

Electronic Theses and Dissertations, 2004-2019

2015

The Effect of Landscape Variables on Adult Mosquito (Diptera:Culicidae) Diversity and Behavior

Caitlyn Debevec
University of Central Florida

 Part of the [Biology Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Debevec, Caitlyn, "The Effect of Landscape Variables on Adult Mosquito (Diptera:Culicidae) Diversity and Behavior" (2015). *Electronic Theses and Dissertations, 2004-2019*. 659.

<https://stars.library.ucf.edu/etd/659>

THE EFFECT OF LANDSCAPE VARIABLES ON ADULT MOSQUITO (DIPTERA:
CULICIDAE) DIVERSITY AND BEHAVIOR

by

CAITLYN DEBEVEC
B.S. University of Central Florida, 2012

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2015

Major Professor: David G. Jenkins

ABSTRACT

Diseases vectored by mosquitoes cause millions of deaths each year. In modern times Florida's disease risk has been reduced due to efforts to lessen the prevalence of mosquitoes through habitat modification of non-adults. With emerging diseases (i.e. Dengue and Chikungunya) encroaching into Florida from the Caribbean, this traditional approach may not be enough. Alternatively, we can better understand the ecology of how disease works in an ecosystem. One possible way is through the Dilution Effect, which states that the more species that are in a system the lower the chance for zoonosis. This project models mosquito diversity across regions, land use, and vegetation height in South-Central Florida, for the purpose of identifying predictors that indicate a higher disease risk using information theory (AICc). The plains and coastal regions as well as the developed areas have a relatively higher risk of disease.

Florida is a fire maintained habitat, but has been fire suppressed for the last century. Archbold Biological Station (ABS) has used prescribed fires since the early 1980s to try and restore a more natural system. This has created a mosaic of different fire histories. Fire affects the structures that mosquitoes rest under during the day (they are vulnerable to desiccation during the day and hide in darker/shady places), therefore there is a high likelihood that fire will have some effect on mosquito assemblages. This project used model selection to determine the most plausible set of predictors that describe the effect of fire on mosquito assemblages at ABS, using information theory (AICc). In general, time of season accounted for the largest proportion of the variation in the data and TSF had negligible effect on adult mosquito assemblages measured as abundance, species richness, and Jost D.

I dedicate this Master's thesis to three major parties: First and foremost, my parents Michael and Rosemarie Debevec. I love you both so much and truly appreciate all that you have done for me my entire life. Second, to Florida Camp for Children and Youth with Diabetes; I know that I would not be the same person I am today without diabetes and would not have survived to this point without the support of my camp friends. I hope that my work as a graduate student helps to inspire young diabetics to know that diabetes does not have to be a hindrance in any way. Finally, I would like to dedicate this project to all the mosquitoes that were sacrificed so that we may learn.

ACKNOWLEDGMENTS

I would like to acknowledge a plethora of people who helped in some capacity on this project, that without which this project would not have been possible. First my committee members: David Jenkins, Raoul Boughton, Joshua King, and Betsie Rothermel. I especially would like to thank Pedro Quintana-Ascencio for all of his 'R' and statistics guidance.

Undergraduates who have been involved: Leah Dash, Chris Grimont, Nate Nafzger. I would like to show my great gratitude for Archbold Biological Station for all the support they provided, specifically I would like to thank: Hillary Swain, Karlie Carman, Kevin Moses, Clay Noss, Rebecca Tucker, and the several others that I bounced ideas off of or helped me carry things in the field on occasion. The D3 lab at UCF (past and present members): Hannah Bevan, Jason Schroeder, and Scott Smith. As well as, the D3 honorary member: Mary Jenkins. Collection sites: Florida State Park Services, St. Lucie County, and all of the families that graciously allowed me to collect mosquitoes near their homes and business while attracting more mosquitoes to their yards unbeknownst to them. Finally, I would like to acknowledge the UCF BGSA who has helped me get through not only this project but get through graduate school and life in general.

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES.....	xii
CHAPTER 1: GENERAL INTRODUCTION	1
CHAPTER 2: MOSQUITO (DIPTERA: CULICIDAE) BIODIVERSITY IN SOUTH- CENTRAL FLORIDA AND THE DILUTION EFFECT FOR EXPOSURE TO ZOONOTIC DISEASE	5
Abstract	5
Introduction	6
Methods.....	8
Study Sites.....	8
Sampling.....	17
Statistical Analysis	19
Results.....	26
Discussion	50
CHAPTER 3: MOSQUITOES AND FIRE IN THE SCRUB	55
Abstract	55
Introduction.....	55
Methods.....	59

Study Sites	59
Sampling.....	61
Statistical Analysis	63
Results	67
Discussion	79
CHAPTER 4: CONCLUDING REMARKS ON SOUTH-CENTRAL FLORIDA	
MOSQUITOES.....	82
APPENDIX A: FIGURES EXTENDED.....	84
APPENDIX B: R CODE	100
REFERENCES	160

LIST OF FIGURES

Figure 1: Each pie represents the proportion of high and low competent species in the community. The left figure represents a system with low diversity where the proportion of highly competent species is high. The right figure represents a system with high diversity where the proportion of highly competent species is low. The number next to each pie is the proportion of highly competent species. The total number is a proportional estimate of the probability of zoonosis based on multiplying the proportion of highly competent species at each level together.

..... 7

Figure 2: Map of south-central Florida indicating sampling site with yellow markers. Each region is outlined by a separate color (Ridge=yellow, Plains=green, Coast=orange). Each region has two land use types: Developed and ‘Natural.’ Each land use type has three vegetative structures: short, medium, and tall..... 9

Figure 3: A Generalized map of Natural communities of ABS from 2004..... 11

Figure 4: Google Earth satellite image of MAERC, which is outlined in red..... 13

Figure 5: Google Earth Satellite image of SPSP (outlined in red) and Walton Scrub Preserve (outlined in purple). 16

Figure 6: Ordination stress plot. The above the red line denotes where stress is poor, above the yellow line is fair, above the green line is good, and below the green line is excellent (McCune and Grace, 2002)..... 24

Figure 7: Ordination distance ~ Observed dissimilarity for the NMDS of both the Jaccard's and Bray-Curtis' indices..... 24

Figure 8: Species richness mean and 95% confidence intervals, per region and land use (black is natural, blue is developed).	30
Figure 9: Jost D means and 95% confidence intervals, per region and land use (black is natural, blue is developed).	34
Figure 10: Vector proportion means and 95% confidence intervals for Dengue/Chikungunya, per land use.	37
Figure 11: Vector proportions for WNV without coast vs. distance coefficient, per sampling event. Each color represents a different sampling event.....	39
Figure 12: Vector proportions for WNV without coast by region and land use (black is natural, blue is developed).	39
Figure 13: Vector proportions for SLE by region and land use and vegetation height (black is natural, blue is developed).	42
Figure 14: Partial mantel correlogram for NMDS of Jaccard and Bray-Curtis index. Closed dots indicate a significant difference ($\alpha=0.05$) and open dots indicate no significant difference.	45
Figure 15: A Generalized map of Natural communities of ABS from 2004.	60
Figure 16: Map of the sampling sites at Archbold Biological Station. Yellow-short TSF, pink -medium TSF, green -long TSf.	62
Figure 17: Macro-plot (in green) with trap in center (diamond). Three 5m diameter micro-plots (in brown) with 3 1m diameter subplots in black. Vegetation and GLA data was taken at each black subplot.	63
Figure 18: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.	85

Figure 19: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.	85
Figure 20: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.	86
Figure 21: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.	86
Figure 22: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.	87
Figure 23: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.	87
Figure 24: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.....	88
Figure 25: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.....	88
Figure 26: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.....	89
Figure 27: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around each sampling event.	89
Figure 28: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around each sampling event.	90
Figure 29: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around each sampling event.	90

Figure 30: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.....	91
Figure 31: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.....	91
Figure 32: Ordination (NMDS 2vs 3) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.....	92
Figure 33: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.	92
Figure 34: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.	93
Figure 35: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.	93
Figure 36: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.....	94
Figure 37: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.....	94
Figure 38: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.....	95
Figure 39: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around sampling events.....	95
Figure 40: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around sampling events.....	96

Figure 41: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around sampling events.	96
Figure 42: Decay of variance for each axis in the vegetation PCA.	97
Figure 43: Biplot for PC1 and PC2.	98
Figure 44: Biplot for PC1 and PC3.	98
Figure 45: Biplot for PC2 and PC3.	99

LIST OF TABLES

Table 1: Universal Transverse Mercator (UTM) locations of all 18 sites.	10
Table 2: Fixed effects structure of models for all sets. Model 12 cannot be used because it creates a singularity.....	21
Table 3: AICc and weights for the five lowest AICc for each species richness model.....	27
Table 4: Estimates for species richness with coast.	27
Table 5: 95% confidence intervals for species richness with coast. Red numbers indicate significantly different from zero.	28
Table 6: Estimates for species richness without coast.....	28
Table 7: 95% confidence intervals for species richness without coast. Red numbers indicate significantly different from zero.	29
Table 8: AICc and weights for the lowest AICc for each Jost D model.....	31
Table 9: Estimates of Jost D model with coast.	31
Table 10: 95% Confidence intervals for Jost D with coast. Red numbers indicate significantly different from zero.	32
Table 11: Estimates of Jost D model without coast.....	32
Table 12: 95% Confidence intervals for Jost D without coast. Red numbers indicate significantly different from zero.	33
Table 13: Estimates of Dengue/Chikungunya vector proportions model without coast.	35
Table 14: Estimates of Dengue/Chikungunya vector proportions model with coast.....	35

Table 15: AICc and weights for the five lowest AICc for the Dengue and Chikungunya vectors models.	36
Table 16: AICc and weight for the five lowest AICc of each WNV models.	38
Table 17: Estimates of WNV model without coast.	38
Table 18: AICc and weights for the five lowest AICc's for each vector proportions model of SLE.	40
Table 19: Estimates of SLE models with coast.	41
Table 20: Estimates of SLE models without coast.	41
Table 21: Adonis output for Jaccard's index. Bright red indicates significance with $\alpha=0.05$. 43	
Table 22: Adonis output for Bray-Curtis index. Bright red indicates significance with $\alpha=0.05$	43
Table 23: Partial mantel test for NMDS of Jaccard's index. Bright red indicates significance with $\alpha=0.05$	44
Table 24: Partial mantel test for NMDS of Bray-Curtis' index. Bright red indicates significance with $\alpha=0.05$	44
Table 25: Indicator species for each region or combination of regions.	47
Table 26: Indicator species for each land use type or combination of land use types.	48
Table 27: Indicator species for each vegetation height or combination of vegetation height.	49
Table 28: Vegetation PCA loadings.	66
Table 29: AICc comparison of Abundance ~ TSF models.	68
Table 30: AICc comparison of Abundance ~ GLA top 5 models.	68
Table 31: AICc comparison of Abundance ~ percent vegetation categories top 5 models.	69

Table 32: AICc comparison of Abundance ~ tall vegetation categories top 5 models.....	69
Table 33: AICc comparison of Abundance ~ vegetation PCA top 5 models.	70
Table 34: AICc comparison of most plausible model from each model set.	70
Table 35: 95% confidence intervals for the most plausible model for abundance (tall vegetation category model).....	71
Table 36: AICc comparison of species richness ~ TSF models.	71
Table 37: AICc comparison of species richness ~ GLA top 5 models.	72
Table 38: AICc comparison of species richness ~ percent vegetation categories top 5 models.	72
Table 39: AICc comparison of species richness ~ tall vegetation categories top 5 models. ..	73
Table 40: AICc comparison of species richness ~ vegetation PCA categories top 5 models.	73
Table 41: AICc comparison of the most plausible models for each set for species richness.	74
Table 42: 95% confidence intervals for the most plausible model for species richness (tall vegetation category model).....	74
Table 43: AICc comparison of jost D ~ TSF models.	75
Table 44: AICc comparison of jost D ~ GLA top 5 models.	75
Table 45: AICc comparison of jost D ~ percent vegetation categories top 5 models.....	76
Table 46: AICc comparison of jost D ~ tallest vegetation categories top 5 models.....	76
Table 47: AICc comparison of jost D ~ vegetation PCA top 5 models.....	77
Table 48: AICc comparison of the most plausible models for each set for Jost D.....	78
Table 49: 95% confidence intervals for the most plausible model for Jost D (Vegetation PCA model).	78

Table 50: Indicator speices values for each TSF. Red indicates significantly different from zero..... 79

Table 51: Full results of vegetation PCA..... 97

CHAPTER 1: GENERAL INTRODUCTION

Mosquitoes are important disease vectors throughout the world, spreading diseases like dengue, malaria, West Nile Virus, etc. For this reason it is important to understand mosquito ecology in order to reduce the risk of zoonosis (transfer of a disease from an animal to a human). Two main factors influence the chance of zoonosis; abundance and diversity. Variables that affect abundance of mosquitoes have been studied extensively (Elnaiem et al., 2008; Barney and Anson, 1912; Wegbreit and Reisen 2000; Bates, 1945; Hayes et. Al. 2005), but diversity is less studied (Chaves et al., 2011; Muturi et al., 2006; Beketov et al. 2010). Moreover, studies about mosquito assemblage are on a single geographic scale, tend to be on larvae, small tree holes, or do not take into account the entire adult assemblage. My studies evaluate abundance and diversity at both local and regional scales.

The term diversity can be ambiguous; diversity is defined for this study as assemblage composition (Anderson et al., 2010), where assemblage is the species of a single group of organisms (here mosquitoes) within the larger community and composition is measured as the species presence and/or abundance.

Mosquitoes are holo-metabolous insects (they go through all four life stages). Eggs are laid in a variety of locations depending on the species including: soil, vegetation, tree holes, containers, rafts on the water surface, swamp margins, etc., and oviposition is influenced by olfactory, physio-chemical, and visual cues (Bently, 1989). Eggs hatch into water as larvae, which then pupate, and finally emerge out of the water as adults.

Generally mosquitoes drink nectar from flowers. Most female adult mosquitoes ingest blood for egg production, while males do not ingest blood. Nocturnal mosquitoes will perch underneath leaves in order to avoid desiccation during the day. A few diurnal mosquitoes exist in the tropics where it is more humid and are actively feeding during the day (Bates, 1949; Silver, 2008; Bates, 1944). Because desiccation is such a risk for mosquitoes the type of vegetation they are found in should be important (Burkett-Cadena et al., 2008). This includes vertical stratification of species due to preference and/or ability to deter desiccation (Bates, 1949). Much of the vegetation in Florida is affected by fire, and/or is fire maintained (Myers and Ewel, 1990). This points to a possible relationship between mosquitoes and fire.

Vegetative structure may also play a role in mating locations for mosquitoes. Remarkably little is known about the mating behavior of mosquitoes, and much of that knowledge is confined to the *Anopheles* genus because of its medical importance as a malaria carrier. In the first 12-36 hours after emergence for all male mosquitoes, their genitalia must complete a full 180° turn, and this seems to limit when they can start to copulate (Clements, 1999). After this period the males will begin to form mating swarms around dusk. The timing of the swarm is mostly light dependent and is different for each species. There is conflicting evidence of how these swarms assemble. Some studies show that males follow certain markers (i.e. overtop of a church steeple (Downes, 1969)); other studies have shown no evidence of such markers (Diabate et al. 2003). Either way it could be surmised that the swarming areas should not be very far from the daily perch sites because of the great metabolic requirements of long flights. Females will join the swarm shortly after, forming copulatory pairs (Yuval, Wekesa, and Washino, 1993).

After the females have been fertilized they search out blood meals in order to gain enough nutrients to make eggs, although some species consume enough iron and protein as larvae to create clutches of eggs without feeding on blood (Bates, 1949). They blood feed in cycles based on gonotrophic state (Provost, 1952). Females may or may not feed on multiple hosts during any one cycle.

Individual mosquito species are not considered to be cosmopolitan because we find varying species throughout the world. This leads to the questions of why we find species in some places and not others, is this pattern predictable, and at what scale does this apply to. A more anthropocentric question is how all of this affects vectoring of diseases by mosquitoes across landscapes. As previously mentioned the effect of abundance on zoonosis is well studied, but the effect of diversity is less so. One of the theories on how diversity affects zoonosis risk is called the Dilution effect (Schmidt and Ostfeld, 2001); which states that few mosquitoes are highly competent at vectoring any disease; therefore the more diverse the system, the lower the proportion of highly competent species will be.

Therefore I addressed the effect of habitat on adult mosquitoes at regional and local spatial scales in Florida. Specifically, I asked two major questions here: “What combination of landscape variables (region, vegetation, and land use) and seasonality best predicts mosquito biodiversity and a potential vector Dilution Effect?” And “How does Time Since Fire (TSF) affect adult mosquito assemblages at Archbold Biological Station?” The former was tested using 18 different sampling sites partitioned orthogonally into 3 regions, 2 land use types, and 3 vegetation structures, over 12 sampling events. The latter was tested using 84 sampling sites at

Archbold Biological Station randomly chosen within 3 different TSF categories. Both were approached from a mixed model perspective with model selection.

This first question addresses a relative probability of zoonosis of the major mosquito vectored diseases in Florida based on the predictors tested, as well as, gain insight into the possibility of the Dilution Effect. The second question addressed if TSF in Florida scrub affects adult mosquito assemblages, and if fire management may also affect zoonotic risk. Comparing the two studies conducted at different scales, allowed me to test similar patterns across those scales.

CHAPTER 2: MOSQUITO (DIPTERA: CULICIDAE) BIODIVERSITY IN SOUTH-CENTRAL FLORIDA AND THE DILUTION EFFECT FOR EXPOSURE TO ZOOONOTIC DISEASE

Abstract

Many efforts have been put into place in Florida to reduce the nuisances of mosquitoes and the risk of mosquito-borne diseases, but much of this effort has been focused on mosquito abundances and control of non-adult individuals. With emerging diseases expanding up from the Caribbean (i.e. dengue and chikungunya) this traditional effort may not be enough. An alternative approach is to understand how the diversity of the mosquito community affects the risk of disease, one idea to explain this is called the Dilution effect. The Dilution effect simply states that the more diverse a system is the more diluted the few individuals that can vector a disease will be. Eighteen sites were sampled to include all combinations of region (3), land use (2), and vegetative height (3) in a factorial design. For 5 months in 2012, 201 samples and ~367,000 individuals were collected. A mixed-effect model approach with AICc and model selection was used in conjunction with indicator species analysis and Non-Metric Multi-Dimensional Scaling. The results do not provide evidence for a simple Dilution Effect because of the complex nature of the system. It does indicate that there is mosquito distribution patterns that can be predicted with region and land use that show that coastal and urban areas have a higher relative mosquito-borne disease risk.

Introduction

Florida has been a hot bed for vector borne diseases throughout its history. In modern times this risk has been reduced due to efforts to reduce the prevalence of mosquitoes which often involves habitat modification (e.g, draining or flooding with salt water), pesticide broadcast spraying and window screens. With the advent of new emerging diseases coming out of the Caribbean, possibly driven by warming temperatures, these efforts may not be enough to keep out diseases like Dengue or Chikungunya. From 2013 to March of 2015 there were 42,994 confirmed cases of Chikungunya in the Caribbean and 2,116,410 suspected cases (PAHO/WHO, March 27, 2015). A local transmission has already been confirmed in Florida at the beginning of 2014 (CDC, Chikungunya virus) and it can only be assumed that with the amount of travel there is between Florida and the Caribbean every day, Chikungunya will be well established soon. Habitat modification and pesticide spraying affect entire ecosystems and window screens reduce nighttime disease risk, but do not affect mosquito diversity or abundance. An alternative approach is to understand the effects that natural diversity has on risk of disease transfer. The effect of diversity on zoonosis is called the Dilution Effect.

The Dilution Effect of zoonosis is based on compounded relative probabilities of diseases, hosts, and vectors coming into contact with each other (Figure 1). Zoonosis risk may be reduced if mosquito species that dominant the assemblage are not competent to vector the disease, diluting the competent species (Schmidt and Ostfeld, 2001). The Dilution Effect can occur at any one layer (or combination of) a multi-layered disease life strategy system. Here I focus on the vector Dilution Effect, assuming this system has a variety of vector and non-vector species.

Given thus Dilution Effect, the more diverse the mosquito assemblage the lower the rate of zoonosis, because there will be proportionally few highly competent species. Because this is only a likelihood based on equiprobable contact with a mosquito species, there will be some systems and situations where the opposite will occur. Few studies have considered a Dilution Effect in any natural system (Swaddle and Calos, 2008; Hall et al., 2009). If the Dilution Effect occurs in nature, and matters for zoonosis, then it is important to understand what factors drive adult mosquito diversity.

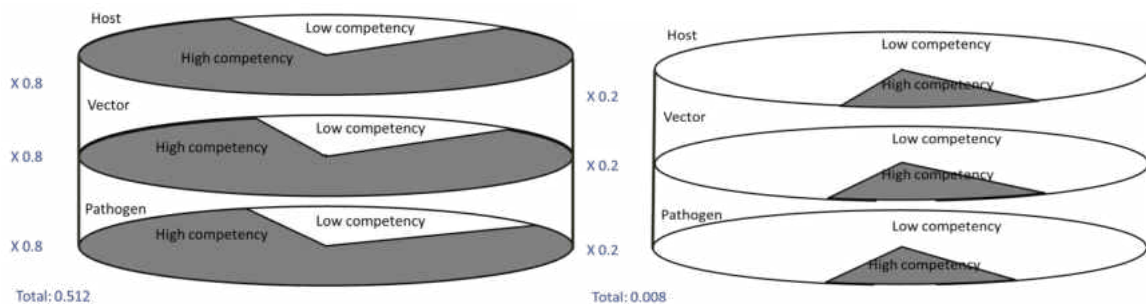


Figure 1: Each pie represents the proportion of high and low competent species in the community. The left figure represents a system with low diversity where the proportion of highly competent species is high. The right figure represents a system with high diversity where the proportion of highly competent species is low. The number next to each pie is the proportion of highly competent species. The total number is a proportional estimate of the probability of zoonosis based on multiplying the proportion of highly competent species at each level together.

Factors that may affect diversity of adult mosquitoes include geography, vegetation, anthropogenic influences, and seasons. Geography affects climate, history, dispersal or available species pools. Vegetation is potentially important for mosquito perch sites, location of breeding swarms, or as an indicator of available vertebrate hosts. Anthropogenic influences matter

because humans change environmental conditions, including the availability of breeding habitats, vegetation, thermal regime, urban homogenization, etc. (Johnson, Gomez, and Pinedo-Vasquez 2008; James and McCulloch, 1990). It has also often been seen that urbanization homogenizes the habitat and would therefore be dominated by only a few species (McKinny, 2006). Finally, seasons matter because different mosquito species may have different life cycle timings throughout a year, related to thermal and hydrological changes.

This information leads me to my first question: What combination of landscape variables (region, vegetation, and land use) and seasonality best predicts mosquito biodiversity and a potential vector Dilution Effect? I hypothesize that if phenology, geography, and land use are large scale drivers of mosquito diversity, then seasonality, region, and land use will significantly predict mosquito assemblages. This was tested in south-central Florida where the mosquito assemblages include species that vector existing diseases (i.e. West Nile virus and St. Luis Encephalitis), as well as emerging diseases (i.e. dengue and Chikungunya). This studies goal is not to obtain an effect size, but to identify predictors of mosquito diversity that could later inform a more focused study on each predictor.

Methods

Study Sites

Eighteen sites were chosen to include all combinations of region (3), land use (2), and vegetative height (3) in a factorial design. The three regions chosen were the south end of the Lake Wales ridge (ridge ~65m above sea level), the plains area just East of the Lake Wales ridge

(plains ~9m above sea level), and the coast directly east of the plains site (coast ~4m above sea level). The two land use types chosen were developed and 'natural', classified by apparent density of buildings or concrete. The three vegetative heights chosen were short, medium, and tall (Figure 2 and Table 1).

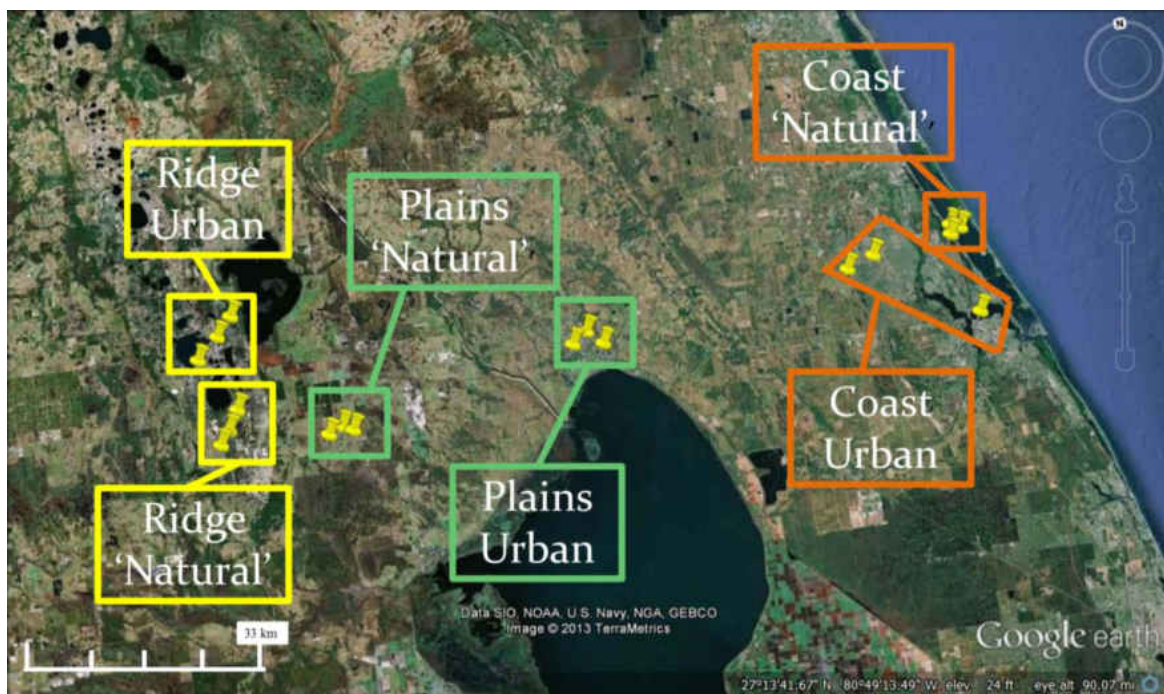


Figure 2: Map of south-central Florida indicating sampling site with yellow makers. Each region is outlined by a separate color (Ridge=yellow, Plains=green, Coast=orange). Each region has two land use types: Developed and 'Natural.' Each land use type has three vegetative structures: short, medium, and tall.

Table 1: Universal Transverse Mercator (UTM) locations of all 18 sites.

Region	Land use	Vegetation	UTM Northing	UTM Easting	PDOP
Ridge	Urban	Short	3016448.95	460829.47	5.52
Ridge	Urban	Medium	3019457.82	463919.77	3.32
Ridge	Urban	Tall	3022178.20	466119.71	6.36
Ridge	Natural	Short	3004412.67	463261.99	3.31
Ridge	Natural	Medium	3005890.30	464240.74	3.22
Ridge	Natural	Tall	3009770.20	465654.69	4.74
Plains	Urban	Short	3010964.01	519150.78	4.14
Plains	Urban	Medium	3012980.06	516925.13	4.35
Plains	Urban	Tall	3011407.06	514787.66	4.21
Plains	Natural	Short	3004425.20	480409.48	5.77
Plains	Natural	Medium	3003213.53	478447.06	3.86
Plains	Natural	Tall	3003562.11	481937.37	4.23
Coast	Urban	Short	3016811.49	555871.32	5.37
Coast	Urban	Medium	3008190.02	574034.72	5.52
Coast	Urban	Tall	3018353.84	559899.81	4.88
Coast	Natural	Short	3021313.41	571768.73	3.79
Coast	Natural	Medium	3020660.08	573245.67	2.07
Coast	Natural	Tall	3019879.57	571663.92	7.99

The Lake Wales ridge is part of the Florida ridge system which is comprised of upland areas that have remained above sea-level since the Pleistocene (James, 1961; Myers and Ewel, 1990-pgs 70-100). The ridge is home to many endemic species that have been isolated on ancient islands. Some of the different natural habitats found on the southern portion of the ridge include

several types of scrub, flatwoods, southern ridge sandhill, and bayheads (Abrahamson et al., 1984a). Much of the natural areas of the ridge have been cultivated for citrus and cattle pasture.

The natural land use type in this region is represented by Archbold Biological Station (ABS) including pasture lands in the Archbold Reserve (hereafter referred to collectively as ABS) (Figure 3). The station was farmed for turpentine and lumber in the past but has been managed for biological research since 1941 (Main and Menges, 1997). Only a handful of permanent structures exist in the 8,841 acres that make up ABS.

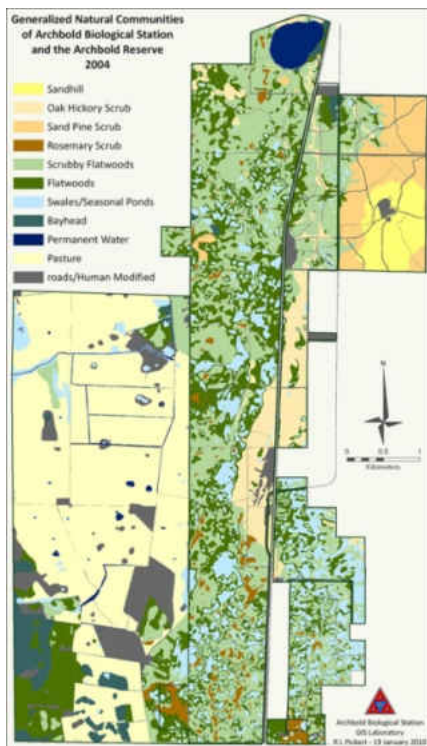


Figure 3: A Generalized map of Natural communities of ABS from 2004.

Short vegetative structure within the ABS site occurred in two intensively managed (Bahia grass seeded and ditched) pastures, located on the north end of the reserve. Cattle were recently excluded from these pastures. The medium vegetative structure is located inside a large scrubby Flatwoods patch in the north part of the station. This area contains many ephemeral ponds and is dominated by typical scrub plants including: palmettos, a mixture of oaks and sparse slash pines. The tall vegetative structure site is in the northeastern most corner of the station inside a bayhead. The bayhead is bordered on one side by railroad tracks, one side by Lake Annie, and one side by a nursery business. State road 70 passes through part of the bayhead but, there is dense bayhead on either side of the road. The dominant plants are of mixed bay-tree species (i.e. Red Bay, Sweet Bay, Cinnamon Bay).

The developed area on the ridge was in and around the town of Lake Placid, Florida which is located just north of ABS and has a population of only 2,223 in 2010 (U.S. Census Bureau). Even though this is a fairly small town, it is concentrated with a city center composed of commercial and residential land use. I selected sites within Lake Placid with relatively higher density of housing and businesses based on satellite images and local appearance.

The short vegetation site was located outside of Lake Placid in a community called Placid Lakes and is characterized by a network of open canals among residential properties, where most vegetation is ornamental but several housing lots are partially overgrown. The medium vegetation site was located in downtown Lake Placid, characterized by businesses, a satellite campus of a state college, a few parks, and many roads. Most vegetation is ornamental. The sampling site was located between the local police station/jail and the Chamber of Commerce.

The tall vegetation site was located in a community called Sylvan Shores on residential property, where many properties have ornamental plants, but various lots also contain vegetation that are common in bayheads.

The plains region was east of the Lake Wales ridge and is of generally lower elevation and flatter. Three dominant natural habitats in this region include: prairie grasses, pine flatwoods, and cabbage palm forests (Davis, 1967). Much of this area is now cultivated for citrus and cattle ranches.

The natural land use type on the plains was at MacArthur Agro-Ecology Research Center (MAERC) a division of ABS. This property is also known as Buck Island Ranch, and is 10,500 acres with ~3,150 cattle year round (Figure 4) (www.Maerc.org). The ranch contains both intensively managed and semi-native pastures (to be explained later).



Figure 4: Google Earth satellite image of MAERC, which is outlined in red.

The short vegetation site was along the main access road through MAERC between intensively managed pastures. These pastures have been tilled, seeded with Bahia grass, fertilized, and ditched. The dominant vegetation is Bahia grass. Cattle are also rotated through these pastures regularly, with greatest densities in the summer months (Swain et al., 2013). The medium vegetation site is in the middle of a semi-native pasture. This pasture is tilled irregularly and rarely seeded, has less ditches and has never been fertilized. Cattle are rotated into semi-native pastures in winter months. The dominant vegetation includes some non-native grasses like Bahia, but is mixed with native grasses (i.e. *Andropogon*) and shrubs (i.e. *Hypericum*). There are also some large oaks. The tall vegetation site is located within a stand of forest within an improved pasture. The dominant vegetation are sabal palm, cabbage palm, and oak. This tall vegetation patch has less sub-canopy structure than other tall vegetation sites in this study.

The developed land use sites in the plains region are in Okeechobee, Florida. Located east of MAERC with a population of 5,621 (U.S. Census Bureau) in 2010, with a central downtown surrounded by residential areas.

The short vegetation site was located in a residential area that has canals behind many homes. Most vegetation in the area is ornamental; with many lots having nearly exclusively short grass. The medium vegetation site was located on a business property one block south of the main downtown area, and well within the downtown area with abundant impervious surfaces. The dominant vegetation is ornamental and some overgrown properties. The tall vegetation site was located on a residential property within an oak hammock. Vegetation is dominated by

ornamentals, oaks, and houses are somewhat spread apart but are still much denser than 'natural' areas.

The coastal region was located on the east coast in Port St. Lucie, Fort Pierce and Stuart, Florida. This region is much more densely populated than the other regions. The dominant habitat here is pine flatwoods and scrub, with marsh directly next to the ocean (Davis, 1967). Some parts of this region contain salt water.

The natural land use area was in Savanna's Preserve State Park (SPSP) (Port St. Lucie) and Walton Scrub Preserve (Fort Pierce). SPSP is over 6,000 acres and includes pine flatwoods, freshwater marsh, scrub, and this entire habitat together has been called Savanna since the early 1800's by Lt. Colonel Benjamin Pierce (<http://www.floridastateparks.org/history/parkhistory.cfm?parkid=159>)(Figure 5). The Walton Scrub Preserve is owned by St. Lucie County and is 33 acres of endangered scrub habitat (Figure 5) (<http://www.stlucieco.gov/beaches/66.htm>).

