( -format => 'fasta' , -file =>
">$basefilename.singlets.fasta");
my $q_singlets_out = Bio::SeqIO->new( -format => 'qual' , -file =>
">$basefilename.singlets.fasta.qual") if $opt_Q;
#Output a fasta format file containing the forward and/or reverse read for
each clone
#Then phrap, then collate into single fasta file
foreach (sort keys %keylist) {
        print "phrapping $_...\n" if $opt_v;
        my $out = Bio::SeqIO->new( -format => 'fasta' , -file => ">$_.fasta");
        if (exists $Fseqs{$_}) { $out->write_seq($Fseqs{$_}); }
```

```
        if (exists $Rseqs{$_}) { $out->write_seq($Rseqs{$_}); }
        if ($opt_Q) {
                my $Qout = Bio::SeqIO->new( -format => 'qual' , -file =>
">$_.fasta.qual");
                if (exists $Fquals{$_}) { $Qout->write_seq($Fquals{$_}); }
                if (exists $Rquals{$_}) { $Qout->write_seq($Rquals{$_}); }
        }

# run phrap to attempt to assemble the F/R reads
# my $info = `phrap  $_.fasta `;
    my $info = `phrap  $_.fasta -retain_duplicates 2> /dev/null`;

        #get rid of extraneous files created by phrap
        foreach my $suffix (qw(singlets log problems problems.qual)) {unlink
"$_.fasta.$suffix";}
        if ((not -e "$_.fasta.contigs") or (-z "$_.fasta.contigs")) { #phrap
failed to merge the two
                #un-merged reads go into the singlets files:
                if (exists $Fseqs{$_}) { $singlets_out-
>write_seq($Fseqs{$_}); }
                if (exists $Rseqs{$_}) { $singlets_out-
>write_seq($Rseqs{$_}); }
                if ($opt_Q) {
                        if (exists $Fquals{$_}) { $q_singlets_out-
>write_seq($Fquals{$_}); }
                        if (exists $Rquals{$_}) { $q_singlets_out-
>write_seq($Rquals{$_}); }
                }
        } else {        #phrap did merge the files, so read in the contig and
write it into the merged file
                my $in_contig = read_sequence("$_.fasta.contigs",'fasta');
                my ($new_id) = $in_contig->display_id =~ /^(.*?)\./;
#Ohm1_1_A02.fasta.Contig1 : want only Ohm1_1_A02
                $in_contig->display_id($new_id);
                my $Qin_contig = read_sequence("$_.fasta.contigs.qual", 'qual')
```

```
if $opt_Q;
                $Qin_contig->display_id($new_id) if $opt_Q;
                $merged_out->write_seq($in_contig);
                $q_merged_out->write_seq($Qin_contig) if $opt_Q;
        }
        unlink "$_.fasta.contigs";
        unlink "$_.fasta.contigs.qual" if $opt_Q;
        unlink "$_.fasta";
        unlink "$_.fasta.qual" if $opt_Q;
}
sub seqhash {
        my ($fname,$seqtype) = @_;
        my %seqhash;
        foreach (read_all_sequences($fname,$seqtype)) {
                #Try to make the index just the clone name (otherwise just use
whole display_id):
                #>Library_PlateName_Well_Primer
                #is the presumed format of display_id
        #But want to allow:
        #>Library_PlateName_Well.Primer_etc
        #as well
                my $key;
                if ( $_->display_id =~ /^[^_]+_[^_]+_[^_.]+/ ) {
                        ($key) = $_->display_id =~ /([^_]+_[^_]+_[^_.]+)/;
                } elsif ( $_->display_id =~ /^[^_]+_[^_]+/ ) {
                        ($key) = $_->display_id =~ /^([^_]+)/;
                } else {($key) = $_->display_id;}

                $seqhash{$key} = $_;
        }
        return \%seqhash;
}
```

Appendix C Landscape genetics codes

The "ssr3.pl" script for the microsatellites.

