THESIS

POTENTIAL ENVIRONMENTAL FACTORS ASSOCIATED WITH THE NEWLY EMERGING BAT WHITE-NOSE SYNDROME IN THE NORTHEASTERN UNITED STATES: AN EXPLORATORY MODELING APPROACH AND CASE-CONTROL STUDY

Submitted by

Abigail R. Flory

Department of Environmental and Radiological Health Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2010

COLORADO STATE UNIVERSITY

July 8, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY ABIGAIL R. FLORY ENTITLED POTENTIAL ENVIRONMENTAL FACTORS ASSOCIATED WITH THE NEWLY EMERGING BAT WHITE-NOSE SYNDROME IN THE NORTHEASTERN UNITED STATES: AN EXPLORATORY MODELING APPROACH AND CASE-CONTROL STUDY BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

Colleen G. Duncan

John S. Reif

Thomas J. Stohlgren

Advisor: Annette M. Bachand

Department Head: Jac A. Nickoloff

ABSTRACT OF THESIS

POTENTIAL ENVIRONMENTAL FACTORS ASSOCIATED WITH THE NEWLY EMERGING BAT WHITE-NOSE SYNDROME IN THE NORTHEASTERN UNITED STATES: AN EXPLORATORY MODELING APPROACH AND CASE-CONTROL STUDY

The emergence of mortality-causing bat White-Nose Syndrome (WNS) in the Northeastern United States during 2006 prompted an immediate need for research surrounding possible causation factors influencing its spread. Due to the mysterious nature of fungal pathogens, it has been very difficult to determine how the WNS-related *Geomyces destructans* fungus is causing bat mortality. Several different hypotheses have been formulated by bat and ecological experts in the field, but major influencing factors remain undetermined. To initiate WNS environmental research, this study utilizes a new machine-learning modeling technique, Maxent modeling, along with a case-control study to assess the hypothesis that certain environmental variables may be associated with the occurrence and distribution of bat WNS. Maxent data uses presence-only data and bases its algorithms on the principal of maximum entropy. Maxent results using 58 environmental predictor variables revealed Slope, Growing-Degree Days, Annual Temperature Range, and Land-Cover as the top four predicting variables for WNS-infected bat hibernacula locations. Similarly, the case-control study showed that two of these top four predictor variables (Growing-Degree Days and Annual Temperature Range) were statistically significantly associated with a hibernacula's WNS infection status. Cases had a slightly higher mean Average Temperature Range than controls (Cases=38.0, Controls=36.0) and lower mean Growing-Degree Days than controls (Cases=3419.1, Controls=3838.1). Both of these variables, along with their correlated terms, are largely temperature-dependent, suggesting a need for further research on the role of temperature in predicting the occurrence and distribution of *Geomyces destructans*. As a starting point for future research, this study has identified the most likely environmental variables related to the potential devastating ecological consequences of WNS-related bat mortality.

Abigail R. Flory
Department of Environmental and Radiological Health Sciences
Colorado State University
Fort Collins, CO 80523
Summer 2010

TABLE OF CONTENTS	PAGE

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

I. Background

The emergence of bat white-nose syndrome (WNS) in the northeastern United States has been recently associated with an unprecedented increase in bat mortality (**Figure 1**); (Blehert *et al.* 2009). Thirty-three surveyed hibernacula (a place for hibernating animals to shelter during winter) sites in Massachusetts, Connecticut, New York, and Vermont have confirmed cases of WNS with a 75% or greater decline in bat populations (Blehert *et al.* 2009). Two of these sites are shown in **Figure 2** (Blehert *et al.* 2009). The first documented case of WNS was recorded by the New York State Department of Environmental Conservation and occurred on February 16th, 2006 in Howes Cave, 52 km west of Albany, New York (**Figure 3**); (Blehert *et al.* 2009; USGS White 2009). Originating cases presented the now common visual characteristics of

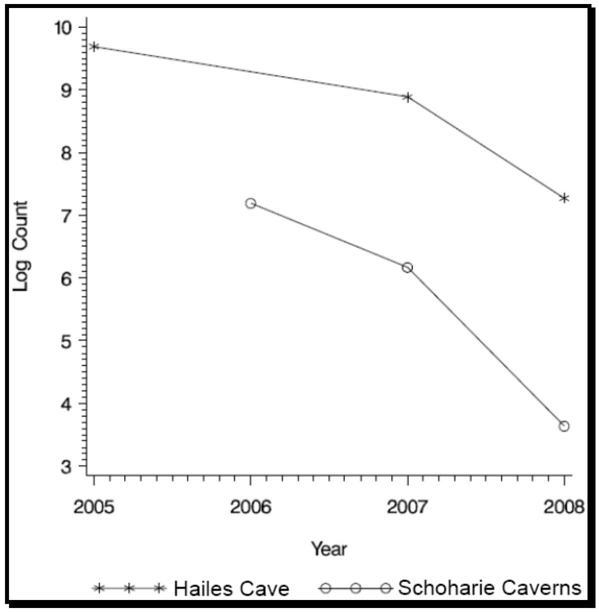
Figure 1: WNS Bat Mortality and Unusual Behavior

a.

b.

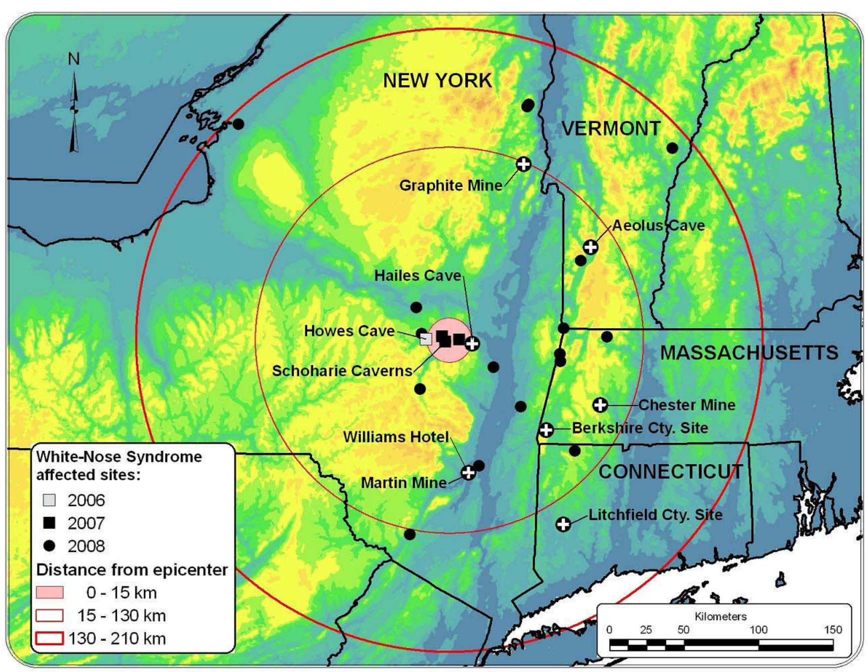
(a) "Remains of bats on the floor of a hibernaculum affected by white-nose syndrome (WNS). Wings of some bats have presumably been removed and left by scavengers."
(b) "A little brown bat (*M Lucifugus*) outside of a hibernaculum during daytime in winter – and unusual behavior, observed immediately before death in some bats affected by WNS. Note the snow in the background."
Source: Boyles and Willis 2009, Photos by: Alan Hicks

Figure 2: Bat Population Trends for Hailes Cave and Schoharie Caverns in New York, United States



Source: Blehert *et al.* 2009

Figure 3: Positive Hibernacula Locations for WNS, Years 2006-2008 (Confirmed by Survey)



Source: Blehert *et al.* 2009

WNS, which include white fungal growth (*Geomyces destructans*) on the muzzle, wing membranes, and ears of hibernating bats (Blehert *et al.* 2009). Through culture analyses, Blehert *et al.* (2009) determined that bats affected by WNS have a psychrophilic (thriving at relatively low temperatures) fungus colonized on their skin that is phylogenetically (historically, in terms of evolution) related to *Geomyces* ssp., but morphologically distinct from other members of the genus. Despite this discovery, the role of the WNS fungi has not yet been identified because little is known about the genetic and molecular basis of fungal pathogenicity (Hajek 1994; Blehert *et al.* 2009).

II. <u>Symptoms</u>

In a recent study, Blehert *et al.* (2009) showed that 105 out of 117 necropsied bats in the northeastern United States revealed fungal hyphae (threads making up the vegetative body of a fungus) that replaced the bat's sweat glands and hair follicles; invading regional tissue, eroding ear and wing epidermis, and reducing fat reserves that are required for hibernation survival. Courtin et al. (2009) conducted a study with 38 infected bats and concluded that WNS-infected bats have very little (if any) fat stores but show no evidence of major organ failure or consistent element toxicity. According to the National Wildlife Center, emaciation and poor body condition have been the most prominent characteristics among bats presenting WNS-related fungal hyphae on their bodies. These characteristics seem to be the ultimate cause of their mortality (Boyles and Willis 2009; USGS White 2009).

III. <u>Hypotheses</u>

According to Boyles and Willis (2009), researchers are unsure if the WNS-related *Geomyces* fungus is even a direct cause of the high bat mortality, or if it is simply a symptom of some larger issue affecting the species. Due to this uncertainty, various different causal hypotheses have been formulated by experts in the field (Boyles and Willis 2009).

One such hypothesis emphasizes the importance of fat reserves, also known as white adipose tissue (WAT), prior to hibernation (i.e., if sufficient fat reserves are not present by the time of hibernation, bats will have inadequate fat stores to survive the winter) (Kunz et al. 2008, Boyles and Willis 2009). Typically, bats will forage for insect biomass, accumulate WAT, and participate in courtship and mating activities in autumn before their hibernation onset (Kunz 1982, Kunz et al. 2008). A lack of adequate WAT in WNS-affected bats prior to hibernation could result from several of the following factors: environmental stressors (such as a decrease in the amount of food available to bats), altered composition of the insect community (e.g., unbalanced proportions of the most needed insects), or some unknown effect on the feeding patterns of bats (Kunz 1982, Kunz et al. 2008, Boyles and Willis 2009).

All current bat species affected by WNS follow a winter hibernation pattern of approximately 12-15 days in torpor (an inactive physiological state) followed by a short stint of arousal (usually a few hours); (Kunz 1982, Zimmerman 2009). During torpor, bat body temperatures do not stray more than 2 degrees Celsius from the hibernacula's ambient air temperature unless the ambient air temperature is at or below freezing (Kunz 1982, Kunz et al. 2008). This minimizes their energy expenditure (Kunz 1982, Kunz et al. 2008). During arousal (a time of mating and often relocation, but generally not fully understood), bats must triple their body temperature, requiring a significant increase in energy use (Kunz 1982, Kunz et al. 2008, Boyles and Willis 2009). Raising the number of arousals during one winter hibernation season may be a second factor influencing the depleted WAT found in WNS-infected bats (Boyles and Willis 2009, WNS Science Strategy Group 2008). This hypothesis suggests that bats may arrive at hibernacula with ample fat stores, but become affected by WNS post-arrival and end up using fat stores too quickly during their hibernation (Boyles and Willis 2009, WNS Science Strategy Group 2008). Suspicions that WNS causes an increased disturbance during hibernation have been suggested by several researchers and scientists within the field (Kunz et al. 2008, Blehert et al. 2009, Boyles and Willis 2009, USGS Investigating 2009, Zimmerman 2009). According to Kunz et al. (2008), increased torpor disruption is most likely being caused by one of the following: an awareness within the bat of fungal growth on their bodies, creating a need for the bats to clean and rid themselves of the *Geomyces* fungus (using additional waking energy), or an increased immune response to WNS, using more of the bat's energy stores to fight the invader. Both disruptions increase energy use and result in hunger, causing a mid-hibernation awakening and search for very small or non-existent food supplies (Boyles and Willis 2009, Kunz et al. 2008).

A third hypothesis, but not mutually exclusive of those previously mentioned, suggests that WNS may be a secondary infection that attacks animals already compromised by another pathogen (Zimmerman 2009). An examination of WNS-infected bats by Meteyer et al. (2009) showed that fungal hyphae filled hair follicles and sebaceous glands of examined bats, but did not typically cause inflammation or immune response in the tissue of the bat (Gargas et al. 2009). These results suggest that some additional un-tested variable or underlying factor may be unknowingly contributing to the large masses of WNS mortalities.