The short vegetation site was located in SPSP along the freshwater marsh north of the Walton road entrance, where dominant vegetation is tall reed plants and phragmite grasses. This marsh had no detectable salt water intrusion at the time of study. The medium vegetation site was located in Walton Scrub Preserve. The dominant vegetation includes typical scrub species like palmetto, oak, and cactus. This site was somewhat over grown due to fire suppression. The tall vegetation site was located in SPSP in a dense tree stand near the southwestern edge of the portion of the park north of the Walton road entrance. The dominant vegetation included pines, palmettoes, and large shrub species. This site was closer to an urban area than other natural sites

(as is true in coastal Florida in general), but the placement of the site was chosen to maximize the amount of tall trees with minimal amount of housing nearby.

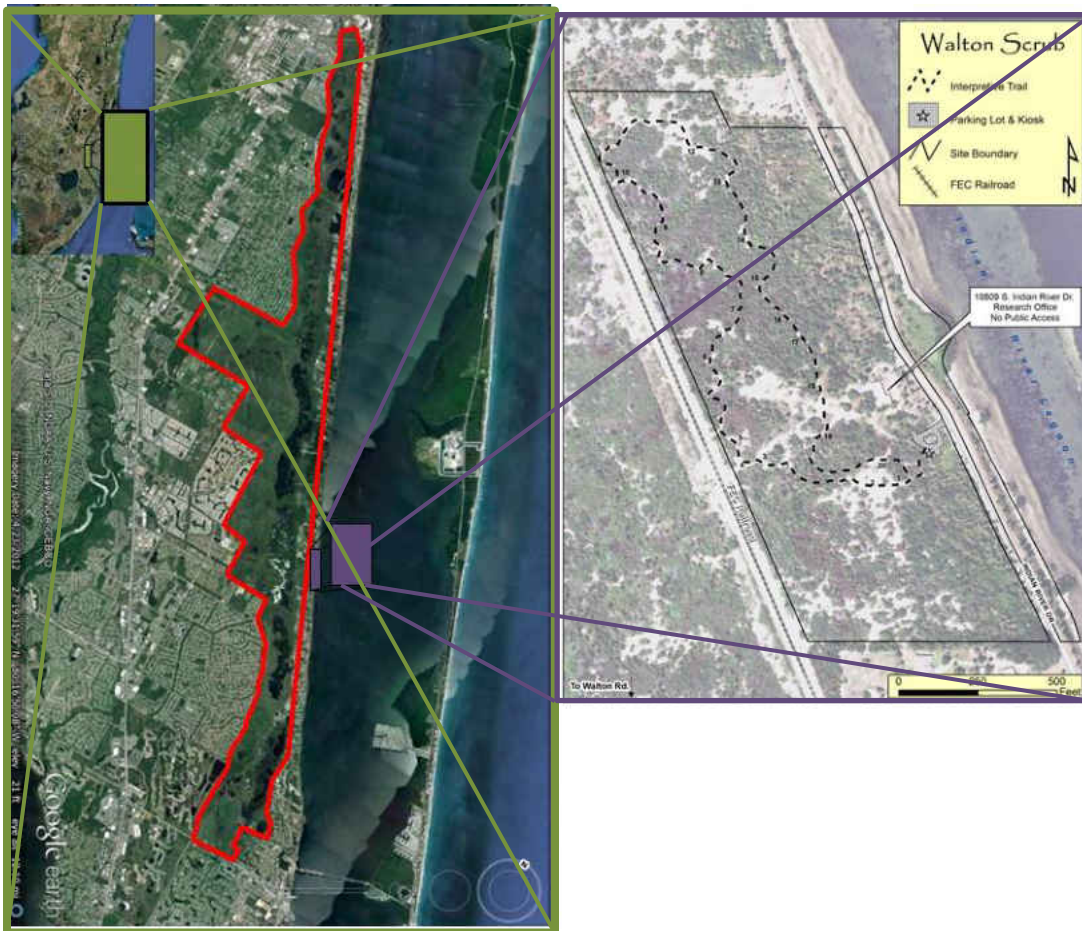


Figure 5: Google Earth Satellite image of SPSP (outlined in red) and Walton Scrub Preserve (outlined in purple).

The urban land use is located in Port St. Lucie and Stuart, Florida. Port St. Lucie has a population size of 166,603 (U.S. Census Bureau) in 2010 and Stuart has a population size of 15,593 (U.S. Census Bureau) in 2010.

The short vegetation site was located in a community called Traditions in the west part of Port St. Lucie. This area is characterized by wide streets, with large grass areas or small parks on most streets. A few blocks from the site is a community with canals behind the homes. The vegetation is almost completely ornamental with few, if any, overgrown lots. The medium vegetation site was located in Stuart in an area comprised predominantly of office buildings and many small shops. The dominant vegetation is ornamental and overgrown alleys. The tall vegetation site was located in Port St. Lucie in an area with many overgrown lots and dense forest. A few medium drainage canals pass through this site. The dominant vegetation is ornamentals, pines, and Brazilian pepper.

Sampling

Adult mosquitoes were trapped at each site weekly from June 20 to July 20, 2012 and nearly biweekly from August 8 to November 10, 2012. An entire region's sites were sampled on any one night, and the regions were sampled on consecutive nights starting with the ridge and ending with the coast (three consecutive nights). Exceptions to the above sampling regime included the following: Two sites located inside the SPSP, were not sampled until July 13 due to permitting restriction (sampling events 1-3); no coastal sites were sampled on Aug 10, due to illness (sampling event 6); three samples were lost in transport, (plains developed short on July 12, plains natural tall and a plains natural short on September 28). Despite sampling gaps, 201 samples were collected during the 5 month sampling interval.

In order to address sampling gaps, analyses requiring balanced designs were analyzed two separate ways. Before any analysis, to fill the gaps of the three lost data points the average of the

same sample point from the directly previous and directly proceeding sampling events. The first analysis omitted all coastal sites and hereafter is called the ‘without coast’ analysis. Second, the coastal sites were used, but the first 3 sampling events and sampling event 6 were omitted from the analysis entirely so that all data were balanced. Hereafter this will be called the ‘with coast’ analysis.

Mosquitoes were trapped with a John W. Hock CDC light trap with CO₂ attractant. The CO₂ attractant used was ~0.68kg of dry ice placed in a cooler hung next to the trap. Each trap was hung ~1m above the ground. The traps were set out just before sunset and retrieved the next morning. The samples were then frozen to kill all invertebrates trapped inside and remained frozen prior to processing. All non-mosquito invertebrates were excluded and then the sample was weighed to the nearest 0.0001 g. If the sample was over 0.5g then a sub-sample of ~0.5g (~1,000 individuals) was haphazardly taken. To evaluate the efficacy of the subsampling, ten samples that were subsampled were randomly chosen to enumerate all mosquitoes in the sample. I then regressed the subsample against the entire sample to come up with a formula to correct for subsampling:

$$Y = -0.32 + 1.04x \tag{1}$$

Where ‘*x*’ is the subsample abundance of each species in the sample and ‘*Y*’ is the corrected abundance. This formula was then used to correct all samples that were subsampled and were used from here on. This regression had an $R^2 = 0.97$ and the intercept and (log) subsample had p-values < 0.000. The correction was minor for the slope and slightly adjusted the intercept (i.e.

minimum N of a species in a sample). The minor adjustments by the equation indicate that subsampling did not strongly bias sampling and this minor bias was handled by the regression.

Subsamples that were <30% of the total sample were unrepresentative of the total species richness of the sample. To address this, if the subsample constituted less than 30% of the total sample, another subsample was taken until approximately 30% of the total sample was sampled.

Females were identified to species when possible using *Identification and Geographical Distribution of the Mosquitoes of North America, North of Mexico* (2005).

Statistical Analysis

The study used a factorial repeated measures design (region X land use) with vegetation height as a covariate, where each “treatment” was fixed and categorical. The terms used in every model were region (ridge, plains, coast), land use type (natural, developed), and vegetation height (short, medium, tall). Time (sampling event) was a random factor in the mixed effects models. This study was conducted across landscapes with spatial autocorrelation of patterns; distance among sites was accounted for with a second covariate ('dist'). 'Dist' was also used as a random factor in the mixed effects models when appropriate.

I used a hybrid of methods from Zuur et. al (2009) and model selection with AICc and weights (Burnham and Anderson, 2002) on all possible models. Where AICc is AIC with a correction for finite or small sample size compared to number of parameters (Hurvich and Tsai, 1989). An optimal set of random terms was selected by using the most complex fixed factors (dist + region*land use + vegetation height) in combination with random structures composed of

none, random intercept only, and random slope and intercept (Table 2). Models were constructed using Restricted Maximum Likelihood (REML). Sampling event was treated as the random intercept to address the repeated measures design. Geographical distance was addressed as a distance-weighted auto-covariate (Augustin, Muggleston, and Buckland. 1996; Bivand et. al., 2014) and was treated as both a fixed effect and the random slope. Significance of estimates were then determined using 95% confidence intervals obtained using 10,000 bootstraps in 'confint' (lme4, 2014). When two models were equally plausible via AICc comparison, either both models were evaluated, or the greatest hierarchical model as the most plausible.

Models were then validated using the function 'r.squaredGLMM' (MuMIn, 2015) to obtain pseudo R squared values. This function provides a marginal R squared value which represents variance of the fixed effects, and a conditional R squared value that represents the variance of both fixed and random effects. Both the confidence intervals and R squared function could not be applied to logistic models. Graphs were created to visually inspect the assumptions of homogeneity of factors and normal residuals. If models did not meet these assumptions then the model was rejected. In order to interpret the models means and 95% confidence intervals were obtained for all significant factors. These were then compared to each other to evaluate the models.

Table 2: Fixed effects structure of models for all sets. Model 12 cannot be used because it creates a singularity.

Models	Dist	Region	Land use	Vegetation Height	Region:Land use
1	X	X	X	X	X
2	X	X	X	X	
3	X	X	X		
4	X	X		X	
5	X		X	X	
6	X			X	
7	X	X			
8	X		X		
9	X	X	X		X
10	X	X		X	X
11	X		X	X	X
12	X			X	X
13	X	X			X
14	X		X		X

Both the Shapiro-Wilk test (Shapiro and Wilk, 1965; R core team, 2014) and the Levene test (Levene, 1960; Fox and Weisberg, 2011) were employed to evaluate parametric statistical assumptions.

Samples generated abundance and presence/absence data for multiple mosquito species, so that data comprised a species X sites X time matrix. Those data were evaluated in several ways, summarized below.

Diversity Models

Mosquito assemblages' diversity was assessed with two different response variables. First, using species richness because the dilution effect was proposed with that measure and it is a classic estimator of diversity. Secondly, using Jost D, or the effective diversity of a sample (Jost, 2006) and was calculated as:

$$Jost D = e^{H'} \quad (2)$$

where H' is Shannon's diversity:

$$H' = - \sum_i p_i \ln p_i \quad (3)$$

Where p_i is the proportional abundance of species i .

Logistic Models of Vector Porportions

The relative probability of vectoring diseases in each sample was calculated as the proportion of individuals that are members of possible vectors species for a disease to the total number of individuals caught in a sample. I determined which diseases to use and what species vector them from the Florida Medical Entomology Lab website (11/25/2014). There are three major diseases in Florida with fairly well known vectors. Dengue is vectored by *Aedes albopictus* and *Aedes aegypti*, West Nile Virus (WNV) is vectored by all *Culex* species, and St. Louis Encephalitis (SLE) is vectored primarily by *Culex nigripalpus*. There are a few other diseases of note in Florida. One is chikungunya, which has not yet established in Florida, but is prevalent throughout the Caribbean and may establish in the near future. Chikungunya is vectored by the same mosquito species as dengue and therefore easy to add in to this analysis. Two others disease, Eastern Equine Encephalitis and Highlands J virus, are not included in the analysis because there is not general consensus on the vectors in Florida.

Models were assessed similarly to approaches described above except that the distribution family used is binomial, or logistic and glmmADMB (Skaug et al., 2015) was used. This function allowed me to incorporate zero inflation when necessary.

Non-Metric Multi-Dimensional Scaling (NMDS)

NMDS (Kruskal, 1964; Minchin, 1987; Faith, Minchin and Belbin, 1987) was used to evaluate compositional data based on the Bray-Curtis' (Bray and Curtis, 1957) and Jaccard's indices (McCune and Grace, 2002). These two approaches are used because Bray-Curtis is widely used to evaluate abundance based differences between samples, while Jaccard is widely used to evaluate presence/ absence differences between samples.

For each ordination a scree plot was created in order to determine the optimal number of dimensions to use (Figure 6). Three dimensions were chosen, because it gave a stress value that was fair with a drastic improvement over 2 dimensions, but allowed for easy visualization (McCune and Grace, 2002). This also created ordinations that fit the data very well with linear $R^2 > 0.92$ and non-linear $R^2 > 0.98$ (Figure 7).

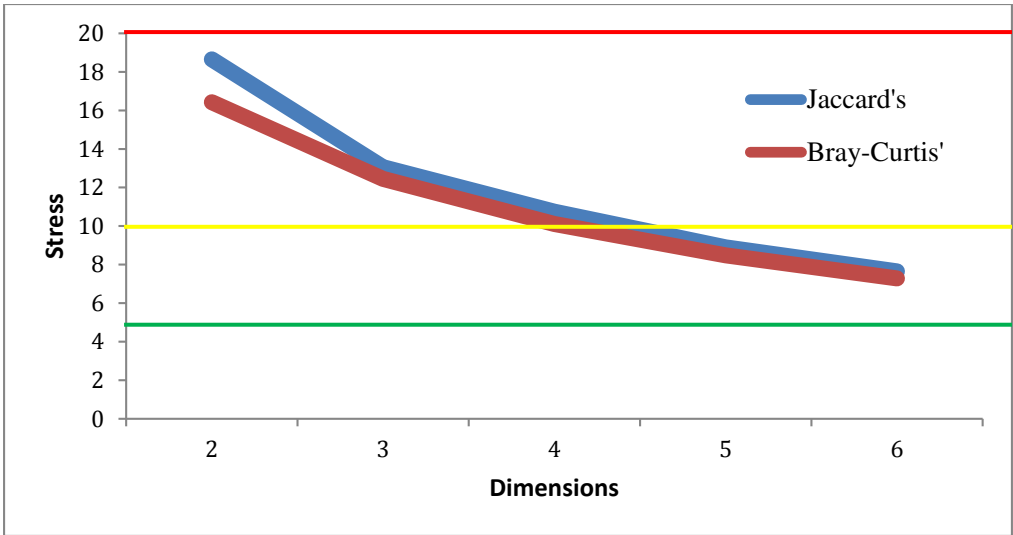


Figure 6: Ordination stress plot. The above the red line denotes where stress is poor, above the yellow line is fair, above the green line is good, and below the green line is excellent (McCune and Grace, 2002).

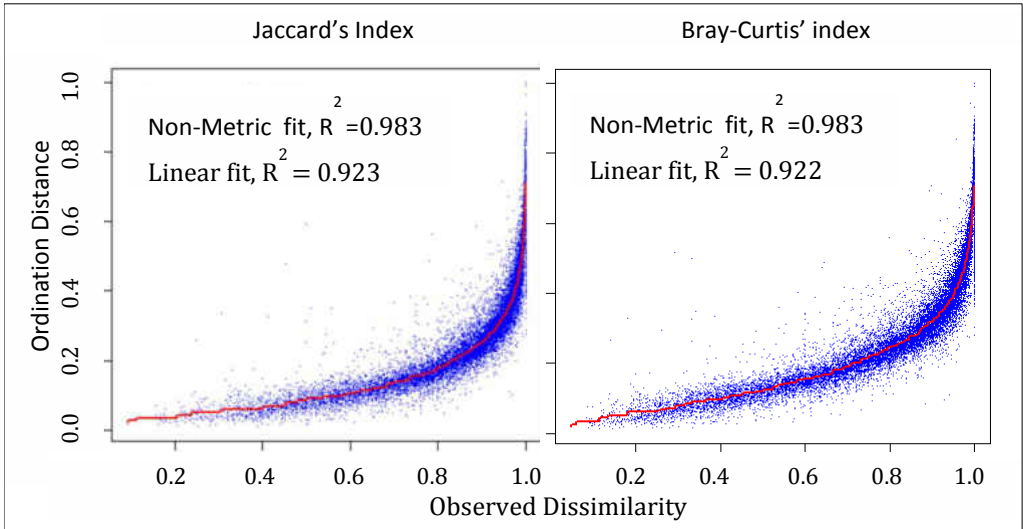


Figure 7: Ordination distance ~ Observed dissimilarity for the NMDS of both the Jaccard's and Bray-Curtis' indices.

The ‘MetaMDS’ command in the vegan package (Oksanen et al., 2013) was used to perform ordinations because it uses several random starts to find a stable solution. Random starts were used until two convergent solutions were obtained. The ‘Ordihull’ command in the vegan package (Oksanen et.al., 2013) was then used to draw convex hulls around different groups of data separated by region, land use type, vegetation height, or sampling event. Permutational multivariate analysis of variance using distance matrices was performed using the ‘adonis’ command in the vegan package (Oksanen et. al., 2013). This is a permutation-based MANOVA (999 permutations) using pseudo-F ratios to determine significance. This test is less sensitive to dispersion effects than anosim or MRPP and is therefore favored in this analysis (Oksanen, 2013-adonis).

Indicator Species

The R-package indicpecies (DeCaceres and Legendre, 2009) uses the group Indicator value (DeCaceres, Legendre, and Moretti, 2010) to produce a set of species that indicate a particular environmental variable, with a permutation test to test significance. Groups consistent to the highest association, of each level or combination of levels for region, land use type, and vegetation height were obtained with 999 permutations. Species component parts of the indicator value, were used to determine species fidelity (A), and species specificity (B). A species was considered an indicator species at an alpha of 0.05, but all species associations were reported.

All analyses were completed in R and all code is provided in the appendix.

Results

In total 201 samples were collected from June 20 to November 10 2012, comprising approximately 0.208kg or 367,000 mosquitoes, including 28 species.

Diversity Models

The most plausible model for species richness with or without the coast is:

Species Richness \sim Dist + Region + Landuse + Region: Landuse + (Dist|Sampling event)
(4)

They had an AICc of 669.2 and 662.4 respectively, with the same weight of 0.297 (Table 3, Table 4, and Table 6). Two other models had the same AICc and weight for each, but they were subsets of these models. This weight is over 8 times higher than the next best model. Distance coefficient, region, and region:land use are significantly different from 0 based on the 95% confidence intervals for both models (Table 5 and Table 7). The marginal and conditional R^2 is 0.28 and 0.55 respectively for with coast, and 0.28 and 0.62 for without coast.

Table 3: AICc and weights for the five lowest AICc for each species richness model.

Species Richness with coast		
	AICc	weight
m14	669.2	0.2
m9	669.2	0.2
m13	669.2	0.2
m10	673.5	0.0
m11	673.5	0.0
Species Richness without coast		
m9	662.4	0.2
m14	662.4	0.2
m13	662.4	0.2
m1	666.6	0.0
m10	666.6	0.037

Table 4: Estimates for species richness with coast.

Random Effects			
	Variance	Standard Deviation	Correlation
Sampling Event	50.9	7.1	
Dist	1.1	1.0	-0.9
Residual	4.6	2.1	
Fixed Effects			
	Estimate	Standard Error	t-value
Coast:Developed	26.8	4.1	6.4
Dist	-2.9	0.6	-4.8
Region-Plains	4.2	0.7	5.8
Region-Ridge	-0.9	0.6	-1.5
Land use-Natural	0.4	0.6	0.7
Region:Land use-Plains:Natural	1.4	0.9	1.6
Region:Land use-Ridge:Natural	4.9	0.9	5.1

Table 5: 95% confidence intervals for species richness with coast. Red numbers indicate significantly different from zero.

	2.5%	97.5%
SD Sampling Event	0.8	12.6
correlation Dist sampling event	-1.0	-0.8
SD Dist Sampling event	0.1	1.8
residual standard deviation	1.8	2.4
Region:Land use-Coast:Developed	18.7	34.8
Dist	-4.0	-1.7
Region-Plains	2.8	5.6
Region-Ridge	-2.2	0.2
Land use-Natural	-0.7	1.6
Region:Land use-Plains:Natural	-0.2	3.3
Region:Land use-Ridge: Natural	3.1	6.8

Table 6: Estimates for species richness without coast.

Random Effects:			
	Variance	Standard Deviation	Correlation
Sampling Event	79.8	8.9	
Dist	0.9	0.9	-0.9
Residual	4.1	2.0	
Fixed Effects:			
	Estimate	Standard Error	t-value
Region:Land use-Plains:Developed	33.4	5.3	6.2
Dist	-2.9	0.6	-4.6
Region-Ridge	-8.7	1.0	-8.0
Land use-Natural	-0.9	0.5	-1.8
Region:Land use-Ridge:Natural	6.4	0.9	7.0

Table 7: 95% confidence intervals for species richness without coast. Red numbers indicate significantly different from zero.

	2.5%	97.5%
SD Sampling Event	3.0	13.9
correlation Dist sampling event	-1.0	-0.9
SD Dist Sampling event	0.2	1.6
residual standard deviation	1.7	2.2
Region:Land use-Plains:Developed	22.9	43.7
Dist	-4.1	-1.7
Region-Ridge	-10.8	-6.6
Land use-Natural	-1.9	0.0
Region:Land use-Ridge:Natural	4.6	8.1

In natural land use the coast is different from the plains and/or ridge which have similarly the highest species richness. In developed land use the plains regions is different than the coast and/or ridge and the plains has the highest species richness. And only the ridge shows differences due to land use for species richness. This is why region, land use and the interaction of the two are significantly important predictors for species richness (Figure 8).

The most plausible model for Jost D with or without the coast is:

$$\text{Jost D} \sim \text{Dist} + \text{Region} + \text{Landuse} + \text{Region: Landuse} + (1|\text{Sampling event}) \quad (5)$$

They had an AICc of 401.3 and 428, with weights of 0.284 and 0.279 respectively (Table 8, Table 9, and Table 11). Two other models had the same AICc and weight for each, but they were subsets of these models. This weight is over 5 times higher than the next most plausible model. Distance coefficient, region, land use, and region:land use were significantly different from 0 based on the 95% confidence intervals for the with coast model. The without coast model is the

same, except region is not significant (Table 10 and Table 12). The marginal and conditional R^2 is 0.25 and 0.39 respectively for with coast, and 0.19 and 0.31 for without coast.

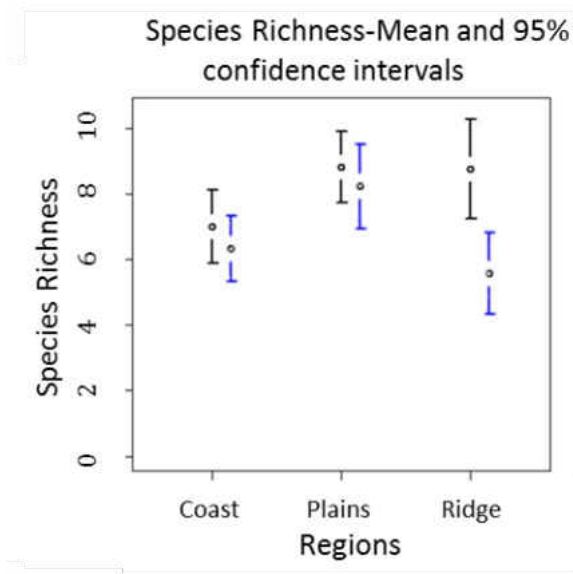


Figure 8: Species richness mean and 95% confidence intervals, per region and land use (black is natural, blue is developed).

Table 8: AICc and weights for the lowest AICc for each Jost D model.

Jost D with Coast		
m9	401.3	0.2
m13	401.3	0.2
m14	401.3	0.2
m1	404.8	0.0
m10	404.8	0.0
Jost D without Coast		
m13	428	0.2
m14	428	0.2
m9	428	0.2
m10	431.3	0.0
m1	431.3	0.0

Table 9: Estimates of Jost D model with coast.

Random Effects			
	Variance	Standard Deviation	
Sampling event	0.1	0.4	
Residual	0.7	0.8	
Fixed Effects			
	Estimate	Standard Error	t-value
Region:Land use-Coast:Developed	10.2	1.62	6.3
Dist	-2.5	0.5	-4.7
Region-Plains	-0.8	0.3	-2.9
Region-Ridge	-0.9	0.2	-3.1
Land use-Natural	1.0	0.2	3.8
Region:Land use-Plains:Natural	1.8	0.4	4.1
Region:Land use-Ridge:Natural	-0.2	0.3	-0.6

Table 10: 95% Confidence intervals for Jost D with coast. Red numbers indicate significantly different from zero.

	2.5%	97.5%
SD Sampling event	0.1	0.6
Residual Standard deviation	0.7	1.0
Region:Land use-Coast:Developed	7.0	13.4
Dist	-3.5	-1.4
Region-Plains	-1.4	-0.2
Region-Ridge	-1.4	-0.3
Land use-Natural	0.5	1.6
Region:Land use-Plains:Natural	0.9	2.7
Region:Land use-Ridge: Natural	-0.9	0.4

Table 11: Estimates of Jost D model without coast.

Random Effects			
	Variance	Standard Deviation	
Sampling event	0.1	0.4	
Residual	0.9	0.9	
Fixed Effects			
	Estimate	Standard Error	t-value
Region:Land use-Plains:Developed	9.4	1.7	5.3
Dist	-2.4	0.6	-3.8
Region-Ridge	-0.1	0.2	-0.4
Land use-Natural	2.3	0.4	5.3
Region:Land use-Ridge:Natural	-1.5	0.4	-3.6

Table 12: 95% Confidence intervals for Jost D without coast. Red numbers indicate significantly different from zero.

	2.5%	97.5%
SD Sampling Event	0.0	0.6
residual standard deviation	0.8	1.1
Region:Land use-Plains:Developed	5.9	12.9
Dist	-3.7	-1.2
Region-Ridge	-0.5	0.3
Land use-Natural	1.4	3.2
Region:Land use-Ridge:Natural	-2.4	-0.7

Natural land use indicate that the ridge is different from the plains and/or coast, but all are borderline different with plains have on average higher Jost D. For developed land use none of the regions indicate differences. Within each region both the coast and plains show effects of land use (Figure 9).

It should be noted that all models were assessed with and without spatial autocorrelation for the most complex model and in all cases including distance gave a smaller AICc and therefore the 'Dist' factor was included in all subsequent models.

In summary, both measures of diversity (species richness and Jost D) obtained consistent results, whether analyzed with or without coast data. In both diversity measures, region, land use, and region: land use interaction affects adult mosquito diversity. Models did have different random effects (species richness- random slope/intercept, Jost D-random intercept only), indicating the measures are differently sensitive to the effect of distance. Only vegetation height does not seem to have an effect at this scale for either measure.

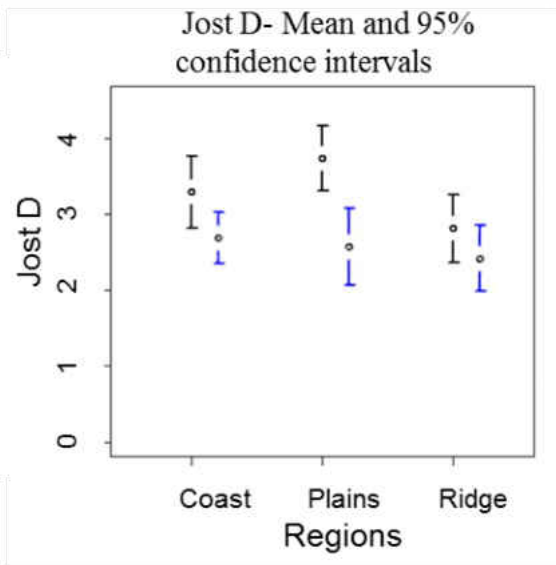


Figure 9: Jost D means and 95% confidence intervals, per region and land use (black is natural, blue is developed).

Logistic Models For Vector Porportions

Two models were equally the most plausible model to predict the proportion of vectors for Dengue and Chikungunya without the coast and they are:

$$\text{Dengue(Chikungunya)} \sim \text{Region OR Land use} \quad (6)$$

They each had an AICc of 4.1, with the same weight of 0.294 (Table 15). This weight is nearly 3 times higher than the next most plausible model (Table 13).

Table 13: Estimates of Dengue/Chikungunya vector proportions model without coast.

Region model				
	Estimate	Standard Error	Z value	Pr(> Z)
(Intercept)	-12.4	57.4	-0.2	0.8
Region-Ridge	0.0	81.1	0.0	1.0
Land use model				
(Intercept)	-12.4	57.4	-0.2	0.8
Land use-Natural	0.0	81.1	0.0	1.0

The most plausible model for Dengue and Chikungunya with coast:

$$Dengue(chikungunya) \sim Land\ use$$

(7)

With an AICc of 4.1 and weight of 0.487, which is nearly 3 times higher than the next most plausible model (Table 15 and Table 14).

Table 14: Estimates of Dengue/Chikungunya vector proportions model with coast.

	Estimate	Standard Error	Z value	Pr(> Z)
(Intercept)	-12.4	57.4	-0.2	0.8
Land use-Natural	0.0	81.1	0.0	1.0

Table 15: AICc and weights for the five lowest AICc for the Dengue and Chikungunya vectors models.

Dengue/Chikungunya w/out coast		
Model	AICc	weight
7	4.1	0.2
8	4.1	0.2
3	6.2	0.1
6	6.2	0.1
4	8.3	0.0
Dengue and Chikungunya w/ coast		
8	4.1	0.4
6	6.2	0.1
7	6.2	0.1
3	8.3	0.0
5	8.3	0.0

Dengue and Chikungunya were predicted most readily by land use and surprisingly, natural lands tend to have a higher proportion of vector mosquitoes. Both *Ae.albopictus* and *Ae. aegypti* are container breeders and containers are common in developed habitats (i.e. gutters, tires, flower pots, cemetery vases, etc.) the opposite pattern was expected (Figure 10). One possible explanation for this result is that *Ae. ablopictus* and *Ae. aegypti* were so rare except for at the coast that the land use effect is dominated by coastal data where ‘natural’ areas are very close to developed.

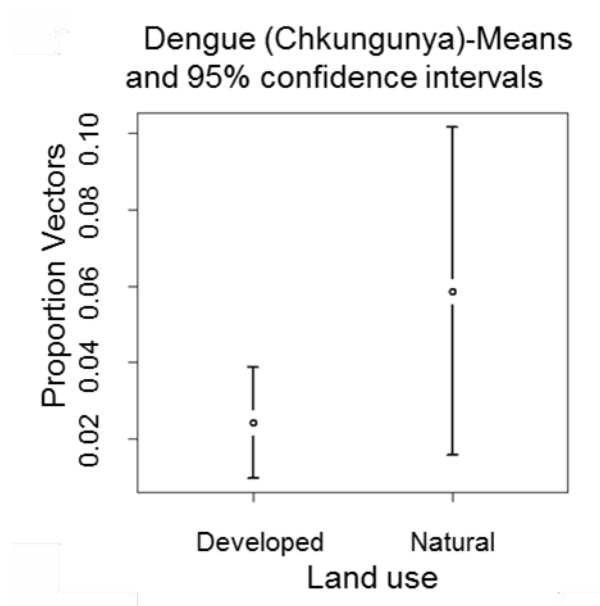


Figure 10: Vector proportion means and 95% confidence intervals for Dengue/Chikungunya, per land use.

The most plausible model for WNV with and without the coast is:

$$\text{WNV} \sim \text{Dist} + \text{Region} + \text{Land use} + \text{Region: Land use} + (1|\text{sampling event}) \quad (8)$$

With an AICc of 136.5 and 162.2 and weights of 0.309 and 0.292, which are over 9 and 7 times higher than the next most plausible models respectively (Table 16 and Table 17). However, the residuals for the with-coast model indicate that it does not represent the data appropriately.

Table 16: AICc and weight for the five lowest AICc of each WNV models.

WNV with coast		
Model	AICc	weight
14	136.5	0.3
13	136.5	0.3
9	136.5	0.3
11	141	0.0
1	141.1	0.0
WNV w/out coast		
9	162.2	0.2
14	162.2	0.2
13	162.2	0.2
11	166.1	0.0
1	166.1	0.0

Table 17: Estimates of WNV model without coast.

Random Effects				
	Variance	Standard. Deviation		
Sampling Event	0.3	0.6		
Fixed Effects				
	Estimate	Standard Error	Z value	Pr(> Z)
(Intercept)	18.9	3.6	5.1	0.0
Dist	-27.3	5.3	-5.1	0.0
Region-Ridge	-1.8	0.7	-2.4	0.0
Landuse-Natural	-6.2	1.2	-4.9	0.0
Region:Land use-Ridge:Natural	5.0	1.2	3.9	0.0

WNV cannot be well predicted when coast is included and therefore I will discuss only the model without coast. There is a general and slight distance decay in WNV vector proportion (Figure 11). Natural areas in the plains region tends to have lower vector proportions than in the ridge, but in developed areas they are not significantly different (Figure 12).

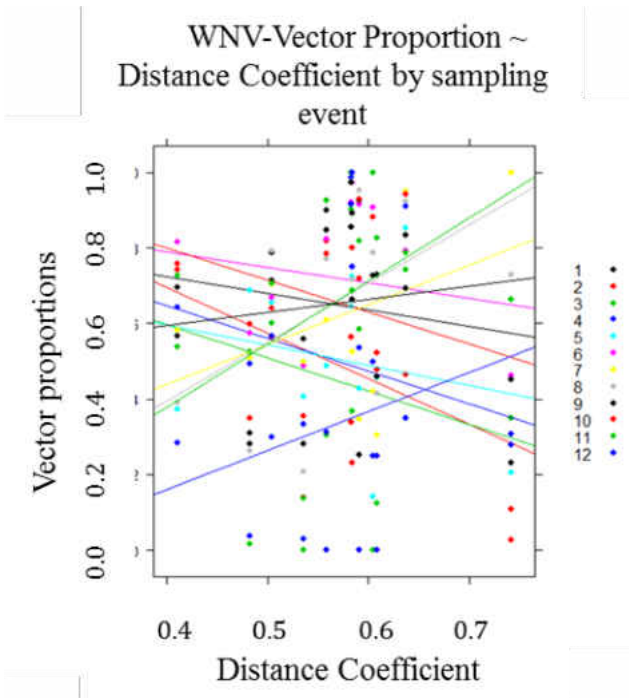


Figure 11: Vector proportions for WNV without coast vs. distance coefficient, per sampling event. Each color represents a different sampling event.