```
####################################

#!/usr/bin/perl
#Program to find Simple Sequence Repeats (SSRs also known as
"microsatellites").
#Reads a fasta formatted sequence file into a hash
#Each sequence is searched for short nucleotide repeats
use strict;
use warnings;
use Getopt::Std;
use vars qw($opt_h $opt_s $opt_r $opt_m $opt_x);

my $usage_string =  "Usage: $0 [-hs] [-r minimum_repeat_number]"
                 .  " [-m minimum_repeat_length] [-x maximum_repeat_length]"
                 .  " <FASTA_SEQUENCE_FILE>\n";

if (! getopts('hsr:m:x:')) {
        die $usage_string;
}
help() if $opt_h;
if (! $ARGV[0]) {
        die $usage_string;
}

my $min_repeat_num      = $opt_r || 5;
my $min_repeat_unit_len = $opt_m || 2;
my $max_repeat_unit_len = $opt_x || 10;
$min_repeat_num--; #Decrement so $min_repeat_num of repeats will be found
using regex capture and backreference.

my %sequences = %{get_fasta()};     #Now put into hash. Names are keys
#print "Number of sequences: $#sequences\n";
print "Name\tSeq Len\tRange\t# of repetitions of sub unit\tSub unit\n";
foreach my $x (sort(keys(%sequences))) {
    my $flag = 0;
    while ( $sequences{$x}
```

```
            =~ m/    #Capture each ssr sub-unit within tolerance
                     #Note "?" for lazy capture. Ensures "AC" is
                     #the repeat unit instead of "ACAC" for example

            ([ACGT]{$min_repeat_unit_len,$max_repeat_unit_len}?)
             \1{$min_repeat_num,}    #(Backref) find minimum number of repeats
of sub-unit
            /gix
          ) {
        my $repeat_unit     =   $1;
        my $start_of_ssr    =   $-[0]+1;
        my $end_of_ssr      =   $+[0];
        my $ssr             =   $&;
        my $ssr_length = length($ssr)/length($repeat_unit);
=pod
        Don't print ssr if repeat unit can be decomposed into sub-repeats
        That is, with $minimum_repeat_unit set to "2" the regex capture above
        will avoid "A" as a repeat unit, but not "AA" So a ssr "AAAAAAAA"
would
        be found even though we don't want it. This check closes that loophole.
=cut
        unless ($repeat_unit =~
            m/^             #Consider entire repeat unit from beginning (to
end--below)
                ([ACGT]+)   #Capture putative sub-repeat unit
                \1+         #Look for at least one extra copy of sub-repeat
unit using back reference
            $/ix) {         #End anchor--whole string must be decomposable.
Skip "AAA" but print "AAC", eg.

            print $x,"\t",length $sequences{$x};
            print "\t$start_of_ssr-$end_of_ssr";
            print "\t$ssr_length of repeat";
            print "\t\"$repeat_unit\"";
            print "\t$ssr" if $opt_s;   #Generally don't want to print the
whole SSR-composed sequence.
            print "\n";
            $flag = 1;
        }
    }
    if ($flag ) { print "----------------------------\n"}
}

sub get_fasta {
    my $seq;
    my $seqname;
```

```
    my $gotOne=0;
    my %seqHash;
    while (<>) {
        chomp;
        if (/^>/) {#Start up new record
            if ($gotOne) {#But first save old one (if there is an old one.)
                storeSeq(\%seqHash,$seqname,$seq);
            }
            $gotOne++;
            $seq = '';
            ($seqname)= $_ =~ /^>(\S+)/ or die "Illegal sequence name \"$_\"";
        } else {
            $seq .= $_;
        }
    }
    storeSeq(\%seqHash,$seqname,$seq) if $gotOne;
    return \%seqHash;
}
sub storeSeq {  #0 reference to hash of sequences to be added to
        #1 new seqname
        #2 new sequence
    my ($seqHashRef,$seqname,$seq) = @_;
    $seq =~ s/\s//g;                #strip out any whitespace
    $seq = uc $seq;         #Uppercase all sequence bases
    if (exists ${$seqHashRef}{$seqname}) {
        warn "Duplicate sequence name \"$seqname\". Will discard the first
one!\n";
    }
    ${$seqHashRef}{$seqname}=$seq;
}
sub help {
die "Usage: $0 [-hs] [-r ordinal] [-m ordinal] [-x ordinal]
FASTA_SEQUENCE_FILE
    -h This help message
    -r repeat minimum. Minimum number of repeats
       of an SSR repeat unit to accept. Default is 5.
    -m mimimum repeat subunit length. E.g, \"2\"
       for \"AC\" or 3 for \"GGA\", etc. Default is 2.
    -x maximum repeat subunit length. Default is 10.
    -s Include whole SSR sequence in output.\n"
}
```

Geneland Analysis.