At this point, the exact role of WNS-related *Geomyces* fungi in bat mortality has yet to be determined. Current research has focused on narrowing down the wide variety of possible factors influencing their mortality, but research has only recently begun and many results have not been ascertained (Boyles and Willis 2009). One environmental factor under investigation is temperature because psychrophilic fungi grow best at 5-10°C, which is the average body temperature of hibernating bats in WNS-affected locations (Boyles and Willis 2009). Other possible factors include climate (specifically extremes of drought, heat, cold, and precipitation), water quality, agricultural intensification, pesticide use, deforestation, urbanization, land cover, elevation, and other various topographic variables (Jaberg and Guisan 2001; Lamb et al. 2008; Jones et al. 2009). With large winter roosting areas containing hundreds to thousands of bats, and annual summer trips to maternity colonies for breeding, these bats may be a perfect host for contagious infectious disease (Fenton 1969; Thomas et al. 1979). Although contagion has been speculated, it has not yet been definitively reported in any publications. However, the U.S. Geological Survey states that WNS is most likely spread by: 1) contact among bats, 2) contact with their environment, 3) human movement between caves, 4) other animal movement between caves, or 5) some combination of these (USGS Investigating 2009). Regardless of how WNS spreads, environmental predictors may have the potential to provide fundamental staging for the advancement of further research on WNS, which would be a necessary step for changing the current pattern of WNS spread.

IV. Strategic Approaches

Strategies for WNS management cannot be developed until a thorough understanding of the epidemiology, ecology, and etiology of WNS has been established (Blehert *et al.* 2009). According to Huq and Colwell (1996), environmental factors can directly or indirectly affect a pathogen's ability to survive, persist, and produce disease. The emergence of disease can therefore be directly influenced by environmental factors and may partially or entirely depend on physical, biological, and chemical states within the potential disease's surrounding environment (Huq and Colwell 1996). Recognition of environmental links potentially affecting disease emergence has catalyzed the introduction of paired environmental variable and emerging disease research (Daszak 2001, McMichael 2004, Wilcox and Gubler 2009). Under such interest, this study hypothesizes that certain environmental factors may be linked with the occurrence and distribution of bat White-Nose Syndrome. Detection of these factors can be aided by advances in GIS (geographic information system) technologies that have provided us with new approaches to predictive modeling and tools for integrating environmental monitoring data into the analysis of health outcomes (Nuckols 2004).

V. Current Status

As of March 2009, WNS was documented in the following nine states: New York, Vermont, Massachusetts, Connecticut, Pennsylvania, New Jersey, Virginia, West Virginia, and New Hampshire (USGS Investigating 2009). Six out of nine known bat species occupying the northeastern U.S. have been found with the WNS-related *Geomyces* fungus on their bodies. These include the little brown bat (*Myotis lucifugus*), the tricolored bat (*Pipistrellus subflavus*), the northern long-eared bat (*Myotis*

Figure 4: Six Known Bat Species Currently Affected by the WNS-Related *Geomyces* Fungus

The little brown bat (*Myotis lucifugus*)

Image Source:
http://mariewin.server304.com/marieblog/uploaded_images/BROWN%20BAT-768944.jpg

The tricolored bat (*Pipistrellus subflavus*)

Image Source:
http://www.biokids.umich.edu/files/29588/P_subflavus1_large.jpg

The northern long-eared bat (*Myotis septentrionalis*)

Photo by: Michael Patrikeev
Image Source:
http://wildnatureimages.org/sitebuildercontent/sitebuilderpictures/n_long-eared_bat_12926.jpg

The endangered Indiana bat (*Myotis sodalis*)

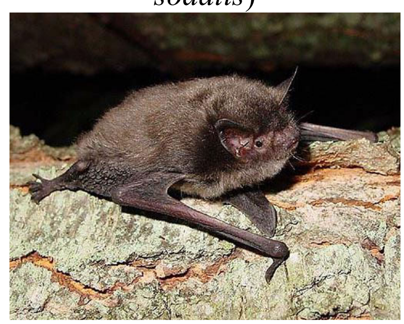


Photo by: Adam Mann
Image Source:
http://www.columbiamissourian.com/media/multimedia/2009/07/07/media/Bats03_t_w600_h600.jpg

The big brown bat (*Eptesicus fuscus*)

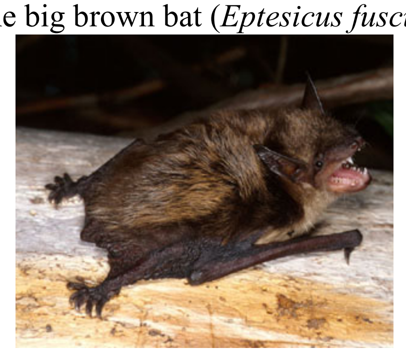


Photo by: Michael Patrikeev
Image Source:
http://wildnatureimages.org/sitebuildercontent/sitebuilderpictures/big_brown_bat_8543.jpg

The eastern small-footed bat (*Myotis Leibii*)

Photo by: Dr. J. Scott Altenbach
Image Source:
http://www.dnr.state.md.us/wildlife/bats/eastsmallfoot.jpg

septentrionalis), the endangered Indiana bat (*Myotis sodalis*), the big brown bat (*Eptesicus fuscus*), and the eastern small-footed bat (*Myotis Leibii*); (**Figure 4**); (Meteyer et al. 2009). The three unaffected species have most likely avoided WNS infection due to their winter migratory patterns of flying south and roosting in trees (Findley 1993). The change in temperature (warmer southern air), roosting area (tree versus enclosed hibernacula), and proximity to other bats are three factors that could be affecting the WNS variation between hibernating and migratory bats (Findley 1993). Of the six species found with WNS-related fungus on their bodies, the little brown bat is the most prevalent and the most affected by WNS (Meteyer et al. 2009). Recovery of this species, along with the five other species, will be difficult due to a bat's average lifespan of 6-7 years and average off-spring count of one bat pup per year (Cockrum 1956; Fenton 1970; Humphrey and Cope 1976).

Bat populations worldwide participate in various key ecologic processes that are essential for plant pollination, insect control, and seed dispersion (L. F. Skerratt *et al.* 2007, Blehert *et al.* 2009). Decreases in North American bat populations will present various percolating consequences to several vital ecologic systems within our environment (Blehert *et al.* 2009).

VI. Objectives

Objectives include the application of a GIS-related spatial modeling technique and a case-control study to identify environmental factors potentially associated with the occurrence and distribution of bat White-Nose Syndrome.

Specific Aims:

1. Download the publicly free Maxent Software.
2. Obtain geographic coordinates for WNS-infected hibernacula from the U.S. Fish and Wildlife Service (verified positive by state agencies via field surveys and/or laboratory testing).
3. Remove spatially correlated hibernacula sites.
4. Select a set of relevant bioclimatic and environmental variables from originating 58 available GIS layers by reducing correlated terms, removing variables that contribute 0% to the model, and testing the importance of phenologic variables' contribution.
5. Transform all bioclimatic and environmental variables into the same geographic projection, geographic coordinate system, resolution size, data format type, and coverage area for data consistency purposes.
6. Run Maxent with all non-correlated hibernacula sites and remaining (i.e., relevant) bioclimatic/environmental variables to determine the ranking order of each variable's contribution to WNS prediction.
7. Evaluate whether Maxent's top four contributing variables are actually predictive of WNS-infected caves specifically (not just bat hibernacula in general) by conducting a case-control study that examines the association between hibernacula infection status and each top predictor variable.
8. Retrieve geographic coordinate locations for non-infected hibernacula (controls) within the study area by contacting individual state agencies for data (sites declared negative via laboratory testing, field inspection, or failure to prove positive infection status).

9. Transform, when needed, all control data to match that of existing data layers for consistency purposes.
10. Evaluate statistical assumptions and perform variable-appropriate case-control statistical analyses for each predictor variable ranked as the top four Maxent results.
11. Perform a logistic regression evaluation using case-control data.

CHAPTER 2: METHODS

I. Maxent Software

The Maxent modeling software used in this study predicts probability distribution of a given occurrence based on presence-only data points and certain given environmental risk factors (Phillips et al. 2006; Maxent Software version 3.1, www.cs. princeton.edu/~schapire/maxent/). This modeling method is considered a machine-learning technique based on the probability distribution of maximum entropy (Phillips et al. 2006; Kumar et al. 2009; www.cs.princeton.edu/~schapire/maxent/). Entropy, or randomness, approaches a maximum value when an isolated system (i.e., one free of external influences) reaches equilibrium (Phillips et al. 2004, Phillips et al. 2006, Pearson 2007). An isolated system that is not in equilibrium will approach equilibrium over time, and will therefore approach maximum entropy over time (i.e., without external influences, a species will spread to fill all areas with suitable conditions and will approach a uniform distribution); (Pearson 2007, Phillips et al. 2004, Phillips et al. 2006). According to Sukumar, the maximum entropy principle "provides a means to obtain least-biased statistical inference when insufficient information is available" (2008). The concept of modeling maximum entropy distributions for presence-only data has been successfully applied to a number of fields, including statistics, ecology, computer system modeling, and general probabilistic problem solving (Shore 1980).

Maxent's ability to model ecologic niches[1] has been considered one of the best among several different modeling methods (Elith et al. 2006; Ortega-Huerta and Peterson 2008, Kumar and Stohlgren 2009), and has been proven effective despite small sample sizes (Hernandez et al. 2006; Pearson et al. 2007; Papes and Gaubert 2007; Wisz et al. 2008; Benito et al. 2009, Kumar and Stohlgren 2009). A recent comparison of 16 different modeling methods encompassing data on birds, terrestrial plants, bats, and reptiles resulted in the ranking of Maxent as the best-performing model algorithm (Elith et al. 2006, Kumar et al. 2009). Jaberg and Guisan (2001) also conducted a study that modeled the relationship between landscape structure and community composition in bats. Through their analyses, Jaberg and Guisan (2001) were able to determine species-specific relationships to resources and found that predictive GIS modeling techniques can provide researchers with an essential conservation tool, especially when a thorough species census is near impossible to obtain.

The Maxent software program was written by Steven Phillips, Miroslav Dudik, and Rob Schapire, with support from AT&T Labs-Research, Princeton University, and the Center for Biodiversity and Conservation, American Museum of Natural History (Phillips et al. 2006). The free, publicly available Maxent software, version 3.3.2, can be found at http://www.cs.princeton.edu/~schapire/maxent/ along with step-by-step tutorials and instructions for input data requirements. Maxent analysis uses species presence-only data and a variety of environmental variable layers to produce a predicted continuous probability distribution ranging 0 to 1. Maxent's output probability distribution uses nonparametric linear, quadratic, product, binary

---

[1] In ecology, niche is a term used to describe an organism's place in the ecosystem (home.mira.net/~gnb/caving/glossary/N.htm). A niche model is one that aims to accurately predict an organism's place in the ecosystem via input data information.

(categorical), threshold, and hinge functions on the environmental variables for optimal WNS prediction. These provide constraints for the output distribution, including mean, variance, covariance, proportion in each category, proportion above threshold, and mean above threshold, respectively (Phillips et al. 2006).

II. Target Species of Analysis

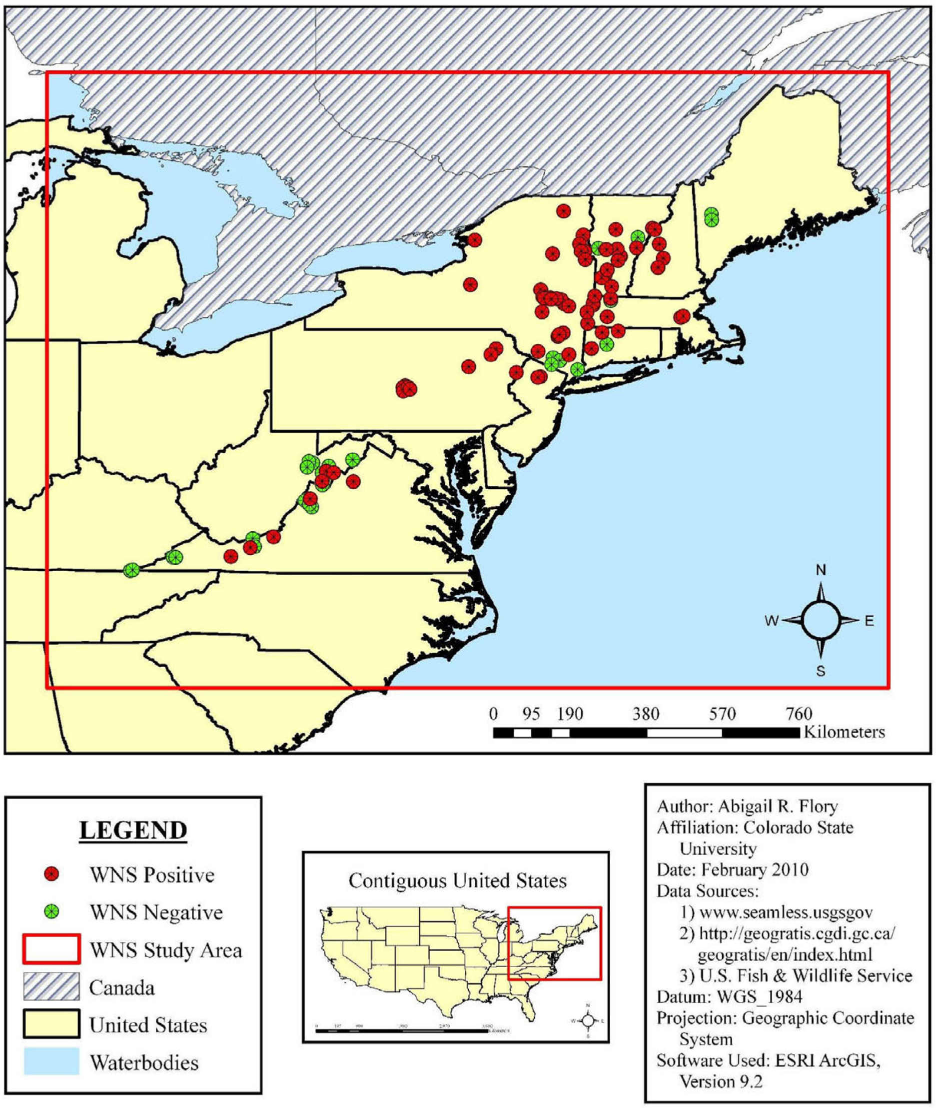
A. *Description*:

Target species included all bats in the northeastern U.S. affected by bat White-Nose Syndrome: The little brown bat (*Myotis lucifugus*), the tricolored bat (*Pipistrellus subflavus*), the northern long-eared bat (*Myotis septentrionalis*), the endangered Indiana bat (*Myotis sodalis*), the big brown bat (*Eptesicus fuscus*), and the eastern small-footed bat (*Myotis Leibii*); (Meteyer et al. 2009).