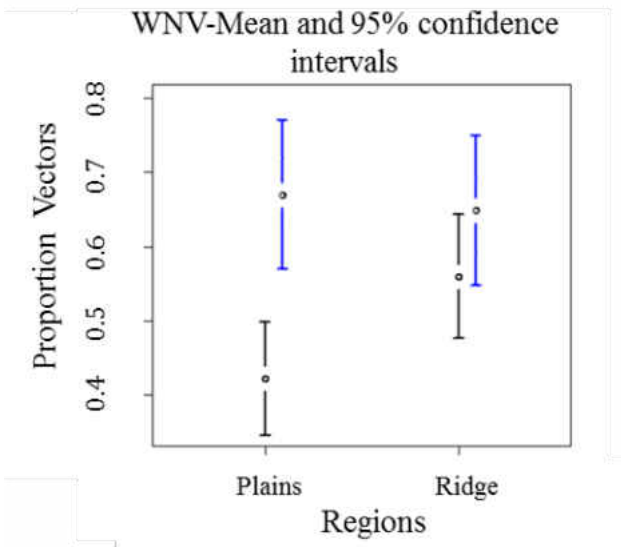


Figure 12: Vector proportions for WNV without coast by region and land use (black is natural, blue is developed).

The most plausible model for SLE with and without the coast is:

$$\text{SLE} \sim \text{Region} * \text{Landuse} + \text{Vegetation height} + (1|\text{sampling event}) \quad (9)$$

With an AICc of 161.8 and 153.5 and weights of 0.322 and 0.288, which are over 10 and 6 times higher than the next most plausible models, respectively (Table 18, Table 19, and Table 20).

Table 18: AICc and weights for the five lowest AICc's for each vector proportions model of SLE.

SLE with coast		
Model	AICc	weight
15	161.8	0.3
11	161.8	0.3
10	161.9	0.3
2	166.5	0.0
13	171.8	0.0
SLE w/out coast		
14	153.5	0.2
9	153.5	0.2
13	153.5	0.2
10	157.2	0.0
11	157.2	0.0

Table 19: Estimates of SLE models with coast.

Random Effects				
	Variance	Standard Deviation		
Sampling Event	0.9	0.9		
Fixed Effects				
	Estimate	Standard Error	Z value	Pr(> Z)
(Intercept)	0.0	0.6	0.0	0.9
Region-Plains	1.1	0.7	1.6	0.0
Region-Ridge	0.6	0.6	0.9	0.3
Land use-Natural	-2.9	0.9	-3.1	0.0
Vegetation height-Short	-1.0	0.5	-2.0	0.0
Vegetation height-Tall	0.9	0.5	1.7	0.0
Region:Land use-Plains:Natural	0.0	1.1	0.0	0.9
Region:Land use-Ridge:Natural	2.7	1.1	2.3	0.0

Table 20: Estimates of SLE models without coast.

Random Effects				
	Variance	Standard Deviation		
Sampling Event	1.0	1.0		
Fixed Effects				
	Estimate	Standard Error	Z value	Pr(> Z)
(Intercept)	18.9	3.5	5.3	0
Dist	-31.0	5.8	-5.2	0
Region-Ridge	-3.0	0.8	-3.6	0
Land use-Natural	-7.7	1.4	-5.3	0
Region:Land use-Ridge:Natural	8.1	1.6	5.0	0

The SLE model is much more complicated and the only model to include vegetation height. For short vegetation height the developed coast has a significantly higher proportion of vectors than natural coast; the developed coast also has a marginally significantly higher proportion than the plains natural and ridge natural. In the medium vegetation height the developed plains have a higher proportion than natural; coast natural has a significantly lower proportion of than coast

developed. In the tall vegetation height there is not distinction of land use alone, but natural sites on the coast have a lower proportion of vectors than coast developed (Figure 13).

These models vary in the list of predictors more than the diversity models, although they are mostly consistent between with- and without-coast. Dengue/Chikungunya is most predictable as a function of land use. WNV's model is very similar to the diversity models and this is most likely because the more diverse list of vectors includes species that were often the most abundant species. SLE is particularly interesting because it is the only model that includes vegetation height as a predictor.

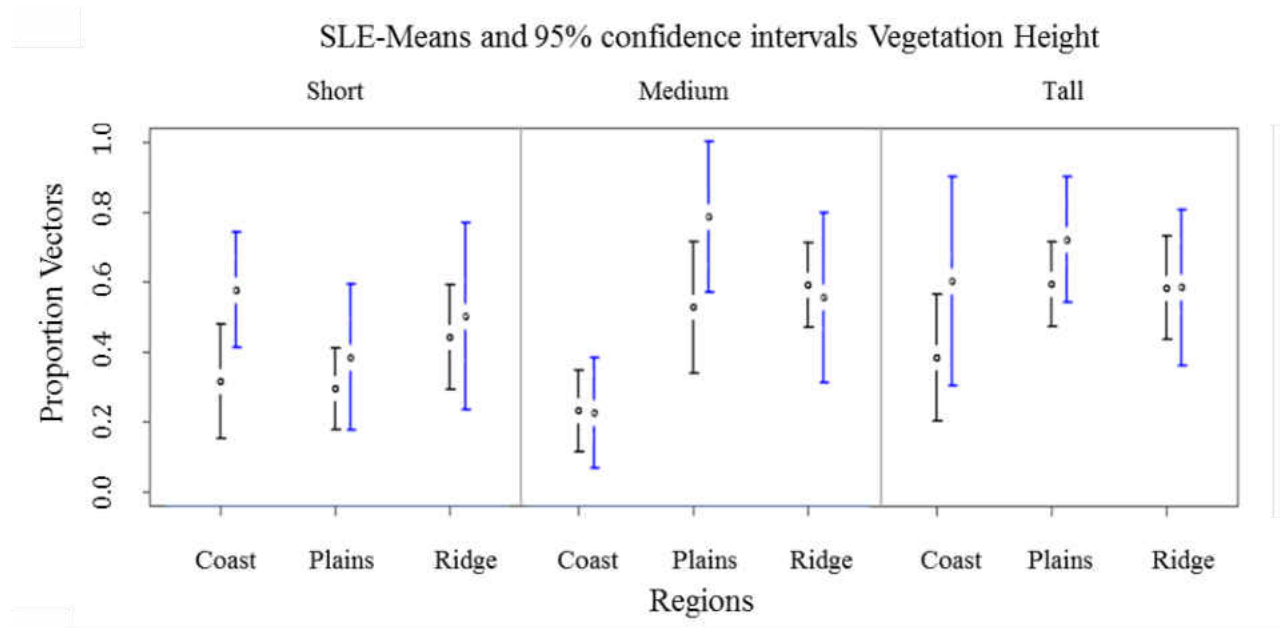


Figure 13: Vector proportions for SLE by region and land use and vegetation height (black is natural, blue is developed).

Compositional Models (NMDS)

The stressplot for NMDS ordination for Jaccard's and Bray-Curtis' indices both had a non-metric fit $R^2 = 0.983$ and a linear fit $R^2 = 0.923$ and $R^2 = 0.922$ respectively, indicating that the ordination fits the data well in ordination space (Figure 7). The final stress value was 12.99 and 13.00 respectively. The Adonis function also indicated all terms of the most complex model (region*landuse + vegetation height) were significant $\alpha=0.05$ (Table 21 and

Table 22). However, spatial autocorrelation was not included in NMDS though a partial Mantel test showed a correlation between Jaccard index and distance, as well as, Bray-Curtis index and distance (Table 21, Table 22, and Figure 14). As a result Adonis results are consistent with prior analysis but are not a definitive analysis.

Table 21: Adonis output for Jaccard's index. Bright red indicates significance with $\alpha=0.05$.

	Df	Sums Of Squares	Mean Squares	F.Model	R2	Pr(>F)
Region	2	3.7	1.8	5.2	0.0	0.0
Land use	1	2.2	2.2	6.4	0.0	0.0
Vegetation height	2	2.1	1.0	2.9	0.0	0.0
Region:Land use	2	1.8	0.9	2.6	0.0	0.0
Residuals	191	67.6	0.3	0.8		
Total	198	77.6	1			

Table 22: Adonis output for Bray-Curtis index. Bright red indicates significance with $\alpha=0.05$.

	Df	Sums Of Squares	Mean Squares	F.Model	R2	Pr(>F)
Region	2	4.2	2.1	7.1	0.06	0.0
Land use	1	2.8	2.8	9.7	0.04	0.0
Vegetation height	2	2.2	1.1	3.8	0.03	0.0
Region:Land use	2	1.9	0.9	3.2	0.02	0.0
Residuals	191	56.3	0.2	0.8		
Total	198	67.6	1			

Table 23: Partial mantel test for NMDS of Jaccard's index. Bright red indicates significance with $\alpha=0.05$.

	lag	ngroup	piecer	pval
[1,]	6500	3046	0.1	0.0
[2,]	19500	3703	0.0	0.8
[3,]	32500	914	0.3	0.0
[4,]	45500	2042	0.2	0.0
[5,]	58500	2784	-0.0	0.0
[6,]	71500	242	-0.2	0.0
[7,]	84500	770	-0.1	0.0
[8,]	97500	2556	-0.1	0.0
[9,]	110500	2627	-0.3	0.0

Table 24: Partial mantel test for NMDS of Bray-Curtis' index. Bright red indicates significance with $\alpha=0.05$.

	lag	ngroup	piecer	pval
[1,]	6500	3046	0.1	0.0
[2,]	19500	3703	0.0	0.9
[3,]	32500	914	0.3	0.0
[4,]	45500	2042	0.2	0.0
[5,]	58500	2784	-0.0	0.0
[6,]	71500	242	-0.2	0.0
[7,]	84500	770	-0.1	0.0
[8,]	97500	2556	-0.1	0.0
[9,]	110500	2627	-0.3	0.0

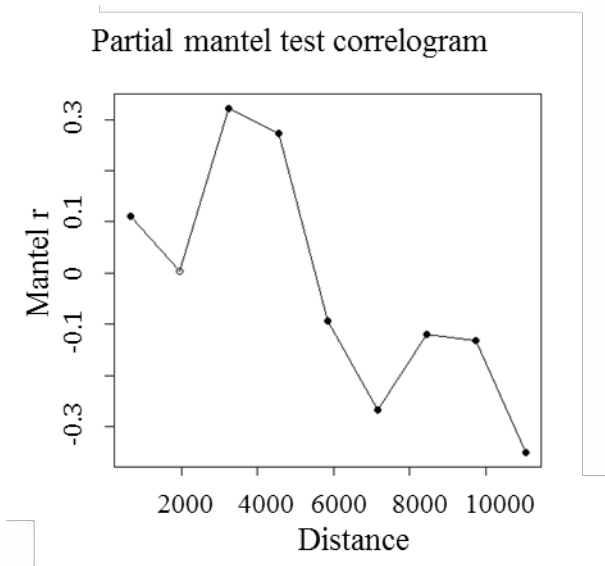


Figure 14: Partial mantel correlogram for NMDS of Jaccard and Bray-Curtis index. Closed dots indicate a significant difference ($\alpha=0.05$) and open dots indicate no significant difference.

The multivariate adonis analysis indicated that region*land use + vegetation height are significant. Ordination graphs with convex hulls around each grouping (region, land use, and vegetation height), showed substantial overlap among groups but temporal changes also were important. The ordination also shows that all sampling events occupy similar ordination space, but if you add each convex hull in sequence, you can see some of the nuances. The first event occupies a larger ordination space than the ones that follow. After that the hulls occupy approximately the same space until a noticeable shift at event 8 and 9. Finally the space shifts back ending with a hull that is of similar size as the first but shifted over. Event 8 and 9 corresponds to 2 and 4 weeks after tropical storm Isacc landed August 26, 2012. This kind of shift in composition could indicate a relatively higher risk of zoonosis often associated with storms (Ahern et al., 2005) (for all ordination graphs see appendix Figures 18-41).

Indicator Speices

Using the indicator species function in R, the coast can be characterized by three species, while the plains and ridge are characterized by one each. The combinations of plains:coast and plains:ridge have three species that characterize each. No species characterize coast:ridge combination, which is expected (Table 25). This suggests that the coast is a very different habitat to a mosquito compared to more inland regions.

Table 25: Indicator species for each region or combination of regions.

	A	B	stat	p.value
Coast				
<i>Ae_albopictus</i>	0.94	0.49	0.68	0.0
<i>Cu_restuans</i>	0.96	0.23	0.47	0.0
<i>Ae_taeniorhynchus</i>	1	0.05	0.22	0.0
<i>Ae_vexans</i>	0.97	0.03	0.18	0.2
<i>Cu_coronator</i>	0.90	0.03	0.17	0.2
<i>Ae_sollicitans</i>	0.70	0.03	0.15	0.3
<i>Ae_fulvus_pallens</i>	1	0.01	0.13	0.2
<i>Ae_tortilus</i>	1	0.01	0.13	0.2
Plains				
<i>M_dyari</i>	0.85	0.88	0.87	0.0
<i>U_sapharrina</i>	1	0.04	0.20	0.0
Ridge				
<i>Ae_atlanticus</i>	0.78	0.26	0.46	0.0
Coast and Plains				
<i>Ae_aegypti</i>	0.98	0.22	0.47	0.0
<i>Cu_declarator</i>	0.96	0.20	0.44	0.0
<i>Cu_salinarius</i>	0.99	0.17	0.42	0.0
<i>Ae_triseriatus</i>	1	0.03125	0.177	0.331
<i>C_melanura</i>	1	0.03125	0.177	0.32
Plains and Ridge				
<i>P_columbiae</i>	0.98799	0.83571	0.909	0.001
<i>P_ciliata</i>	0.99342	0.52143	0.72	0.001
<i>U_lowii</i>	0.90981	0.37143	0.581	0.001
<i>An_walkerii</i>	0.96884	0.15	0.381	0.168
<i>P_ferox</i>	0.99639	0.09286	0.304	0.793

Developed sites are characterized by three species while natural sites by ten. This suggests that developed areas have a more narrow range of habitat for mosquitoes than natural areas (Table 26).

Table 26: Indicator species for each land use type or combination of land use types.

	A	B	stat	p.value
Developed				
<i>Ae_aegypti</i>	0.98	0.29	0.53	0.00
<i>Cu_restuans</i>	0.98	0.21	0.46	0.00
<i>Cu_declarator</i>	0.74	0.22	0.40	0.01
<i>Cu_coronator</i>	0.88	0.01	0.13	0.74
Natural				
<i>An_crucians</i>	0.96	0.91	0.94	0.00
<i>M_titillians</i>	0.87	0.79	0.83	0.00
<i>P_ciliata</i>	0.95	0.57	0.73	0.00
<i>Co_perturbans</i>	0.90	0.53	0.69	0.00
<i>Ae_infirmatus</i>	0.87	0.37	0.57	0.00
<i>U_lowii</i>	0.85	0.36	0.55	0.00
<i>Ae_atlanticus</i>	0.99	0.25	0.49	0.00
<i>An_walkerii</i>	0.93	0.18	0.41	0.01
<i>An_quadrimaculatus</i>	0.98	0.16	0.40	0.03
<i>P_ferox</i>	0.99	0.14	0.38	0.00
<i>Ae_sollicitans</i>	1	0.03	0.17	0.11
<i>Ae_taeniorhynchus</i>	1	0.03	0.17	0.08
<i>C_melanura</i>	0.94	0.03	0.17	0.22
<i>A_fulvus_pallens</i>	1	0.01	0.10	0.47
<i>A_tortilus</i>	1	0.01	0.10	0.48

Medium and tall vegetation heights are characterized by 2 and 3 species respectively, while short does not have any species that explicitly characterizes it. The combination of medium/tall has one species that characterizes it (Table 27).

Table 27: Indicator species for each vegetation height or combination of vegetation height.

	A	B	stat	p.value
Medium				
<i>A_aegypti</i>	0.84	0.28	0.49	0.00
<i>Cu_restuans</i>	0.96	0.21	0.45	0.00
Short				
<i>A_fulvus_pallens</i>	1	0.01	0.12	0.34
<i>A_tortilus</i>	1	0.01	0.12	0.32
Tall				
<i>A_infirmitus</i>	0.91	0.43	0.63	0.00
<i>U_lowii</i>	0.66	0.40	0.52	0.02
<i>A_atlanticus</i>	0.93	0.24	0.47	0.00
<i>Cu_declarator</i>	0.78	0.25	0.44	0.00
<i>A_vexans</i>	0.99	0.04	0.21	0.09
<i>Cu_coronator</i>	0.89	0.03	0.16	0.41
Medium and Short				
<i>P_columbiae</i>	0.73	0.72	0.72	0.56
<i>M_fitillians</i>	0.72	0.66	0.69	0.69
<i>P_ciliata</i>	0.76	0.43	0.57	0.32
Medium and tall				
<i>Cu_nigripalpus</i>	0.83	0.94	0.88	0.03
<i>Cu_erraticus</i>	0.73	0.78	0.75	0.35
<i>An_crucians</i>	0.73	0.72	0.73	0.66
<i>M_dyari</i>	0.74	0.61	0.67	0.67
<i>Co_perturbans</i>	0.70	0.38	0.51	0.75
<i>Ae_albopictus</i>	0.79	0.27	0.47	0.32
<i>Cu_salinarius</i>	0.97	0.13	0.37	0.33
<i>P_ferox</i>	1	0.13	0.36	0.10
<i>An_quadrimaculatus</i>	0.97	0.12	0.35	0.55
<i>An_walkeri</i>	0.55	0.16	0.3	0.91
<i>A_triseriatus</i>	1	0.02	0.17	0.47
<i>C_melanura</i>	1	0.02	0.17	0.45
<i>Ae_sollicitans</i>	1	0.02	0.14	0.58
<i>Ae_teniorhynchus</i>	1	0.02	0.14	0.72
<i>U_sapharrina</i>	1	0.02	0.14	0.65

Discussion

This study established that region, land use, the interaction of region and land use, seasonality and geographic distance all significantly affect mosquito diversity. My hypothesis that seasonality, geography, and land use are large scale drivers of mosquito diversity and thus seasonality, region, and land use can predict mosquito assemblages is supported. If diversity translates to differential disease risk via the Dilution Effect, these predictor variables should be able to translate to predicting disease risk probability.

However, this study does not support a simple vector Dilution Effect (Schmidt and Ostfeld, 2001) because the study system is complex; multiple factors across regions through time/seasons, affect mosquito assemblages. In addition, multiple mosquito species can vector multiple viral diseases in Florida and nearby regions (Caribbean islands). However, we can predict areas with relative higher risk of zoonosis in general because variables predict mosquito composition and thus vectors.

This study establishes some of the important predictors of mosquito diversity in south-central Florida, but does not give effect size or estimates of these factors. A more focused study on individual predictors would be needed to obtain an effect size of each predictor variable. Species richness as a summary measure reveals little useful information about mosquito diversity overall. It does show consistency of urban homogenization, with the coast and developed land use overall having lower mean species richness (McKinney, 2006). A more detailed view of diversity is Jost D, which indicates the effective diversity and accounts for abundances. Jost D shows a starker

difference between natural and developed land use that is again consistent with urban homogenization.

The most detailed analysis (multivariate adonis) indicated that region*land use + vegetation height are significant, illustrating the value of more detailed analysis over marginal sums (i.e. species richness) to understand the system. Without this analysis this study would have suggested that vegetation height has little importance to mosquito composition, as well as revealed the effect of a major weather event on the system. It is important to incorporate these types of analysis in appropriate ways that can be descriptive and exploratory to inform future more focused studies (James and McCulloch, 1990).

Vector models differed; it is difficult to generalize about how mosquito borne disease works as a whole in the study region, but we can gain some understanding of how each disease can be predicted in the landscape. Each disease in this study had a higher proportion in either certain developed areas or in the coastal region.

Among all the varied results here, one of the most useful for evaluating mosquito-borne disease risk was indicator species analysis. Dengue and Chikungunya are vectored by *Ae. albopictus* and *Ae. aegypti* which are most closely associated with the coast or a combination of the coast and plains. Using the component parts of the indicator species analysis, every coastal sample has a probability of 0.49 containing *Ae. albopictus*, and in the combination of the coast and plains a probability of 0.23 containing *Ae. aegypti*. Assuming these probabilities can be extrapolated to the area as a whole, then any point within these regions have a probability of up to 0.72 of containing at least one individual that can vector Dengue and Chikungunya (see

component B in Table 25). It is difficult to say which of these diseases or vectors imposes a bigger risk because the relative biting rates among vectors and/or disease transfer probabilities for these diseases are not known in the study area. In addition, *Ae. aegypti* is also associated with developed sites with a probability of 0.29 containing it (see component B in Table 26). Regardless of further details on relative biting rates and disease transfer, coastal and plains regions clearly are at greater risk of Dengue and Chikungunya exposure than the ridge.

West Nile Virus is vectored by species in the *Culex* genera; three of the six species in this study are significantly associated with either the coast or the combination of coast and plains. Again, if these probabilities can be extrapolated to the area as a whole, then any point within these regions have a probability of up to 0.64 of containing at least one individual that can vector WNV. This is relatively higher than other areas, but does not mean there is no or little risk on the ridge. *Culex* species that are not associated with any area are more evenly spread throughout South-Central Florida, so depending on relative abundances those species could potentially pose a higher risk of WNV. Interestingly, *Culex* species are associated with a land use are also associated with only developed land use, giving developed areas a relatively higher risk than natural areas for WNV.

Culex nigripalpus, is the sole vector of SLE in this study but, is not associated with any particular region or land use type. It is significantly associated with the combination of short and medium vegetation height, with a staggering probability of 0.94 sites in that class containing *C. nigripalpus*. This suggests that *C. nigripalpus* selects areas with short and medium height vegetation, but can be just about anywhere.

None of the zoonosis analysis above suggests any area has no risk of zoonosis; just that those mentioned above have a relatively higher risk and should therefore be monitored more closely. As such results here should help direct attention to Florida regions that bear most attention regarding zoonosis. Also, thus prioritization was separate from human population density and considered only mosquito assemblages. For example the Ridge is not associated with any particular disease vector, yet many vectors are found there. Overall the coast and plains, and developed areas tend to have higher relative risk of zoonosis for all of the diseases evaluated here. In fact no vectors of these diseases are associated with natural sites. In addition short vegetation height has a relatively low risk of zoonosis for all of these diseases. But diversity patterns also vary among regions and land use so that greater zoonosis chance due to vector distributions is not a simple function of a Dilution Effect.

Florida had a fire maintained landscape that would once have naturally kept vegetation shorter and with larger gaps. Fire suppression in developed areas reduces the dangers of fire but an unintentional consequence may be more suitable mosquito habitat and greater zoonosis risk. Fire managed areas could not only restore the ecology of a system, but reduce the relative disease risk to humans. Thus, fire management may be considered a public health issue as well as an ecosystem issue. The most troubling part of this is that the coast and developed areas also have the highest human population coupled with the highest relative risk of zoonosis.

Region, land use, and vegetation height, contribute to the overall mosquito assemblage, which result in a variance in disease risk among those conditions. Because so many factors contribute to each of these predictors individually, a simple Dilution Effect model does not

adequately describe this system. Yet, we still have the ability to understand our relative mosquito-borne disease risk in South-Central Florida and the potential to do something about it (i.e. manage with prescribed fires).

CHAPTER 3: MOSQUITOES AND FIRE IN THE SCRUB

Abstract

Florida is well known for mosquitoes, and mosquito control has been implemented in Florida for a long time. With emerging disease (i.e. dengue and chikungunya) expanding into Florida it is important to understand the factors that play a role in adult mosquito assemblage patterns. Florida scrub is a fire maintained habitat and for the past century or so there has been extreme fire suppression throughout the United States. Archbold Biological Station (ABS) has used fire return cycles since the 1980s' to try and return its scrub habitat to a more natural state. This study sampled 84 sites for 5 months at ABS with CDC light traps and collected 9,037 individual mosquitoes and 14 species. A mixed-effect model selection approach using AICc was used in conjunction with indicator species analysis to determine how Time Since Fire (TSF) and vegetation affect mosquito assemblages at ABS on an approximately 4km scale. In general, time of season accounted for the largest proportion of the variation in the data and TSF had negligible effect on adult mosquito assemblages measured as abundance, species richness, and Jost D.

Introduction

Humans have altered even what is often considered 'natural' habitat to a great extent. In the United States fire suppression was the standard for a long time and is still common (Wagtendonk, 2007). The lack of natural fires has changed ecosystems including species that are fire-adapted (Pausas and Keeley, 2009; Weekley and Menges, 2003). In Florida, fire is a natural occurrence in scrub habitat (Abrahamson and Abrahamson, 1996), and some scrub plants require

fire to germinate or re-sprout. The natural intervals between fires in Florida varied depending on habitat from 2-5 years for sandhill to 20-59 years for rosemary scrub (Main and Menges, 1997). This variation both causes the difference in habitat as well as maintains it, therefore disturbance in these cycles cause difficulties in returning to the original state of the habitat (Abrahamson, 1984c). Fire is directly related to plant cover; the longer it has been since a fire the more plant cover there is (Abrahamson, 1984a). Because adult mosquitoes use vegetation for daily perch sites, fire should indirectly affect mosquito presence by affecting vegetation. Surprisingly, no information exists in the literature on how fire affects adult mosquitoes, directly or indirectly. After an extensive literature search I was only able to find three studies on fire effects on mosquitoes and none for adults (Janousek and Olson, 1994; Wallace et al., 1990; Whittle et al., 1993). In addition, little is also known about mosquito assemblages in Florida scrub, probably because scrub habitat is generally considered dry and historically moderately populated by humans (which is changing with urban sprawl). However, many adult mosquitoes inhabit scrub habitats, and many potential breeding sites exist there. Understanding how fire affects mosquitoes in the scrub could inform prescribed fire management and help support its use in areas where fire is still suppressed (e.g. near urban land use).

Some land management areas have a strong record of attempting to mimic natural fire cycles. Archbold Biological Station (ABS) in south-central Florida (Highlands county) is one such place. After a short experimental period, ABS began a fire management plan in the early 1980's, which includes a patch work system of fire units that differ in fire return intervals. Using fire return intervals instead of strict cycles helps to allow for better success of the system and

flexibility due to natural events that may occur at the station. As a result, ABS now includes a mosaic of different fire histories that is close to natural systems and rarely has any two adjacent units with the exact same fire history (Main and Menges, 1997). Because this system has been so closely studied and monitored it is a great setting to understand how any set of organisms respond to fire or fire effects.

Not much is known about mosquitoes at ABS and majority of the information is from recent studies. No study has tested how fire affects mosquitoes. Depending on the species, adult mosquitoes may migrate in the first few days of life anywhere from 0.5 km up to 3+ km against the wind (Bates, 1949), meaning that they are free to select suitable habitats among a variety of fire-affected units at ABS. Mosquitoes fly in three ways, migratory, appetential and consumatory. Migratory flight lacks a goal, has no return flight, and only occurs in newly emerged adults. Appetential flight occurs after the first 24 hrs of emergence, in response to physiological stimulus, such as searching for a perch site, breeding site, food source, etc. Consumatory flight often, but not always, occurs after appetential flight and is triggered by visual and biochemical cues in a short flight from a perch site, to a food source or breeding site (Bidlingmayer, 1985, Provost, 1952). All of these flight patterns can be affected by a wide range of variables. Given the complexity of mosquito flight behaviors and the mosaic of burn units at ABS, the study system provided a unique opportunity to investigate the effects of fire management on mosquito habitat selection in Florida scrub.

The two ideas of what impacts adult mosquito distributions are perch site or breeding site preferences and food availability. Each compete to explain which cues have the biggest impact

on dictating if an adult mosquito is found in any particular place. Each of these factors is affected by fire which leads me to my second question: How does Time Since Fire (TSF) affect adult mosquito assemblages at Archbold Biological Station?

I addressed this question by a sampling of adult mosquitoes using a CDC light trap at 84 sites over 5 months, at ABS, with 3 different TSF categories: short (<1yr), medium (1-5yrs), and long (>5yrs). Mosquitoes may respond to fire and its effects at different scales due to the patchy nature of fire, therefore several predictors were selected to determine the most plausible model of mosquito assemblages at ABS. Using a mixed-effects model approach we collected data on each site for Time Since Fire (TSF), Gap Light Analysis (GLA-a measurement of the percent canopy cover) at different heights, 7 percent vegetation categories, 5 of those had the height of the tallest individual in that category recorded, the first 3 Principle Components (PC) of a Principle Components Analysis (PCA) for all 12 vegetation categories), and a distance coefficient (Dist- a distance weighted auto-covariate). Several alternative hypotheses were formed with sets of each of these predictors that were then compared using model selection with AICc. The response variable as a function of 'Dist': time since fire, percent canopy cover at 0m, 1m, and 2m, percent vegetation categories, tallest vegetation categories, vegetation categories PCA, and random effects.

My hypotheses are:

Response Variable ~ *Dist* + *TSF* + (*Random effects*)

(10)

$$\text{Response Variable} \sim \text{Dist} + \text{Percent Canopy Cover} + (\text{Random effects}) \quad (11)$$

$$\text{Response Variable} \sim \text{Dist} + \text{precent vegetation} + (\text{Random effects}) \quad (12)$$

$$\text{Response Variable} \sim \text{Dist} + \text{tallest vegetation} + (\text{Randome effects}) \quad (13)$$

$$\text{Response Variable} \sim \text{Dist} + \text{Vegetation PCA} + (\text{Random effects}) \quad (14)$$

$$\text{Response Variable} \sim \text{Dist} + (\text{Random effects}) \quad (15)$$

Methods

Study Sites

This entire study was conducted at ABS, located in south-central Florida at the southern end of the Lake Wales Ridge. The Lake Wales Ridge is part of the Florida ridge system which is comprised of upland areas that have remained above sea-level since the Pleistocene (James, 1961; Myers and Ewel, 1990-pgs 70-100). The ridge is home to many endemic species that have been isolated on ancient islands. Some of the different natural habitats found on the southern portion of the ridge include several types of scrub, flatwoods, southern ridge sandhill, and bayheads (Abrahamson et al., 1984a). Much of the natural areas of the ridge have been cultivated

for citrus and cattle pasture. The station was farmed for turpentine and lumber in the past but has been managed for biological research since 1941 (Main and Menges, 1997).

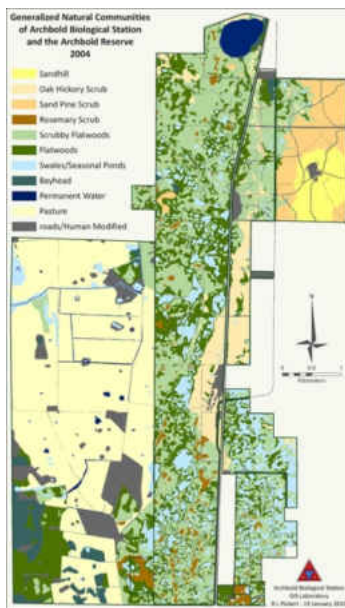


Figure 15: A Generalized map of Natural communities of ABS from 2004.

Most sites were located in scrub or scrubby flatwoods, but some were in rosemary balds. This area contains many ephemeral ponds and is dominated by typical scrub plants including: palmettos, a mixture of oaks and sparse slash pines. The short TSF sites, are most often still covered in charcoal, with stark green newly sprouting plants, with little to no canopy or sub-canopy. The medium TSF sites, have denser ground cover, some have sub-canopy and sparse canopy. The long TSF varied much more greatly; often these sites had thick sub-canopy and canopy, but if located in or near a rosemary bald, then there was no canopy.

Sampling

Eighty-four sites were sampled using John C. Hock light traps. Each trap was placed a minimum of 50m from any road or fire lane to avoid edge effects. The sampling sites were chosen to be associated with a wetland for mosquito emergence sampling, those data are not considered here. Sites can thus be considered to be haphazardly distributed in burn units of the three different TSF categories: short (<1yr), medium (1-5yrs), and long (>5yrs) (Abrahamson and Abrahamson, 1996) (Figure 16). Proximity to a wetland was irrelevant in this sampling year because extensive flooding occurred; wetlands were combined and "uplands" were flooded for extended periods. No trap was placed in standing water. Twelve traps were placed out during any one sampling event, 4 in each TSF category, and no site was ever sampled twice. Sampling occurred from July 2013 until November 2013 when mosquitoes trapped had dwindled to very low numbers, for a total of 7 events. Each female mosquito was identified to species when possible.

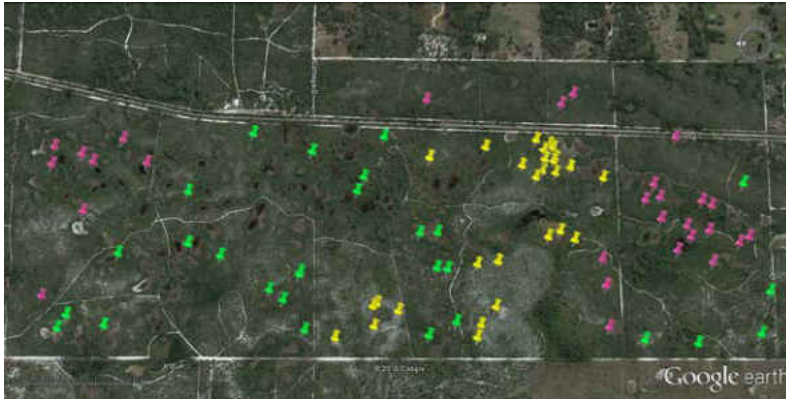


Figure 16: Map of the sampling sites at Archbold Biological Station. Yellow-short TSF, pink -medium TSF, green - long TSf.

Each sample site included a light trap, hung ~1m above the ground in the center of a 30m diameter macro-plot. Traps were set out just before sunset and retrieved the next morning. The samples were then frozen to kill all invertebrates trapped inside and remained frozen prior to processing. All non-mosquito invertebrates were excluded and then the sample was weighed to the nearest 0.0001 g.

Each light trap was at the center of a macroplot (30 m radius) (A CDC light trap has an estimated non-baited effective distance of approximately 5 m (Odetoyinbo, 1969)), and three micro-plots (5 m in diameter) were randomly placed within each macro-plot using a random direction and random distance. Each micro-plot had 3 subplots selected haphazardly by throwing a 1m diameter circular PVC frame (Figure 17). Vegetation was evaluated to indicate the relative number of perch and vegetative structure sites in that area. The percent cover of palmetto, grasses, herbs, sub-shrub (≤ 50 cm tall), woody shrubs (≥ 50 cm and ≤ 200 cm, including scrub oak, live oak, *Lyonia lucida*, etc.), bare sand, and litter were estimated in each subplot. The

tallest height (cm) of each of those categories was also recorded when applicable (5 categories). At each subplot a vertically –oriented picture was taken at 0, 1, and 2 meters using a 180° fish eye lens. Digital photographs were analyzed with Gap Light Analysis software where data was recorded as percent canopy cover as an indicator of perch site availability and relative trap visibility (Frazer et al., 2001). Data from the nine subplots around a sampling point were averaged to represent vegetative structure of the area around the sampling point.

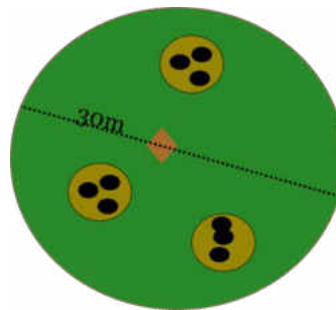


Figure 17: Macro-plot (in green) with trap in center (diamond). Three 5m diameter micro-plots (in brown) with 3 1m diameter subplots in black. Vegetation and GLA data was taken at each black subplot.