```
#################

library(Geneland)

coord<-read.csv("coords_ind.csv", header=T)
coord.ind<- coord[,3:4]
dim(coord.ind) # Dimensions should be 453 rows (individuals) x 2 columns (NE)

msat1<-read.csv("SSR10_454_geneland.csv", header=T)
msat1<- msat1[,-1]
nrow(msat1) # Number of rows should be 453 (individuals)
ncol(msat1) # Number of columns should be 20 (diploid individuals scored at 10
microsatellite markers)

plot(coord.ind, xlab="Eastings", ylab="Northings", asp=1) # plot geo-refrenced
individuals

msat1_format<-FormatGenotypes(msat1,2) # 2 indicates ploidy status ==2
(diploid)
geno1<-msat1_format$genotypes

allele.no1<-msat1_format$allele.numbers

pop.mbrship1<-read.csv("pop_454_assign.csv", header=T)# created a vector
numerical values corresponding to population membership

pop.mbrship1<- pop.mbrship1[,-1]

MCMC(coordinates=coord.ind, geno.dip.codom=geno1, varnpop=TRUE, npopmax=17,
spatial=TRUE, freq.model="Correlated", nit=100000, thinning=100,
path.mcmc="/Volumes/My Book/sites/geneland_MCMC2/")

PostProcessChain(coordinates=coord.ind,path.mcmc="/Volumes/My
Book/sites/geneland_MCMC2/", nxdom=100, nydom=100, burnin=20)

Plotnpop(path.mcmc="/Volumes/My Book/sites/geneland_MCMC2/", burnin=20,
printit=TRUE,
file="/Volumes/My Book/sites/geneland_MCMC2/No_Clusters1.pdf", format="pdf")

PosteriorMode(coordinates=coord.ind, path.mcmc="/Volumes/My
Book/sites/geneland_MCMC2/", printit=TRUE,
file="/Volumes/My Book/sites/geneland_MCMC2/map1.pdf", format="pdf")

PlotTessellation(coordinates=coord.ind, path.mcmc="/Volumes/My
Book/sites/geneland_MCMC2/", printit=TRUE)
```

GAMM with surface metrics.

```
#########################

library(stats)
library(vegan)
library(mgcv)
library(AED)
library(MASS)

full_mat<-read.table(file='TV_gene_flow.csv',header=T, sep=',')
names(full_mat)

mat_stand<- decostand(full_mat[,3:13], "standardize")
names(mat_stand)

all_mat<- cbind(full_mat[,2],mat_stand,full_mat[,14])
names(all_mat)
colnames(all_mat)[1] <- "Rst"
colnames(all_mat)[13] <- "Sites"

attach(all_mat)

##################
# Transformation #
##################

Mod.GF <- lm(abs(Rst) ~ 1+ (Sa+ S10z+ Ssk+ Sku+ Sdr+ Sbi+ Std+ Stdi+ Sfd+ Srwi)
* Geo_dist, data=all_mat)
boxcox(Mod.GF)
boxcox(Mod.GF, lambda = seq(0.1, 0.5, 0.1)) # best lambda=0.28
resp<- ((Rst)^0.28)/0.28



#######################################
# GAMM regression with surface metrics #
#######################################

fsites <- factor(all_mat$Sites)
Mod.GF <- gamm(resp ~ 1+ (Sa+ S10z+ Ssk+ Sku+ Sdr+ Sbi+ Std+ Stdi+ Sfd+ Srwi)
* Geo_dist, random = list(fsites=~1) , method = "REML", data=all_mat)
Mod.GF
summary(Mod.GF$gam)

############################
# GAMM regression with IBD #
############################