B. *Data Preparation and Parameters*:

For modeling purposes, WNS point data were formatted to coincide with Maxent's input requirements. Briefly, three columns (species name, longitude, and latitude, both in decimal degrees) were generated in Microsoft Office Excel 2007 (Microsoft Corporation, Redmond, WA). WNS-infected bat hibernacula locations were provided by Jeremy Coleman of the United States Fish and Wildlife Service and entered into Microsoft Office Excel 2007 under the appropriate column headings (USFWS 2010). Spreadsheet information was then saved in two separate file formats: .xls and .csv. Comma delimited, .csv, is the format required by Maxent software parameters for data input. For data privacy purposes, exact WNS-infected bat hibernacula locations can not be listed in this document. Infected bat hibernacula are generally confined to the northeastern United States (**Figure 5)**.

Figure 5:
WNS-Positive and WNS-Negative Data Points
for the Study Area of Interest (AOI)



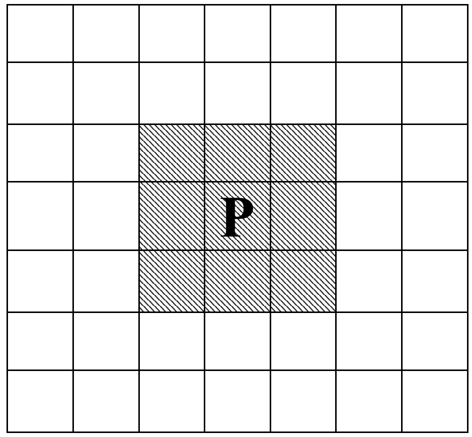
Original 84 WNS-positive points provided by the U.S. Fish and Wildlife Service were reduced to 74 points due to duplication and clustering (four data points were duplicates and six data points were highly spatially autocorrelated). Duplicates were found by sorting the numerical decimal degree values of all WNS-infected bat hibernacula data and visually checking for identical latitude/longitude points. Visual duplicate checking was possible because of the small sample size. Spatial autocorrelation was accounted for by manually selecting and removing all data points in neighboring cells of a given WNS point. Each grid cell measured 1km x 1km, making a total neighboring grid cell area of 9,000-meters^2. Any points within this area were removed (see **Figure 6)**.

Case data for 2009 were confirmed positive by either, 1) laboratory testing, 2) field surveying (masses of dead bats presenting white fungal growth found within, beside, or at the entrance of hibernacula sites), 3) continuing infection from a previous year, or 4) some combination of these.

III. <u>Environmental Variables</u>

Fifty-eight environmental variables were placed in the Maxent software program for analysis (**Appendix 1**). These included 39 spatially explicit features (phenology, topography, hydrology, climate, and land-cover) and 19 bioclimatic variables, all of which spanned across the northeastern United States study region. Although described separately below, all datasets were entered into the model at the same time. A description of these layers along with their data preparation procedures follow.

Figure 6: Process of Removing Spatially Correlated Points



P = WNS-Positive Point

= Neighboring Cells

Each cell measures 1km x 1km (total neighboring cell area = $9km^2$, or $9{,}000m^2$). Any points within the neighboring cells were removed.

A. <u>*Phenology Variables*</u>

i. Description

In a recent study, Morisette et al. determined that phenological responses are becoming increasingly relevant for applied environmental issues (2009). Phenology (the study of recurring life-cycle events) examines the timing of biological events (such as animal migration or plant maturation) (Morisette et al. 2009). Previous measurements for estimating or describing such biological events have been unsatisfactory, but current advances in MODIS' phenologic satellite data collection have allowed researchers to more accurately measure and predict phenologic events (Tan et al. 2008, Morisette et al. 2009).

Due to the recent availability of this improved data, both MODIS satellite products (each measuring different variation in vegetation indices capturing biomass) were utilized for this study (Kumar et al. 2009). These temporally smoothed and spatially gap-filled vegetation metrics included NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index), with 15 metrics each (30 total); (Tan et al. 2008).

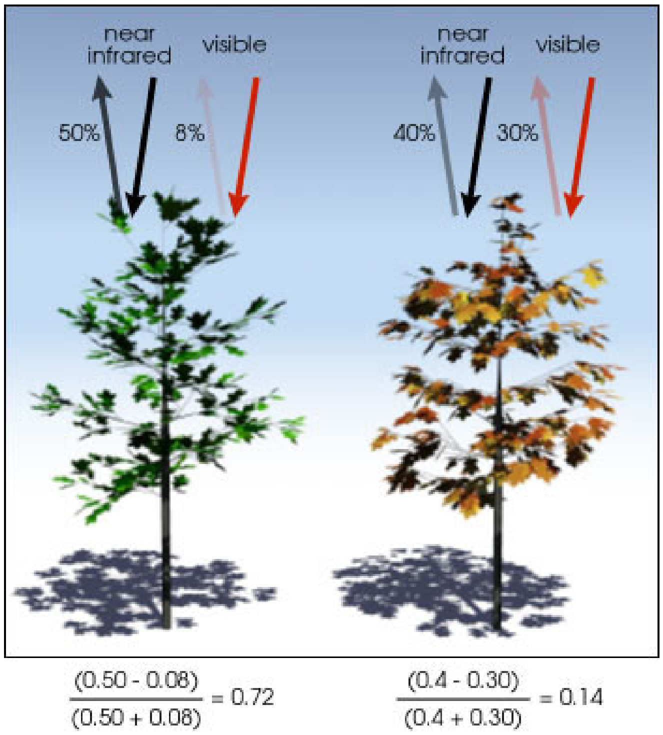
NDVI metrics are a result of near-infrared radiation minus visible radiation divided by near-infrared radiation plus visible radiation (Weier and Herring 2010). **Figure 7a.** contains a visual depiction of the NDVI pixel value collection process, and **Figure 7b.** displays the formula used for this process (Weier and Herring 2010).

EVI metrics are calculated similarly to NDVI, but correct for some distortions in the reflected light caused by particles in the air and ground cover below vegetation (Weier and Herring). Due to this, EVI metrics are considered an improvement over the quality of NDVI metrics because they can capture a better measure of vegetation variation in densely vegetated areas.

Despite the accuracy benefits of EVI metrics over NDVI metrics, both phenologic measures are often used in conjunction to optimize study results (Weier and Herring 2010, Li and Weng 2005, Nagler et al. 2007, Tan et al. 2008). For example, Tan et al. showed in their study that NDVI metrics were more sensitive to small vegetation variations and performed better in sparse vegetation regions than EVI metrics (2008). Due to the separate phenologic detection advantages of each metric type, both NDVI and EVI vegetation indices were used in this study as Maxent inputs.

Figure 7 (a. & b.): MODIS' NDVI Data Collection Process

a. NDVI Data Collection Process



"NDVI is calculated from the visible and near-infrared light reflected by vegetation. Healthy vegetation (left) absorbs most of the visible light that hits it, and reflects a large portion of the near-infrared light. Unhealthy or sparse vegetation (right) reflects more visible light and less near-infrared light. The numbers on the figure above are representative of actual values, but real vegetation is much more varied" (Weier and Herring, n.d.). Illustration by Robert Simmon.

b. NDVI Formula

NDVI = (NIR – VIS) / (NIR + VIS)

where:
NIR = Near-Infrared light
VIS = Visible light

Source: Weier and Herring 2010

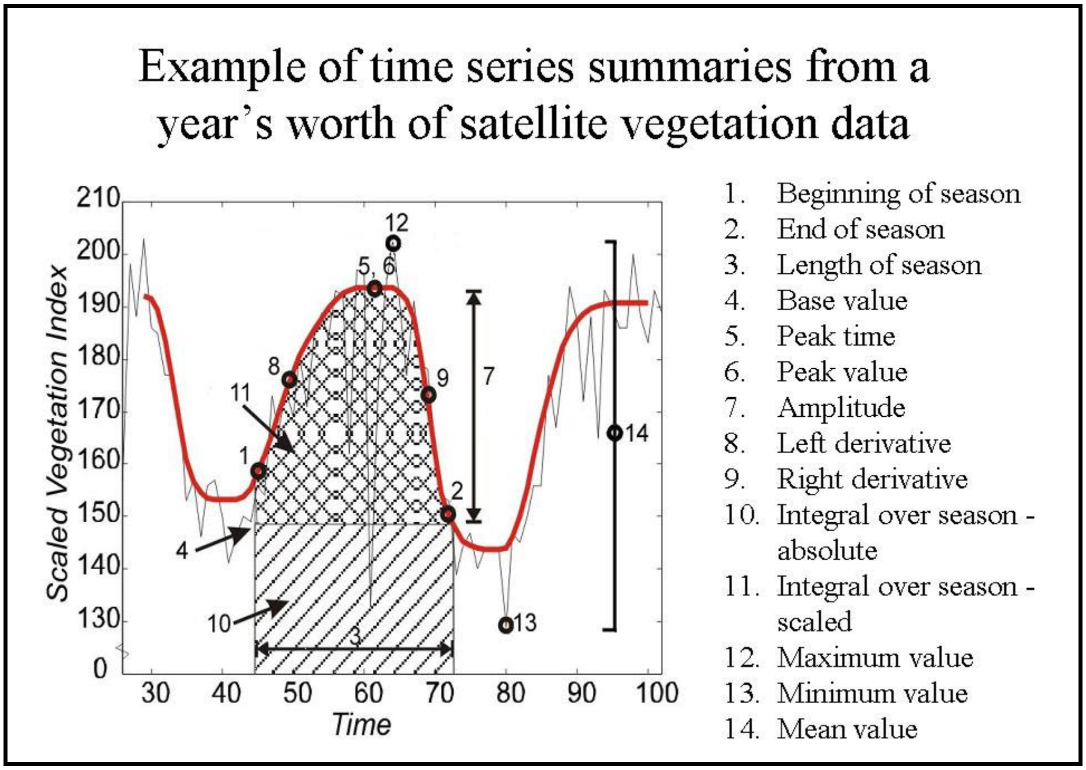
ii. Data Preparation

MODIS' NDVI and EVI data layers were downloaded (with help from Jeff Morisette, a U.S. Geological Survey Invasive Species Science specialist) from the National Aeronautics and Space Administration's (NASA's) FTP (file transfer protocol) site for the years 2006-2007 (FTP website was provided via email after data was ordered from the following NASA website: http://accweb.nascom.nasa. gov/; data only available for five days). These were the latest years available to download. NDVI and EVI metrics were formed by NASA using the TIMESAT (program for analyzing time-series of satellite sensor data) phenology algorithm that produces a spatial coverage more complete than any other remotely sensed data-based phenology product (Jonsson and Eklundh 2004, Tan et al. 2008). The attributes of TIMESAT data can be found in **Figure 8.**

Each NDVI and EVI layer was then gap-filled using Environmental Systems Research Institute's (ESRI's[2]) 'Nibble' function (ESRI, Redlands, California, USA). The 'Nibble' function generalizes scores for missing values within the phenology layers based on a nearest neighbor analysis (ESRI 2008). Metrics were then averaged over their two year period using ERSI's 'Raster Calculator' function, such that each of the NDVI and EVI continuous metrics represented a two-year averaged value. All NDVI and EVI metrics were clipped to the study area using ESRI's 'Analysis Mask' function (ESRI, Redlands, California, USA). The finest imagery produced by MODIS has an approximate resolution of 250-meters (Weier and Herring 2010). NASA MODIS data downloaded for this study was received in this resolution. Since MODIS data is not available at a resolution finer than ~250m, all other Maxent input data layers were resampled to ~250m (most data layers were of a finer resolution, mainly 30-meter or 90-meter resolution) using ESRI's 'Cell Size' function (ESRI, Redlands, California, USA). The cell size used for ~250-meter resolution resampling was 0.0020833333 decimal degrees, matching that of

[2] Unless otherwise stated, the term 'ESRI' refers to ESRI's ArcGIS software version 9.2.