Light trap samples generated abundance and presence/absence data for multiple mosquito species, so that data comprised a species X sites X time matrix. Those data were evaluated in several ways, summarized below.

Statistical Analysis

I used a hybrid of methods from Zuur et. al (2009) and model selection with AICc and weights (Burnham and Anderson, 2002) on all models. Using the same fixed effects structure,

the random structure was determined using Restricted Maximum Likelihood (REML), with possible random structures composed of none, random intercept only, or random slope and intercept. Sampling event was treated as the random intercept to address the repeated measures design. Geographical distance between sampling points was addressed as a distance-weighted auto-covariate (Augustin, Muggleston, and Buckland. 1996; Bivand et. al., 2014) and was treated as both a fixed effect and the random slope.

Models were then validated using the function 'r.squaredGLMM' (MuMIn, 2015) to obtain pseudo R squared values. This function provides a marginal R squared value which represents variance of the fixed effects, and a conditional R squared value that represents the variance of both fixed and random effects. Graphs were created to visually inspect the assumptions of homogeneity of factors and normal residuals. Both the Shapiro-Wilk test (Shapiro and Wilk, 1965; R core team, 2014) and the Levene test (Levene, 1960; Fox and Weisberg, 2011) were employed to evaluate parametric statistical assumptions. If models did not meet these assumptions then the model was rejected. In order to interpret the models, means and 95% confidence intervals were obtained for all significant factors.

Models were assessed using three different response variables: abundance, species richness and Jost D, or the effective diversity of a sample (Jost, 2006) and was calculated as:

$$Jost D = e^{H'} \quad (16)$$

where H' is Shannon's diversity:

$$H' = - \sum_i p_i \ln p_i \quad (17)$$

where p_i is the proportional abundance of species i .

Predictors were fixed and/or random effects for 5 different sets of variables: Time Since Fire (TSF), canopy cover, vegetation (percent cover predictors), vegetation (tallest predictors), and Principle Components Analysis (PCA) of vegetation. All fixed effects portions were analyzed with and without the distance coefficient when possible.

For model sets with many possible models the function 'expand.grid' (R core team, 2014) was used to obtain a matrix of all possible models in conjunction with excel. Each model set had a different number of possible models, for which the most plausible was derived with AICc. Then each of the most plausible was included in a final comparison to find the most plausible model of all the sets. Each sample was one of the three (short, medium, or long) TSF categories creating 3 possible models. Percent canopy cover models consisted of three continuous variables derived from the GLA pictures averaged at each point at 0 m, 1 m and 2 m. All interactions were included in the most complex canopy cover model which makes 127 models. Vegetation (percent cover) consisted of the 7 continuous predictors, listed above; only additive effects were evaluated for vegetation which makes 127 possible models. Vegetation (tallest) consisted of 5 continuous predictors; only additive effects were evaluated for vegetation which makes 31 possible models. Also to simplify the vegetation predictors, PCA was used with the function 'prcomp' (R core team, 2014). The first 3 axes were used as 3 continuous variables yielding 8 models. These axes

explain 73.3% of the variation in the vegetation variables (Table 28). Only additive effects of PCA axis scores were included, because all vegetation variables are included in each axis. Finally, the most plausible model of each of these sets and a model with only geographic distance coefficient were compared using AICc to determine which predictors gave the most plausible model.

Table 28: Vegetation PCA loadings.

	PC1	PC2	PC3
Standard Deviation	2.01	1.62	1.46
Percent Variation	34	22	18
Cumulative Variation	34	55	73
Eigenvalues	4.04	2.61	2.15
Percent grass	0.09	0.08	0.12
Tallest grass	0.14	0.00	0.11
Percent subshrub	0.05	0.20	0.08
Tallest subshrub	0.04	0.21	0.05
Percent trees	0.09	0.07	0.16
Tallest trees	0.15	0.04	0.06
Percent herbs	0.05	0.10	0.00
Tallest herbs	0.15	0.01	0.00
Percent palmetto	0.05	0.11	0.16
Tallest palmetto	0.08	0.11	0.12
Percent sand	0.05	0.07	0.11
Percent litter	0.06	0.00	0.04

The R-package *indicspecies* (DeCaceres and Legendre, 2009) uses the group Indicator value (DeCaceres, Legendre, and Moretti, 2010) to produce a set of species that indicate a particular environmental variable, with a permutation test to test significance. Groups most consistent to

the highest association, of each category or combination of categories of TSF were obtained with 999 permutations. Species component parts of the indicator value, were used to determine species fidelity (A), and species specificity (B) (DeCaceres, Legendre, and Moretti, 2010). A species was considered an indicator species at an alpha of 0.05, but all species associations were reported.

All analysis was completed in R and all code is provided in the appendix.

Results

Eighty-four samples were collected over 5 months and consisted of 9,037 individual mosquitoes and 14 species including: *Oclerotatus atlanticus*, *Oclerotatus infirmatus*, *Anopheles crucians*, *Anopheles quadrimaculatus*, *Mansonia dyari*, *Mansonia titillans*, *Coquillettidia perturbans*, *Culex erraticus*, *Culex nigripalpus*, *Culex declarator*, *Psorophora ciliata*, *Psorophora columbiae*, *Uranotaenia lowii*, and *Uranotaenia sapphirina*.

Mosquito Abundance

The most plausible model for abundance by TSF includes only TSF and the random intercept and the weight is nearly triple the next most plausible model:

(log) *Abundance* ~ *TSF* + (1|*Sampling event*) (Table 29)

(18)

Table 29: AICc comparison of Abundance ~ TSF models.

	AICc	dAICc	df	weight
TSF + (1 Sampling event)	213.7	0	5	0.59
Dist + TSF + (1 Sampling event)	215.7	2	6	0.22
Dist + (1 Sampling event)	216	2.3	4	0.19

The most plausible model for abundance by percent canopy cover includes the percent canopy at 1m and random intercept, but the weight is not drastically higher than the next most plausible model:

$$(\log) \text{ Abundance} \sim - \% \text{ Canopy at } 1\text{m} + (1|\text{Sampling event}) \text{ (Table 30)}$$

(19)

Table 30: AICc comparison of Abundance ~ GLA top 5 models.

	AICc	dAICc	df	weight
% canopy1m + (1 Sampling event)	207.9	0	4	0.07
% canopy1m: % canopy2m + (1 sampling_event)	208.3	0.4	4	0.05
% canopy0m + % canopy0m: % canopy1m + (1 sampling_event)	209.7	1.8	5	0.03
% Canopy2m + % canopy1m: % canopy2m + (1 sampling_event)	209.9	2	5	0.02
% canopy0m: % canopy2m + % canopy0m: % canopy1m: % canopy2m + (1 sampling_event)	209.9	2	5	0.02

The most plausible model for abundance by percent vegetation categories includes percent grass cover, percent cover of woody shrubs and random intercept and the weight is nearly double the next most plausible model:

$$(\log) \text{ Abundance} \sim \% \text{ grass} + \% \text{ woody shrubs} + (1|\text{Sampling event}) \text{ (Table 31)}$$

(20)

Table 31: AICc comparison of Abundance ~ percent vegetation categories top 5 models.

	AICc	dAICc	df	weight
% grass + % woody shrubs + (1 Sampling event)	204.3	0	5	0.11
% woody shrubs + (1 Sampling event)	205.2	0.9	4	0.07
% woody shrubs + % palmetto + (1 Sampling event)	206	1.8	5	0.04
% grass + % woody shrubs + % palmetto + (1 Sampling event)	206.3	2.1	6	0.04
% grass + % woody shrubs + % sand + (1 Sampling event)	206.4	2.1	6	0.04

The most plausible model for abundance by tall vegetation categories includes the tallest woody shrub, tallest palmetto, and random intercept and the weight is nearly double the next most plausible model:

$(\log) Abundance \sim \text{tallest woody shrub} - \text{Tallest palmetto} + (1|sampling\ event)$ (Table 32)

(21)

Table 32: AICc comparison of Abundance ~ tall vegetation categories top 5 models.

	AICc	dAICc	df	weight
Tallest woody shrub + tallest palmetto + (1 sampling event)	202.2	0	5	0.27
Tallest grass + tallest woody shrub + tallest palmetto + (1 sampling event)	203.2	1	6	0.16
Tallest subshrub + tallest woody shrub + tallest palmetto + (1 sampling event)	203.4	1.3	6	0.14
Tallest tree + tallest herb + tallest palmetto + (1 sampling event)	204.5	2.3	6	0.08
Tallest grass + tallest subshrub + tallest woody shrub + + tallest palmetto + (1 sampling event)	204.8	2.6	7	0.07

The most plausible model for abundance by vegetation PCA includes PC1, PC2, PC3 and random intercept but the weight is not much more than the next most plausible model:

$(\log) Abundance \sim - PC1 - PC2 - PC3 + (1|Sampling\ event)$ (Table 33)

(22)

Table 33: AICc comparison of Abundance ~ vegetation PCA top 5 models.

	AICc	dAICc	df	weight
PC1 + PC2 + PC3 + (1 sampling event)	206.4	0	6	0.30
PC2 + PC3 + (1 sampling event)	206.8	0.4	5	0.25
Dist + PC1 + PC2 + PC3 + (1 sampling event)	207.7	1.3	7	0.16
PC1 + PC2 + (1 sampling event)	208	1.6	5	0.14
PC2 + (1 sampling event)	208.3	1.9	4	0.12

Among the above models for mosquito abundance, the most plausible is the tall vegetation with an AICc 202.2 of and weight of 0.65 which is 3 times higher than the next most plausible model (Table 34). Notably here geographic distance did not contribute to, any of the plausible models. Based off the 95% confidence intervals all terms in the tall vegetation category model are significantly different from zero (Table 35). Although this model shows significant predictors the pseudo R^2 values indicate that nearly all of the variation explained in the model is due to sampling event, the random factor. The variation explained by the predictors is only 2.6%, while the combined fixed and random factors explain 88.5% of the variation.

Table 34: AICc comparison of most plausible model from each model set.

	dAICc	df	weight
Tallest woody shrub + tallest palmetto + (1 sampling event)	0	5	0.65
% grass + % woody shrubs + (1 Sampling event)	2.1	5	0.22
PC1 + PC2 + PC3 + (1 sampling event)	4.3	6	0.07
% canopy 1m + (1 Sampling event)	5.8	4	0.03
TSF + (1 Sampling event)	11.5	5	0.00
Dist only	13.8	4	<0.00

Table 35: 95% confidence intervals for the most plausible model for abundance (tall vegetation category model).

	2.5%	97.5%
SD Sampling event	0.839	2.937
Residual Standard Deviation	0.531	0.733
Intercept	1.810	4.690
Tallest woody shrub	0.003	0.008
Tallest palmetto	-0.011	-0.001

Mosquito Species Richness

The most plausible model for species richness by TSF only includes the distance coefficient and random intercept and the weight is over 5.5 times greater than the next most plausible model:

Species Richness ~ *Dist* + (1|*Sampling event*) (Table 36)

(23)

Table 36: AICc comparison of species richness ~ TSF models.

	AICc	dAICc	df	weight
Dist + (1 sampling event)	294.7	0	4	0.79
Dist + TSF (1 sampling event)	298.2	3.5	6	0.14
TSF + (1 sampling event)	299.7	5	5	0.06

The most plausible model for species richness by percent canopy cover includes the distance coefficient and the interaction between the percent canopy cover at 1m and 2m and random intercept but the weight is only slightly higher than the next most plausible model:

Species Richness ~ - *Dist* - %*canopy1m*: %*canopy2m* + (1|*Sampling event*) (Table 37)

(24)

Table 37: AICc comparison of specie richness ~ GLA top 5 models.

	AICc	dAICc	df	weight
Dist + %canopy1m:%canopy2m + (1 sampling event)	296.3	0	5	0.05
Dist + %canopy1m + (1 sampling event)	296.3	0	5	0.05
Dist + %canopy0m:%canopy1m:%canopy2m + (1 sampling event)	296.4	0.1	5	0.05
Dist + %canopy0m:%canopy1m + (1 sampling event)	296.5	0.2	5	0.04
Dist + %canopy0m:%canopy2m + (1 sampling event)	296.5	0.3	5	0.04

The most plausible model for species richness by percent vegetation categories includes only percent grass and random intercept but the weight is not drastically higher than the next most plausible model:

$$\text{Species richness} \sim \% \text{grass} + (1|\text{sampling event}) \text{ (Table 38)}$$

(25)

Table 38: AICc comparison of species richness ~ percent vegetation categories top 5 models.

	AICc	dAICc	df	weight
%grass + (1 sampling event)	295.2	0	4	0.06
%grass + %woody trees + (1 sampling event)	295.7	0.4	5	0.05
%palmetto + (1 sampling event)	296.9	1.7	4	0.02
%grass + %palmetto + (1 sampling event)	296.9	1.7	5	0.02
%grass + %subshrub + (1 sampling event)	297	1.8	5	0.02

The most plausible model for species richness by tall vegetation categories includes distance coefficient, tallest grass, and tallest woody shrub and random intercept, but the weight is not much higher than the next most plausible model:

$$\text{Species Richness} \sim - \text{Dist} + \text{Tallest grass} + \text{Tallest woody shrub} + (1|\text{Sampling event}) \text{ (Table 39)}$$

(26)

Table 39: AICc comparison of species richness ~ tall vegetation categories top 5 models.

	AICc	dAICc	df	Weight
Dist + tallest grass + tallest woody shrub + (1 sampling event)	294.3	0	6	0.13
Dist + tallest subshrub + (1 sampling event)	294.9	0.7	5	0.09
Dist + tallest grass + tallest subshrub + tallest woody shrub + (1 sampling event)	295.9	1.6	7	0.06
Dist + tallest grass + tallest subshrub + (1 sampling event)	296	1.7	6	0.05
Dist + tallest grass + (1 sampling event)	296.2	1.9	5	0.05

The most plausible model for species richness by vegetation PCA includes the distance coefficient and PC2 and the random intercept and the weight is nearly three times higher than the next most plausible model:

$$\text{Species Richness} \sim - \text{Dist} - \text{PC2} + (1|\text{Sampling event}) \text{ (Table 40)}$$

(27)

Table 40: AICc comparison of species richness ~ vegetation PCA categories top 5 models.

	AICc	dAICc	df	weight
Dist + PC2 + (1 sampling event)	294.6	0	5	0.41
Dist + PC2 + PC3 + (1 sampling event)	296.6	2	6	0.15
Dist + PC1 + PC2 + (1 sampling event)	296.9	2.3	6	0.13
Dist + PC1 + (1 sampling event)	297	2.4	5	0.12
Dist + PC1 + PC2 + PC3 + (1 sampling event)	298.9	4.3	6	0.04

Among the species richness models, the most plausible is the tall vegetation category model, but the vegetation PCA and the TSF model are not much less plausible based on the AICc weights (Table 41). Notably the TSF model is actually the same as the distance only model therefore this model was only used once in the AICc sets comparison. In the tall vegetation category model both the coefficients for tallest grass and tallest woody shrub were significantly

different from zero but distance was not (Table 42). The vegetation PCA and TSF models both have no coefficients that significantly differ from zero. The pseudo R^2 for the tall vegetation category model for the fixed effects is 0.05 and for the fixed and random effects together was 0.54, indicating that the majority of the variation was again due to sampling event.

Table 41: AICc comparison of the most plausible models for each set for species richness.

	AICc	dAICc	df	weight
Dist + tallest grass + tallest woody shrub + (1 sampling event)	294.3	0	6	0.30
Dist + PC2 + (1 sampling event)	294.6	0.3	5	0.26
Dist + (1 sampling event)	294.7	0.4	4	0.24
Dist + %canopy1m:%canopy2m + (1 sampling event)	296.3	2	5	0.11
%grass + (1 sampling event)	296.9	2.6	4	0.08

Table 42: 95% confidence intervals for the most plausible model for species richness (tall vegetation category model).

	2.5%	97.5%
SD sampling event	0.48	2.02
Residual Standard deviation	0.99	1.37
Intercept	1.59	11.02
Dist	-2.63	0.12
Tallest grass	0.00	0.02
Tallest woody shrub	0.00	0.01

Mosquito Jost D

The most plausible model for Jost D by TSF only includes the distance coefficient and the random intercept and the weight is nearly twice the next most plausible model:

$$(\log) Jost D \sim - Dist + (1|Sampling event) \text{ (Table 43)}$$

(28)

Table 43: AICc comparison of jost D ~ TSF models.

	AICc	dAICc	df	weight
Dist + (1 sampling event)	62.3	0	4	0.56
Dist + TSF + (1 sampling event)	63.3	1	6	0.33
TSF + (1 sampling event)	65.5	3.2	5	0.11

The most plausible percent canopy model for Jost D includes the distance coefficient and the interaction between the percent canopy at 0m and 1m and the random intercept, but the weight is only slightly better than the next most plausible model:

$$(log)Jost D \sim - Dist + \%Canopy0m:\%Canopy1m + (1|Sampling event) \text{ (Table 44)} \quad (29)$$

Table 44: AICc comparison of jost D ~ GLA top 5 models.

	AICc	dAICc	df	weight
Dist + %canopy0m:%canopy1m + (1 sampling event)	61.6	0	5	0.05
Dist + %canopy0m + (1 sampling event)	61.6	0	5	0.05
Dist + %canopy0m:%canopy2m + (1 sampling event)	61.8	0.2	5	0.05
Dist + %canopy0m:%canopy1m:%canopy2m + (1 sampling event)	61.9	0.3	5	0.05
Dist + %canopy0m:%canopy1m + %canopy0m:%canopy1m:%canopy2m + (1 sampling event)	63.3	1.7	6	0.02

The most plausible percent vegetation categories model for Jost D includes the distance coefficient, percent grass, and percent woody shrubs and random intercept, but the weight is only slightly higher than the next most plausible model:

$$(log)Jost D \sim - Dist + \%grass - \%woody shrub + (1|sampling event) \text{ (Table 45)} \quad (30)$$

Table 45: AICc comparison of jost D ~ percent vegetation categories top 5 models.

	AICc	dAICc	df	weight
Dist + %grass + % woody shrub + (1 sampling event)	51.9	0	6	0.08
Dist + %grass + (1 sampling event)	52.2	0.3	5	0.06
Dist + %grass + %palmetto + (1 sampling event)	52.5	0.6	6	0.06
Dist + %grass + %subshrub + %woody shrubs + (1 sampling event)	53.5	1.6	7	0.03
Dist + %grass + % woody shrub + %palmetto + (1 sampling event)	53.6	1.7	7	0.03

The most plausible model for the vegetation tallest categories for Jost D model includes tallest grass and tallest subshrub and random intercept but again the weight is only slightly higher than the next most plausible model:

$$(log)Jost D \sim - Dist + Tallest grass - Tallest subshrub + (1|Sampling event) \text{ (Table 46)}$$

(31)

Table 46: AICc comparison of jost D ~ tallest vegetation categories top 5 models.

	AICc	dAICc	df	weight
Dist + tallest grass + tallest subshrub + (1 sampling event)	51.9	0	6	0.18
Dist + tallest grass + (1 sampling event)	52.6	0.7	5	0.13
Dist + tallest grass + tallest subshrub + tallest palmetto + (1 sampling event)	53.1	1.1	7	0.10
Dist + tallest grass + tallest subshrub + tallest herb + tallest palmetto + (1 sampling event)	53.7	1.8	8	0.07
Dist + tallest grass + tallest subshrub + tallest herb + (1 sampling event)	53.7	1.8	7	0.07

The most plausible PCA model for Jost D includes the distance coefficient, PC1, PC3 and random intercept and the weight is 2.5 times higher than the next most plausible model:

$(\log)Jost D \sim -Dist + PC1 + PC3 + (1|Sampling\ event)$ (Table 47)

(32)

Table 47: AICc comparison of jost D ~ vegetation PCA top 5 models.

	AICc	dAICc	df	weight
Dist + PC1 + PC3 + (1 sampling event)	51	0	6	0.58
Dist + PC1 + PC2 + PC3 + (1 sampling event)	52.8	1.8	7	0.23
Dist + PC3 + (1 sampling event)	54.5	3.4	7	0.10
Dist + PC2 + PC3 + (1 sampling event)	56.1	5	6	0.04
Dist + PC1 + (1 sampling event)	58.2	7.1	6	0.01

Among the Jost D models the most plausible is the vegetation PCA model with a weight of 0.58 which is 1.5 times higher than the next most plausible model (Table 48). Notably the TSF model is actually the same as the distance only model therefore this model was only used once in the AICc sets comparison. Also, the top four models for GLA had similar weights and were therefore all included in the final comparison. Based off the 95% confidence intervals for the vegetation PCA model all fixed effects are significantly different from zero (Table 49). The pseudo R^2 for the vegetation PCA model for the fixed effects is 0.11 and for the fixed and random effects together is 0.55, indicating that the majority of the variation is again due to sampling event.

Table 48: AICc comparison of the most plausible models for each set for Jost D.

	AICc	dAICc	df	weight
Dist + PC1 + PC3 + (1 sampling event)	51	0	6	0.43
Dist + tallest grass + tallest subshrub + (1 sampling event)	51.9	0.9	6	0.27
Dist + %grass + % woody shrub + (1 sampling event)	51.9	0.9	6	0.27
Dist + %canopy0m:%canopy1m + (1 sampling event)	61.6	10.6	5	0.00
Dist + %canopy0m + (1 sampling event)	61.8	10.8	5	0.00
Dist + %canopy0m:%canopy2m + (1 sampling event)	61.9	10.9	5	0.00
Dist + (1 sampling event)	62.3	11.3	4	0.00

Table 49: 95% confidence intervals for the most plausible model for Jost D (Vegetation PCA model).

	2.5%	97.5%
SD sampling event	0.11	0.47
Residual standard deviation	0.23	0.32
Intercept	0.89	2.48
Dist	-3.00	-0.57
PC1	0.00	0.06
PC3	0.03	0.11

In summary, a few recurring themes occurred in the variety of plausible models. Sampling event as a random intercept term (i.e., repeated measures model) was always included in the most plausible model and accounted for a majority of variation. Other vegetation-related variables accounted for less variation but were significant. Woody shrub was always included in the most plausible percent vegetation category models. The most plausible model of the sets always includes vegetation in some way, either the vegetation model or the PCA vegetation model.

Indicator Species

Using the indicator species function in R, only (*Psorophora ciliata*) significantly indicated one category, short TSF, all other indicators are not significant (Table 50). This set of results

indicates that mosquito species and TSF are not well correlated, and is consistent with relatively weak TSF and vegetation effects in analyses above.

Table 50: Indicator species values for each TSF. Red indicates significantly different from zero.

	A	B	stat	p.value
Long TSF				
<i>U.sappharina</i>	1	0.07	0.26	0.31
Medium TSF				
<i>A.quadrinaculatus</i>	1	0.03	0.18	1
<i>C.declarator</i>	1	0.03	0.18	1
Short TSF				
<i>P.ciliata</i>	0.74	0.25	0.43	0.01
Long and Medium TSF				
<i>O.atlanticus</i>	1	0.14	0.37	0.11
<i>O.infirmatus</i>	0.93	0.08	0.28	0.49
Medium and Short TSF				
<i>P.columbiae</i>	0.98	0.23	0.47	0.15
<i>C.perturbans</i>	1	0.03	0.18	1

Discussion

This study establishes that the fire management regime at Archbold Biological Station, as measured by TSF and related vegetation variables, has little effect on mosquito assemblages. Time or season has the largest effect on mosquito assemblages in this study, and mosquito assemblages are not spatially patterned within the study area. Nearly all the variation in the models was due to the random effect of sampling event and only a single species was a significant indicator of only one TSF category or combination.

This is the first study of adult mosquito assemblages in a scrub habitat. Other studies have demonstrated some distribution patterns for mosquitoes at comparable spatial scales, but these

have been in habitats with more topographic variation (Mutuari et al., 2006). Adult mosquito assemblages within ABS seem to be well-mixed with a spatial extent of approximately 4km. The previous chapter demonstrated that there was some differentiation of mosquito assemblages within a spatial extent of approximately 110km. This indicates that the mosquitoes respond to spatial differentiation on a scale somewhere between these two distances, in south-central Florida. This spatial differentiation will probably change in different locations.

Also, there is no evidence from these results here that fire management is expected to have a substantial local affect on adult mosquito distributions or abundances via effects on vegetation. Therefore, fire management in exurban and rural areas cannot be expected to modify exposure to mosquitoes or risk of exposure to mosquito-borne disease. Other factors (i.e. host availability or hydrology) may be a more fruitful path of study as modifiers of mosquito composition and abundance. In this study year, much of the study area was flooded creating some unique habitat homogeneity, which may have caused some of the mosquito assemblage homogenization. A study of similar nature in a drier year may yield greatly different results.

There has been a debate on whether, vegetation or host availability is more important to mosquitoes (Burkett-Cadena et. al., 2013). Because there was no major effect of vegetation, it is unlikely that vegetation is playing a major role on mosquito assemblage at this scale. Host availability was not described in this study, but may play a larger role for mosquitoes in a scrub habitat. The weak role of vegetation here may be due to the fact that most scrub plants are xerophytic, with small leaves that provide little shade during the day. Another factor may be the brightness of the sand in this scrub habitat and sparse forest canopy. Even an area with more

vegetative structure, the sandy soil reflects so much light, that mosquitoes may not detect strong differences among TSF units.

It is often thought that Florida doesn't have much seasonality, but Florida has distinct wet and dry seasons, and ABS has extreme hydrological differences between seasons. Within only a week or so, it can go from a dry desert-like habitat, to nearly a swamp where there is almost no differentiation between wetlands and much of the "upland" areas are flooded. As a result the strong effect of season on mosquito assemblages, and the relatively weak effect of TSF and vegetation were observed here.

Unfortunately, this study does not provide evidence for using fire to help control mosquitoes or diseases vectored by mosquitoes at this scale. If much larger fire units were implemented then there is a greater possibility that mosquitoes may respond to TSF. This study does not address the immediate effects of fire on mosquitoes, which could be great. I would suggest a study investigating before and after difference in mosquito assemblage in this habitat, that could show how fire could be used as a short term fix for mosquito control. This type of control could be implemented just before or just after an event (i.e. hurricane) that would increase mosquito abundance and risk of mosquito vectored diseases.

CHAPTER 4: CONCLUDING REMARKS ON SOUTH-CENTRAL FLORIDA MOSQUITOES

The first study here was a large scale analysis of mosquito distributions and diversity in a Dilution Effect framework, based on several spatial-environmental factors including: region, land use, vegetation height, and time. A mixed-model approach was implemented in conjunction with indicator species, with emphasis on disease vector species. It was discovered that different measures of diversity are sensitive to different predictors, although not with great variation. There was no evidence for a simple vector Dilution Effect in play in this study. The coast seems to have the greatest risk of disease, which is alarming because this is also where there is the largest human population.

The second study was a more intensively sampled small scale investigation on the effects of Time Since Fire (TSF) on mosquito assemblages in a Florida scrub habitat. Again, a mixed effects model approach was used in conjunction with indicator species. Nearly all the variation in mosquito assemblage for any measure, abundance, species richness, or Jost D, was due to seasonality with a trivial amount due to TSF or vegetation. This was the first study about adult mosquitoes' response to fire, as well as, the most extensive mosquito study in Florida scrub.

These two studies in conjunction were very different in spatial scale and indicate that mosquito assemblages are separated on large scales, but very little on smaller scales in South-Central Florida. This is maybe not surprising because they are flying insects that can fly 3+km in a single night. The first study provided some evidence that vegetation plays some role in

mosquito diversity, but the second study did not provide the same evidence. This difference suggests that the more extreme differences in vegetation in the first study (and greater spatial scales) are necessary to detect an effect on adult mosquitoes. It also suggests that metacommunity analyses of mosquitoes must be conducted at scales exceeding several km.

APPENDIX A: FIGURES EXTENDED

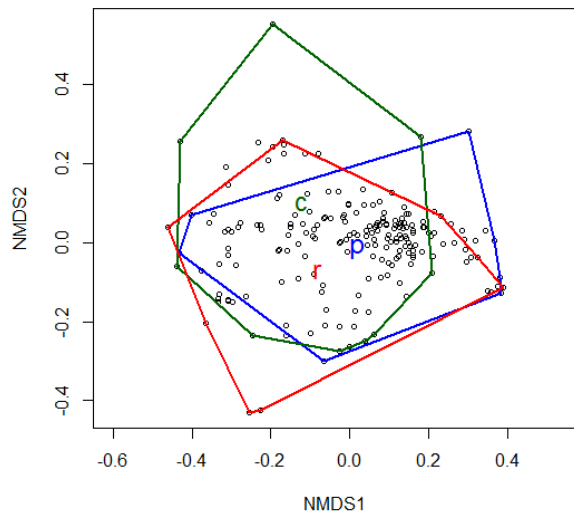


Figure 18: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.

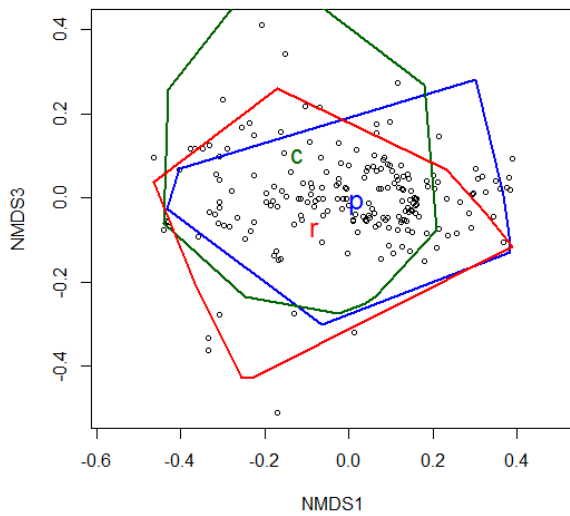


Figure 19: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.

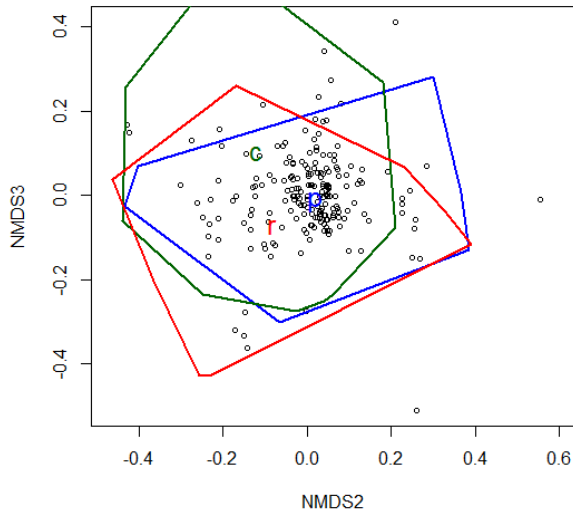


Figure 20: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around ridge (red), plains (blue), and coast (green) sites.

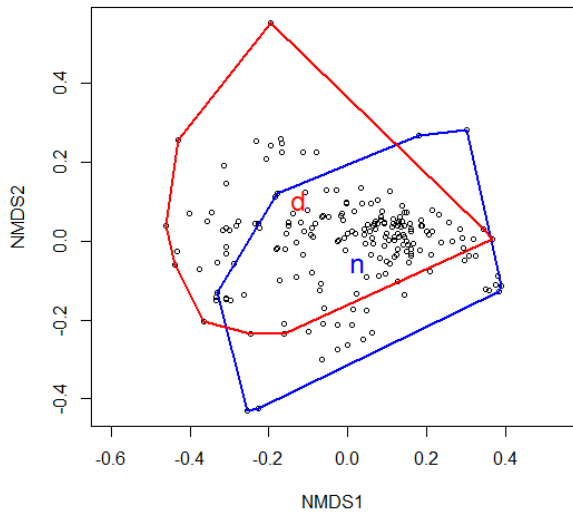


Figure 21: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.

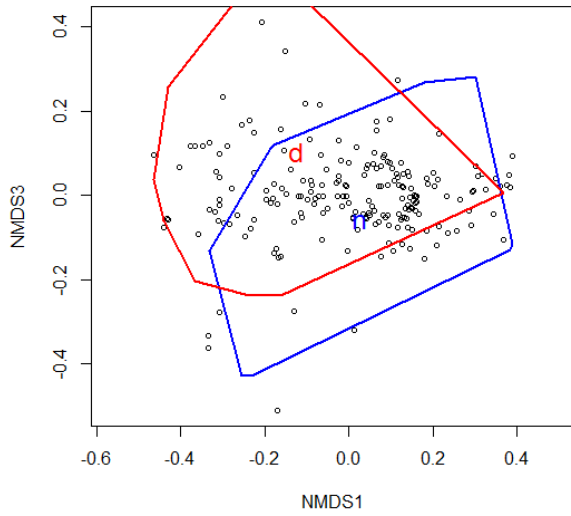


Figure 22: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.

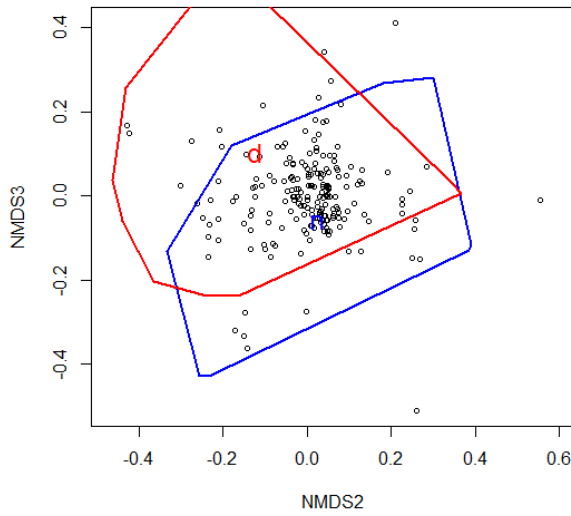


Figure 23: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around developed (red) and natural (blue) sites.

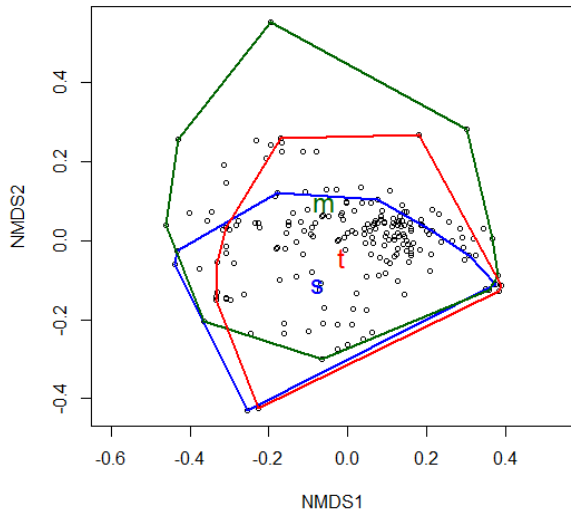


Figure 24: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.