Mod.GF_geo <- gamm(resp ~ 1+ Geo_dist, random = list(fsites=~1) , method =
"REML", data=all_mat)
Mod.GF_geo
summary(Mod.GF_geo$gam)
```

VITA

VITA

Hossam Eldien M. Abdel Moniem

Office: 765-496-4601
Fax: 765-494-0535
E-mail: hmoniem@purdue.edu (and hossamesapres@gmail.com)
Address: 901 West State Street, West Lafayette, IN 47907, USA.

a. Professional Preparation

Institution	Major	Degree
Suez Canal University, Egypt.	Zoology	B.Sc., 1997
Suez Canal University, Egypt.	Zoology/Ecology	M.Sc., 2005
Purdue University, IN, USA.	Entomology/Landscape genetics	Ph.D., 2014

b. Appointments

2008–2013 Ph.D. Graduate Research Associate, Department of Entomology, Purdue University.

2006–2008 GIS and Data Management Consultant, BioMap of Egypt project, Egyptian Ministry of State for Environmental Affairs (EMSEA) and UNDP, Cairo, Egypt.

2005–2008 Associate Lecturer, Department of Zoology, School of Science, Suez Canal University.

1999–2005 Assistant Lecturer, Department of Zoology, School of Science, Suez Canal University.

1997–1999 Research Assistant, Department of Zoology, School of Science, Suez Canal University.

c. Publications

(i) *Peer-Reviewed Articles*

Abdel Moniem, H. M., Schemerhorn, B. J., DeWoody, J. A., Holland, J. D. (In Review). Landscape genetics of a pollinator longhorn beetle (*Typocerus v. velutinus* Olivier): A surface metrics approach. Molecular Ecology.

Abdel Moniem, H. M., Schemerhorn, B. J., DeWoody, J. A., Holland, J. D. (In Review). Phylogeography and demographic history of *Typocerus v. velutinus* (Olivier) as shaped by the Quaternary. Insect Molecular Biology.

Abdel Moniem, H. M., Holland, J. D. (2013). Habitat Connectivity for pollinator beetles using surface metrics. Landscape Ecology. 28: 1251–1267.

Holland, J. D., Shukle, J. T., **Abdel Moniem, H. M.,** Mager, T. W., Raje, K. R., Schnepp, K., Yang, S. (2013). Pre-treatment assemblages of wood-boring beetles (Coleoptera: Buprestidae, Cerambycidae) of the hardwood ecosystem experiment. In: Swihart, R.K., Saunders, M.R., Kalb, R.A., Haulton, G.S., Michler, C.H. (eds.) The Hardwood Ecosystem Experiment: a framework for studying responses to forest management. General Technical Report. NRS-P-108. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 218–236.

Raje, K. R., **Abdel Moniem, H. M.**, Farlee, L., Ferris, V. R., Holland, J. D. (2012). Pest and benign wood-borers both decrease with site index. Agricultural and Forest Entomology. 14: 165–169.

Abdel Moniem, H. M., Zalat, S. M. (2006) Illustrated field guide to spiders of south Sinai. Product of BioMap of Egypt project. EEAA press, Cairo, Egypt.

Abdel Moniem, H. M., Zalat, S. M., El–Naggar, M. H., Ghobashy, A. M. (2003). Habitat heterogeneity and altitudinal gradients in relation to spiders' diversity in south Sinai-Egypt. Egyptian Journal of Biology. 5: 129–137.

(ii) *Conference Proceedings /Presentations*

Abdel Moniem, H. M., Holland, J. D. (2012) Habitat Connectivity for pollinator beetles using surface metrics. US-International Association of Landscape Ecology (US-IALE) "Informing Decisions in a Changing World," Newport, RI, USA. April 8–12. *Oral presentation*

Abdel Moniem, H. M., Holland, J.D. (2009) Biodiversity of wood boring beetles and their host trees in Hoosier national forests. The Hardwood Ecosystem Experiment (HEE) annual research meeting. Bloomington, IN, USA. *Poster presentation*

Abdel Moniem, H. M., Holland, J.D. (2009) Is beetle diversity correlated with host tree diversity? International conference of the Entomological Society of America (ESA). November 5–9, Indianapolis, USA. *Poster presentation*

Abdel Moniem, H. M. (2009) The Geographic Information System day (GIS day) "Mash something." Stewart center, Purdue University. *Poster presentation*

Abdel Moniem, H. M., Zalat, S. M. (2004) Taxonomic and phylogenetic studies on Sinai spiders. The 3rd international conference of the Egyptian British Biological Soc. "Chemical Ecology" Cairo University, Egypt. *Poster presentation & Organizer*

Abdel Moniem, H. M., Zalat, S. M. (2003) Biodiversity of spiders in south Sinai-Egypt. The 2nd international conference. The Egyptian British Biological Society. "Challenges in biodiversity and biological conservation of endangered fauna and flora of Egypt" Cairo University, Egypt. *Poster presentation & Organizer*

d. Synergistic Activities

(i) Grant writing

Project	Role	Fund	years