Figure 8: Attributes of TIMESAT Data



TIMESAT derives 11 phenology parameters, or metrics, for up to two separate seasons each year. The phenology parameters are 1) time for the start of the season, which is when the left edge of the fitted function has increased to a user-defined level, often 20% of the seasonal amplitude, 2) time for the end of the season, or when the right edge has decreased to a user-defined level, 3) length of the season, 4) base level, or the average of left and right minimum values, 5) time for the middle of the season, 6) peak value of the fitted function, 7) seasonal amplitude, or the difference between the peak value and base level, 8) left slope, or rate of increase at the beginning of the season, 9) right slope, or the rate of decrease at the end of the season, 10) large seasonal integral, from season start to season end, and 11) small seasonal integral, relative to the base level. Numbers 12-14 are maximum, minimum, and mean values for all of the data collected.

Source: Jonsson and Eklundh 2004

the MODIS data. All MODIS data layers were converted to ASCII format (a Maxent input requirement) using ESRI's 'Raster to ASCII' function (ESRI, Redlands, California, USA).

B. *Topography Variables*

i. Description

National elevation data (the National Elevation Dataset, NED) with 30-meter spatial resolution was obtained from The National Map Seamless Server and used for Maxent input (USGS 2010; http://seamless.usgs.gov/website/seamless/viewer.htm). Two additional rasters, aspect and slope (both in degrees), were generated from the NED. Elevation and slope, along with various other topographic variables, have been listed in recent publications as significant contributors to potential environmental factors affecting bat mortality (Jaberg and Guisan 2001; Lamb et al. 2008; Jones et al. 2009). See **Figures 9-11** for maps of elevation, aspect, and slope.

The Compound Topographic Index (CTI), a continuous topographic variable measuring wetness, was also added to the Maxent analysis. CTI is a function of both upstream contribution area and slope of the landscape and is calculated with the following equation: CTI = ln [Flow Accumulation / tan(slope)] (Moore et al. 1991). The CTI dataset was obtained from the U.S. Geological Survey's Earth Resources Observation and Science data center (USGS HYDRO 2009, http://eros.usgs.gov/).

ii. Data Preparation

NED data preparation began with downloading all elevation blocks covering the study area of interest from The National Map Seamless Server at http://seamless.usgs.gov/website/seamless/viewer.htm. Elevation blocks were then mosaicked into one raster GRID using ESRI's 'Mosaic to New Raster' function

Figure 9:
National Elevation Dataset (NED) Within the WNS Study Area

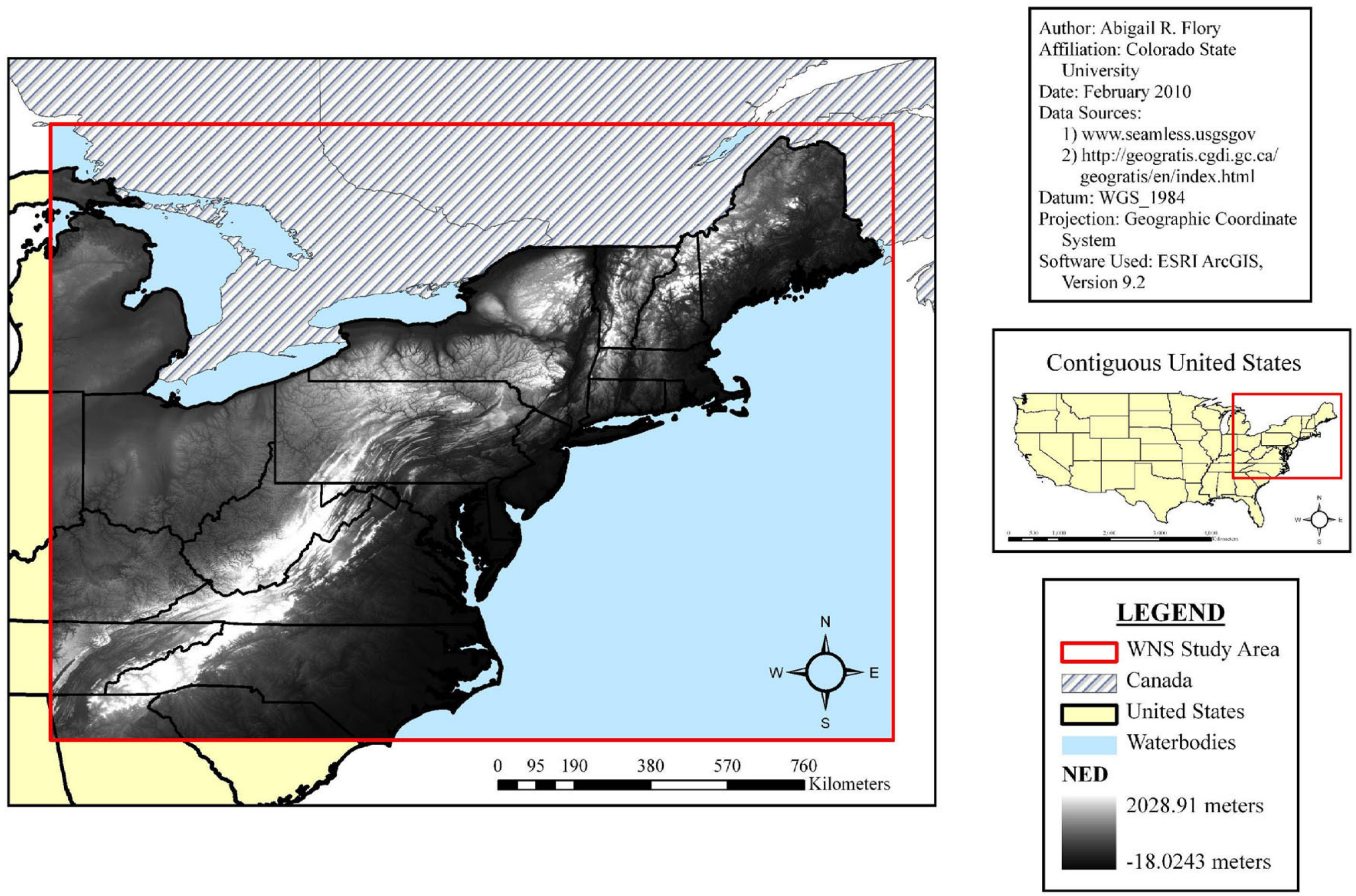


Figure 10:
Aspect Data Layer Within the WNS Study Area

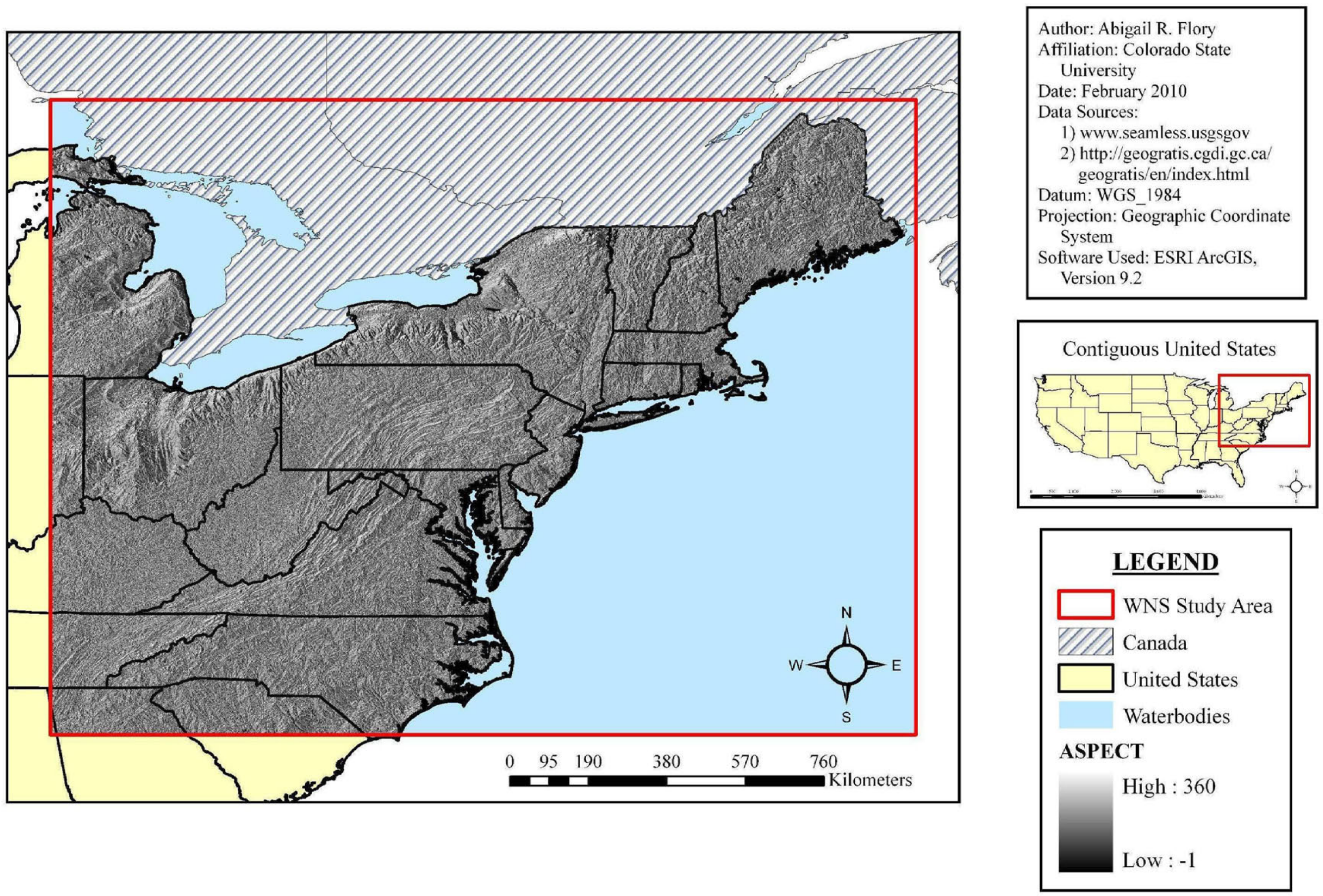
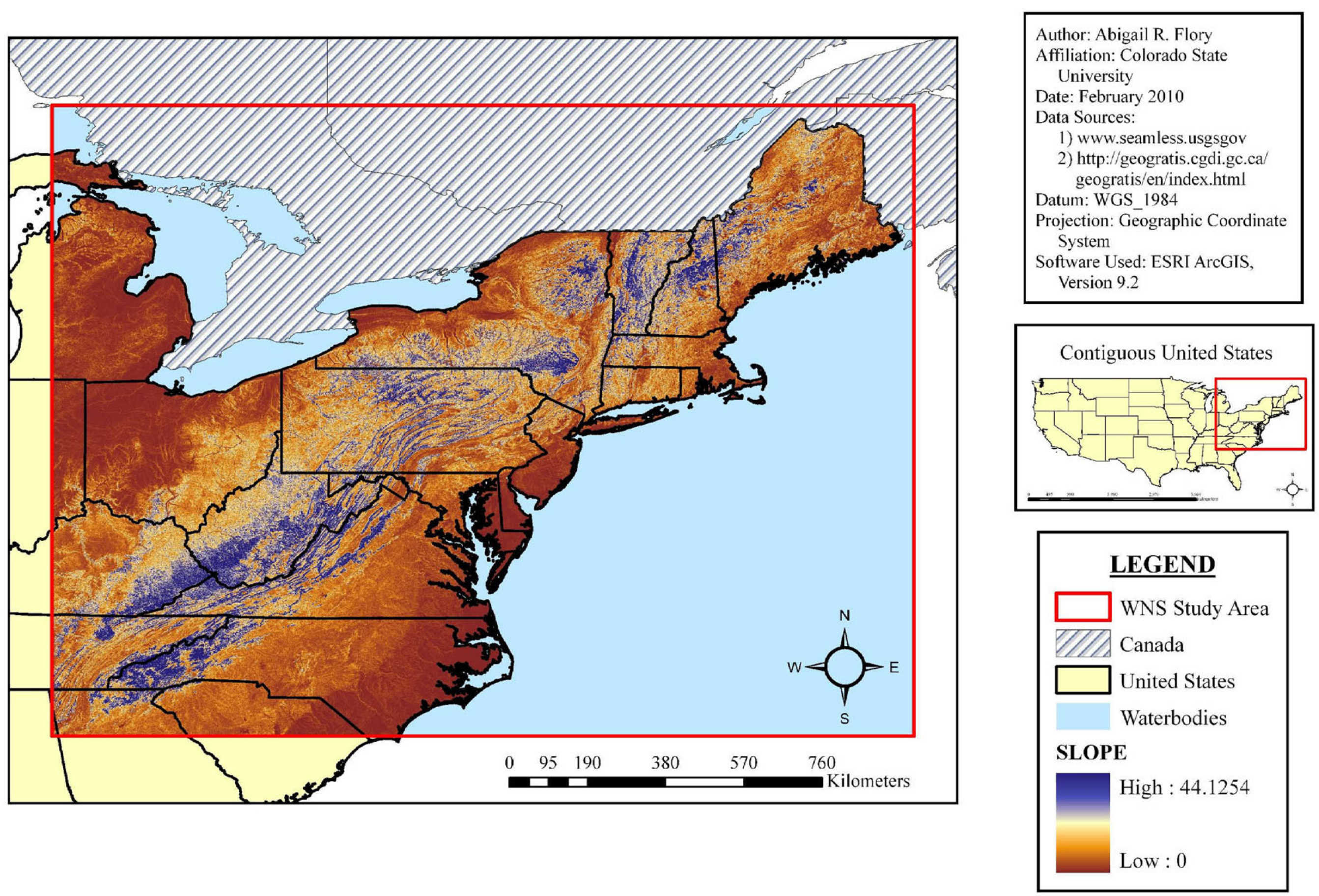