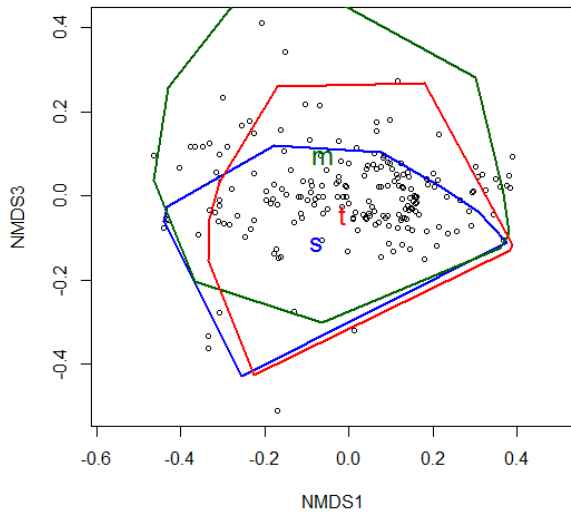


Figure 25: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.

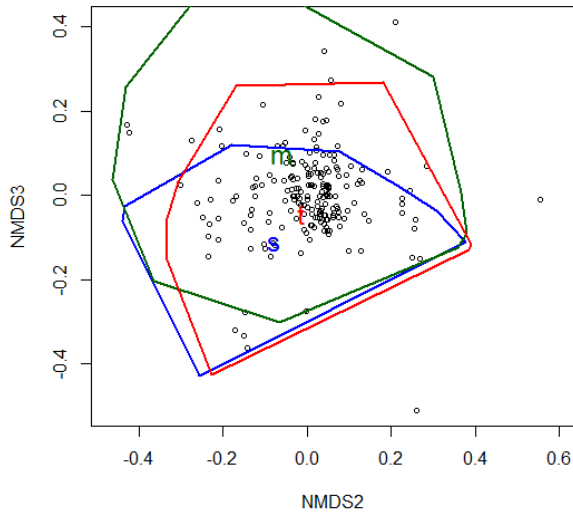


Figure 26: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around short (blue), medium (green), and tall (red) sites.

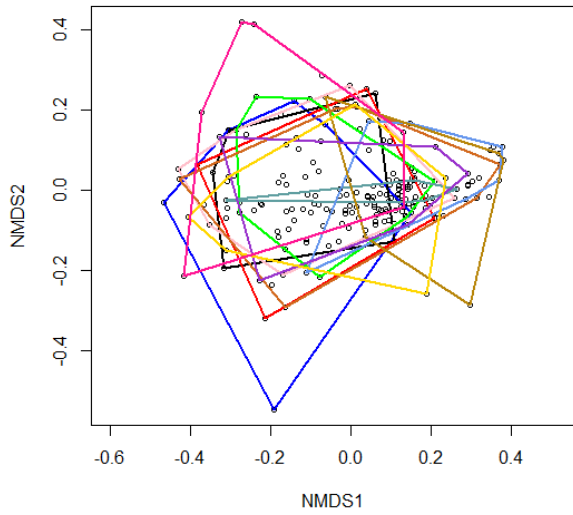


Figure 27: Ordination (NMDS 1 vs 2) of Jaccard's index with convex hulls around each sampling event.

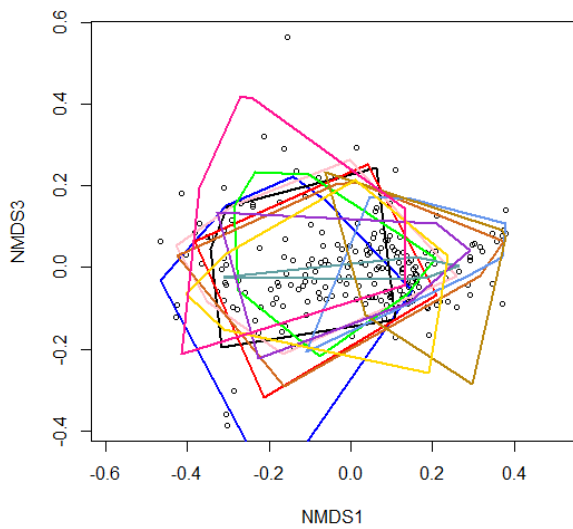


Figure 28: Ordination (NMDS 1 vs 3) of Jaccard's index with convex hulls around each sampling event.

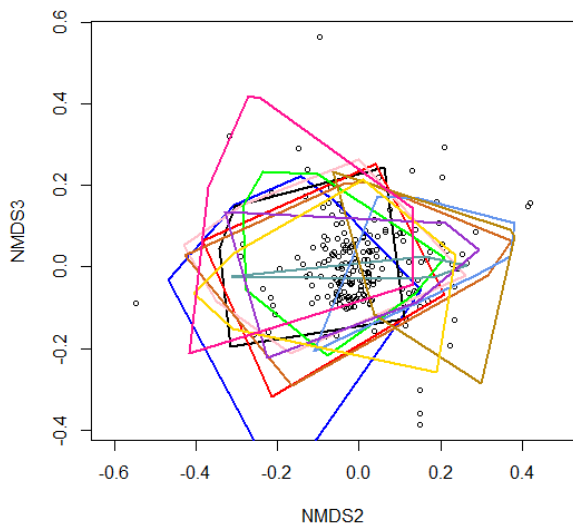


Figure 29: Ordination (NMDS 2 vs 3) of Jaccard's index with convex hulls around each sampling event.

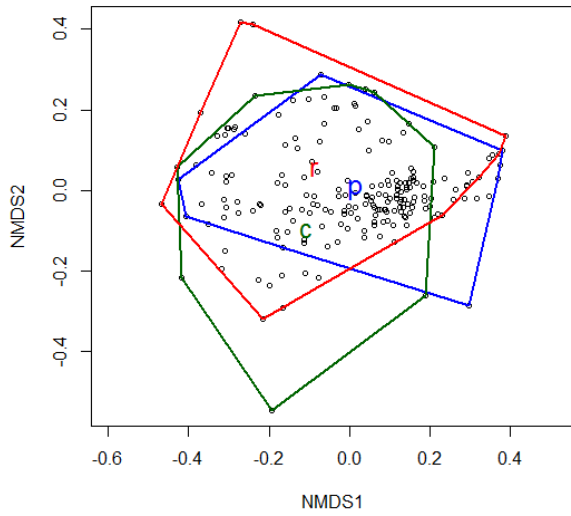


Figure 30: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.

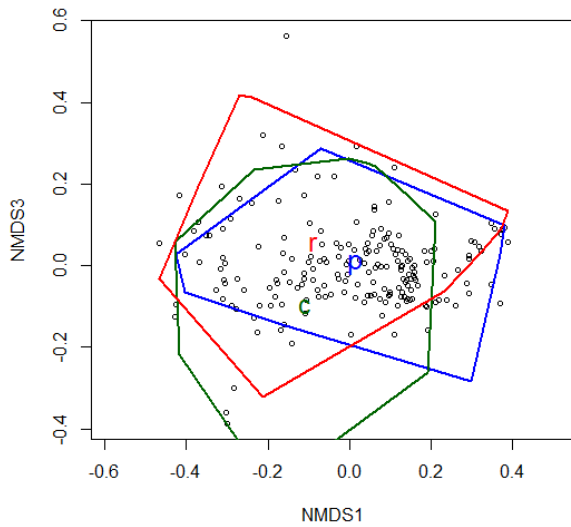


Figure 31: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.

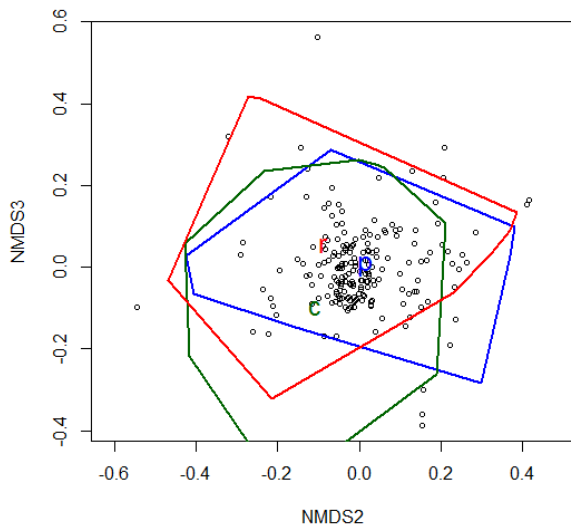


Figure 32: Ordination (NMDS 2vs 3) of Bray-Curtis with convex hulls around ridge (red), plains (blue), and coast (green) sites.

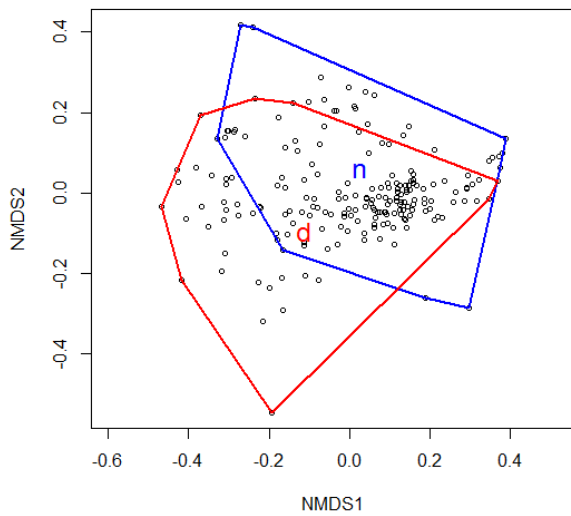


Figure 33: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.

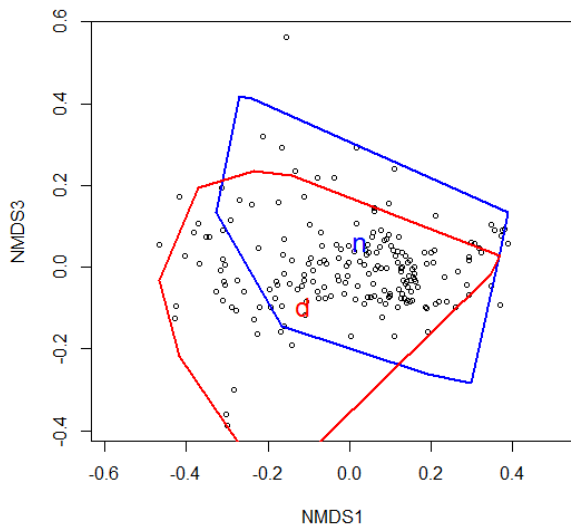


Figure 34: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.

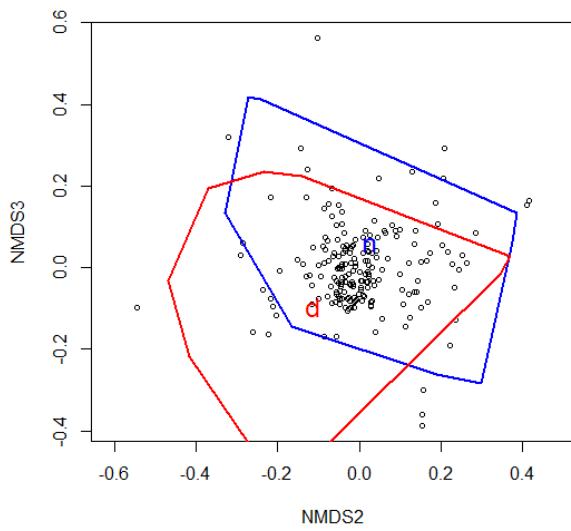


Figure 35: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around developed (red) and natural (blue) sites.

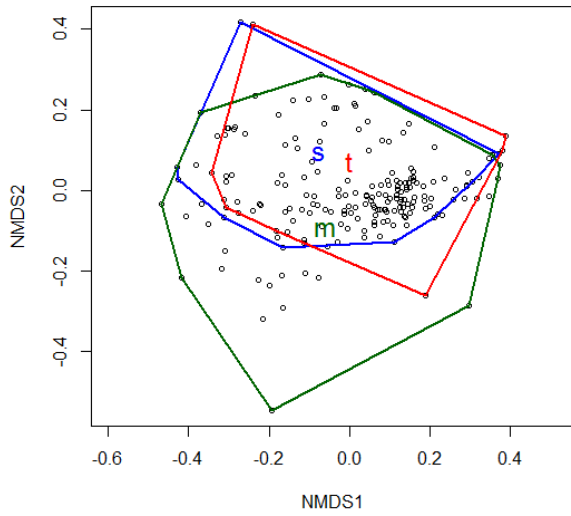


Figure 36: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.

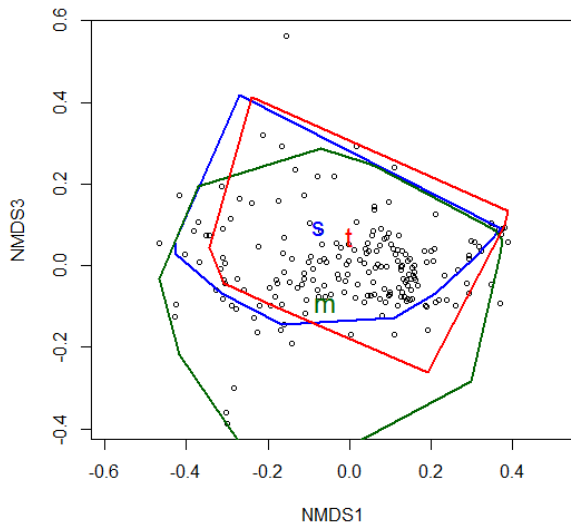


Figure 37: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.

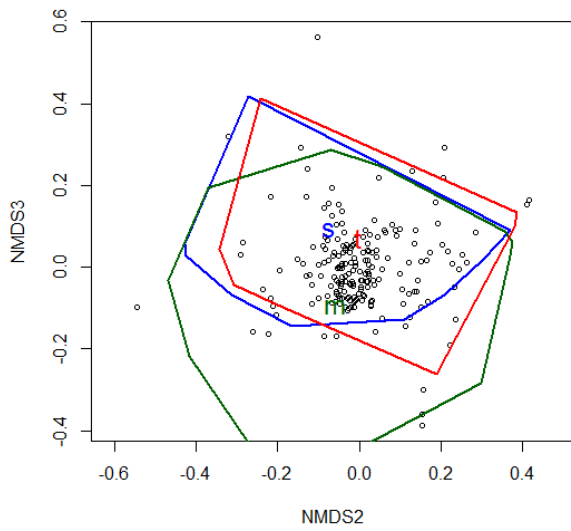


Figure 38: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around short (blue), medium (green), and tall (red) sites.

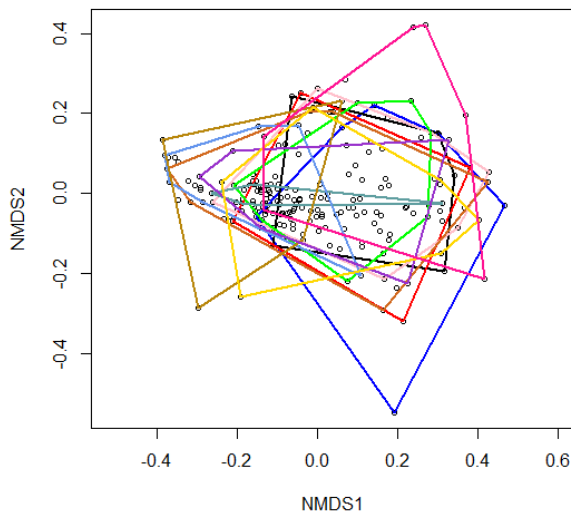


Figure 39: Ordination (NMDS 1 vs 2) of Bray-Curtis with convex hulls around sampling events.

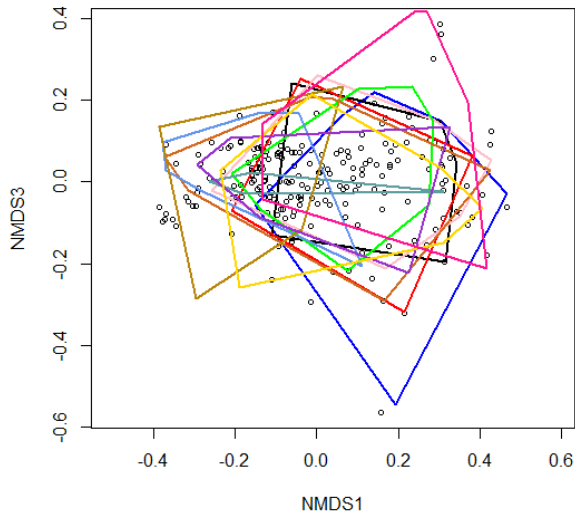


Figure 40: Ordination (NMDS 1 vs 3) of Bray-Curtis with convex hulls around sampling events.

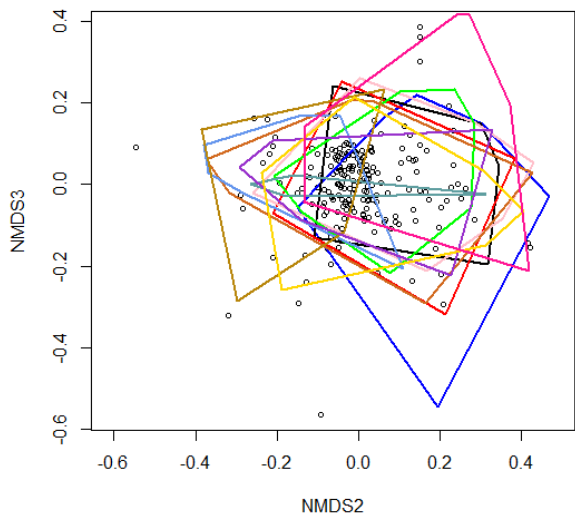


Figure 41: Ordination (NMDS 2 vs 3) of Bray-Curtis with convex hulls around sampling events.

Table 51: Full results of vegetation PCA.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
Std. Deviation	2.01	1.62	1.46	1.01	0.83	0.76	0.64	0.49	0.34	0.30	0.24	0.02
Prop. of Variance	0.34	0.22	0.18	0.08	0.06	0.05	0.03	0.02	0.01	0.01	0.00	0.00
Cum. Prop.	0.34	0.55	0.73	0.82	0.88	0.92	0.96	0.98	0.99	1.00	1.00	1.00

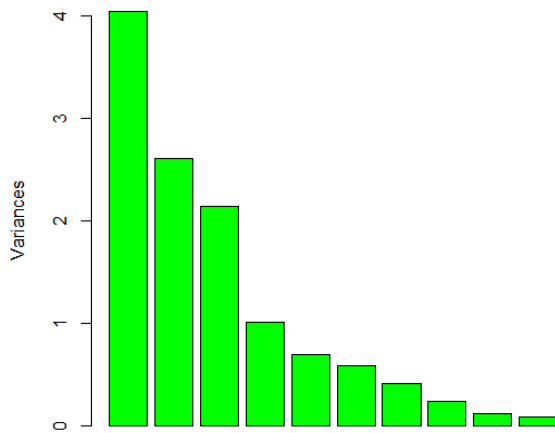


Figure 42: Decay of variance for each axis in the vegetation PCA.

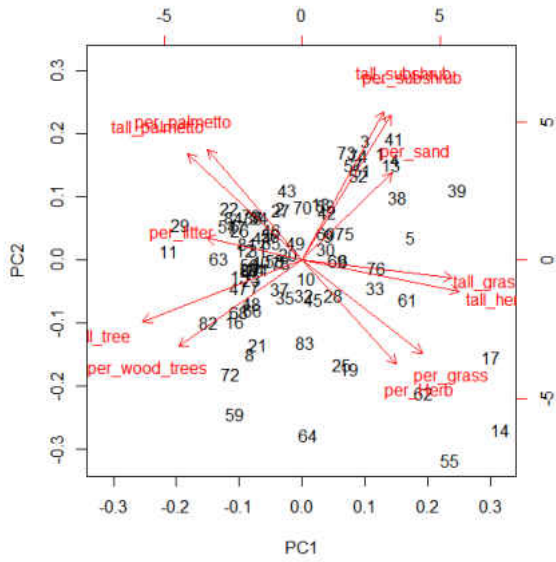


Figure 43: Biplot for PC1 and PC2.

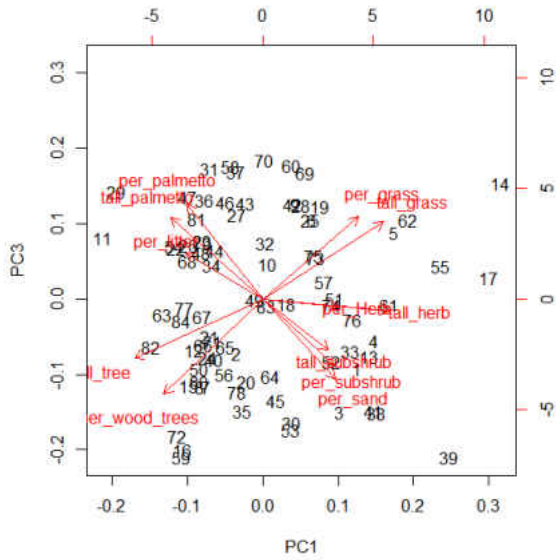


Figure 44: Biplot for PC1 and PC3.

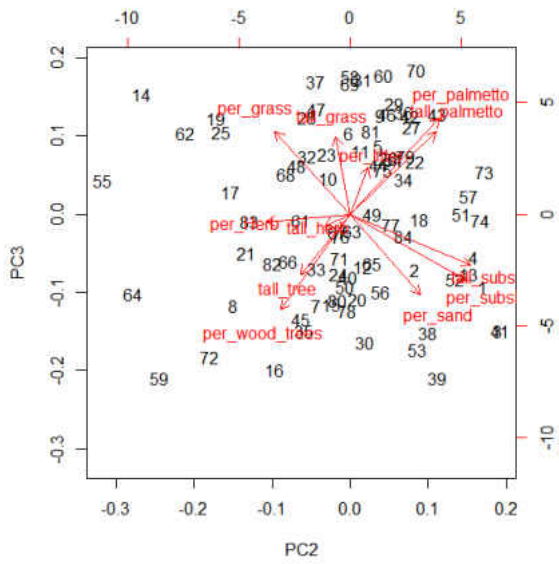


Figure 45: Biplot for PC2 and PC3.

APPENDIX B: R CODE

Chapter 2

Species richness without coast

```
rm(list=ls(all=TRUE))

##models for species richness no coast
modpr<-read.table("only plains and ridge.txt",header=T)
#packages
library(spdep)
library(nlme)
library(bbmle)
library(car)
library(MuMIn)

hist(modpr$spec_rich)

###testing for normality of the response variable
shapiro.test(modpr$spec_rich)

###subsetting to each region to see if this makes a difference

newdata<-modpr[which(modpr$region=='r'),]
shapiro.test(newdata$spec_rich)

newdata2<-modpr[which(modpr$region=='p'),]
shapiro.test(newdata2$spec_rich)

#subsetting to landuse
newdata3<-modpr[which(modpr$landuse=='n'),]
shapiro.test(newdata3$spec_rich)

newdata4<-modpr[which(modpr$landuse=='d'),]
shapiro.test(newdata4$spec_rich)

#subsetting to veg
newdata5<-modpr[which(modpr$veght=='s'),]
shapiro.test(newdata5$spec_rich)

newdata6<-modpr[which(modpr$veght=='m'),]
shapiro.test(newdata6$spec_rich)

newdata7<-modpr[which(modpr$veght=='t'),]
shapiro.test(newdata7$spec_rich)
```

```

####testing for homegeneity of variance across groups
leveneTest(modpr$spec_rich,modpr$region,center=median)
leveneTest(modpr$spec_rich,modpr$landuse,center=median)
leveneTest(modpr$spec_rich,modpr$veght,center=median)

#subset to region
leveneTest(newdata$spec_rich,newdata$landuse,center=median)
leveneTest(newdata$spec_rich,newdata$veght,center=median)

leveneTest(newdata2$spec_rich,newdata2$landuse,center=median)
leveneTest(newdata2$spec_rich,newdata2$veght,center=median)

#subset to landuse
leveneTest(newdata3$spec_rich,newdata3$veght,center=median)

leveneTest(newdata4$spec_rich,newdata4$veght,center=median)

####create the autocorrelation scores
modprcord<-read.table("utm plains and ridge.txt",header=T)

xy3<-SpatialPoints(modprcord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared
distspecr<-autocov_dist(modpr$spec_rich,xy3,nbs=59000)

modprd<-cbind(modpr,distspecr)

plot(modprd$distspecr,modprd$spec_rich)
plot(modprd$samp_event,modprd$spec_rich)

#####
####random factors
m1<-gls(jost_ind~distspecr+region*landuse+veght,data=modprd,method="REML")
m2<-
lme(jost_ind~distspecr+region*landuse+veght,random=~1|samp_event,data=modprd,method="R
EML")
lmc <- lmeControl(niter=5200,msMaxIter=5200)
m3<-
lme(jost_ind~distspecr+region*landuse+veght,random=~distspecr|samp_event,data=modprd,met
hod="REML",control=lmc)

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 3 wins

```

```
##cant do the with or without dist comparison because of the random structure
```

```
#####model selection time
m31<-lme(
spec_rich~distspecr+region*landuse+veght,random=~distspecr|samp_event,control=lmc,data=m
odprd,method="ML")
m32<-lme(
spec_rich~distspecr+region+landuse+veght,random=~distspecr|samp_event,control=lmc,data=m
odprd,method="ML")
m33<-lme(
spec_rich~distspecr+region+landuse,random=~distspecr|samp_event,control=lmc,data=modprd,
method="ML")
m34<-lme(
spec_rich~distspecr+region+veght,random=~distspecr|samp_event,control=lmc,data=modprd,me
thod="ML")
m35<-lme(
spec_rich~distspecr+landuse+veght,random=~distspecr|samp_event,control=lmc,data=modprd,
method="ML")
m36<-lme(
spec_rich~distspecr+veght,random=~distspecr|samp_event,control=lmc,data=modprd,method="
ML")
m37<-lme(
spec_rich~distspecr+region,random=~distspecr|samp_event,control=lmc,data=modprd,method="
ML")
m38<-lme(
spec_rich~distspecr+landuse,random=~distspecr|samp_event,control=lmc,data=modprd,method
="ML")
m39<-lme(
spec_rich~distspecr+region+landuse+region:landuse,random=~distspecr|samp_event,control=lmc,
data=modprd,method="ML")
m310<-lme(
spec_rich~distspecr+region+veght+region:landuse,random=~distspecr|samp_event,control=lmc,
data=modprd,method="ML")
m311<-lme(
spec_rich~distspecr+landuse+veght+region:landuse,random=~distspecr|samp_event,control=lmc
,data=modprd,method="ML")
#m312<-lme(
spec_rich~distspecr+veght+region:landuse,random=~distspecr|samp_event,control=lmc,data=m
odprd,method="ML")
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
m313<-lme(
spec_rich~distspecr+region+region:landuse,random=~distspecr|samp_event,control=lmc,data=m
odprd,method="ML")
```

```
m314<-lme(
spec_rich~distspecr+landuse+region:landuse,random=~distspecr|samp_event,control=lmc,data=
modprd,method="ML")
```

```
AICctab(m31,m32,m33,m34,m35,m36,m37,m38,m39,m310,m311,m313,m314,weights=TR
UE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#best model
mn39r<-lmer(
spec_rich~distspecr+region+landuse+region:landuse+(distspecr|samp_event),control=lmerContr
ol(optCtrl=list(maxfun=20000) ),data=modprd,method="REML")
summary(mn39r)
confmn39r<-confint(mn39r,method="boot",nsim=10000,level=0.95)
confmn39r

r.squaredGLMM(mn39r)
```

Species richness with coast

```
rm(list=ls(all=TRUE))
###models including coast, but excluding events 1-3 and 6
modc<-read.table("not 1-3 or 6.txt",header=T)
#packages
library(spdep)
library(nlme)
library(bbmle)
library(car)
library(MuMIn)

hist(modc$spec_rich)

###testing for normality of the response variable
shapiro.test(modc$spec_rich)

#subsetting to each region to see if this makes a difference
newdata<-modc[which(modc$region=='r'),]
shapiro.test(newdata$spec_rich)

newdata2<-modc[which(modc$region=='p'),]
shapiro.test(newdata2$spec_rich)

newdata3<-modc[which(modc$region=='c'),]
shapiro.test(newdata3$spec_rich)
```



```

#subsetting to landuse
newdata4<-modc[which(modc$landuse=='n'),]
shapiro.test(newdata4$spec_rich)

newdata5<-modc[which(modc$landuse=='d'),]
shapiro.test(newdata5$spec_rich)

#subsetting to veght
newdata6<-modc[which(modc$veght=='s'),]
shapiro.test(newdata6$spec_rich)

newdata7<-modc[which(modc$veght=='m'),]
shapiro.test(newdata7$spec_rich)

newdata8<-modc[which(modc$veght=='t'),]
shapiro.test(newdata8$spec_rich)

###testing for homogeneity of variance across groups

leveneTest(modc$spec_rich,modc$region,center=median)
leveneTest(modc$spec_rich,modc$landuse,center=median)
leveneTest(modc$spec_rich,modc$veght,center=median)

#subset to region
leveneTest(newdata$spec_rich,newdata$landuse,center=median)
leveneTest(newdata$spec_rich,newdata$veght,center=median)

leveneTest(newdata2$spec_rich,newdata2$landuse,center=median)
leveneTest(newdata2$spec_rich,newdata2$veght,center=median)

leveneTest(newdata3$spec_rich,newdata3$landuse,center=median)
leveneTest(newdata3$spec_rich,newdata3$veght,center=median)

#subset to landuse
leveneTest(newdata4$spec_rich,newdata4$veght,center=median)

leveneTest(newdata5$spec_rich,newdata5$veght,center=median)

###create the autocorrelation scores
modcord<-read.table("coast utms.txt",header=T)

xy3<-SpatialPoints(modcord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nobs part of the following line is the largest difference between points aht you want to be
compared
distcspecr<-autocov_dist(modc$spec_rich,xy3,nbs=114000)

```

```

modcd<-cbind(modc,distspecr)

plot(modcd$distspecr,modcd$spec_rich)
plot(modcd$samp_event,modcd$spec_rich)

#####
####random factors
m1<-gls(spec_rich~distspecr+region*landuse+veght,data=modcd,method="REML")
m2<-
lme(spec_rich~distspecr+region*landuse+veght,random=~1|samp_event,data=modcd,method="
REML")
lmc <- lmeControl(niter=5200,msMaxIter=5200)
m3<-
lme(spec_rich~distspecr+region*landuse+veght,random=~distspecr|samp_event,data=modcd,
method="REML",control=lmc)

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 3 wins

m21<-
lme(spec_rich~distspecr+region*landuse+veght,random=~distspecr|samp_event,control=lmc,d
ata=modcd,method="ML")
m22<-
lme(spec_rich~distspecr+region+landuse+veght,random=~distspecr|samp_event,control=lmc,
data=modcd,method="ML")
m23<-
lme(spec_rich~distspecr+region+landuse,random=~distspecr|samp_event,control=lmc,data=m
odcd,method="ML")
m24<-
lme(spec_rich~distspecr+region+veght,random=~distspecr|samp_event,control=lmc,data=mod
cd,method="ML")
m25<-
lme(spec_rich~distspecr+landuse+veght,random=~distspecr|samp_event,control=lmc,data=mo
dcd,method="ML")
m26<-
lme(spec_rich~distspecr+veght,random=~distspecr|samp_event,control=lmc,data=modcd,met
hod="ML")
m27<-
lme(spec_rich~distspecr+region,random=~distspecr|samp_event,control=lmc,data=modcd,met
hod="ML")

```

```

m28<-
lme(spec_rich~distcspecr+landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
m29<-
lme(spec_rich~distcspecr+region+landuse+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
m210<-
lme(spec_rich~distcspecr+region+veght+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
m211<-
lme(spec_rich~distcspecr+landuse+veght+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
#m212<-
lme(spec_rich~distcspecr+veght+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
#this model can not work because of the singularity issue with veght having only one per region/landuse
m213<-
lme(spec_rich~distcspecr+region+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")
m214<-
lme(spec_rich~distcspecr+landuse+region:landuse,random=~distcspecr|samp_event,control=lmc,data=modcd,method="ML")

AICctab(m21,m22,m23,m24,m25,m26,m27,m28,m29,m210,m211,m213,m214,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#best model
mn29r<-
lmer(spec_rich~distcspecr+region+landuse+region:landuse+(distcspecr|samp_event),control=lmerControl(optCtrl=list(maxfun=20000)),data=modcd)
summary(mn29r)

confmn29r<-confint(mn29r,method="boot",nsim=10000,level=0.95)
confmn29r
r.squaredGLMM(mn29r)

##validation
par(mfrow=c(2,2))
resm29r<-residuals(m29r,type="normalized")
F29r<-fitted(m29r,level=0)
plot(fitted(m29r),resm29r,xlab="fitted values",ylab="norm resids")
plot(distcspecr,resm29r,xlab="distance coefficient",ylab="norm resids")
boxplot(resm29r~region,data=modcd,main="regions",ylab="norm resids")
boxplot(resm29r~landuse,data=modcd,main="landuse",ylab="norm resids")