Governmental Ph.D. scholarship programs. Funded by: The Egyptian government. Ministry of High Education. Administrated by the Egyptian Cultural and Educational Bureau, Washington DC.	Developed the idea, wrote the research and grant proposal, and linked to collaborators	**$** 100,000	4 years 2008–2012
BioMap of Egypt. Funded by: United Nations Development Program (UNDP) and the Italian Cooperation.	Data collection and helped writing parts of the research proposal	**€.** 1 million	3 years 2006–2009
Botanical Garden Project – British Council. Funded by: The British Council, Cairo, Egypt.	Data collection and helped writing parts of the research proposal	**€.** 5000	1 year 2004
Masters Students Research Funding. Funded by: Suez Canal University, Ismailia, Egypt.	Developed the idea, wrote the research and grant proposal and linked to collaborators	**£.E.** 9000	2 years 2003–2004

(ii) Other Significant Activities

[1] Produced an illustrated field guide to spiders of south Sinai and participated in some flyers, short stories and videos to educate public people and children about biodiversity and conservation.

[2] Trained and mentored some students, rangers and researchers in different projects and workshops:
Mentored a capstone project in the Department of Entomology at Purdue University (2011). Mentored two B.Sc. projects for undergraduate student from the Department of Zoology at SCU (2003 & 2005). Trainee in Saloga and Ghazal Protectorates Expedition; (EEAA) and BioMap of Egypt, Aswan, Egypt (2005). Trainee in Biostatistics and field data management workshop for rangers of Egyptian national protectorates; EMSEA and UNDP, Dahab, Egypt (2004). Trainee and site coordinator; Operation Wallacea (OPWAL) Sinai-Egypt (2004).

[3] Volunteered in a number of scientific events and filed expeditions: The HEE bard owl and screech owl winter surveys (2011–2012). The 74th, 75th, and 76th Pest Management Conference at Purdue University (2011 –2013). Judge for Summer Undergraduate Research Fellowship (SERF) Students' posters symposium, Purdue University (2012).

[4] Teaching Experience: Currently, I am involved in teaching three classes (General Entomology 207, Forest Entomology 441, and Ecological data analysis in R ENTM 692) in the Department of Entomology at Purdue University. I am also a TA and a grader for another class (Insects friend and foe 105) in the same department. In addition to that I have nine years of teaching experience in the Zoology Department at the school of Science at SCU. During this period I taught 8 different classes for undergraduate and graduate students.

e. Collaborators & Other Affiliations

(i) Collaborators

Raje, K.; Ferris, V.; Zaspel, J.; Scott, C.; San Miguel, F. *and others* (Purdue University). Shuey, J. (The Nature Conservancy Indiana Field Office, USA). Koh, I. (University of Vermont). Gilbert, F. (University of Nottingham, UK). James, M. (University of Nottingham, UK). Reiss, M. (University of London, UK).

(ii) Graduate and Postdoctoral Advisors

[1] Ph.D. Holland J. D.; Schemerhorn, B.; DeWoody, J. A. (Purdue University). [2] M.Sc. Zalat, S. M. (University of Taibah, KSA); El-Naggar, M. H. (K. Abdulaziz University, KSA); Ghobashi, A. M. Suez Canal University, Egypt)

(iii) Affiliation with Professional Organizations

US–International Association of Landscape Ecology (US-IALE); Entomological Society of North America (ESA); American Association for the Advancement of Science (AAAS); Egyptian British Biological Society (EBB Soc.); Egyptian Society of Ecological and Conservational Biology.

f. Honors and Awards

The most recent and notable of these honors and awards are:

1. Outstanding Doctoral student award by the Department of Entomology, Purdue University, 2014.

2. Student presentation award, US–IALE conference “Informing decisions in a changing world,” April 8–12,Newport, RI, USA (2012)
3. Bilsland dissertation fellowship (Fall 2012) Purdue University
4. William L. Brehm memorial scholarship (2010) Purdue University
5. Governmental scholarship from the Egyptian government, administrated by the Egyptian Cultural and Educational Bureau (ECEB), Washington DC (2008–2011)
6. University honors for M.Sc. students, Suez Canal University (2005)
7. Outstanding teaching award by the Faculty of Science for academic professionals (2004)
8. 1st place student competition award for posters session. The 2nd international conference of the EBB Soc. "Challenges in biodiversity and biological conservation of endangered fauna and flora of Egypt. Cairo University, Egypt (2003)
9. University honors for B.Sc. students. Suez Canal University (1997)