Figure 11:
Slope Data Layer Within the WNS Study Area



(ESRI, Redlands, California, USA). Once mosaicked, the raster GRID was reprojected (for projection consistency) from geographic coordinate system "North American Datum_1983" to geographic coordinate system "World Geodetic System_1984" using ESRI's 'Project' function (ESRI, Redlands, California, USA). For cell size consistency, the raster's cell sizes were resampled (using the previously described technique) to match that of MODIS' metric resolution (ESRI, Redlands, California, USA). This procedure converted the original decimal-degree cell size of 0.00027777778 (30-meter) to a decimal-degree cell size of 0.0020833333 (~250-meter). Slope and aspect raster GRIDs were created from the resulting elevation raster GRID using ESRI's 'Slope' and 'Aspect' functions, respectively (ESRI, Redlands, California, USA). All three raster GRIDs were then clipped to the study area and converted to ASCII format (per Maxent requirement) using the previously described technique.

Geographic bounds, geographic projection, and cell size for the Complex Topographic Index layer were also converted (as needed) using the previously described techniques in order to match those of all other input data layers. The Complex Topographic Index layer was then saved in ASCII format for Maxent analysis.

C. <u>*Hydrology Variables*</u>

i. Description

United States hydrologic and boundary data were obtained from The National Map Seamless Server at http://seamless.usgs.gov/website/seamless/viewer.htm (USGS 2010). Hydrographic data for Canada were retrieved from

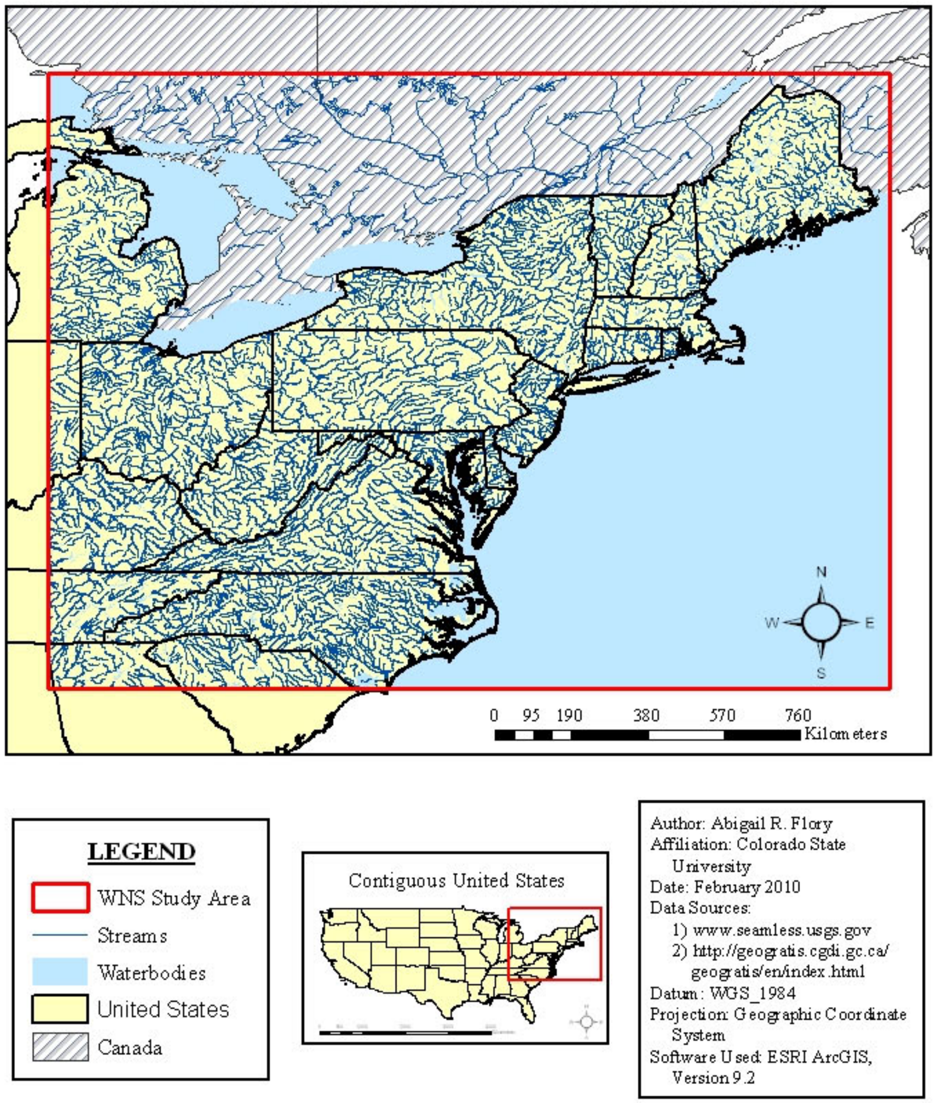
GeoGratis at http://geogratis.cgdi.gc.ca/geogratis/en/download/northamerica.html, a website produced and maintained by the Canadian Government Division Natural Resources, Earth Sciences Sector (GeoGratis 2009). Canadian boundary data were obtained from http://finder.geocommons.com/overlays/2264, a website also containing data produced and maintained by the Government of Canada, Natural Resources Canada, Centre for Remote Sensing (USGS North American 2008). See **Figure 12** for a map of the vector hydrologic network dataset.

ii. Data Preparation

Hydrology data originally came divided by country (U.S. and Canada) and separated into two layers per country: streams and waterbodies. All layers were downloaded in sub-units of blocks (the only format available), and then merged into one appropriate data layer (either streams or waterbodies) using ESRI's 'Merge' function (ESRI, Redlands, California, USA). Both layers were then reprojected from GCS_NAD_1983 to GCS_WGS_1984 using the same projection technique as described for NED data. Once in the correct projection, both vector layers were converted to raster GRID format using ESRI's 'Feature to Raster' function, and then combined into one raster GRID using ESRI's 'Raster Calculator' function (ESRI, Redlands, California, USA). The resulting raster's cell size was resampled (as previously described) to match that of MODIS' metric resolution and then clipped to the study area. This raster was further used to derive a 'distance to water' continuous data layer (**Figure 13**) using ESRI's 'Distance to...(straight line)' Euclidean distance[3] function (ESRI, Redlands, California, USA). The continuous

[3] Euclidean distance uses a triangle's hypotenuse calculation to measure the distance from the center of a source cell to the center of each of its surrounding cells (ESRI How n.d.). This produces a continuous

Figure 12:
North American Hydrologic Networks
(Vector Datasets) Within the WNS Study Area



raster of shortest distance to a given source (ESRI How n.d.). View **Figure 14** to see an illustration on how Euclidean distance works (ESRI How n.d.).

Figure 13:
'Distance to Water' Layer within the WNS Study Area

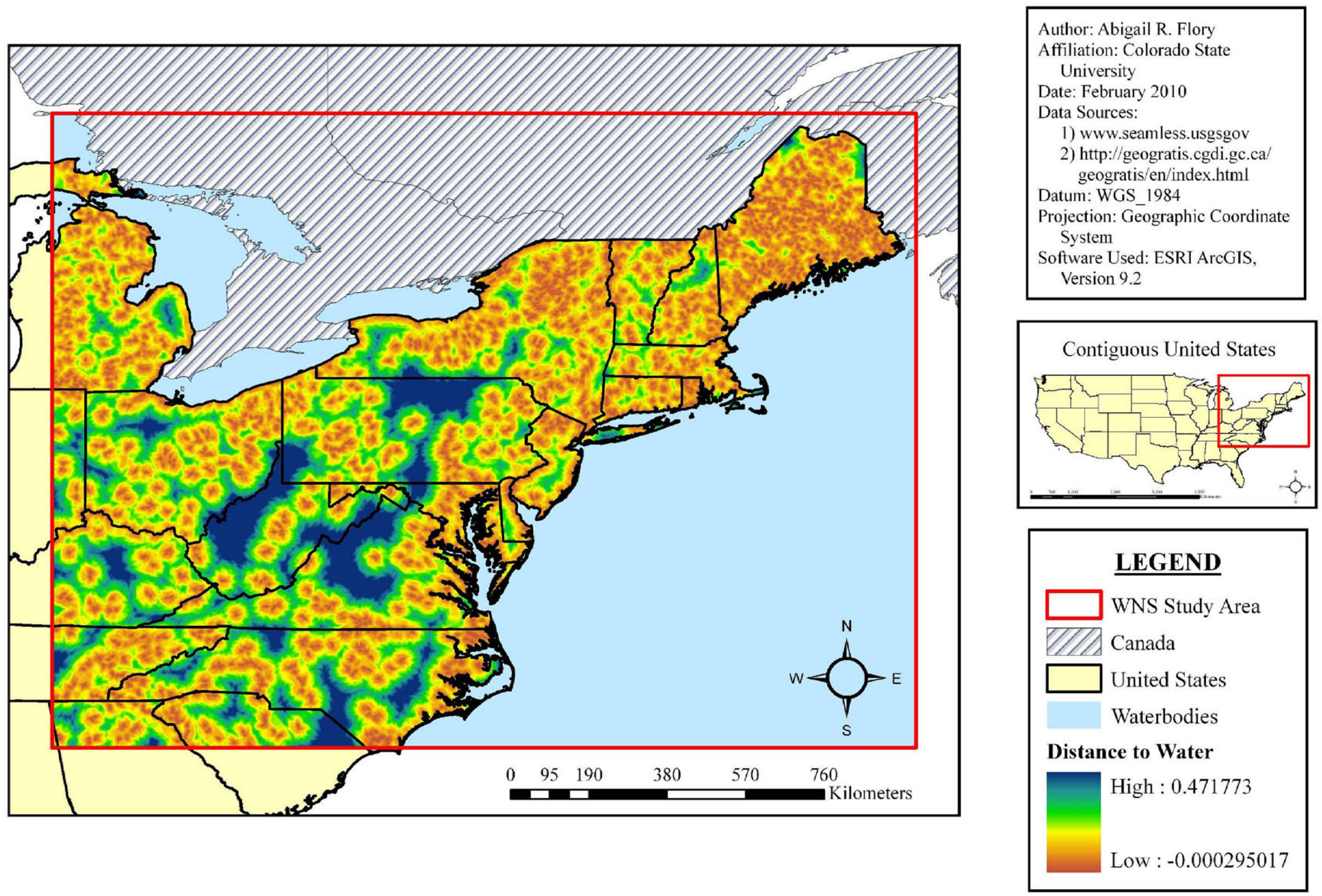
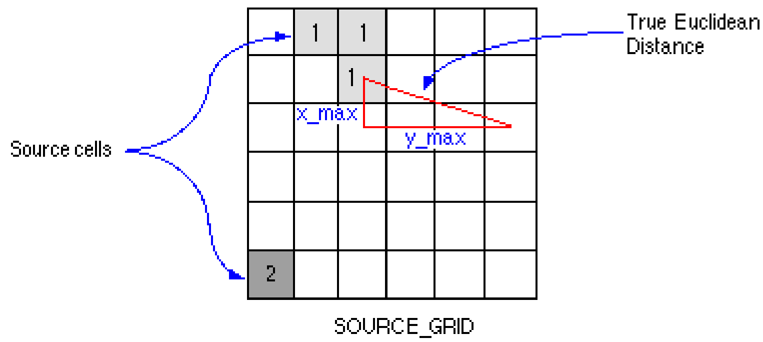


Figure 14: Euclidean Distance Illustration



Source: (ESRI How n.d.)

raster output was converted to ASCII and used as a measure of proximity to water within the Maxent model.

D. *Climatic Variables*

i. Description

Growing-Degree Days, Frostdays, Frequency of Precipitation, Humidity, and Annual Precipitation Event Size were all climatic variables added to the Maxent analysis. Data layers were obtained from the Daymet climate dataset (1-km spatial resolution; 1980-1997; Kumar et al. 2009; www.daymet.org/).

ii. Data Preparation

The geographic bounds, geographic projection, cell size, and data format (i.e., ASCII) for all climatic variables were uniformly defined (as previously described) to match those of all other Maxent input data layers.