```

```
hist(resm29r)
```

Jost D -Coast

```
rm(list=ls(all=TRUE))
```

```
###models including coast, but excluding events 1-3 and 6
```

```
modc<-read.table("not 1-3 or 6.txt",header=T)
```

```
#packages
```

```
library(spdep)
```

```
library(nlme)
```

```
library(bbmle)
```

```
library(car)
```

```
library(lme4)
```

```
library(MuMIn)
```

```
hist(modc$jost_ind)
```

```
###testing for normality of the response variable
```

```
shapiro.test(modc$jost_ind)
```

```
#subsetting to each region to see if this makes a difference
```

```
newdata<-modc[which(modc$region=='r'),]
```

```
shapiro.test(newdata$jost_ind)
```

```
newdata2<-modc[which(modc$region=='p'),]
```

```
shapiro.test(newdata2$jost_ind)
```

```
newdata3<-modc[which(modc$region=='c'),]
```

```
shapiro.test(newdata3$jost_ind)
```

```
#subsetting to landuse
```

```
newdata4<-modc[which(modc$landuse=='n'),]
```

```
shapiro.test(newdata4$jost_ind)
```

```
newdata5<-modc[which(modc$landuse=='d'),]
```

```
shapiro.test(newdata5$jost_ind)
```

```
#subsetting to veght
```

```
newdata6<-modc[which(modc$veght=='s'),]
```

```
shapiro.test(newdata6$jost_ind)
```

```
newdata7<-modc[which(modc$veght=='m'),]
```

```
shapiro.test(newdata7$jost_ind)
```

```
newdata8<-modc[which(modc$veght=='t'),]
```

```
shapiro.test(newdata8$jost_ind)
```

```
###testing for homegeneity of variance across groups
```

```
leveneTest(modc$jost_ind,modc$region,center=median)  
leveneTest(modc$jost_ind,modc$landuse,center=median)  
leveneTest(modc$jost_ind,modc$veght,center=median)
```

```
#subset to region
```

```
leveneTest(newdata$jost_ind,newdata$landuse,center=median)  
leveneTest(newdata$jost_ind,newdata$veght,center=median)
```

```
leveneTest(newdata2$jost_ind,newdata2$landuse,center=median)  
leveneTest(newdata2$jost_ind,newdata2$veght,center=median)
```

```
leveneTest(newdata3$jost_ind,newdata3$landuse,center=median)  
leveneTest(newdata3$jost_ind,newdata3$veght,center=median)
```

```
#subset to landuse
```

```
leveneTest(newdata4$jost_ind,newdata4$veght,center=median)
```

```
leveneTest(newdata5$jost_ind,newdata5$veght,center=median)
```

```
#####
```

```
####random factors
```

```
m1<-gls(jost_ind~distcjost+region*landuse+veght,data=modcd,method="REML")
```

```
m2<-
```

```
lme(jost_ind~distcjost+region*landuse+veght,random=~1|samp_event,data=modcd,method="REML")
```

```
lmc <- lmeControl(niter=5200,msMaxIter=5200)
```

```
m3<-
```

```
lme(jost_ind~distcjost+region*landuse+veght,random=~distcjost|samp_event,data=modcd,method="REML",control=lmc)
```

```
AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#####model selection time
```

```
m21<-
```

```
lme(jost_ind~distcjost+region*landuse+veght,random=~1|samp_event,data=modcd,method="REML")
```

```
m22<-
```

```
lme(jost_ind~distcjost+region+landuse+veght,random=~1|samp_event,data=modcd,method="REML")
```

```

m23<-
lme(jost_ind~distcjost+region+landuse,random=~1|samp_event,data=modcd,method="ML")
m24<-
lme(jost_ind~distcjost+region+veght,random=~1|samp_event,data=modcd,method="ML")
m25<-
lme(jost_ind~distcjost+landuse+veght,random=~1|samp_event,data=modcd,method="ML")
m26<-lme(jost_ind~distcjost+veght,random=~1|samp_event,data=modcd,method="ML")
m27<-lme(jost_ind~distcjost+region,random=~1|samp_event,data=modcd,method="ML")
m28<-lme(jost_ind~distcjost+landuse,random=~1|samp_event,data=modcd,method="ML")
m29<-
lme(jost_ind~distcjost+region+landuse+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m210<-
lme(jost_ind~distcjost+region+veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m211<-
lme(jost_ind~distcjost+landuse+veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")
#m212<-
lme(jost_ind~distcjost+veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
m213<-
lme(jost_ind~distcjost+region+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m214<-
lme(jost_ind~distcjost+landuse+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m215<-lme(jost_ind~
region*landuse+veght,random=~1|samp_event,data=modcd,method="ML")
m216<-lme(jost_ind~
region+landuse+veght,random=~1|samp_event,data=modcd,method="ML")
m217<-lme(jost_ind~ region+landuse,random=~1|samp_event,data=modcd,method="ML")
m218<-lme(jost_ind~ region+veght,random=~1|samp_event,data=modcd,method="ML")
m219<-lme(jost_ind~ landuse+veght,random=~1|samp_event,data=modcd,method="ML")
m220<-lme(jost_ind~ veght,random=~1|samp_event,data=modcd,method="ML")
m221<-lme(jost_ind~ region,random=~1|samp_event,data=modcd,method="ML")
m222<-lme(jost_ind~ landuse,random=~1|samp_event,data=modcd,method="ML")
m223<-lme(jost_ind~
region+landuse+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m224<-lme(jost_ind~
region+veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m225<-lme(jost_ind~
landuse+veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")

```

```

#m226<-lme(jost_ind~
veght+region:landuse,random=~1|samp_event,data=modcd,method="ML")
#this model can not work because of the singularity issue with vecht having only one per
region/landuse
m227<-lme(jost_ind~
region+region:landuse,random=~1|samp_event,data=modcd,method="ML")
m228<-lme(jost_ind~
landuse+region:landuse,random=~1|samp_event,data=modcd,method="ML")

AICctab(m21,m22,m23,m24,m25,m26,m27,m28,m29,m210,m211,m213,m214,m215,m216,
m217,m218,m219,m220,m221,m222,m223,m224,m225,m227,m228,weights=TRUE,delta=TRU
E,base=TRUE,sort=TRUE)
AICctab(m21,m22,m23,m24,m25,m26,m27,m28,m29,m210,m211,m213,m214,weights=TR
UE,delta=TRUE,base=TRUE,sort=TRUE)

#best model
mn29r<-
lmer(jost_ind~distcjost+region+landuse+region:landuse+(1|samp_event),data=modcd)
summary(mn29r)
confmn29r<-confint(mn29r,method="boot",nsim=10000,level=0.95)
confmn29r
r.squaredGLMM(mn29r)
coef(mn29r)

##validation
par(mfrow=c(2,2))
resm29r<-residuals(m29r,type="normalized")
F29r<-fitted(m29r,level=0)
plot(fitted(m29r),resm29r,xlab="fitted values",ylab="norm resids")
plot(distcjost,resm29r,xlab="distance coefficient",ylab="norm resids")
boxplot(resm29r~region,data=modcd,main="regions",ylab="norm resids")
boxplot(resm29r~landuse,data=modcd,main="landuse",ylab="norm resids")
hist(resm29r)

Jost D without coast
rm(list=ls(all=TRUE))
modpr<-read.table("only plains and ridge.txt",header=T)
#packages
library(spdep)
library(nlme)
library(bbmle)
library(car)
library(lme4)
library(MuMIn)

```

```

#####plots
hist(modpr$jost_ind)
hist(sqrt(modpr$jost_ind))

###testing for normality of the response variable
shapiro.test(modpr$jost_ind)
shapiro.test(sqrt(modpr$jost_ind))

#subsetting to each region to see if this makes a difference
newdata<-modpr[which(modpr$region=='r'),]
shapiro.test(newdata$jost_ind)

newdata2<-modpr[which(modpr$region=='p'),]
shapiro.test(newdata2$jost_ind)

#subsetting to landuse
newdata3<-modpr[which(modpr$landuse=='n'),]
shapiro.test(newdata3$jost_ind)

newdata4<-modpr[which(modpr$landuse=='d'),]
shapiro.test(newdata4$jost_ind)

#subsetting to region:landuse
newdata5<-modpr[which(modpr$region=='r' & modpr$landuse=='n'),]
shapiro.test(newdata5$jost_ind)

newdata6<-modpr[which(modpr$region=='r' & modpr$landuse=='d'),]
shapiro.test(newdata6$jost_ind)

newdata7<-modpr[which(modpr$region=='p' & modpr$landuse=='n'),]
shapiro.test(newdata7$jost_ind)

newdata8<-modpr[which(modpr$region=='p' & modpr$landuse=='d'),]
shapiro.test(newdata8$jost_ind)

###testing for homogeneity of variance across groups

leveneTest(modpr$jost_ind,modpr$region,center=median)
leveneTest(modpr$jost_ind,modpr$landuse,center=median)
leveneTest(modpr$jost_ind,modpr$veght,center=median)

#subset to region
leveneTest(newdata$jost_ind,newdata$landuse,center=median)
leveneTest(newdata$jost_ind,newdata$veght,center=median)

```



```

leveneTest(newdata2$jost_ind,newdata2$landuse,center=median)
leveneTest(newdata2$jost_ind,newdata2$veght,center=median)

#subset to landuse

leveneTest(newdata3$jost_ind,newdata3$veght,center=median)

leveneTest(newdata4$jost_ind,newdata4$veght,center=median)

#subset to region:landuse
leveneTest(newdata5$jost_ind,newdata5$veght,center=median)
leveneTest(newdata6$jost_ind,newdata6$veght,center=median)
leveneTest(newdata7$jost_ind,newdata7$veght,center=median)
leveneTest(newdata8$jost_ind,newdata8$veght,center=median)

###create the autocorrelation scores
modprcord<-read.table("utm plains and ridge.txt",header=T)

xy3<-SpatialPoints(modprcord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared
distjost3<-autocov_dist(modpr$jost_ind,xy3,nbs=59000)

modprd<-cbind(modpr,distjost3)

plot(modprd$distjost3,modprd$jost_ind)
plot(modprd$samp_event,modprd$jost_ind)

#####
####random factors
m1<-gls(jost_ind~distjost3+region*landuse+veght,data=modprd,method="REML")
m2<-
lme(jost_ind~distjost3+region*landuse+veght,random=~1|samp_event,data=modprd,method="R
EML")
lmc <- lmeControl(niter=5200,msMaxIter=5200)
m3<-
lme(jost_ind~distjost3+region*landuse+veght,random=~distjost3|samp_event,data=modprd,met
hod="REML",control=lmc)

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
#model 2 wins!

```

```

#####model selection time

m21<-
lme(jost_ind~distjost3+region*landuse+vegght,random=~1|samp_event,data=modprd,method="
ML")
#everything is significant. according to model simplification I am done, but lets see what
happens
m22<-
lme(jost_ind~distjost3+region+landuse+vegght,random=~1|samp_event,data=modprd,method="
ML")
m23<-
lme(jost_ind~distjost3+region+landuse,random=~1|samp_event,data=modprd,method="ML")
m24<-
lme(jost_ind~distjost3+region+vegght,random=~1|samp_event,data=modprd,method="ML")
m25<-
lme(jost_ind~distjost3+landuse+vegght,random=~1|samp_event,data=modprd,method="ML")
m26<-lme(jost_ind~distjost3+vegght,random=~1|samp_event,data=modprd,method="ML")
m27<-lme(jost_ind~distjost3+region,random=~1|samp_event,data=modprd,method="ML")
m28<-lme(jost_ind~distjost3+landuse,random=~1|samp_event,data=modprd,method="ML")
m29<-
lme(jost_ind~distjost3+region+landuse+region:landuse,random=~1|samp_event,data=modprd,m
ethod="ML")
m210<-
lme(jost_ind~distjost3+region+vegght+region:landuse,random=~1|samp_event,data=modprd,met
hod="ML")
m211<-
lme(jost_ind~distjost3+landuse+vegght+region:landuse,random=~1|samp_event,data=modprd,me
thod="ML")
#m212<-
lme(jost_ind~distjost3+vegght+region:landuse,random=~1|samp_event,data=modprd,method="M
L")
#this model can not work because of the sigularity issue with vegght haveing only one per
region/landuse
m213<-
lme(jost_ind~distjost3+region+region:landuse,random=~1|samp_event,data=modprd,method="
ML")
m214<-
lme(jost_ind~distjost3+landuse+region:landuse,random=~1|samp_event,data=modprd,method="
ML")
m215<-lme(jost_ind~
region*landuse+vegght,random=~1|samp_event,data=modprd,method="ML")
m216<-lme(jost_ind~
region+landuse+vegght,random=~1|samp_event,data=modprd,method="ML")

```

```

m217<-lme(jost_ind~ region+landuse,random=~1|samp_event,data=modprd,method="ML")
m218<-lme(jost_ind~ region+veght,random=~1|samp_event,data=modprd,method="ML")
m219<-lme(jost_ind~ landuse+veght,random=~1|samp_event,data=modprd,method="ML")
m220<-lme(jost_ind~ veght,random=~1|samp_event,data=modprd,method="ML")
m221<-lme(jost_ind~ region,random=~1|samp_event,data=modprd,method="ML")
m222<-lme(jost_ind~ landuse,random=~1|samp_event,data=modprd,method="ML")
m223<-lme(jost_ind~
region+landuse+region:landuse,random=~1|samp_event,data=modprd,method="ML")
m224<-lme(jost_ind~
region+veght+region:landuse,random=~1|samp_event,data=modprd,method="ML")
m225<-lme(jost_ind~
landuse+veght+region:landuse,random=~1|samp_event,data=modprd,method="ML")
#m226<-lme(jost_ind~
veght+region:landuse,random=~1|samp_event,data=modprd,method="ML")
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
m227<-lme(jost_ind~
region+region:landuse,random=~1|samp_event,data=modprd,method="ML")
m228<-lme(jost_ind~
landuse+region:landuse,random=~1|samp_event,data=modprd,method="ML")

```

```

AICctab(m21,m22,m23,m24,m25,m26,m27,m28,m29,m210,m211,m213,m214,m215,m216,
m217,m218,m219,m220,m221,m222,m223,m224,m225,m227,m228,weights=TRUE,delta=TRUE,
base=TRUE,sort=TRUE)

```

```

AICctab(m21,m22,m23,m24,m25,m26,m27,m28,m29,m210,m211,m213,m214,weights=TRUE,
delta=TRUE,base=TRUE,sort=TRUE)

```

```

#best model
mn29r<-
lmer(jost_ind~distjost3+region+landuse+region:landuse+(1|samp_event),data=modprd)
summary(mn29r)

```

```

confmn29r<-confint(mn29r,method="boot",nsim=10000,level=0.95)
confmn29r
r.squaredGLMM(mn29r)

```

Dengue models

```

rm(list=ls(all=TRUE))
###this is for vector anslsis
###Dengue using A. albopictus and A. aegypti as vectors
library(spdep)
library(nlme)
library(bbmle)
library(car)
library(lme4)

```

```

library(glmADMB)
library(MuMIn)
dengmod<-read.table("dengue no coast.txt",header=T)

###autocorrelation stuff--load spdep package
modcord<-read.table("deng_vector utms no coast.txt",header=T)

xy3<-SpatialPoints(modcord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nobs part of the following line is the largest difference between points aht you want to be
compared
distdeng<-autocov_dist(dengmod$Dengue,xy3,nbs=114000)

dengmod<-cbind(dengmod,distdeng)

hist(dengmod$Dengue)
plot(dengmod$distdeng,dengmod$Dengue)
plot(dengmod$samp_event,dengmod$Dengue)

fsamp_event<-factor(dengmod$samp_event)
####random factors
m1<-glmmadmb(Dengue~distsle+region*landuse+veght,data=dengmod,family="binomial")
m1.2<-
glmmadmb(Dengue~distsle+region*landuse+veght,data=dengmod,family="binomial",zeroInflati
on=TRUE)
m2<-glmmadmb(Dengue~distsle+region*landuse+veght+
(1|fsamp_event),data=dengmod,family="binomial")
m2.2<-glmmadmb(Dengue~distsle+region*landuse+veght+
(1|fsamp_event),data=dengmod,family="binomial",zeroInflation=TRUE)
m3<-glmmadmb(Dengue~distsle+region*landuse+veght+
(distsle|fsamp_event),data=dengmod,family="binomial")
m3.2<-glmmadmb(Dengue~distsle+region*landuse+veght+
(distsle|fsamp_event),data=dengmod,family="binomial",zeroInflation=TRUE)

AICctab(m1,m1.2,m2,m2.2,m3,m3.2,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#m1 wins, with no zero inflation or random effects
###not coast
md1<-glmmadmb(Dengue~distcdeng +region*landuse+veght, data=dengmod
,family="binomial" )
md2<-glmmadmb(Dengue~region+landuse+veght, data=dengmod ,family="binomial" )
md3<-glmmadmb(Dengue~region+landuse, data=dengmod ,family="binomial" )
md4<-glmmadmb(Dengue~region+veght, data=dengmod ,family="binomial" )
md5<-glmmadmb(Dengue~landuse+veght, data=dengmod ,family="binomial" )
md6<-glmmadmb(Dengue~veght, data=dengmod ,family="binomial" )

```

```

md7<-glmmadmb(Dengue~region, data=dengmod ,family="binomial" )
md8<-glmmadmb(Dengue~landuse, data=dengmod ,family="binomial" )
md9<-glmmadmb(Dengue~region+landuse+region:landuse, data=dengmod
,family="binomial" )
md10<-glmmadmb(Dengue~region+veght+region:landuse, data=dengmod
,family="binomial" )
md11<-glmmadmb(Dengue~landuse+veght+region:landuse, data=dengmod
,family="binomial" )
#md12<-glmmadmb(Dengue~veght+region:landuse, data=dengmod ,family="binomial" )
#this model can not work because of the sigularity issue with veght haveing only one per
region/landuse
md13<-glmmadmb(Dengue~region+region:landuse, data=dengmod ,family="binomial" )
md14<-glmmadmb(Dengue~landuse+region:landuse, data=dengmod ,family="binomial" )
md15<-glmmadmb(Dengue~ region*landuse+veght, data=dengmod ,family="binomial" )

AICctab(md1,md15,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#better without distance

AICctab(md15,md2,md3,md4,md5,md6,md7,md8,md9,md10,md11,md13,md14,weights=T
RUE,delta=TRUE,base=TRUE,sort=TRUE)

#md8 and md7

summary(md7)
par(mfrow=c(2,2))
resmd7<-residuals(md7,type="pearson")
plot(fitted(md7),resmd7,xlab="fitted values", ylab="norm resid",main="md7")
boxplot(resmd7~landuse,data=dengmod,main="landuse",ylab="norm resid")
hist(resmd7)

summary(md8)
par(mfrow=c(2,2))
resmd8<-residuals(md8,type="pearson")
plot(fitted(md8),resmd8,xlab="fitted values", ylab="norm resid",main="md8")
boxplot(resmd8~landuse,data=dengmod,main="landuse",ylab="norm resid")
hist(resmd8)

R.squaredGLMM(md7)
R.squaredGLMM(md8)

###not 1-3,6
dengcmod<-read.table("dengue 1_3_6_not.txt",header=T)

```

```

hist(dengcmod$Dengue)
####autocorrelation stuff--load spdep package
modcordc<-read.table("deng_vector utms w_coast.txt",header=T)

xy3<-SpatialPoints(modcordc,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared
distcdeng<-autocov_dist(dengcmod$Dengue,xy3,nbs=114000)

dengcmod<-cbind(dengcmod,distcdeng)

#####random effects
m1<-
glmmadmb(Dengue~distcdeng+region*landuse+veght,data=dengcmod,family="binomial")
m1.2<-
glmmadmb(Dengue~distcdeng+region*landuse+veght,data=dengcmod,family="binomial",zeroI
nflation=TRUE)
m2<-glmmadmb(Dengue~distcdeng+region*landuse+veght+
(1|fsamp_event),data=dengcmod,family="binomial")
m2.2<-glmmadmb(Dengue~distcdeng+region*landuse+veght+
(1|fsamp_event),data=dengcmod,family="binomial",zeroInflation=TRUE)
m3<-glmmadmb(Dengue~distcdeng+region*landuse+veght+
(distcdeng|fsamp_event),data=dengcmod,family="binomial")
m3.2<-glmmadmb(Dengue~distcdeng+region*landuse+veght+
(distcdeng|fsamp_event),data=dengcmod,family="binomial",zeroInflation=TRUE)

AICctab(m1,m1.2,m2,m2.2,m3,m3.2,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

modc1<-
glm(Dengue~distcdeng+region*landuse*veght,data=dengcmod,family="binomial",weight=samp
_event)
modc2<-glm(Dengue~distcdeng+region*landuse*veght-
region:landuse:veght,data=dengcmod,family="binomial",weight=samp_event)
modc3<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-
region:landuse,data=dengcmod,family="binomial",weight=samp_event)
modc4<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-
region:veght,data=dengcmod,family="binomial",weight=samp_event)
modc5<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-
landuse:veght,data=dengcmod,family="binomial",weight=samp_event)
modc6<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-
region:landuse-region:veght,data=dengcmod,family="binomial",weight=samp_event)
modc7<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-
region:landuse-landuse:veght,data=dengcmod,family="binomial",weight=samp_event)

```

```

modc8<-glm(Dengue~distcdeng+region*landuse*veght-region:landuse:veght-region:veght-
landuse:veght,data=dengcmod,family="binomial",weight=samp_event)
modc9<-
glm(Dengue~distcdeng+region+landuse+veght,data=dengcmod,family="binomial",weight=samp
_event)
modc10<-
glm(Dengue~distcdeng+region+landuse,data=dengcmod,family="binomial",weight=samp_event
)
modc11<-
glm(Dengue~distcdeng+region+veght,data=dengcmod,family="binomial",weight=samp_event)
modc12<-
glm(Dengue~distcdeng+landuse+veght,data=dengcmod,family="binomial",weight=samp_event)
modc13<-
glm(Dengue~distcdeng+region,data=dengcmod,family="binomial",weight=samp_event)
modc14<-
glm(Dengue~distcdeng+landuse,data=dengcmod,family="binomial",weight=samp_event)
modc15<-
glm(Dengue~distcdeng+veght,data=dengcmod,family="binomial",weight=samp_event)
AICctab(modc1,modc15,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

#better without distance

```

AICctab(modc15,modc2,modc3,modc4,modc5,modc6,modc7,modc8,modc9,modc10,modc1
1,modc13,modc14,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
#mdc8 wins

```

```

summary(mdc8)
par(mfrow=c(2,2))
resmdc8<-residuals(mdc8,type="pearson")
F27<-fitted(mdc8,level=0)
plot(fitted(mdc8),resmdc8,xlab="fitted values",ylab="norm resid",main="md8")
boxplot(resmdc8~landuse,data=dengmod,main="landuse",ylab="norm resid")
hist(resmdc8)

```

```

confdmdc8<-confint(mdc8,method="boot",nsim=1000,level=0.95)
confdmdc8

```

WNV

```

###this is for vector anlysis
###West Nile Virust using culex as vectors
rm(list=ls(all=TRUE))
library(spdep)
library(nlme)

```

```

library(bbmle)
library(car)
library(lme4)
library(glmmADMB)

wnvmod<-read.table("wnv no_coast.txt",header=T)

hist(wnvmod$wnv)

###autocorrelation stuff--load spdep package
modcordw<-read.table("wnv utms no_coast.txt",header=T)

xy3w<-SpatialPoints(modcordw,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared
distwnv<-autocov_dist(wnvmod$wnv,xy3w,nbs=114000)

wnvmod<-cbind(wnvmod,distwnv)

fsamp_event<-factor(wnvmod$samp_event)
###not coast
#random
m1<-glm(wnv~distwnv+region*landuse+veght,data=wnvmod,family="binomial")
m2<-glmer(wnv~distwnv+region*landuse+veght+
(1|samp_event),data=wnvmod,family="binomial")
m3<-glmer(wnv~distwnv+region*landuse+veght+
(distwnv|samp_event),data=wnvmod,family="binomial")

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#fixed
mw1<-glmer(wnv~distwnv +region*landuse+veght+ (1|samp_event), data=wnvmod
,family="binomial" )
mw2<-glmer(wnv~distwnv +region+landuse+veght+ (1|samp_event), data=wnvmod
,family="binomial" )
mw3<-glmer(wnv~distwnv +region+landuse+ (1|samp_event), data=wnvmod
,family="binomial" )
mw4<-glmer(wnv~distwnv +region+veght+ (1|samp_event), data=wnvmod
,family="binomial" )
mw5<-glmer(wnv~distwnv +landuse+veght+ (1|samp_event), data=wnvmod
,family="binomial" )
mw6<-glmer(wnv~distwnv +veght+ (1|samp_event), data=wnvmod ,family="binomial" )
mw7<-glmer(wnv~distwnv +region+ (1|samp_event), data=wnvmod ,family="binomial" )
mw8<-glmer(wnv~distwnv +landuse+ (1|samp_event), data=wnvmod ,family="binomial" )

```



```

mw9<-glmer(wnv~distwnv +region+landuse+region:landuse+ (1|samp_event),
data=wnvmod ,family="binomial" )
mw10<-glmer(wnv~distwnv +region+veght+region:landuse+ (1|samp_event), data=wnvmod
,family="binomial" )
mw11<-glmer(wnv~distwnv +landuse+veght+region:landuse+ (1|samp_event),
data=wnvmod ,family="binomial" )
#mw12<-glmer(wnv~distwnv +veght+region:landuse+ (1|samp_event), data=wnvmod
,family="binomial" )
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
mw13<-glmer(wnv~distwnv +region+region:landuse+ (1|samp_event), data=wnvmod
,family="binomial" )
mw14<-glmer(wnv~distwnv +landuse+region:landuse+ (1|samp_event), data=wnvmod
,family="binomial" )
mw15<-glmer(wnv~region*landuse+veght+ (1|samp_event), data=wnvmod
,family="binomial" )

```

```

mw1L<-refitML(mw1)
mw2L<-refitML(mw2)
mw3L<-refitML(mw3)
mw4L<-refitML(mw4)
mw5L<-refitML(mw5)
mw6L<-refitML(mw6)
mw7L<-refitML(mw7)
mw8L<-refitML(mw8)
mw9L<-refitML(mw9)
mw10L<-refitML(mw10)
mw11L<-refitML(mw11)
mw13L<-refitML(mw13)
mw14L<-refitML(mw14)
mw15L<-refitML(mw15)

```

```
AICctab(mw1L,mw15L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#better with distcance
```

```
AICctab(mw1L,mw2L,mw3L,mw4L,mw5L,mw6L,mw7L,mw8L,mw9L,mw10L,mw11L,mw13L,mw14L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#9,14, and 13 are equally as plausible, but I choose 9 because it includes the others
```

```

summary(mw9)
par(mfrow=c(2,2))
resmw9<-residuals(mw9,type="pearson")
plot(fitted(mw9),resmw9,xlab="fitted values", ylab="norm resids",main="mw9")
boxplot(resmw9~landuse,data=wnvmod,main="landuse",ylab="norm resids")

```

```

boxplot(resmw9~region,data=wnvmod,main="region",ylab="norm resids")
hist(resmw9)

confdmw9<-confint(mw9,method="boot",nsim=1000,level=0.95)
confdmw9

###not 1-3,6
wnvcmmod<-read.table("wnv w_coast.txt",header=T)
hist(wnvcmmod$wnv)
###autocorrelation stuff--load spdep package
modcordcw<-read.table("wnv utm w_coast.txt",header=T)

xy3cw<-SpatialPoints(modcordcw,proj4string=CRS("+proj=utm
+zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared
distcwnv<-autocov_dist(wnvcmmod$wnv,xy3cw,nbs=114000)

wnvcmmod<-cbind(wnvcmmod,distcwnv)

#random
m1<-glm(wnv~distcwnv+region*landuse+veght,data=wnvcmmod,family="binomial")
m2<-glmer(wnv~distcwnv+region*landuse+veght+
(1|samp_event),data=wnvcmmod,family="binomial")
m3<-glmer(wnv~distcwnv+region*landuse+veght+
(distwnv|samp_event),data=wnvcmmod,family="binomial")

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#m2 wins

#fixed
mwc1<-glmer(wnv~distcwnv +region*landuse+veght+ (1|samp_event), data=wnvcmmod
,family="binomial" )
mwc2<-glmer(wnv~distcwnv +region+landuse+veght+ (1|samp_event), data=wnvcmmod
,family="binomial" )
mwc3<-glmer(wnv~distcwnv +region+landuse+ (1|samp_event), data=wnvcmmod
,family="binomial" )
mwc4<-glmer(wnv~distcwnv +region+veght+ (1|samp_event), data=wnvcmmod
,family="binomial" )
mwc5<-glmer(wnv~distcwnv +landuse+veght+ (1|samp_event), data=wnvcmmod
,family="binomial" )
mwc6<-glmer(wnv~distcwnv +veght+ (1|samp_event), data=wnvcmmod ,family="binomial" )
mwc7<-glmer(wnv~distcwnv +region+ (1|samp_event), data=wnvcmmod ,family="binomial"
)

```

```

mwc8<-glmer(wnv~distcwnv +landuse+ (1|samp_event), data=wnvcmod ,family="binomial"
)
mwc9<-glmer(wnv~distcwnv +region+landuse+region:landuse+ (1|samp_event),
data=wnvcmod ,family="binomial" )
mwc10<-glmer(wnv~distcwnv +region+veght+region:landuse+ (1|samp_event),
data=wnvcmod ,family="binomial" )
mwc11<-glmer(wnv~distcwnv +landuse+veght+region:landuse+ (1|samp_event),
data=wnvcmod ,family="binomial" )
#mwc12<-glmer(wnv~distcwnv +veght+region:landuse+ (1|samp_event), data=wnvcmod
,family="binomial" )
#this model can not work because of the sigularity issue with veght haveing only one per
region/landuse
mwc13<-glmer(wnv~distcwnv +region+region:landuse+ (1|samp_event), data=wnvcmod
,family="binomial" )
mwc14<-glmer(wnv~distcwnv +landuse+region:landuse+ (1|samp_event), data=wnvcmod
,family="binomial" )
mwc15<-glmer(wnv~region*landuse+veght+ (1|samp_event), data=wnvcmod
,family="binomial" )

mwc1L<-refitML(mwc1)
mwc2L<-refitML(mwc2)
mwc3L<-refitML(mwc3)
mwc4L<-refitML(mwc4)
mwc5L<-refitML(mwc5)
mwc6L<-refitML(mwc6)
mwc7L<-refitML(mwc7)
mwc8L<-refitML(mwc8)
mwc9L<-refitML(mwc9)
mwc10L<-refitML(mwc10)
mwc11L<-refitML(mwc11)
mwc13L<-refitML(mwc13)
mwc14L<-refitML(mwc14)
mwc15L<-refitML(mwc15)

AICctab(mwc1L,mwc15L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#better with distcance

AICctab(mwc1L,mwc2L,mwc3L,mwc4L,mwc5L,mwc6L,mwc7L,mwc8L,mwc9L,mwc10L,
mwc11L,mwc13L,mwc14L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#9,13, and 14 win, but i choose 9 again

summary(mwc9)
par(mfrow=c(2,2))

```

```

resmw9<-residuals(mwc9,type="pearson")
plot(fitted(mwc9),resmw9,xlab="fitted values", ylab="norm resid",main="mw9")
boxplot(resmw9~landuse,data=wnvmod,main="landuse",ylab="norm resid")
boxplot(resmw9~region,data=wnvmod,main="region",ylab="norm resid")
hist(resmw9)

confdmwc9<-confint(mwc9,method="boot",nsim=1000,level=0.95)
confdmwc9

```

SLE

```

###this is for vector anlysis
###St. Louis Encephalitis using cu. nigripalpus as vector

```

```

rm(list=ls(all=TRUE))
library(spdep)
library(nlme)
library(bbmle)
library(car)
library(lme4)
library(glmmADMB)

```

```

slemod<-read.table("sle no_coast.txt",header=T)

```

```

hist(slemod$sle)
###autocorrelation stuff--load spdep package
modcords<-read.table("sle utm no coast.txt",header=T)

```

```

xy3s<-SpatialPoints(modcords,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nob's part of the following line is the largest difference between points aht you want to be
compared

```

```

distsle<-autocov_dist(slemod$sle,xy3s,nbs=114000)

```

```

slemod<-cbind(slemod,distsle)

```

```

###not coast

```

```

####random factors

```

```

m1<-glm(sle~distsle+region*landuse+vegth,data=slemod,family="binomial")
m2<-glmer(sle~distsle+region*landuse+vegth+
(1|samp_event),data=slemod,family="binomial")
m3<-glmer(sle~distsle+region*landuse+vegth+
(distsle|samp_event),data=slemod,family="binomial")

```

```

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
#m2 wins

```

```

ms1<-glmer(sle~distsle +region*landuse+veght+ (1|samp_event), data=slemod
,family="binomial" )
ms2<-glmer(sle~distsle +region+landuse+veght+ (1|samp_event), data=slemod
,family="binomial" )
ms3<-glmer(sle~distsle +region+landuse+ (1|samp_event), data=slemod ,family="binomial"
)
ms4<-glmer(sle~distsle +region+veght+ (1|samp_event), data=slemod ,family="binomial" )
ms5<-glmer(sle~distsle +landuse+veght+ (1|samp_event), data=slemod ,family="binomial" )
ms6<-glmer(sle~distsle +veght+ (1|samp_event), data=slemod ,family="binomial" )
ms7<-glmer(sle~distsle +region+ (1|samp_event), data=slemod ,family="binomial" )
ms8<-glmer(sle~distsle +landuse+ (1|samp_event), data=slemod ,family="binomial" )
ms9<-glmer(sle~distsle +region+landuse+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
ms10<-glmer(sle~distsle +region+veght+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
ms11<-glmer(sle~distsle +landuse+veght+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
#ms12<-glmer(sle~distsle +veght+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
ms13<-glmer(sle~distsle +region+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
ms14<-glmer(sle~distsle +landuse+region:landuse+ (1|samp_event), data=slemod
,family="binomial" )
ms15<-glmer(sle~region*landuse+veght+ (1|samp_event), data=slemod ,family="binomial"
)

ms1L<-refitML(ms1)
ms2L<-refitML(ms2)
ms3L<-refitML(ms3)
ms4L<-refitML(ms4)
ms5L<-refitML(ms5)
ms6L<-refitML(ms6)
ms7L<-refitML(ms7)
ms8L<-refitML(ms8)
ms9L<-refitML(ms9)
ms10L<-refitML(ms10)
ms11L<-refitML(ms11)
ms13L<-refitML(ms13)
ms14L<-refitML(ms14)
ms15L<-refitML(ms15)

AICctab(ms1L,ms15L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

#better with distcance

AICctab(ms1L,ms2L,ms3L,ms4L,ms5L,ms6L,ms7L,ms8L,ms9L,ms10L,ms11L,ms13L,ms14L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#14,9, 13 are equally plausible, but will use 9 because it encompasses 14 and 13

summary(ms9)
par(mfrow=c(2,2))
resms9<-residuals(ms9,type="pearson")
plot(fitted(ms9),resms9,xlab="fitted values", ylab="norm resid")
boxplot(resms9~landuse,data=slemod,main="landuse",ylab="norm resid")
boxplot(resms9~region,data=slemod,main="region",ylab="norm resid")
hist(resms9)

confdms9<-confint(ms9,method="boot",nsim=1000,level=0.95)
confdms9

#####
#####
###with coast
slecmo<-read.table("sle w_coast.txt",header=T)

hist(slecmo$sle)

###autocorrelation stuff--load spdep package
modcordsc<-read.table("sle utm w coast.txt",header=T)

xy3sc<-SpatialPoints(modcordsc,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nobs part of the following line is the largest difference between points aht you want to be
compared
distslec<-autocov_dist(slecmo$sle,xy3sc,nbs=114000)

slecmo<-cbind(slecmo,distslec)

####random factors
m1<-glm(sle~distslec+region*landuse+veght,data=slecmo,family="binomial")
m2<-glmer(sle~distslec+region*landuse+veght+
(1|samp_event),data=slecmo,family="binomial")
m3<-glmer(sle~distslec+region*landuse+veght+
(distslec|samp_event),data=slecmo,family="binomial")