E. *Land-Cover*

i. Description

The National Land-Cover Database (NLCD) was used to incorporate the effects of different habitat types on WNS occurrence. According to Jaberg and Guisan (2001), landscape structure and land-cover characteristics are both related to the distribution of bat hibernacula, and therefore possibly related to other bat-associated phenomena (Fry et al. 2009, Sattler et al. 2007). NLCD data was obtained from the Multi-Resolution Land Characteristics Consortium (MRLC), a project of the U.S. Geological Survey (USGS Multi 2010, http://www.mrlc.gov/).

ii. Data Preparation

Geographic bounds, geographic projection, cell size, and format for the NLCD dataset were all uniformly defined to match those of the phenologic, topographic, climatic, and hydrographic datasets using previously described methods.

F. *Bioclimatic Variables*

i. Description

Nineteen bioclimatic variables from the WorldClim dataset (Nix 1986; Hijmans et al. 2006; www.worldclim.org/bioclim) representing a better measure for the ecological and physiological tolerance of a species (i.e., more biologically meaningful); (Graham and Hijmans 2006; Kumar et al. 2009; Kumar and Stohlgren 2009, www.worldclim.org/bioclim) were added to the analysis. These variables were derived using ARC AML script (MkBCvars.AML; www.worldclim.org/mkBCvars.aml; Hijmans 2006; Kumar et al. 2009) and the Daymet climate dataset (www.daymet.org/; 1-km spatial resolution; 1980-1997; Kumar et al. 2009). According to Worldclim, bioclimatic variables are very useful when applied to ecological niche modeling (n.d.). They encompass annual trends, seasonality, and extreme or limiting environmental factors (WorldClim n.d.); (**Appendix 1**).

ii. Data Preparation

Geographic bounds, geographic projection, and cell size for each of the bioclimatic layers were all uniformly defined to match those of the phenologic, topographic, climatic, hydrographic, and land-cover datasets (as previously described) and then converted to ASCII format.

IV. Maintent Modeling Procedures

A. *Assessing Multicollinearity*

Prior to running Maxent, multicollinearity among all variables was tested by examining cross-correlations. Correlation coefficient values were calculated with individual hibernacula site data. Each hibernacula site had a total of 58 values, one for each input environmental predictor. Data values were extracted using Hawth's 'Intersect Point' analysis tool (Beyer 2004). Hawth's Tools is an extension for ESRI's ArcMap designed to perform spatial analysis functions that cannot be conveniently accomplished with out-of-the-box ArcGIS (Beyer 2004). These tools are publicly free for download at http://www.spatialecology.com/htools/download.php (Beyer 2004).

i. Correlation Terms

Pearson's correlation coefficient, and in cases of suspected lack of normality, the non-parametric Spearman correlation coefficient, were used to assess correlations within the environmental predictor data (Plonsky Nonparametric 2009, StatSoft n.d.). Statistical tests for correlation were calculated with the statistical software package Stata/SE 10.1 for Windows (Stata/SE 10.1 2007).

ii. Correlated Variable Reduction

Based on potential biological relevance to the distribution of bat WNS and ease of interpretation, only one variable from a set of highly cross-correlated variables was included in the final Maxent model (Kumar et al. 2009; Kumar and Stohlgren 2009). In order to select the most appropriate/WNS-relevant variable for each group of highly correlated variables, a Maxent training model (termed Preliminary Maxent Model, or PMM) was run with all data variables. The most important output from this training model was a ranking of each variable's percent contribution to the overall model (table not shown). The highest ranking variable out of a group of highly correlated variables was used to represent those variables in the final Maxent model (**Table 1**). For example, BIO-12 (Mean Annual Precipitation), BIO-14 (Precipitation of Driest Month), BIO-16 (Precipitation of Wettest Quarter), BIO-17 (Precipitation of Driest Quarter), BIO-18 (Precipitation of Warmest Quarter), and BIO-19 (Precipitation of Coldest Quarter) were all highly correlated. Since Precipitation of Driest Month (BIO-14) had the largest percent contribution (7.5%) within this group of correlated variables, it was kept for the final model. The accompanying correlated variables were dropped from the model, but not excluded from later results and discussion.

Table 1: Highly Correlated Environmental Variables (Pearson's correlation coefficient, r≥±0.90) and Corresponding Maxent Contributions from Training Model [%]

Group	Environmental Variables within Each Correlated Group	Variable Kept in Model
1	Annual Mean Temperature [0.0] Maximum Temp. of Warmest Month [0.0] Minimum Temp. of Coldest Month [3.9] Mean Temp. of Warmest Quarter [0.2] Mean Temp. of Coldest Quarter [0.1] FrostDays [0.1] GrowDays [19.5] Humidity [0.0]	GrowDays [19.5]
2	Temperature Seasonality [0.5] Annual Temperature Range [6.5]	BIO-7 [6.5]
3	Mean Annual Precipitation [0.3] Precipitation of Driest Month [7.5] Precipitation of Wettest Quarter [0.0] Precipitation of Driest Quarter [0.0] Precipitation of Warmest Quarter [0.0] Precipitation of Coldest Quarter [0.2]	BIO-14 [7.5]
4	Precipitation of Wettest Month [2.0] Precipitation of Wettest Quarter [0.0]	BIO-13 [2.0]
5	EVI-2 [0.1] EVI-3 [0.0]	EVI-2 [0.1]
6	EVI-4 [0.1] EVI_Min [0.1]	EVI-4 [0.1]
7	EVI-6 [0.2] EVI_Max [0.0]	EVI-6 [0.2]
8	NDVI-2 [0.2] NDVI-3 [0.0]	NDVI-2 [0.2]
9	NDVI-4 [0.0] NDVI_Min [2.5]	NDVI_Min [2.5]
10	NDVI-6 [0.1] NDVI_Max [0.0]	NDVI-6 [0.1]

B. *Other Variable Reductions*

In addition to correlated variable reduction, fourteen predictor layers were excluded due to a 0.0% contribution to the training model (table not shown). This left 26 environmental variables for the first Maxent model (Maxent Model #1, or MM1).

C. *Running Maxent*

i. Training and Testing Percentages

Before running the first Maxent model, a percentage value (totaling 100%) had to be assigned to training and testing datasets within the sample data (termed a *split-sample* approach) (Guisan et al. 1999, Guisan & Hofer 2003). According to Veloz, the best method for obtaining a valid, accurate niche model is to train a model with one dataset and then test model predictions against an independent dataset (2009). A truly independent dataset is often unavailable, so random subsets of the available data are used for training and testing purposes (Veloz 2009). Typically, presence-only Maxent models will use 60-80% of the input sample data for training purposes, and 40-20% for testing purposes (Phillips 2008). It is best to use the highest training percentage possible within the 60-80% range so that the model has a best estimate of niche prediction. This can be difficult with a small sample size because there may not be enough data leftover to adequately test the trained model. To account for this, 75% of the hibernacula sites (56 points) were used to train Maxent on how to predict a WNS niche model, and 25% (18 points) were used to test this trained model.

ii. First Model Run

Upon running the first model, a 0.0% contribution from BIO-15 (precipitation seasonality) was immediately noted in the results table. This variable was removed due to its lack of contribution, leaving 25 environmental variables, and then the model was run again (Maxent Model #2, or MM2).

iii. Second Model Run

To improve upon Maxent's predictive ability, MM2 was repeated 25 times[4] and averaged over the repetitions to produce overall results. Each of the 25 repeated models chose a different, randomly selected 75% and 25% dataset for training and testing purposes.

iv. Examination of Phenology Metrics

As an additional hypothesis of interest, the following question was addressed: Does including the 250-meter resolution MODIS phenology layers improve Maxent's prediction of bat WNS? To test this hypothesis Maxent was run first with MODIS phenology layers only (Comparison Model #1, or CM1), second with the reduced MODIS, climatic, topographic, bioclimatic, hydrographic, and land-type datasets (Comparison Model #2, or CM2), and third with all reduced variables except for the MODIS phenology layers (Comparison Model #3, or CM3).

v. Final Maxent Model

Maxent Model #3 (MM3), the final Maxent model, was run with a total of 13 environmental predictor variables and 74 hibernacula locations. MM3 also used

[4] Twenty-five repetitions is a typical quantity chosen by researchers using Maxent modeling (Kumar et al. 2006, Kumar et al. 2009, Kumar and Stohlgren 2009).

the same training/testing data percentages and the same number of repeated model runs as MM2; 75%, 25%, and 25, respectively.

V. Case-Control Study

The case-control study was conducted to distinguish predictors of cave locations from predictors of infected cave locations.

A. *Case and Control Data*

The previously mentioned WNS-positive hibernacula locations were used as case data (n=74). Control points for WNS-negative bat hibernacula were obtained from individual state wildlife agencies within the states that had documented WNS-positive hibernacula before January 1, 2010 (n=31). Not every state agency responded to the data request (WNS-negative response rate=69%, WNS-positive response rate=90%), and some state agencies did not have any known WNS-negative hibernacula sites. Two WNS-negative sites for the state of Maine were kept in the dataset even though WNS-positive points were not provided for Maine. This is because very few WNS-negative points were available, and WNS-positive sites have been reported in Maine since 2009, but exact coordinate locations of these sites have not yet been obtained by Maine's state wildlife division (Strickland 2008, Haneisen 2009).

i. *Geographic Coordinate System and Projection Consistency Among Control Data*

Control hibernacula locations were collected using different parameters by differing state agencies. For example, some were collected in a UTM coordinate system, some in decimal degrees, and some in degrees, minutes, seconds. To generate data consistency, the appropriate data transformation was performed on all control hibernacula points not already in decimal degrees through the use of ERSI's ArcGIS software (ESRI, Redlands, California, USA) and an online conversion website (http://www.cellspark.com/UTM.html). Also, some states collected data using the North American Datum 1927 (NAD27); some collected data using the North American Datum 1983 (NAD83); and some collected data using the World Geographic System 1984 (WGS84). Again, the appropriate data transformation was performed on all control data points not already in WGS84.

ii. Duplication and Clustering Among Control Data

Duplication and clustering were assessed for control data points in the same manner as previously described for case data. No duplicates were identified, and no data points were spatially autocorrelated. This left the original control dataset at a total of 31 points.

iii. Control Data Parameters

Control data were considered negative in 2009 if, 1) the site was confirmed negative in 2008 via laboratory testing or field inspection, *and* 2) there was no overwhelming evidence that the site had become infected (via field inspection) by the time the agency was contacted in 2009.

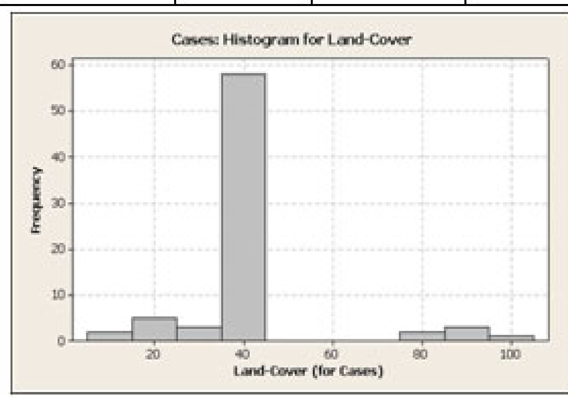
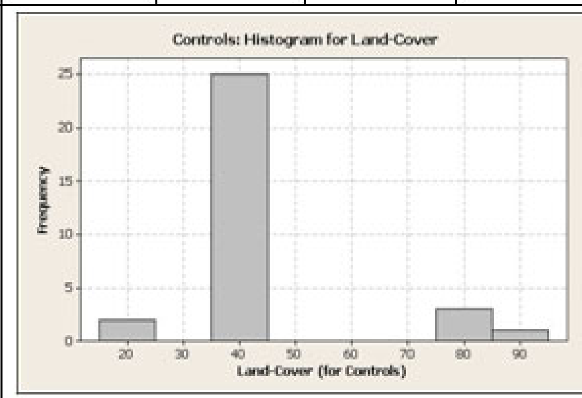
B. <u>*Exposure Assessment*</u>

Exposure assessment for the top four predictor variables among case and control points was determined by extracted data from Hawth's Tools using the previously described 'Intersect Point' analysis tool. This tool provided four data points (a value for each top predictor variable) corresponding with each case and control hibernacula site. All values were extracted from 250-meter resolution data. Descriptive statistics for all four predictor variables can be viewed in **Table 2** below.

C. *Statistical Analyses*

Two-sample t-tests were applied to all datasets that met the required assumptions (Two-Sample Minitab 2010). As a measure of association, this test was used to compare mean values among cases and controls (Two-Sample Minitab 2010). One predictor variable did not meet the required assumptions and therefore had nonparametric statistical tests performed in addition to the two-sample t-test for comparative purposes. The categorical Land-Cover variable used a Chi-Square analysis test to measure association. All analyses were completed using Minitab software, Student Version, Release 14 (Minitab 2003) except for the nonparametric Kolmogorov-Smirnov test, which was computed with an online calculation site (http://www.physics.csbsju.edu/stats/KS-test.n.plot_form.html). In addition, a logistic regression analysis was performed on the top four predictors' case-control data.

Table 2: Descriptive Statistics for Maxent's Top Four Predictor Variables

	WNS Positive Points (n = 74)				WNS Negative Points (n = 31)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Slope	7.2	4.9	0.6	28.5	8.7	6.4	2.2	28.2
Growing Degree Days	3419.1	476.2	2214.7	4441.1	3838.1	497.1	2567.9	4706.0
Annual Temp. Range (BIO-7)	38.0	1.8	33.2	41.5	36.0	2.1	33.1	41.0
Land-Cover								

Note: SD = Standard Deviation

CHAPTER 3: RESULTS

I. Maxent Results

A. *Variable Correlations*

As previously mentioned, all groups of correlated variables were replaced with one surrogate variable from that group, and surrogate variables were chosen based on percent contribution to the Maxent model. **Table 3** shows that most temperature variables were correlated with each other, and most precipitation variables were correlated with each other. Certain additional phenology metrics were also correlated, but these were usually paired variables instead of large groups.

B. *Maxent Model Runs*

The Preliminary Maxent Model (PMM) and Maxent Model #1 (MM1) provided information for running subsequent models. They, individually, did not contribute relevant results to the study and will therefore not be discussed in further detail. A summary of each model run can be found in **Table 4**.

i. Maxent Model #2

Percent contribution results from MM2 (25 averaged models using 74 hibernacula site points and 25 environmental variables) are shown in **Table 5**. Slope [35.1% contribution], Growing-Degree Days [15.2% contribution], Annual Temperature Range [9.6% contribution], and Land-Cover [7.7% contribution] were the top four contributing predictor variables. A continuous prediction map based on

Table 3: Highly Correlated Environmental Variable [Corresponding Maxent Contribution %] Descriptions		
Group	**Correlated Variables**	**Variable Description**
1	BIO-1 [0.0] BIO-5 [0.0] BIO-6 [3.9] BIO-10 [0.2] BIO-11 [0.1] FrostDays [0.1] GrowDays [19.5] Humidity [0.0]	Annual Mean Temp. Max. Temp. of Warmest Month Min. Temp. of Coldest Month Mean Temp. of Warmest Quarter Mean Temp. of Coldest Quarter # of Frost Days # of Growing-Degree Days Humidity
2	BIO-4 [0.5] BIO-7 [6.5]	Temperature Seasonality Annual Temp. Range
3	BIO-12 [0.3] BIO-14 [7.5] BIO-16 [0.0] BIO-17 [0.0] BIO-18 [0.0] BIO-19 [0.2]	Mean Annual Precipitation Precipitation of Driest Month Precipitation of Wettest Quarter Precipitation of Driest Quarter Precipitation of Warmest Quarter Precipitation of Coldest Quarter
4	BIO-13 [2.0] BIO-16 [0.0]	Precipitation of Wettest Month Precipitation if Wettest Quarter
5	EVI-2 [0.1] EVI-3 [0.0]	EVI phenology metric for end of season EVI phenology metric for length of season
6	EVI-4 [0.1] EVI_Min [0.1]	EVI phenology metric for base value EVI phenology metric for minimum value
7	EVI-6 [0.2] EVI_Max [0.0]	EVI phenology metric for peak value EVI phenology metric for maximum value
8	NDVI-2 [0.2] NDVI-3 [0.0]	NDVI phenology metric for end of season NDVI phenology metric for length of season
9	NDVI-4 [0.0] NDVI_Min [2.5]	NDVI phenology metric for base value NDVI phenology metric for minimum value
10	NDVI-6 [0.1] NDVI_Max [0.0]	NDVI phenology metric for peak value NDVI phenology metric for maximum value

Table 4: Model Descriptions & Result Summaries

MODEL	DESCRIPTION & RESULTS
PMM	Beginning # of Variables: 58 Result: Drop 32 variables due to high correlation ($r \geq \pm 0.90$) and 0.0% contribution; repeat with 26 variables.
MM1	Beginning # of Variables: 26 Result: Immediately found 0.0% contribution from BIO-15; remove non-contributing variable and repeat.
MM2	Beginning # of Variables: 25 Result: Variable rankings and model outputs. Comparison Model (CM) assessments showed phenology layers were not increasing the model's prediction; drop phenology layers and repeat.
MM3	Beginning # of Variables: 13 Result: Final model; variable rankings and final model outputs.

Table 5: Percent contribution results from MM2

Variable	*Percent contribution*	*Variable*	*Percent contribution*
Slope	35.1	**CTI**	0.9
Growdays	15.2	**NDVI_6**	0.7
Annual Temp. Range	9.6	**EVI-6**	0.6
Land-Cover	7.7	**NDVI_Mean**	0.3
Precip. Of Wettest Mo.	5.6	**EVI_11**	0.3
Precip. Frequency	5.3	**EVI_2**	0.2
Precip. Of Driest Mo.	4.8	**EVI_9**	0.2
Mean Temp. of Wettest Q.	3.6	**EVI_Min**	0.1
Mean Temp. of Driest Q.	3.5	**NDVI_11**	0.1
NDVI_Min	2.1	**EVI_Mean**	0.1
Precipitation Size	1.5	**NEVI-2**	0.1
Distance to Water	1.4	**EVI-1**	0.1
Aspect	1.0		

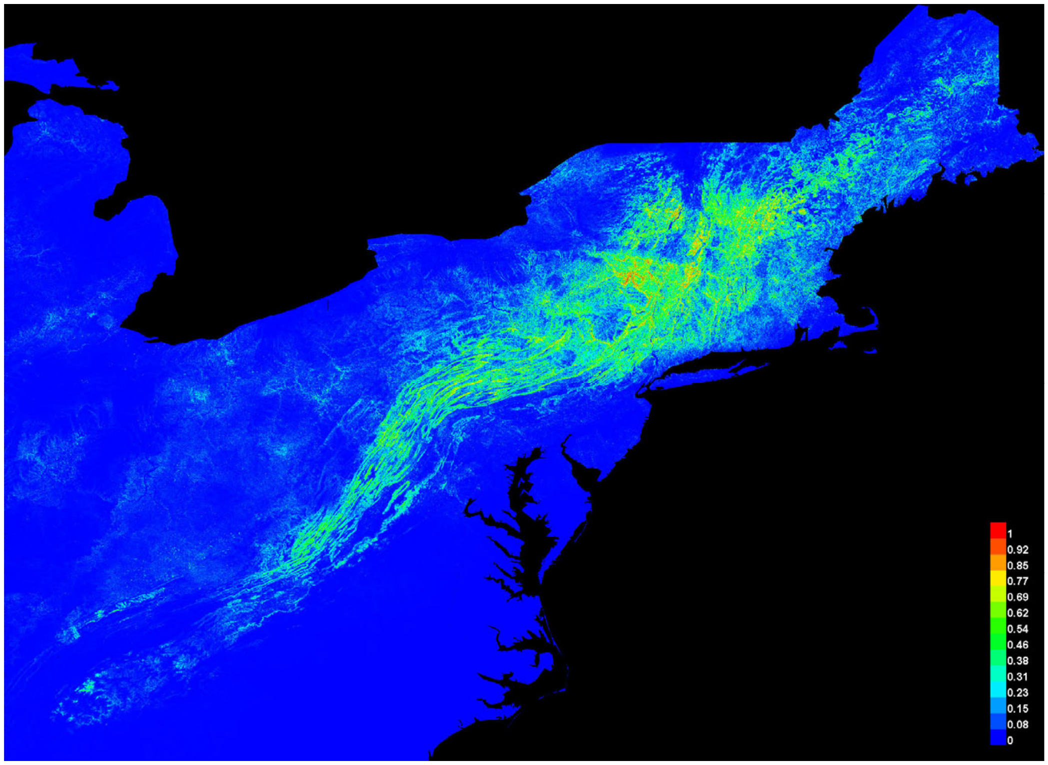
Note: Precip. = Precipitation, Mo. = Month, Temp. = Temperature, Q. = Quarter the averaged 25 MM2 model runs can be found in **Figure 15**, and **Figure 16** shows the receiver operating characteristic (ROC) curve. The ROC curve plots true positive rate against false positive rate for each threshold point (x,y) (Phillips et al. 2004). The area under the ROC curve (the AUC) provides a single measure of model performance, independent of any particular choice of threshold (Phillips et al. 2006). Ideally, the AUC should be as close to one as possible (perfect AUC = 1), while an AUC value of 0.5 indicates that performance is no better than random (Phillips et al. 2004). This measure helps distinguish WNS-presence from random, rather than distinguish WNS-presence from absence (a more common technique, but also requiring presence *and* absence data) (Phillips et al. 2004). For MM2, the AUC was 0.92 and the standard deviation (SD) was approximately ±0.03.

ii. Phenology Metrics

Examination of the phenology metrics showed that CM1 (consisting of MODIS layers alone) had an average mean AUC value of 0.83 (SD = ±0.04), CM2 (analogous to MM2 – all reduced variables included) had an average mean AUC value of 0.92 (SD = ±0.03), and CM3 (13 reduced, non-phenologic environmental variables) had an average mean AUC value of 0.93 (SD = ±0.02).

Thus, the model's predicted AUC increased (indicating better predictability), and the associated SD decreased (indicating less variability; preferred) with each comparison model. The 25 variable model (MM2) was not better than the 13 variable model (MM3) and therefore the more parsimonious model was chosen. Other studies may choose to run a final model with the inclusion of phenology variables. A summary of comparison models can be found below in **Table 6**.

Figure 15: Maxent Model #2 Map Output -
Continuous Predicted Probability of WNS Occurrence
Using 25 Environmental Predictor Variables



Description: Probability of suitable conditions for bat White-Nose Syndrome (Red=high, to Blue=low).

Source: Phillips, n.d.

Figure 16 – ROC graph for MM2

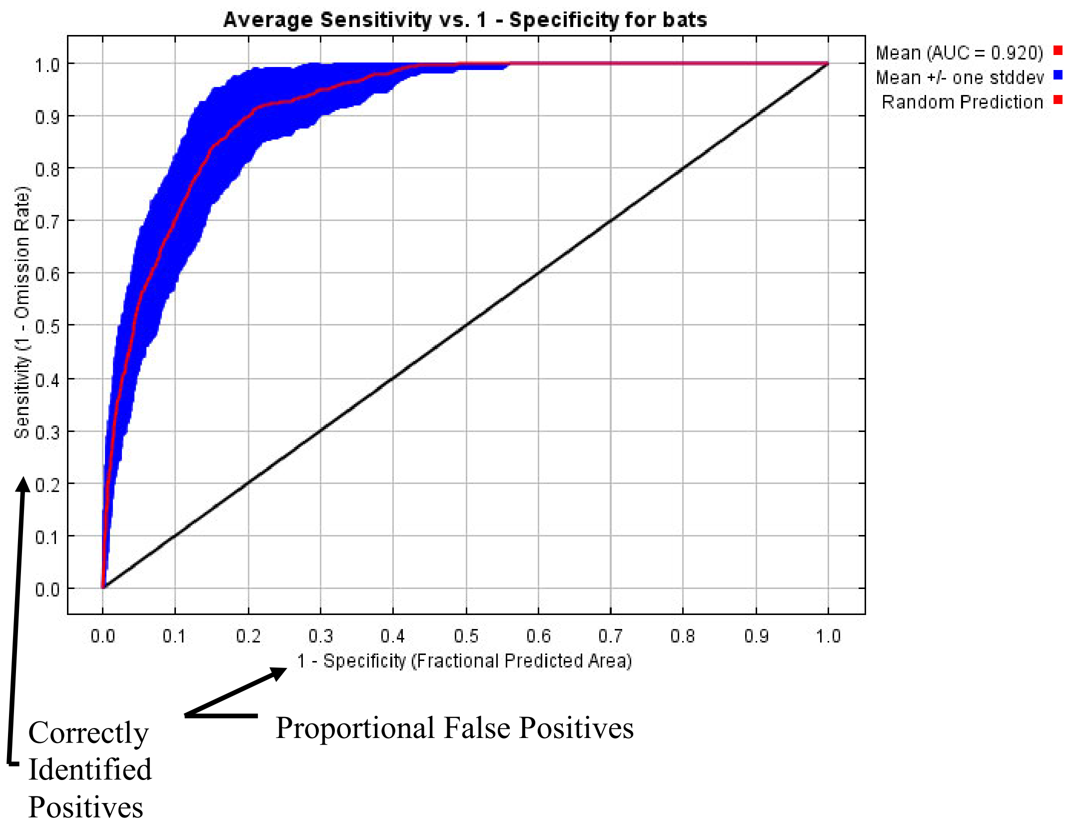


Table 6: Comparison Models (25 replicates for each; 75% training & 25% test data)

Comparison Model (CM)	Environmental Variables Included in Model	Average Mean AUC (±SD)	Rank
CM1	MODIS Layers (Only): 28 total variables	0.83 (±0.04)	**3rd** (Worst Model)
CM2	All Reduced Variables in Model: 25 total variables	0.92 (±0.03)	**2nd** (Decent Model)
CM3	All Reduced Variables in Model minus MODIS Layers: 13 total variables	0.93 (±0.02)	**1st** (Best Model)

i. Final Maxent Model

MM3 variable contribution results from 25 averaged runs can be viewed in **Table 7**. The top four contributing variables were Slope [36.9% contribution], Growing-Degree Days [15.6% contribution], Annual Temperature Range [10.4% contribution], and Land-Cover [8.7% contribution]. In comparison to MM2, MM3's top four variables, including their rank order were the same.