AICctab(m1,m2,m3,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

#model 2 wins

#fixed
msc1<-glmer(sle~distslec +region*landuse+veght+ (1|samp_event), data=slecmod
,family="binomial" )
msc2<-glmer(sle~ region+landuse+veght+ (1|samp_event), data=slecmod
,family="binomial" )
msc3<-glmer(sle~region+landuse+ (1|samp_event), data=slecmod ,family="binomial" )
msc4<-glmer(sle~region+veght+ (1|samp_event), data=slecmod ,family="binomial" )
msc5<-glmer(sle~landuse+veght+ (1|samp_event), data=slecmod ,family="binomial" )
msc6<-glmer(sle~veght+ (1|samp_event), data=slecmod ,family="binomial" )
msc7<-glmer(sle~region+ (1|samp_event), data=slecmod ,family="binomial" )
msc8<-glmer(sle~landuse+ (1|samp_event), data=slecmod ,family="binomial" )
msc9<-glmer(sle~region+landuse+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
msc10<-glmer(sle~region+veght+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
msc11<-glmer(sle~landuse+veght+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
#msc12<-glmer(sle~veght+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
#this model can not work because of the singularity issue with veght haveing only one per
region/landuse
msc13<-glmer(sle~region+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
msc14<-glmer(sle~landuse+region:landuse+ (1|samp_event), data=slecmod
,family="binomial" )
msc15<-glmer(sle~region*landuse+veght+ (1|samp_event), data=slecmod
,family="binomial" )

msc1L<-refitML(msc1)
msc2L<-refitML(msc2)
msc3L<-refitML(msc3)
msc4L<-refitML(msc4)
msc5L<-refitML(msc5)
msc6L<-refitML(msc6)
msc7L<-refitML(msc7)
msc8L<-refitML(msc8)
msc9L<-refitML(msc9)
msc10L<-refitML(msc10)
msc11L<-refitML(msc11)
msc13L<-refitML(msc13)
msc14L<-refitML(msc14)
msc15L<-refitML(msc15)

```

```

AICctab(msc1L,msc15L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#better with distcance

AICctab(msc15L,msc2L,msc3L,msc4L,msc5L,msc6L,msc7L,msc8L,msc9L,msc10L,msc11
L,msc13L,msc14L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#m15/l1 and close followed by 10 are winners
#but 15 includes the others so I will choose that one

summary(msc15)
par(mfrow=c(2,2))
resmsc15<-residuals(msc15,type="pearson")
plot(fitted(msc15),resmsc15,xlab="fitted values",ylab="norm resid")
boxplot(resmsc15~landuse,data=slemod,main="landuse",ylab="norm resid")
boxplot(resmsc15~region,data=slemod,main="region",ylab="norm resid")
hist(resmsc15)

confdmsc15<-confint(msc15,method="boot",nsim=1000,level=0.95)
confdmsc15

#####
###running the models with out distances

###no coast
modn1<-
glm(sle~region*landuse*veght,data=slemod,family="binomial",weight=samp_event)
modn2<-glm(sle~region*landuse*veght-
region:landuse:veght,data=slemod,family="binomial",weight=samp_event)
modn3<-glm(sle~region*landuse*veght-region:landuse:veght-
region:landuse,data=slemod,family="binomial",weight=samp_event)
modn4<-glm(sle~region*landuse*veght-region:landuse:veght-
region:veght,data=slemod,family="binomial",weight=samp_event)
modn5<-glm(sle~region*landuse*veght-region:landuse:veght-
landuse:veght,data=slemod,family="binomial",weight=samp_event)
modn6<-glm(sle~region*landuse*veght-region:landuse:veght-region:landuse-
region:veght,data=slemod,family="binomial",weight=samp_event)
modn7<-glm(sle~region*landuse*veght-region:landuse:veght-region:landuse-
landuse:veght,data=slemod,family="binomial",weight=samp_event)
modn8<-glm(sle~region*landuse*veght-region:landuse:veght-region:veght-
landuse:veght,data=slemod,family="binomial",weight=samp_event)
modn9<-
glm(sle~region+landuse+veght,data=slemod,family="binomial",weight=samp_event)
modn10<-glm(sle~region+landuse,data=slemod,family="binomial",weight=samp_event)

```



```

modn11<-glm(sle~region+veght,data=slemod,family="binomial",weight=samp_event)
modn12<-glm(sle~landuse+veght,data=slemod,family="binomial",weight=samp_event)
modn13<-glm(sle~region,data=slemod,family="binomial",weight=samp_event)
modn14<-glm(sle~landuse,data=slemod,family="binomial",weight=samp_event)
modn15<-glm(sle~veght,data=slemod,family="binomial",weight=samp_event)

```

```

AICctab(modn1,modn2,modn3,modn4,modn5,modn6,modn7,modn8,modn9,modn10,modn11,modn12,modn13,modn14,modn15,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

AICctab(modn1,mod8,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

###with coast

```

```

#models

```

```

modcn1<-

```

```

glm(sle~region*landuse*veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn2<-glm(sle~region*landuse*veght-

```

```

region:landuse:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn3<-glm(sle~region*landuse*veght-region:landuse:veght-

```

```

region:landuse,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn4<-glm(sle~region*landuse*veght-region:landuse:veght-

```

```

region:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn5<-glm(sle~region*landuse*veght-region:landuse:veght-

```

```

landuse:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn6<-glm(sle~region*landuse*veght-region:landuse:veght-region:landuse-

```

```

region:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn7<-glm(sle~region*landuse*veght-region:landuse:veght-region:landuse-

```

```

landuse:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn8<-glm(sle~region*landuse*veght-region:landuse:veght-region:veght-

```

```

landuse:veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn9<-

```

```

glm(sle~region+landuse+veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn10<-glm(sle~region+landuse,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn11<-glm(sle~region+veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn12<-glm(sle~landuse+veght,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn13<-glm(sle~region,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn14<-glm(sle~landuse,data=slecmod,family="binomial",weight=samp_event)

```

```

modcn15<-glm(sle~veght,data=slecmod,family="binomial",weight=samp_event)

```

```

AICctab(modcn1,modcn2,modcn3,modcn4,modcn5,modcn6,modcn7,modcn8,modcn9,modcn10,modcn11,modcn12,modcn13,modcn14,modcn15,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

AICctab(modcn1,modc1,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

```

```

summary(modcn1)

```

Ordination

```
rm(list=ls(all=TRUE))
library(BiodiversityR)
library(vegan)
library(labdsv)
```

```
###This is to get the Jaccard's indices. The txt file associated with this
#has been adjusted for both the regression sampled
mrja<-read.table("abund for jaccards.txt",header=T)
vari<-read.table("ordivari.txt",header=T)
newmrja <- mrja[c(7:34)]
mrjavar<-mrja[c(2:6)]
```

```
###change this line around in order to get different dissimilarity indices
jac<-vegdist(newmrja,method="jaccard")
```

```
###trying to figure out correct number of dimensions for jaccards
```

```
screedim2<-NULL
for(i in 1:30){
  screedim2 [i]= isoMDS(jac,k=2)$stress
}
screedim3<-NULL
for(i in 1:30){
  screedim3 [i]= isoMDS(jac,k=3)$stress
}
screedim4<-NULL
for(i in 1:30){
  screedim4 [i]= isoMDS(jac,k=4)$stress
}
screedim5<-NULL
for(i in 1:30){
  screedim5 [i]= isoMDS(jac,k=5)$stress
}
screedim6<-NULL
for(i in 1:30){
  screedim6 [i]= isoMDS(jac,k=6)$stress
}
scree.frame<-cbind(c(screedim2),c(screedim3),c(screedim4),c(screedim5),c(screedim6))
write.csv(scree.frame,"scree.plotiso4_27_15.csv")
```

```
###MNDS actual ordination--the metaMNDS uses several random starts in order to avoid
```

```

#a local minimum, also you need to change the number of "trymax" until when you call
#the ordination it tells you that you had at least 2 convergent solutions
ordj1<-metaMDS(jac,k=3,trymax=100,engine="isoMDS",autotransform=FALSE)
ordj1

ordj2<-
metaMDS(jac,k=3,trymax=100,engine="isoMDS",autotransform=FALSE,previous.best=ordj1)
ordj2

###plotting the ordination
par(mfrow=c(2,2))
plot(ordj2,display=c("sites"),choices=c(1,2),type="p")
plot(ordj2,display=c("sites"),choices=c(1,3),type="p")
plot(ordj2,display=c("sites"),choices=c(3,2),type="p")
ordiplot3d(ordj2)
stressplot(ordj2)

###goodness of fit
good1<-goodness(ordj1)
plot(good1)

###ordination graphics
#region
par(mfrow=c(1,3))
plot(ordj1,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj1,region,show.groups="p",col="blue",lwd=2,draw="lines",label=TRUE,
cex=1.5))
with(vari,ordihull(ordj1,region,show.groups="c",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))
with(vari,ordihull(ordj1,region,show.groups="r",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

plot(ordj1,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj1,region,show.groups="p",col="blue",lwd=2,draw="lines",label=TRUE,
cex=1.5))
with(vari,ordihull(ordj1,region,show.groups="c",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))
with(vari,ordihull(ordj1,region,show.groups="r",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

plot(ordj1,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj1,region,show.groups="p",col="blue",lwd=2,draw="lines",label=TRUE,
cex=1.5))
with(vari,ordihull(ordj1,region,show.groups="c",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))

```

```

with(vari,ordihull(ordj1,region,show.groups="r",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

#landuse
par(mfrow=c(2,2))
plot(ordj1,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj1,landuse,show.groups="n",col="blue",lwd=2,draw="lines",label=TR
UE,cex=1.5))
with(vari,ordihull(ordj1,landuse,show.groups="d",col="red",lwd=2,draw="lines",label=TRU
E,cex=1.5))

plot(ordj1,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj1,landuse,show.groups="n",col="blue",lwd=2,draw="lines",label=TR
UE,cex=1.5))
with(vari,ordihull(ordj1,landuse,show.groups="d",col="red",lwd=2,draw="lines",label=TRU
E,cex=1.5))

plot(ordj1,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj1,landuse,show.groups="n",col="blue",lwd=2,draw="lines",label=TR
UE,cex=1.5))
with(vari,ordihull(ordj1,landuse,show.groups="d",col="red",lwd=2,draw="lines",label=TRU
E,cex=1.5))

#veg height
par(mfrow=c(2,2))
plot(ordj1,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj1,veght,show.groups="s",col="blue",lwd=2,draw="lines",label=TRUE
,cex=1.5))
with(vari,ordihull(ordj1,veght,show.groups="m",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))
with(vari,ordihull(ordj1,veght,show.groups="t",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

plot(ordj1,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj1,veght,show.groups="s",col="blue",lwd=2,draw="lines",label=TRUE
,cex=1.5))
with(vari,ordihull(ordj1,veght,show.groups="m",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))
with(vari,ordihull(ordj1,veght,show.groups="t",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

plot(ordj1,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj1,veght,show.groups="s",col="blue",lwd=2,draw="lines",label=TRUE
,cex=1.5))

```

```

with(vari,ordihull(ordj1,veght,show.groups="m",col="dark
green",lwd=2,draw="lines",label=TRUE,cex=1.5))
  with(vari,ordihull(ordj1,veght,show.groups="t",col="red",lwd=2,draw="lines",label=TRUE,
cex=1.5))

#by sampling event
plot(ordj1,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj1,samp_event,show.groups="1",col="blue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="2",col="black",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="3",col="red",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="4",col="pink",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="5",col="green",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="6",col="cadetblue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="7",col="chocolate",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="8",col="cornflowerblue",lwd=2,draw="li
nes"))
  with(vari,ordihull(ordj1,samp_event,show.groups="9",col="darkgoldenrod",lwd=2,draw="li
nes"))
  with(vari,ordihull(ordj1,samp_event,show.groups="10",col="darkorchid",lwd=2,draw="lines
"))
  with(vari,ordihull(ordj1,samp_event,show.groups="11",col="gold",lwd=2,draw="lines"))
  with(vari,ordihull(ordj1,samp_event,show.groups="12",col="deeppink",lwd=2,draw="lines")
)

plot(ordj1,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj1,samp_event,show.groups="1",col="blue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="2",col="black",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="3",col="red",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="4",col="pink",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="5",col="green",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="6",col="cadetblue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="7",col="chocolate",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="8",col="cornflowerblue",lwd=2,draw="li
nes"))
  with(vari,ordihull(ordj1,samp_event,show.groups="9",col="darkgoldenrod",lwd=2,draw="li
nes"))
  with(vari,ordihull(ordj1,samp_event,show.groups="10",col="darkorchid",lwd=2,draw="lines
"))
  with(vari,ordihull(ordj1,samp_event,show.groups="11",col="gold",lwd=2,draw="lines"))
  with(vari,ordihull(ordj1,samp_event,show.groups="12",col="deeppink",lwd=2,draw="lines")
)

plot(ordj1,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj1,samp_event,show.groups="1",col="blue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="2",col="black",lwd=2,draw="lines"))

```

```

with(vari,ordihull(ordj1,samp_event,show.groups="3",col="red",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="4",col="pink",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="5",col="green",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="6",col="cadetblue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="7",col="chocolate",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="8",col="cornflowerblue",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="9",col="darkgoldenrod",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="10",col="darkorchid",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="11",col="gold",lwd=2,draw="lines"))
with(vari,ordihull(ordj1,samp_event,show.groups="12",col="deeppink",lwd=2,draw="lines"))
)

```

```

#region:land
par(mfrow=c(2,2))
plot(ordj2,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj2,reg_land,draw="lines",col=c(1,2,3,4,5,6), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj2,reg_land,draw="lines",col=c(1,2,3,4,5,6), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj2,reg_land,draw="lines",col=c(1,2,3,4,5,6), label=TRUE))

```

```

#region:veg
par(mfrow=c(2,2))
plot(ordj2,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj2,reg_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj2,reg_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj2,reg_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

```

```

#land:veg
par(mfrow=c(2,2))
plot(ordj2,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj2,land_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj2,land_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

```

```

plot(ordj2,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj2,land_veg,draw="lines",col=c(1,2,3,4,5,6,7,8,9), label=TRUE))

#reg:land:veg
par(mfrow=c(2,2))
plot(ordj2,display=c("sites"),choices=c(1,2),type="p")
with(vari,ordihull(ordj2,reg_1_v,draw="lines",col=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18), label=TRUE))

plot(ordj2,display=c("sites"),choices=c(1,3),type="p")
with(vari,ordihull(ordj2,reg_1_v,draw="lines",col=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18), label=TRUE))

plot(ordj2,display=c("sites"),choices=c(2,3),type="p")
with(vari,ordihull(ordj2,reg_1_v,draw="lines",col=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18), label=TRUE))

###MRPP
#mrpp1<-mrpp(jac,mrjavar$region,permutations=999,weight.type=1)
#mrpp1

#mrpp2<-mrpp(jac,mrjavar$landuse,permutations=999,weight.type=1)
#mrpp2

#mrpp3<-mrpp(jac,mrjavar$veght,permutations=999,weight.type=1)
#mrpp3

#mrpp4<-mrpp(jac,mrjavar$Loc,permutations=999,weight.type=1)
#mrpp4

#mrpp5<-mrpp(jac,mrjavar$samp_event,permutations=999,weight.type=1)
#mrpp5

###adonis stuff
ad1<-
adonis(jac~region*landuse+veght,data=mrjavar,permutations=999,strata=mrjavar$samp_event)
ad1

```

Indicator species

```

rm(list=ls(all=TRUE))
library(indicspecies)

```

```

####adjusted for both the regression sampled
mrja<-read.table("abund for jaccards.txt",header=T)
vari<-read.table("ordivari.txt",header=T)
newmrja <- mrja[c(7:34)]

####indicator species analysis
#use func="IndVal.g" because it equalizes the statistic for grouping together groups
ind1<-multipatt(newmrja,vari$region,func="IndVal.g",control=how(nperm=999))
summary(ind1,indvalcomp=TRUE,alpha=1)
summary(ind1)

ind2<-multipatt(newmrja,vari$landuse,func="IndVal.g",control=how(nperm=999))
summary(ind2,indvalcomp=TRUE,alpha=1)
summary(ind2)

ind3<-
multipatt(newmrja,vari$veght,func="IndVal.g",control=how(nperm=999),restcomb=c(1,2,3,4,5))
summary(ind3,indvalcomp=TRUE,alpha=1)
summary(ind3)

```

Chapter 3

Abundance models

```

rm(list=ls(all=TRUE))
## models for abundance
library(spdep)
library(lme4)
library(bbmle)
library(car)
library(lattice)
library(influence.ME)
library(nlme)
mosfire<-read.table("variables.txt",header=T)

labd<-log(mosfire$abundance+1)

mosfireabd<-cbind(mosfire,labd)

##### create distance variable

cord<-read.table("ptutms.txt",header=T)

xy1<-SpatialPoints(cord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))

```


#the nob's part of the following line is the largest difference between points aht you want to be compared

```
distlabd<-autocov_dist(mosfireabd$labd,xy1,nbs=4000)
```

```
mosfireabd<-cbind(mosfireabd,distlabd)
```

```
### testing for normality of the response variable
```

```
shapiro.test(mosfireabd$labd)
```

```
newdata<-mosfire[which(mosfireabd$tsf=='s'),]
```

```
shapiro.test(log(newdata$abundance+1))
```

```
newdata1<-mosfire[which(mosfire$tsf=='l'),]
```

```
shapiro.test(log(newdata1$abundance+1))
```

```
newdata2<-mosfire[which(mosfire$tsf=='m'),]
```

```
shapiro.test(log(newdata2$abundance+1))
```

```
#### look normal
```

```
#testing for hoeogeneity of variance across groups
```

```
leveneTest(log(mosfire$abundance+1),mosfire$tsf,center=median)
```

```
#looks good
```

```
hist(mosfireabd$abundance)
```

```
hist(mosfireabd$labd)
```

```
plot(mosfireabd$labd,mosfireabd$distlabd,main="log abundance v dist coef")
```

```
boxplot(mosfireabd$labd~mosfireabd$sampling_event,main="log abundance v sampling event")
```

```
boxplot(mosfireabd$labd~mosfireabd$location,main="log abundance v location")
```

```
#####
```

```
boxplot(mosfireabd$labd~mosfireabd$tsf,main=" log abundance v tsf")
```

```
###models for abundance and tsf
```

```
#random structure
```

```
###v tsf random factors
```

```
m1<-gls(labd~distlabd+tsf,data=mosfireabd,method="REML")
```

```
m2<-lmer(labd~distlabd+tsf+(1|sampling_event),data=mosfireabd)
```

```
m3<-lmer(labd~distlabd+tsf+(distlabd|sampling_event), data= mosfireabd)
```

```

m4<-lmer(labd~distlabd+tsf+(1|location),data=mosfireabd)
m5<-lmer(labd~distlabd+tsf+(distlabd|location),data=mosfireabd)
m6<-lmer(labd~distlabd+tsf+(1|sampling_event) + (1|location),data=mosfireabd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins

##fixed sturcture
m2a1<-lmer(labd~distlabd+tsf+(1|sampling_event),data=mosfireabd)
m2a1L<-refitML(m2a1)
m2a2<-lmer(labd~distlabd+(1|sampling_event),data=mosfireabd)
m2a2L<-refitML(m2a2)
m2a3<-lmer(labd~tsf+(1|sampling_event),data=mosfireabd)
m2a3L<-refitML(m2a3)

AICctab(m2a1L,m2a2L,m2a3L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model m2a3L wins

summary(m2a3)
confm2a3<-confint(m2a3,method="boot",nsim=1000,level=0.95)
confm2a3

#validation
par(mfrow=c(2,2))
resm2a3<-residuals(m2a3,type="pearson")
F2a3r<-fitted(m2a3,level=0)
plot(fitted(m2a3),resm2a3,xlab="fitted values", ylab="norm resid")
boxplot(resm2a3~mosfireabd$tsf,xlab="time since fire", ylab="norm resid")
hist(resm2a3)

#####
#gla pictures
plot(mosfireabd$cnpy0,mosfireabd$cnpy1)
plot(mosfireabd$cnpy0,mosfireabd$cnpy2)
plot(mosfireabd$cnpy1,mosfireabd$cnpy2)

#random sturcture
m1<-gls(labd~distlabd+cnpy0*cnpy1*cnpy2,data=mosfireabd,method="REML")
m2<-lmer(labd~distlabd+cnpy0*cnpy1*cnpy2+(1|sampling_event),data=mosfireabd)
m3<-lmer(labd~distlabd+cnpy0*cnpy1*cnpy2+(distlabd|sampling_event), data= mosfireabd)
m4<-lmer(labd~distlabd+cnpy0*cnpy1*cnpy2+(1|location),data=mosfireabd)

```

```

m5<-lmer(labd~distlabd+cnpy0*cnpy1*cnpy2+(distlabd|location),data=mosfireabd)
m6<-lmer(labd~distlabd+cnpy0*cnpy1*cnpy2+(1|sampling_event) +
(1|location),data=mosfireabd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins

#find mods in "gla abund mods.txt"
bm3<-lmer(labd~ +cnpy1 + + + + + (1|sampling_event),data=mosfireabd)
lbm3<-refitML(bm3)
#model m2b32wins
summary(m2b32)
confm2b32<-confint(m2b32,method="boot",nsim=1000,level=0.95)
confm2b32

par(mfrow=c(2,2))
resm2b32<-residuals(m2b32,type="pearson")
Fm2b32<-fitted(m2b32,level=0)
plot(fitted(m2b32),resm2b32,xlab="fitted values", ylab="norm resid")
plot(mosfireabd$cnpy1,resm2b32,xlab="% canopy at 0",ylab="norm resid")
hist(resm2b32)

#####
###full percent veg data
#random structure
m1<-gls( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter,data=mosfireabd,method="REML")
m2<-lmer( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(1|sampling_event),data=mosfireabd)
m3<-lmer( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(distlabd|sampling_event), data= mosfireabd)
m4<-lmer( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(1|location),data=mosfireabd)
m5<-lmer( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(distlabd|location),data=mosfireabd)
m6<-lmer( labd~distlabd+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(1|sampling_event) + (1|location),data=mosfireabd)

```

```

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
#model 2 wins

n <- 7
l <- rep(list(0:1), n)

vegmat<-expand.grid(l)
write.csv(vegmat,"pervegmat.csv")

m6<-lmer(labd~ per_grass +per_wood_trees + (1|sampling_event),data=mosfireabd)
lm6<-refitML(m6)

#vegmods are in the "per veg abund mods.txt" file

summary(m6)
confm6<-confint(m6,method="boot",nsim=10000,level=0.95)
confm6

par(mfrow=c(2,2))
resm6<-residuals(m6,type="pearson")
Fm6<-fitted(m6,level=0)
plot(fitted(m6),resm6,xlab="fitted values", ylab="norm resid")
plot(mosfireabd$per_grass,resm6,xlab="percent grass",ylab="norm resid")
plot(mosfireabd$per_wood_trees,resm6,xlab="percent woody trees",ylab="norm resid")
hist(resm6)

#####
#tallest veg mods
m1<-gls(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto,data=mosfireabd,method="REML")
m2<-lmer(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|sampling_event),data=mosfireabd)
m3<-lmer(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(distlabd|sampling_event), data= mosfireabd)
m4<-lmer(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|location),data=mosfireabd)
m5<-lmer(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(distlabd|location),data=mosfireabd)

```

```

m6<-lmer(labd~distlabd+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|sampling_event) + (1|location),data=mosfireabd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#mod 2 wins
n <- 5
l <- rep(list(0:1), n)

tvegmat<-expand.grid(l)
write.csv(tvegmat,"tallvegmat.csv")

m21<-lmer(labd~ + +tall_tree + +tall_palmetto +(1|sampling_event),data=mosfireabd)
lm21<-refitML(m21)
#vegmods are in the "per veg abund mods.txt" file

summary(m21)
confm21<-confint(m21,method="boot",nsim=10000,level=0.95)
confm21

par(mfrow=c(2,2))
resm21<-residuals(m21,type="pearson")
Fm21<-fitted(m21,level=0)
plot(fitted(m21),resm21,xlab="fitted values", ylab="norm resid")
plot(mosfireabd$tall_tree,resm21,xlab="tallest tree",ylab="norm resid")
plot(mosfireabd$tall_palmetto,resm21,xlab="tallest palmetto",ylab="norm resid")
hist(resm21)

#####
###just distance v abundance
### random structure
d1<-gls(labd~distlabd ,data=mosfireabd,method="REML")
d2<-lmer(labd~distlabd+(1|sampling_event),data=mosfireabd)
d3<-lmer(labd~distlabd+(distlabd|sampling_event), data= mosfireabd)
d4<-lmer(labd~distlabd+(1|location),data=mosfireabd)
d5<-lmer(labd~distlabd+(distlabd|location),data=mosfireabd)
d6<-lmer(labd~distlabd+(1|sampling_event) + (1|location),data=mosfireabd)

AICctab(d1,d2,d3,d4,d5,d6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
#2 wins again

ld2<-refitML(d2)

```

```

summary(d2)
confd2<-confint(d2,method="boot",nsim=1000,level=0.95)
confd2

par(mfrow=c(2,2))
resd2<-residuals(d2,type="pearson")
Fm2<-fitted(d2,level=0)
plot(fitted(d2),resd2,xlab="fitted values", ylab="norm resid")
plot(distlabd,resd2,xlab="distance coefficient",ylab="norm resid")
hist(resd2)

#####
#pca veg

vegdata<-mosfire[,15:26]

summary(model1<-prcomp(vegdata,scale=TRUE))
plot(model1,main="",col="green")
biplot(model1)

yv<-predict(model1)[,1]
yv2<-predict(model1)[,2]
yv3<-predict(model1)[,3]

### random structure
m1<-gls(labd~distlabd+yv+yv2+yv3,data=mosfireabd,method="REML")
m2<-lmer(labd~distlabd+yv+yv2+yv3+(1|sampling_event),data=mosfireabd)
m3<-lmer(labd~distlabd+yv+yv2+yv3+(distlabd|sampling_event), data= mosfireabd)
m4<-lmer(labd~distlabd+yv+yv2+yv3+(1|location),data=mosfireabd)
m5<-lmer(labd~distlabd+yv+yv2+yv3+(distlabd|location),data=mosfireabd)
m6<-lmer(labd~distlabd+yv+yv2+yv3+(1|sampling_event) + (1|location),data=mosfireabd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#fixed structure
m21<-lmer(labd~distlabd+yv+yv2+yv3+(1|sampling_event),data=mosfireabd)
m22<-lmer(labd~yv+yv2+yv3+(1|sampling_event),data=mosfireabd)
m23<-lmer(labd~yv+yv2+(1|sampling_event),data=mosfireabd)
m24<-lmer(labd~yv+yv3+(1|sampling_event),data=mosfireabd)
m25<-lmer(labd~yv2+yv3+(1|sampling_event),data=mosfireabd)
m26<-lmer(labd~yv+(1|sampling_event),data=mosfireabd)
m27<-lmer(labd~yv2+(1|sampling_event),data=mosfireabd)
m28<-lmer(labd~yv3+(1|sampling_event),data=mosfireabd)

```

```

lm21<-refitML(m21)
lm22<-refitML(m22)
AICctab(lm21,lm22,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

lm22<-refitML(m22)
lm23<-refitML(m23)
lm24<-refitML(m24)
lm25<-refitML(m25)
lm26<-refitML(m26)
lm27<-refitML(m27)
lm28<-refitML(m8)
AICctab(lm21,lm22,lm23,lm24,lm25,lm26,lm27,lm28,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

summary(m22)

confm22<-confint(m22,method="boot",nsim=1000,level=0.95)
confm22

par(mfrow=c(3,2))
resm22<-residuals(m22,type="pearson")
Fm22<-fitted(m22,level=0)
plot(fitted(m22),resm22,xlab="fitted values",ylab="norm resid")
plot(yv,resm22,xlab="PCA Axis 1",ylab="norm resid")
plot(yv2,resm22,xlab="PCA Axis 2",ylab="norm resid")
plot(yv3,resm22,xlab="PCA axis 3",ylab="norm resid")
hist(resm22)

#####
#AIC comparisons

AICctab(lm22,lm6,lm21,ld2,m2b32l,m2a3L,sort=T,weights=T,delta=T,nobs=nrow(mosfireabd))

library(MuMIn)
r.squaredGLMM(lm6)

Jost D

rm(list=ls(all=TRUE))
mosfire<-read.table("variables.txt",header=T)
library(spdep)

```

```

library(lme4)
library(bbmle)
library(car)
library (lattice)
library(influence.ME)
library (nlme)
#####
check for normality

shapiro.test(mosfire$jost_ind)
shapiro.test(log(mosfire$jost_ind))

leveneTest(mosfire$jost_ind,mosfire$tsf,center=median)

leveneTest(log(mosfire$jost_ind),mosfire$tsf,center=median)

#####
ljd<-(log(mosfire$jost_ind))
mosfirejd<-cbind(mosfire,ljd)

cord<-read.table("ptutms.txt",header=T)

xy1<-SpatialPoints(cord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nobs part of the following line is the largest difference between points aht you want to be
compared
distljd<-autocov_dist(mosfirejd$ljd,xy1,nbs=4000)

mosfirejd<-cbind(mosfirejd,distljd)

hist(mosfirejd$jost_ind)
hist(mosfirejd$ljd)

plot(mosfirejd$ljd,mosfirejd$distljd,main="log jost v dist coef")
boxplot(mosfirejd$ljd~mosfirejd$sampling_event,main="log jost v sampling event")
boxplot(mosfirejd$ljd~mosfirejd$location,main="log jost v location")

#####
boxplot(mosfirejd$ljd~mosfirejd$tsf,main=" log jost v tsf")

###v tsf random factors
m1<-gls(ljd~distljd+tsf,data=mosfirejd,method="REML")
m2<-lmer(ljd~distljd+tsf+(1|sampling_event),data=mosfirejd)
m3<-lmer(ljd~distljd+tsf+(distljd|sampling_event), data= mosfirejd)

```



```

m4<-lmer(ljd~distljd+tsf+(1|location),data=mosfirejd)
m5<-lmer(ljd~distljd+tsf+(distljd|location),data=mosfirejd)
m6<-lmer(ljd~distljd+tsf+(1|sampling_event) + (1|location),data=mosfirejd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins
#####fixed factors

m2a1<-lmer(ljd~distljd+tsf+(1|sampling_event),data=mosfirejd)
m2a1L<-refitML(m2a1)
m2a2<-lmer(ljd~distljd+(1|sampling_event),data=mosfirejd)
m2a2L<-refitML(m2a2)
m2a3<-lmer(ljd~tsf+(1|sampling_event),data=mosfirejd)
m2a3L<-refitML(m2a3)

AICctab(m2a1L,m2a2L,m2a3L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

plot(m2a2)
summary(m2a2)
confm2a2<-confint(m2a2,method="boot",nsim=1000,level=0.95)
confm2a2

##validation
par(mfrow=c(2,2))
resm2a2<-residuals(m2a2,type="pearson")
F2a2r<-fitted(m2a2,level=0)
plot(fitted(m2a2),resm2a2,xlab="fitted values", ylab="norm resid")
plot(distljd,resm2a2,xlab="distance coefficient",ylab="norm resid")
hist(resm2a2)

#####
### jost v gla pics
plot(mosfirejd$cnpy0,mosfirejd$cnpy1)
plot(mosfirejd$cnpy0,mosfirejd$cnpy2)
plot(mosfirejd$cnpy1,mosfirejd$cnpy2)

### random structure
m1<-gls(ljd~distljd+cnpy0*cnpy1*cnpy2,data=mosfirejd,method="REML")
m2<-lmer(ljd~distljd+cnpy0*cnpy1*cnpy2+(1|sampling_event),data=mosfirejd)
m3<-lmer(ljd~distljd+cnpy0*cnpy1*cnpy2+(distljd|sampling_event), data= mosfirejd)

```

```

m4<-lmer(ljd~distljd+cnp0*cnp1*cnp2+(1|location),data=mosfirejd)
m5<-lmer(ljd~distljd+cnp0*cnp1*cnp2+(distljd|location),data=mosfirejd)
m6<-lmer(ljd~distljd+cnp0*cnp1*cnp2+(1|sampling_event) +
(1|location),data=mosfirejd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#mod 2 wins

###fixed structure
#find mods "gla jost mods.txt"

bm9<-lmer(ljd~distljd+ + + +cnp0:cnp1 + + + (1|sampling_event),data=mosfirejd)
bm2<-lmer(ljd~distljd+cnp0 + + + + + + (1|sampling_event),data=mosfirejd)
bm17<-lmer(ljd~distljd+ + + + +cnp0:cnp2 + + + (1|sampling_event),data=mosfirejd)
bm65<-lmer(ljd~distljd+ + + + + + +cnp0:cnp1:cnp2 +
(1|sampling_event),data=mosfirejd)
lbm9<-refitML(bm9)
lbm2<-refitML(bm2)
lbm17<-refitML(bm17)
lbm65<-refitML(bm65)

summary(bm9)
confbm9<-confint(bm9,method="boot",nsim=1000,level=0.95)
confbm9

##validation
par(mfrow=c(3,2))
resbm9<-residuals(bm9,type="pearson")
Fbm9<-fitted(bm9,level=0)
plot(fitted(bm9),resbm9,xlab="fitted values", ylab="norm resids")
plot(distljd,resbm9,xlab="distance coefficient",ylab="norm resids")
plot(mosfirejd$cnp0,resbm9,xlab="% canopy at 0",ylab="norm resids")
plot(mosfirejd$cnp1,resbm9,xlab="% canopy at 1",ylab="norm resids")
hist(resbm9)

#####
###veg
vegie<-mosfirejd[,15:26]
pairs(vegie)

### random structure
m1<-glsl(ljd~distljd+per_grass+per_subshrub +

```

```

    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter,data=mosfirejd,method="REML")
m2<-lmer(ljd~distljd+per_grass+per_subshrub +
    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter+(1|sampling_event),data=mosfirejd)
m3<-lmer(ljd~distljd+per_grass+per_subshrub +
    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter+(distljd|sampling_event), data= mosfirejd)
m4<-lmer(ljd~distljd+per_grass+per_subshrub +
    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter+(1|location),data=mosfirejd)
m5<-lmer(ljd~distljd+per_grass+per_subshrub +
    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter+(distljd|location),data=mosfirejd)
m6<-lmer(ljd~distljd+per_grass+per_subshrub +
    per_wood_trees+per_Herb+per_palmetto+
    +per_sand+per_litter+(1|sampling_event) + (1|location),data=mosfirejd)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins

#n <- 6
#l <- rep(list(0:1), n)

#vegmat<-expand.grid(l)
#write.csv(vegmat,"vegmatsmall.csv")

#find veg mods at "per veg jost mods.txt"

#used vegadd mods to do the rest of the models
m6<-lmer(ljd~distljd+ per_grass +per_wood_trees + (1|sampling_event),data=mosfirejd)
lm6<-refitML(m6)

summary(m6)
confm6<-confint(m6,method="boot",nsim=10000,level=0.95)
confm6

##validation
par(mfrow=c(3,2))
resm6<-residuals(m6,type="pearson")
Fm6<-fitted(m6,level=0)
plot(fitted(m6),resm6,xlab="fitted values", ylab="norm resid")
plot(distljd,resm6,xlab="distance coefficient",ylab="norm resid")

```

```

plot(mosfirejd$per_grass, resm6, xlab="percent grass", ylab="norm resids")
plot(mosfirejd$per_wood_trees, resm6, xlab="percent woody trees", ylab="norm resids")
hist(resm6)

#####
### random structure
m1 <- gls(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto, data = mosfirejd, method = "REML")
m2 <- lmer(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto + (1|sampling_event), data = mosfirejd)
m3 <- lmer(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto + (distljd|sampling_event), data = mosfirejd)
m4 <- lmer(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto + (1|location), data = mosfirejd)
m5 <- lmer(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto + (distljd|location), data = mosfirejd)
m6 <- lmer(ljd ~ distljd + tall_grass +
  tall_subshrub + tall_tree + tall_herb +
  tall_palmetto + (1|sampling_event) + (1|location), data = mosfirejd)

AICctab(m1, m2, m3, m4, m5, m6, weights = TRUE, delta = TRUE, base = TRUE, sort = TRUE)

#model 2 wins
#find mods in "tall veg jost mods.txt"
m4 <- lmer(ljd ~ distljd + tall_grass + tall_subshrub + + +
+(1|sampling_event), data = mosfirejd)
lm4 <- refitML(m4)

summary(m4)
confm4 <- confint(m4, method = "boot", nsim = 10000, level = 0.95)
confm4

par(mfrow = c(3, 2))
resm4 <- residuals(m4, type = "pearson")
Fm6 <- fitted(m4, level = 0)
plot(fitted(m4), resm4, xlab = "fitted values", ylab = "norm resids")
plot(distljd, resm4, xlab = "distance coefficient", ylab = "norm resids")
plot(mosfirejd$tall_grass, resm4, xlab = "tall grass", ylab = "norm resids")
plot(mosfirejd$tall_subshrub, resm4, xlab = "tall subshrub", ylab = "norm resids")
hist(resm4)

```

```

#####
###just distance

### random structure
d1<-gls(ljd~distljd ,data=mosfirejd,method="REML")
d2<-lmer(ljd~distljd+(1|sampling_event),data=mosfirejd)
d3<-lmer(ljd~distljd+(distljd|sampling_event), data= mosfirejd)
d4<-lmer(ljd~distljd+(1|location),data=mosfirejd)
d5<-lmer(ljd~distljd+(distljd|location),data=mosfirejd)
d6<-lmer(ljd~distljd+(1|sampling_event) + (1|location),data=mosfirejd)

AICctab(d1,d2,d3,d4,d5,d6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#m2 wins
ld2<-refitML(d2)

plot(distljd,ljd)

##validation
par(mfrow=c(2,2))
resd2<-residuals(d2,type="pearson")
Fm2<-fitted(d2,level=0)
plot(fitted(d2),resd2,xlab="fitted values", ylab="norm resids")
plot(distljd,resd2,xlab="distance coefficient",ylab="norm resids")
hist(resd2)

confd2<-confint(d2,method="boot",nsim=1000,level=0.95)
confd2

summary(d2)

#####
##pca veg
vegdata<-mosfire[,15:26]

summary(model1<-prcomp(vegdata,scale=TRUE))
plot(model1,main="",col="green")
biplot(model1)

yv<-predict(model1)[,1]
yv2<-predict(model1)[,2]
yv3<-predict(model1)[,3]

```

```
mosfirejd<-cbind(mosfirejd,yv,yv2,yv3)
```

```
### random structure
```

```
m1<-glsl(ljd~distljd+yv+yv2+yv3,data=mosfirejd,method="REML")
```

```
m2<-lmer(ljd~distljd+yv+yv2+yv3+(1|sampling_event),data=mosfirejd)
```

```
m3<-lmer(ljd~distljd+yv+yv2+yv3+(distljd|sampling_event), data= mosfirejd)
```

```
m4<-lmer(ljd~distljd+yv+yv2+yv3+(1|location),data=mosfirejd)
```

```
m5<-lmer(ljd~distljd+yv+yv2+yv3+(distljd|location),data=mosfirejd)
```

```
m6<-lmer(ljd~distljd+yv+yv2+yv3+(1|sampling_event) + (1|location),data=mosfirejd)
```

```
AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#model 2 wins again
```

```
m21<-lmer(ljd~distljd+yv+yv2+yv3+(1|sampling_event),data=mosfirejd)
```

```
m22<-lmer(ljd~yv+yv2+yv3+(1|sampling_event),data=mosfirejd)
```

```
m23<-lmer(ljd~distljd+yv+yv2+(1|sampling_event),data=mosfirejd)
```

```
m24<-lmer(ljd~distljd+yv+yv3+(1|sampling_event),data=mosfirejd)
```

```
m25<-lmer(ljd~distljd+yv2+yv3+(1|sampling_event),data=mosfirejd)
```

```
m26<-lmer(ljd~distljd+yv+(1|sampling_event),data=mosfirejd)
```

```
m27<-lmer(ljd~distljd+yv2+(1|sampling_event),data=mosfirejd)
```

```
m28<-lmer(ljd~distljd+yv3+(1|sampling_event),data=mosfirejd)
```

```
lm21<-refitML(m21)
```

```
lm22<-refitML(m22)
```

```
lm23<-refitML(m23)
```

```
lm24<-refitML(m24)
```

```
lm25<-refitML(m25)
```

```
lm26<-refitML(m26)
```

```
lm27<-refitML(m27)
```

```
lm28<-refitML(m28)
```

```
AICctab(lm21,lm22,lm23,lm24,lm25,lm26,lm27,lm28,weights=TRUE,delta=TRUE,base=T  
RUE,sort=TRUE)
```

```
summary(m24)
```

```
confm24<-confint(m24,method="boot",nsim=10000,level=0.95)
```

```
confm24
```

```
par(mfrow=c(3,2))
```

```
resm24<-residuals(m24,type="pearson")
```

```
Fm24<-fitted(m24,level=0)
```

```
plot(fitted(m24),resm24,xlab="fitted values", ylab="norm resid")
```

```

plot(distljd,resm24,xlab="distance coefficient",ylab="norm resids")
plot(yv,resm24,xlab="PCA Axis 1",ylab="norm resids")
plot(yv3,resm24,xlab="PCA axis 3",ylab="norm resids")
hist(resm24)

```

```

#####
#AIC comparisons

```

```

AICctab(m2a2L,lm24,lm4,lm6,lbm9,lbm2,lbm17,lbm65,sort=T,weights=T,delta=T,base=T)

```

```

library(MuMIn)
r.squaredGLMM(lm24)

```

Species richness models

```

rm(list=ls(all=TRUE))
mosfire<-read.table("variables.txt",header=T)
##### create distance variable
library(spdep)
library(lme4)
library(bbmle)
library(car)
library(lattice)
library(influence.ME)
library(nlme)

```

```

cord<-read.table("ptutms.txt",header=T)

```

```

xy1<-SpatialPoints(cord,proj4string=CRS("+proj=utm +zone=17+datum=WGS84"))
#the nobs part of the following line is the largest difference between points aht you want to be
compared

```

```

distspr<-autocov_dist(mosfire$spec_rich,xy1,nbs=4000)

```

```

mosfirespr<-cbind(mosfire,distspr)

```

```

### testing for normality of the response variable
shapiro.test(mosfirespr$spec_rich)
shapiro.test((log(mosfirespr$spec_rich+1)))

```

```

###

```

```

newdata1<-mosfirespr[which(mosfirespr$tsf=='l'),]
shapiro.test(newdata1$spec_rich)

newdata2<-mosfirespr[which(mosfirespr$tsf=='m'),]
shapiro.test(newdata2$spec_rich)

##with log trans
newdata<-mosfirespr[which(mosfirespr$tsf=='s'),]
shapiro.test(newdata$l spr)

newdata1<-mosfirespr[which(mosfirespr$tsf=='l'),]
shapiro.test(newdata1$l spr)

newdata2<-mosfirespr[which(mosfirespr$tsf=='m'),]
shapiro.test(newdata2$l spr)

#testing for hoeogeneity of variance across groups
library(car)

leveneTest(mosfirespr$spec_rich,mosfirespr$tsf,center=median)

leveneTest(mosfirespr$l spr,mosfirespr$tsf,center=median)

hist(mosfirespr$spec_rich)
hist(mosfirespr$l spr)

plot(mosfirespr$spec_rich,mosfirespr$distspr,main=" species richness v dist coef")
boxplot(mosfirespr$spec_rich~mosfirespr$sampling_event,main=" species richness v
sampling event")
boxplot(mosfirespr$spec_rich~mosfirespr$location,main=" species richness v location")

#####
###v tsf random factors--use the models with b in it becasue it takes out the location part
m1<-gls(spec_rich~distspr+tsf,data=mosfirespr,method="REML")
m2<-lmer(spec_rich~distspr+tsf+(1|sampling_event),data=mosfirespr)
m3<-lmer(spec_rich~distspr+tsf+(distspr|sampling_event), data= mosfirespr)
m4<-lmer(spec_rich~distspr+tsf+(1|location),data=mosfirespr)
m5<-lmer(spec_rich~distspr+tsf+(distspr|location),data=mosfirespr)
m6<-lmer(spec_rich~distspr+tsf+(1|sampling_event) + (1|location),data=mosfirespr)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 6 wins

#fixed structure

```



```

m2a1<-lmer(spec_rich~distspr+tsf+(1|sampling_event) + (1|location),data=mosfirespr)
m2a1L<-refitML(m2a1)
m2a2<-lmer(spec_rich~distspr+(1|sampling_event) + (1|location),data=mosfirespr)
m2a2L<-refitML(m2a2)
m2a3<-lmer(spec_rich~tsf+(1|sampling_event) + (1|location),data=mosfirespr)
m2a3L<-refitML(m2a3)

AICctab(m2a1L,m2a2L,m2a3L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model m2a2L wins

bm2a1<-lmer(spec_rich~distspr+tsf+(1|sampling_event) ,data=mosfirespr)
bm2a1L<-refitML(bm2a1)
bm2a2<-lmer(spec_rich~distspr+(1|sampling_event),data=mosfirespr)
bm2a2L<-refitML(bm2a2)
bm2a3<-lmer(spec_rich~tsf+(1|sampling_event),data=mosfirespr)
bm2a3L<-refitML(bm2a3)

AICctab(bm2a1L,bm2a2L,bm2a3L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

AICctab(m2a2L,bm2a2L,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

summary(bm2a2)
confbm2a2<-confint(bm2a2,method="boot",nsim=10000,level=0.95)
confbm2a2

##validation
par(mfrow=c(2,2))
resbm2a2<-residuals(bm2a2,type="pearson")
Fb2a2r<-fitted(bm2a2,level=0)
plot(fitted(bm2a2),resbm2a2,xlab="fitted values", ylab="norm resid")
plot(distspr,resbm2a2,xlab="distance coefficient",ylab="norm resid")
hist(resbm2a2)

#####
#####GLA pics
### random structure
m1<-glms(spec_rich~distspr+cnp0*cnp1*cnp2,data=mosfirespr,method="REML")
m2<-lmer(spec_rich~distspr+cnp0*cnp1*cnp2+(1|sampling_event),data=mosfirespr)
m3<-lmer(spec_rich~distspr+cnp0*cnp1*cnp2+(distspr|sampling_event), data=
mosfirespr)
m4<-lmer(spec_rich~distspr+cnp0*cnp1*cnp2+(1|location),data=mosfirespr)
m5<-lmer(spec_rich~distspr+cnp0*cnp1*cnp2+(distspr|location),data=mosfirespr)
m6<-lmer(spec_rich~distspr+cnp0*cnp1*cnp2+(1|sampling_event) +
(1|location),data=mosfirespr)

```

```

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#mod 2 wins
###fixed structure

#find mods in "gla spec rich mods.txt"
bm33<-lmer(spec_rich~distspr+ + + + +cnp1:cnp2 + +
(1|sampling_event),data=mosfirespr)
lbm33<-refitML(bm33)

summary(bm33)
confbm33<-confint(bm33,method="boot",nsim=10000,level=0.95)
confbm33

##validation
par(mfrow=c(3,2))
resbm33<-residuals(bm33,type="pearson")
Fbm33<-fitted(bm33,level=0)
plot(fitted(bm33),resbm33,xlab="fitted values", ylab="norm resid")
plot(distspr,resbm33,xlab="distance coefficient",ylab="norm resid")
plot(mosfirespr$cnp1,resbm33,xlab="% canopy at 1",ylab="norm resid")
plot(mosfirespr$cnp2,resbm33,xlab="% canopy at 2",ylab="norm resid")
hist(resbm33)

#####
##
#####full veg data
### random structure
m1<-gls(spec_rich~distspr+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter,data=mosfirespr,method="REML")
m2<-lmer(spec_rich~distspr+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(1|sampling_event),data=mosfirespr)
m3<-lmer(spec_rich~distspr+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(distspr|sampling_event), data= mosfirespr)
m4<-lmer(spec_rich~distspr+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(1|location),data=mosfirespr)
m5<-lmer(spec_rich~distspr+per_grass+per_subshrub +
per_wood_trees+per_Herb+per_palmetto+
+per_sand+per_litter+(distspr|location),data=mosfirespr)

```

```

m6<-lmer(spec_rich~distspr+per_grass+per_subshrub +
  per_wood_trees+per_Herb+per_palmetto+
  +per_sand+per_litter+(1|sampling_event) + (1|location),data=mosfirespr)

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins

#for fixed look in the file "per veg spec rich mods.txt"

#lm 2 wins
m2<-lmer(spec_rich~ tall_grass + + + + +(1|sampling_event),data=mosfirespr)
lm2<-refitML(m2)
summary(m2)
confm2<-confint(m2,method="boot",nsim=10000,level=0.95)
confm2

par(mfrow=c(2,2))
resm2<-residuals(m2,type="pearson")
Fm2<-fitted(m2,level=0)
plot(fitted(m2),resm2,xlab="fitted values", ylab="norm resid")
plot(mosfirespr$per_grass,resm2,xlab="percent grass",ylab="norm resid")
hist(resm2)

#####
#tall veg mods
m1<-gls(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto,data=mosfirespr,method="REML")
m2<-lmer(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|sampling_event),data=mosfirespr)
m3<-lmer(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(distspr|sampling_event), data= mosfirespr)
m4<-lmer(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|location),data=mosfirespr)
m5<-lmer(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(distspr|location),data=mosfirespr)
m6<-lmer(spec_rich~distspr+tall_grass+
  tall_subshrub+tall_tree+tall_herb+
  tall_palmetto+(1|sampling_event) + (1|location),data=mosfirespr)

```

```

AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#model 2 wins
m6<-lmer(spec_rich~distspr+ tall_grass + +tall_tree + +
+(1|sampling_event),data=mosfirespr)
lm6<-refitML(m6)

#for fixed look in the file "veg mods for spec rich.txt"

#lm 6 wins

summary(m6)
confm6<-confint(m6,method="boot",nsim=10000,level=0.95)
confm6

par(mfrow=c(3,2))
resm6<-residuals(m6,type="pearson")
Fm6<-fitted(m6,level=0)
plot(fitted(m6),resm6,xlab="fitted values", ylab="norm resid")
plot(distspr,resm6,xlab="distance coefficient",ylab="norm resid")
plot(mosfirespr$tall_grass,resm6,xlab="tallest grass",ylab="norm resid")
plot(mosfirespr$tall_tree,resm6,xlab="tallest tree",ylab="norm resid")
hist(resm6)

#####
#just distance
### random structure/fixed
d1<-glms(spec_rich~distspr ,data=mosfirespr,method="REML")
d2<-lmer(spec_rich~distspr+(1|sampling_event),data=mosfirespr)
d3<-lmer(spec_rich~distspr+(distspr|sampling_event), data= mosfirespr)
d4<-lmer(spec_rich~distspr+(1|location),data=mosfirespr)
d5<-lmer(spec_rich~distspr+(distspr|location),data=mosfirespr)
d6<-lmer(spec_rich~distspr+(1|sampling_event) + (1|location),data=mosfirespr)

AICctab(d1,d2,d3,d4,d5,d6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)

#mod 6 wins

ld6<-refitML(d6)
summary(d6)
confd6<-confint(d6,method="boot",nsim=1000,level=0.95)
confd6

##validation

```

```

par(mfrow=c(2,2))
resd6<-residuals(d6,type="pearson")
Fm6<-fitted(d6,level=0)
plot(fitted(d6),resd6,xlab="fitted values", ylab="norm resid")
plot(distspr,resd6,xlab="distance coefficient",ylab="norm resid")
hist(resd6)

```

```
#####
```

```

###PCA for veg models
vegdata<-mosfire[,15:26]

```

```

summary(model1<-prcomp(vegdata,scale=TRUE))
plot(model1,main="",col="green")
biplot(model1, pc1,pc3)
biplot(model1, choices=c(1,3))
biplot(model1, choices=c(2,3))

```

```

yv<-predict(model1)[,1]
yv2<-predict(model1)[,2]
yv3<-predict(model1)[,3]

```

```

mosfirespr<-cbind(mosfirespr,yv,yv2,yv3)
### random structure
m1<-gls(spec_rich~distspr+yv+yv2+yv3,data=mosfirespr,method="REML")
m2<-lmer(spec_rich~distspr+yv+yv2+yv3+(1|sampling_event),data=mosfirespr)
m3<-lmer(spec_rich~distspr+yv+yv2+yv3+(distspr|sampling_event), data= mosfirespr)
m4<-lmer(spec_rich~distspr+yv+yv2+yv3+(1|location),data=mosfirespr)
m5<-lmer(spec_rich~distspr+yv+yv2+yv3+(distspr|location),data=mosfirespr)
m6<-lmer(spec_rich~distspr+yv+yv2+yv3+(1|sampling_event) +
(1|location),data=mosfirespr)

```

```
AICctab(m1,m2,m3,m4,m5,m6,weights=TRUE,delta=TRUE,base=TRUE,sort=TRUE)
```

```
#mod 2 wins
```

```

#fixed sturcture
m21<-lmer(spec_rich~distspr+yv+yv2+yv3+(1|sampling_event),data=mosfirespr)
m22<-lmer(spec_rich~yv+yv2+yv3+(1|sampling_event),data=mosfirespr)
m23<-lmer(spec_rich~distspr+yv+yv2+(1|sampling_event),data=mosfirespr)
m24<-lmer(spec_rich~distspr+yv+yv3+(1|sampling_event),data=mosfirespr)

```

```

m25<-lmer(spec_rich~distspr+yv2+yv3+(1|sampling_event),data=mosfirespr)
m26<-lmer(spec_rich~distspr+yv+(1|sampling_event),data=mosfirespr)
m27<-lmer(spec_rich~distspr+yv2+(1|sampling_event),data=mosfirespr)
m28<-lmer(spec_rich~distspr+yv3+(1|sampling_event),data=mosfirespr)

lm21<-refitML(m21)
lm22<-refitML(m22)
lm23<-refitML(m23)
lm24<-refitML(m24)
lm25<-refitML(m25)
lm26<-refitML(m26)
lm27<-refitML(m27)
lm28<-refitML(m8)
AICctab(lm21,lm22,lm23,lm24,lm25,lm26,lm27,lm28,weights=TRUE,delta=TRUE,base=T
RUE,sort=TRUE)

summary(m27)

confm27<-confint(m27,method="boot",nsim=10000,level=0.95)
confm27

par(mfrow=c(2,2))
resm27<-residuals(m27,type="pearson")
Fm27<-fitted(m27,level=0)
plot(fitted(m27),resm27,xlab="fitted values",ylab="norm resid")
plot(distspr,resm27,xlab="distance coefficient",ylab="norm resid")
plot(yv2,resm27,xlab="PCA Axis 2",ylab="norm resid")
hist(resm27)
#####
#AIC comparisons

AICctab(bm2a2L,lbm33,lm2,lm6,lm27,sort=T,weights=T,delta=T,base=T)
library(MuMIn)
r.squaredGLMM(lm6)

PCA

rm(list=ls(all=TRUE))
mosfire<-read.table("variables.txt",header=T)
vegdata<-mosfire[,15:26]

summary(model1<-prcomp(vegdata,scale=TRUE))
plot(model1,main="",col="green")

```

```
biplot(model1)
```

```
loading<-(model1$rotation)
```

```
loading
```

```
(loading)^2
```

```
yv<-predict(model1)[,1]
```

```
yv2<-predict(model1)[,2]
```

```
yv3<-predict(model1)[,3]
```

```
window(7,4)
```

```
par(mfrow=c(1,3))
```

```
plot(mosfire$tsf,yv,pch=16,xlab="time since fire",ylab = "pc1",col="red")
```

```
plot(mosfire$tsf,yv2,pch=16,xlab="time since fire",ylab = "pc2",col="red")
```

```
plot(mosfire$tsf,yv3,pch=16,xlab="time since fire",ylab = "pc3",col="red")
```

Indicator species

```
rm(list=ls(all=TRUE))
```

```
library(indicspecies)
```

```
mosfireind<-read.table("mosfire_siteid_abund.txt.",header=T)
```

```
newmosind<-mosfireind[c(5:18)]
```

```
###indicator species analysis
```

```
#use func="IndVal.g" because it equalizes the statistic for grouping together groups
```

```
ind1<-multipatt(newmosind,mosfireind$tsf,func="IndVal.g",control=how(nperm=999))
```

```
summary(ind1,indvalcomp=TRUE,alpha=1)
```

```
summary(ind1)
```

REFERENCES

- “2010 Population Finder.” United States Census Bureau. Updated January 15, 2013. Web. March 20, 2013. < <http://www.census.gov/popfinder/>>.
- Abrahamson, W.G. 1984. Species responses to fire on the Florida Lake Wales ridge. *American journal of botany*. 71:35-43.a
- Abrahamson, W.G., A.F. Johnson, J.N. Layne and P.A. Peroni. 1984. Vegetation of the Archbold Biological Station, Florida: An example of the southern Lake Wales ridge. *Florida Scientist*. 47: 211-250.b
- Abrahamson, W.G. 1984. Post-fire recovery of Florida Lake Wales Ridge vegetation. *American journal of botany*. 71:9-21.c
- Abrahamson, W.G. and C.R. Abrahamson. 1996. Effects of fire on long-unburned Florida uplands. *Journal of vegetation science*. 7:565-574.
- Ahern, M., R.S. Kovats, P. Wilkinson, R. Few, and F. Matthies. 2005. Global health impacts of floods: Epidemiologic evidence. *Epidemiologic Reviews*. 27:36-46.
- Anderson, M.J., T.O. Crist, J.M. Chase, M. Vellend, B.D. Inouye, A.L. Freestone, N.J. Sanders, H.V. Cornell, L.S. Comita, K.F. Davies, S.P. Harrison, N.J.B. Kraft, J.C. Stegen, and N.G. Swenson. 2010. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*. 14:19-28.

- Augustin, N.H., M.A. Muggleston, and S.T. Buckland. An autologistic model for the spatial distribution of wildlife. *Journal of applied ecology*. 33:339-347.
- Barney, R.L. and B.J. Anson. 1921. Seasonal abundance of the mosquito destroying Top-Minnow, *Gambusia Affinis*, Especially in relation to male frequency. *Ecological Society of America*. 2:53-69.
- Barton, K. 2015. MuMIn: Multi-Model Inference. R package version 1.13.4. <http://CRAN.R-project.org/package=MuMIn>.
- Bates, M. 1944. Observations on the distribution of diurnal mosquitoes in a tropical forest. *Ecology*. 25:159-170.
- Bates, M. 1945. Observations on climate and seasonal distribution of mosquitoes in eastern Colombia. *Journal of Animal Ecology*. 14: 17-25.
- Bates, M. 1949. *The natural history of mosquitoes*. Harper & Row. New York, NY (1965).
- Bates, D., M. Maechler, B. Bolker and S. Walker. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, <URL: <http://CRAN.R-project.org/package=lme4>>.
- Beketov, M.A., Y.A. Yurchenko, O.E. Belevich, and M. Liess. 2010. What environmental factors are important determinants of structure, species richness, and abundance of mosquito assemblages? *Journal of medical entomology*. 47:129-139.

- Bently, M.D. and J.F. Day. 1989. Chemical ecology and behavioral aspects of mosquito oviposition. *Annual review of entomology*. 34:401-421.
- Bidlingmayer, W.L. 1985. The measurement of adult mosquito population changes—some considerations. *Journal of American Mosquito Control Association*. 1:328-348.
- Bivand, R. 2014. Spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-77. <http://CRAN.R-project.org/package=spdep>.
- Bolker, B. and R development core team. 2014. Bbmle. Tools for general maximum likelihood estimation. R package version 1.0.17. <http://CRAN.R-project.org/package=bbmle>.
- Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs*. 27:325-349.
- Burkett-Cadena, N.D., M. D. Eubanks, and T.R. Unnasch. Preference of femal mosquitoes for natural and artificial resting sites. *Journal of American Mosquito control association*. 2008. 24: 228-235.
- Burkett-Cadena, N.D., C.J.W. McClure, L.K. Estep, and M.D. Eubanks. 2013. Hosts or habitats: What drives the spatial distribution of mosquitoes? *Ecosphere*. 4:1-16.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media.

- Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Disease, and Division of Vector-Borne Disease. Chikungunya virus in the United States. <http://www.cdc.gov/chikungunya/geo/united-states.html>. April 1, 2015.
- Chaves, L.F., G.L. Hamer, E.D. Walker, W.M. Brown, M.O. Ruiz, and D. Kitron. 2011. Climatic variability and landscape heterogeneity impact urban mosquito diversity and vector abundance and infection. *Ecosphere*. 2: 1-21.
- Clements, A.N. 1999. "The biology of mosquitoes volume 2: Sensory reception and behavior." CABI publishing. New York, NY.
- Courvoisier, D.S., C. Combescure, T. Agoritsas, A. Gayet-Ageron, and T.V. Perneger. 2011. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. *Journal of Clinical Epidemiology*. 64: 993-1000.
- Crawley, M.J. *Statistics: An introduction using R*. John Wiley & Sons Ltd. West Sussex, England. 2005.
- Darsie, R.F. Jr. and R.A. Ward. 2005. "Identification and geographical distribution of the mosquitoes of North America, north of Mexico". University Press of Florida. Gainesville, FL.
- Davis, J.H. 1967. *General map of natural vegetation of Florida*. Institute of Food and Agricultural Sciences University of Florida, Gainesville.
- DeCaceres, M., P. Legendre. 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology*, URL <http://sites.google.com/site/miqueldecaceres/>.

- DeCaceres, M., P.Legendre, and M. Moretti. 2010. Improving indicator species analysis by combining groups of sites. *Oikos*. 119:1674-1684.
- Diabate, A., T. Baldet, C. Brengues, P. Kengne, K.R. Dabire, F. Simard, F. Chandre, J.M. Hougard, J. Hemingway, J.B. Quedraogo, and D. Fontenille. 2003. Natural swarming behavior of the molecular M form of *Anophles gambiae*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 97: 713-716.
- Downes, J.A. 1969. The Swarming and mating flight of Diptera. *Annual review of Entomology*. 4:271-298.
- Elnaiem, D.A., K.Kelley, S. Wright, R. Laffey, G. Yoshimura, M. Reed, G. Goodman, T. Thiemann, L. Reimer, W. K. Reisen, and D. Brown. 2008. Impact of aerial spraying of pyrethrin insecticide on *Culex pipiens* and *Culex tarsalis* (Diptera: Culicidae) abundance and West Nile Virus infection rates in an urban/suburban area of Sacramento County, California. *Journal of Medical Entomology*. 45: 751-757.
- “Ecosystem of Floria.” 1990. Edited by R.L. Myers and J.J. Ewel. University Presses of Florida. Gainesville, Fl.
- Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*. 69:57-68.
- Fox, J. and S. Weisberg. 2011. An {R} companion to applied regression, second edition. Thousand Oaks Ca.: Sage. URL:
<http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.

- Frazer, G.W., R.A. Fournier, J.A. Trofymow, and R.J. Hall. 2001. A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. *Agricultural and Forest Meteorology*. 109:249-263.
- Hall, S.R., C.R. Becker, J.L. Simonis, M.A. Duffy, A.J. Tessier, and C.E. Caceres. 2009. Friendly competition: evidence for a Dilution Effect among competitors in a planktonic host-parasite system. *Ecology*. 90: 791-801.
- Hayes, E.B., N. Komar, R.S. Nasci, S.P. Montgomery, D.R. O'Leary, and G.L. Campbell. 2005. Epidemiology and transmission dynamics of West Nile Virus disease. *Emerging Infectious Diseases*. 11: 1167-1173.
- Hurvich, C.M. and C. Tsai. 1989. Regression and Time Series Model Selection in Small Samples. *Biometrika*. 76:297-307.
- James, C.W. 1961. Endemism in Florida. *Brittonia*. 13:225-244.
- James, F.C. and C.E. McCulloch. 1990. Multivariate analysis in ecology and systematics: Panacea or Pandora's box? *Annual review of ecology and systematics*. 21:129:166.
- Janousek, T.E. and J.K. Olson. 1994. Effects of natural marsh fire on larval populations of *Culex salinarius* in east Texas. *Journal of the American Mosquito Control Association*. 10:233-235.
- Johnson, M.F., A. Gomez, and M. Pinedo-Vasquez. 2008. Land use and mosquito diversity in the Peruvian Amazon. *Journal of Medical Entomology*. 45:1023-1030.

- Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*. 29:1-27.
- Levene, H. (1960). Robust test for equality of variances. In I. Olkin et al. (EDS.), *Contributions to probability and statistics: Essay in honor of Harold Hotelling* (pp. 278-292). Standfor, Ca: Standfor University Press.
- Main, K. and E. Menges. 1997. Station fire management plan. Archbold Biological Station. 97-1.
- “MacArthur Agro-Ecology Research Center.” A division of Archbold Biological Station.
Updated Febuary 15, 2013. Web March 20, 2013. <www.maerc.org>.
- McCune, B. and J.B. Grace. 2002. *Analysis of Ecological Communities*. USA.
- McKinney, M.L. 2006. Urbanization as a major cause of biotic homogenization. *Biological conservation*. 27:247-260.
- Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Vegetation*. 69:89-107.
- Muturi, J.E., J. Shililu, B. Jacob, W. Gu, J. Githure, and R. Novak. 2006. Mosquito species diversity and abundance in relation to land use in a Riceland agroecosystem im Mwea, Kenya. *Journal of Vector Ecology*. 31: 129-137.
- Odetoyinbo, J.A. 1969. Preliminary investigation on the use of a light-trap for sampling Malaria vectors in the Gambia. *Bulliten of the World Health Organization*. 40: 547-560.

- Oksanen, J., F.G. Blanchet, R. Kindt, R. Legendre, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.H. Stevens, and H. Wagner. 2013. *Vegan: Community ecology package*. R package version 2.0-10. <http://CRAN.R-project.org/package=vegan>.
- PAHO and WHO. March 27, 2015. Number of reported cases of Chikungunya fever in the Americas, by country or territory 2013-2015 (to week noted) *Epidemiological week/EW* 12. <http://www.paho.org/hq/index.php?Itemid=40931>. April 4, 2015.
- Pausas, J.G. and J.E. Keeley. 2009. A burning story: The role of fire in the history of life. *BioScience*. 59:593-601.
- Provost, M.W. 1952. The dispersal of *Aedes Taeniorhynchus* I. Preliminary studies. *Mosquito news*. 12:174-190.
- R core team. 2014. *R: A language and environment for statistical computing*. R foundation for statistical computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Sanderson, E.W., M. Jaiteh, M.A. Levy, K.H. Redford, A.V. Wannebo, and G. Woolmer. 2002. The human footprint and the last of the wild. *Bioscience*. 52: 891-904.
- "Savannas Preserve State Park." Florida State Parks. Web. March 20, 2013. <http://www.floridastateparks.org/history/parkhistory.cfm?parkid=159>.
- Shapiro, S.S. and M.B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika*. 52:591-611.
- Schmidt, K.A. and R.S. Ostfeld. 2001. Biodiversity and the dilution effect in disease ecology. *Ecology*. 82: 609-619.

- Silver, J.B. 2008. "Mosquito Ecology: Field sampling methods third edition". Springer. New York, NY (2008).
- Skaug H, D. Fournier, B. Bolker, A. Magnusson and A. Nielsen. 2015. Generalized Linear Mixed Models using AD Model Builder. R package version 0.8.1.
- Swaddle, J.P., and S.E. Calos. 2008. Increased avian diversity is associated with lower incidence of human West Nile infection: observation of the Dilution Effect. PLoS one. 3:e2488.
- Swain, H., E.H. Boughton, P.J. Bohlen, and L.O. Lollis. Trade-offs among ecosystem services and disservices on a Florida ranch. Rangelands. 35: 75-87.
- Wagtendonk, J.W. 2007. The history and evolution of wildland fire use. Fire ecology special issue. 3:3-17.
- Wallace, F.L., M.A. Tidwell, D.C. Williams and K.A. Jackson. 1990. Effects of controlled burning on *Aedes taeniorhynchus* eggs in an abandoned rice impoundment in South Carolina. The Journal of American Mosquito Control Association. 6:525-529.
- "Walton Scrub Preserve." St. Lucie County. 2013. Web. March 20, 2013.
<<http://www.stlucieco.gov/beaches/66.htm>>.
- Weekley, C.W. and E.S. Menges. 2003. Species and vegetation responses to prescribed fire in a long-unburned, endemic-rich Lake Wales Ridge scrub. Journal of the Torrey Botanical Society. 130: 265-282.

Wegbreit, J. and W.H. Reisen. 2000. Relationships among weather, mosquito abundance, and encephalitis virus activity in California: Kern County 1990-98. *Journal of American Mosquito Control Association*. 16: 22-27.

Whittle, R.K., K.J. Linthicum, P.C. Thande, J.N. Wagati, C.M. Kamau and C.R. Roberts. 1993. Effect of controlled burning on survival of floodwater *Aedes* eggs in Kenya. *The Journal of American Mosquito Control Association*. 9:72-77.

Yuval, B. 2006. Mating systems of blood-feeding flies. *Annual review of Entomology*. 51:413-440.

Yuval, B., J.W. Wekesa, and R.K. Washino. 1993. Effect of body size on swarming behavior and mating success of male *Anopheles freeborni* (Diptera: Culicidae). *Journal of insect behavior*. 6: 333-342.

Zuur, A.F., E.N. Ieno, N.J. Walker, A.A. Saveliev, G.M. Smith. 2009. *Mixed effects models and extensions in Ecology with R*. Springer. New York, New York USA.