Table 7: Variable Contributions from MM3
(All reduced variables without phenology layers)

Variable	Percent Contribution
Slope	36.9
Growing-Degree Days	15.6
Annual Temp. Range	10.4
Land-Cover	8.7
Precipitation Of Wettest Month	6.1
Frequency of Precipitation	5.5
Precipitation Of Driest Month	4.7
Mean Temp. of Driest Quarter	4.6
Mean Temp. of Wettest Quarter	3.8
Distance to Water	1.3
Precipitation Event Size	0.9
CTI	0.9
Aspect	0.7

Individual response curves for MM3's top variables can be found in **Figure 17 (a.-d.).** The response curve for Slope (**Fig. 17a.**) shows increasing probability of WNS presence with increasing slope degrees (dramatic increase found from 0-6°, slower incline after 6° degrees). The highest WNS probability is found at approximately 28°, and the lowest is found on flat terrain. The response curve for

Growing-Degree Days can be found in **Figure 17b**. Probability of WNS presence based on Growing-Degree Days peaks at about 3,700 degree-days and then sharply drops. The curve suggests WNS occurrence is most probable between approximately 1,400 and 3,700 degree-days, and most unlikely at 6,400+ degree-days. **Figure 17c.** displays Annual Temperature Range. This response curve looks somewhat bell-shaped, beginning its climb at approximately 28.5ºC, peaking around 38ºC, and ending its bell-shaped decent at about 46.5ºC. An individual response curve for Land-Cover is presented in **Figure 17d**. Based on the figure, category 31 (Barren Land) has the largest mean response. Based on frequencies, category 41 (Deciduous Forest) occurs the most.

A continuous prediction map of WNS occurrence based on MM3's results can be found in **Figure 18**. This map shows that WNS occurrence is predicted in several areas of southern Maine (avoiding it's eastern-most coasts), scattered areas of Pennsylvania and Ohio, mildly around eastern Kentucky and eastern Tennessee, along the Ohio and Michigan border, within parts of east-central Indiana, and possibly even stretching up to the northern tip of Michigan's 'glove'.

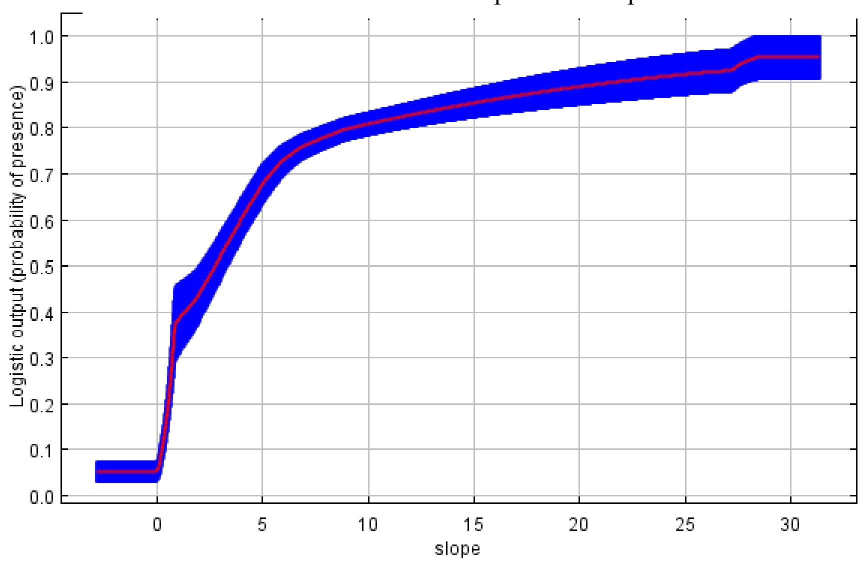
The ROC curve is shown in **Figure 19**. This graph demonstrates the mean AUC value averaged over replicated runs and represents the likihood that a presence will have a higher predicted value than an absence (Hosmer and Lemeshow 2000, Phillips et al. 2004, Phillips et al. 2006). The blue band surrounding the mean (red line) is the standard deviation (± one SD) and shows that most variability occurs around the curved portion of the line. A mean AUC value of 0.93 shows this model's predictive ability, but examination of the top Maxent

Figure 17 (a.-d.): Individual Response Curves for the Top Four Contributing Variables from Maxent Model #4

These curves represent the effect of each variable on Maxent prediction when all other variables are at their average sample value. The mean response of 25 replicate Maxent runs is shown in red, and the mean +/- one standard deviation is shown blue.

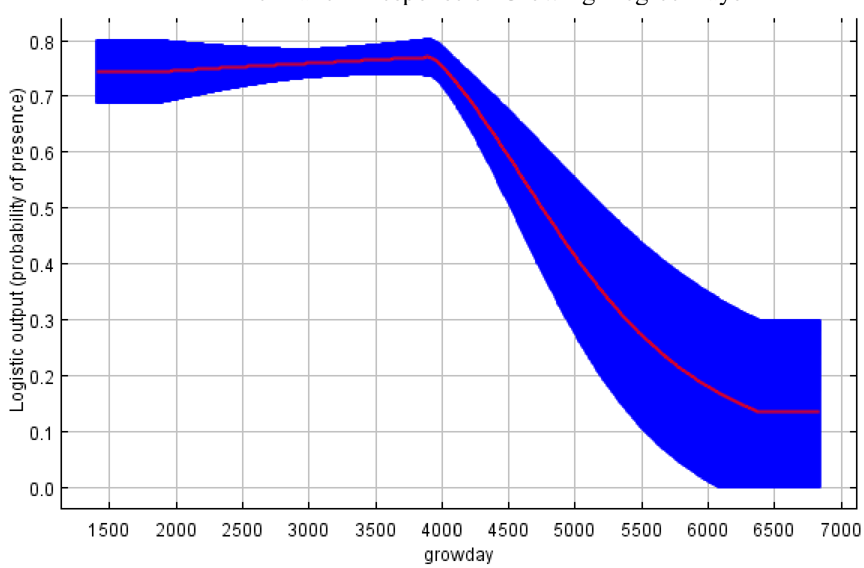
a. Slope

The Maxent Response of Slope

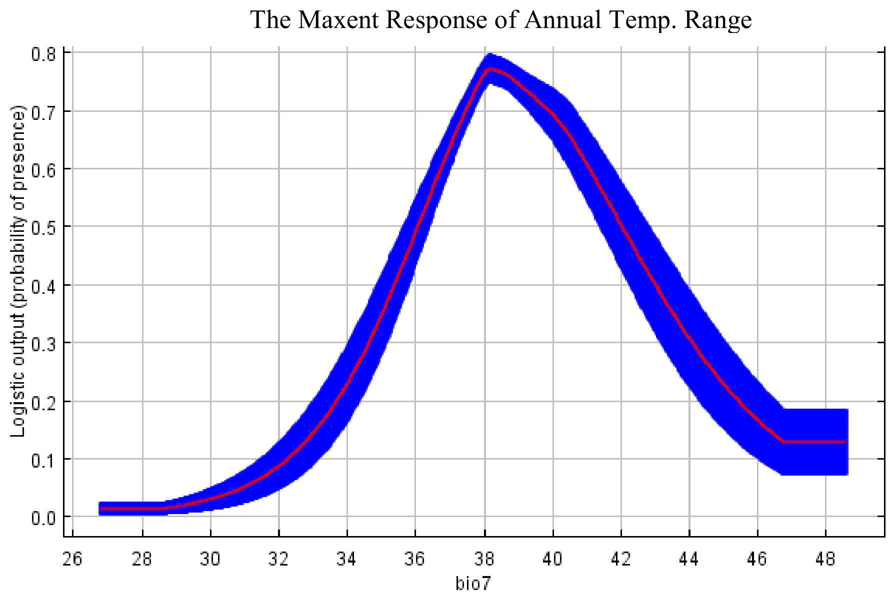


b. Growing-Degree Days

The Maxent Response of Growing-Degree Days



c. Annual Temperature Range



d. Land-Cover

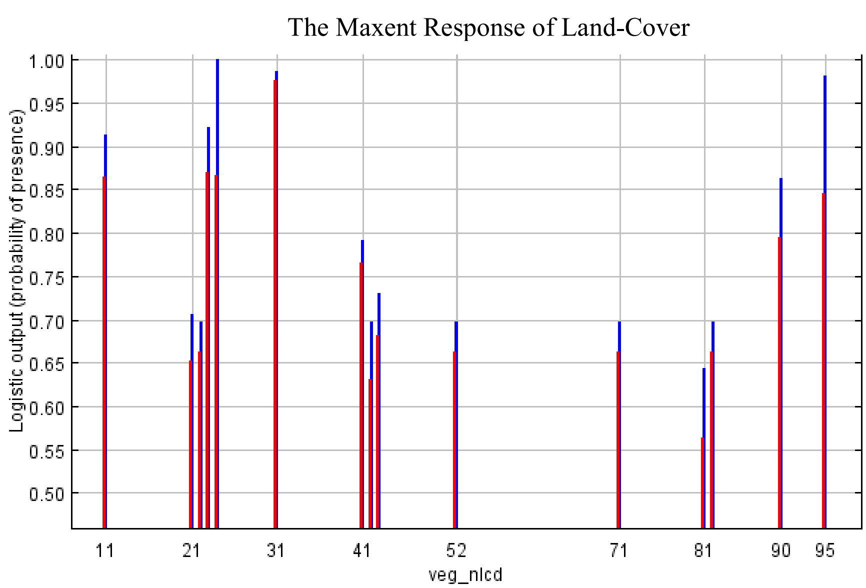
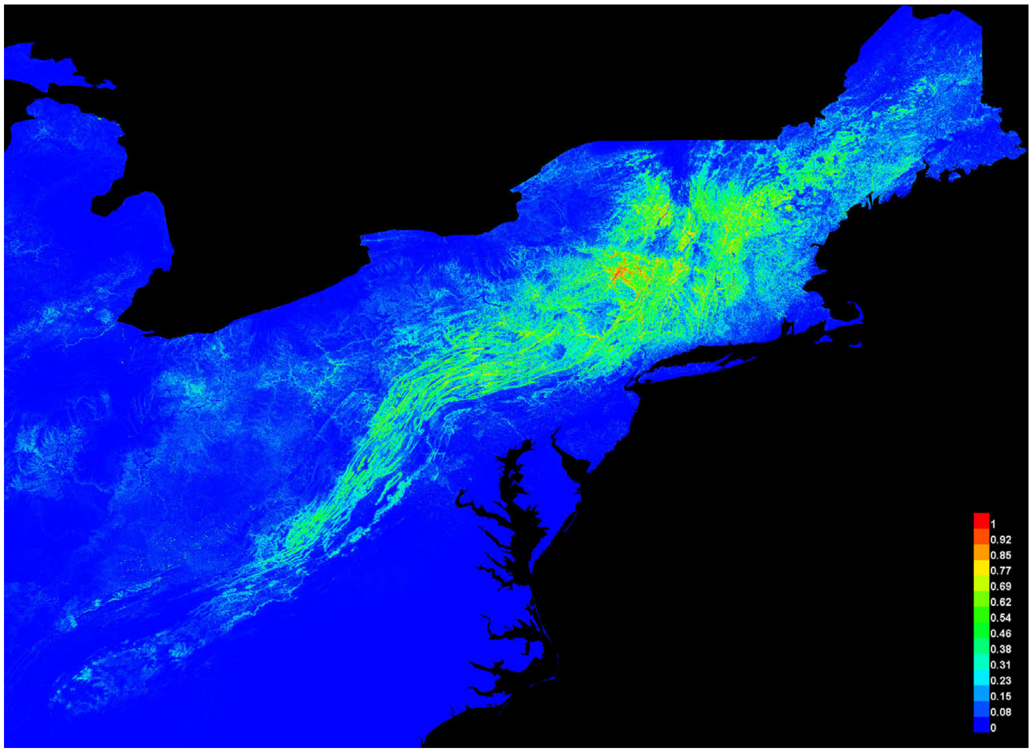


Figure 18: Maxent Model #3 Map Output -
Continuous Predicted Probability of WNS Occurrence
Using 13 Environmental Predictor Variables



Description: Probability of suitable conditions for bat White-Nose Syndrome (Red=high, to Blue=low).

Source: Phillips, n.d.

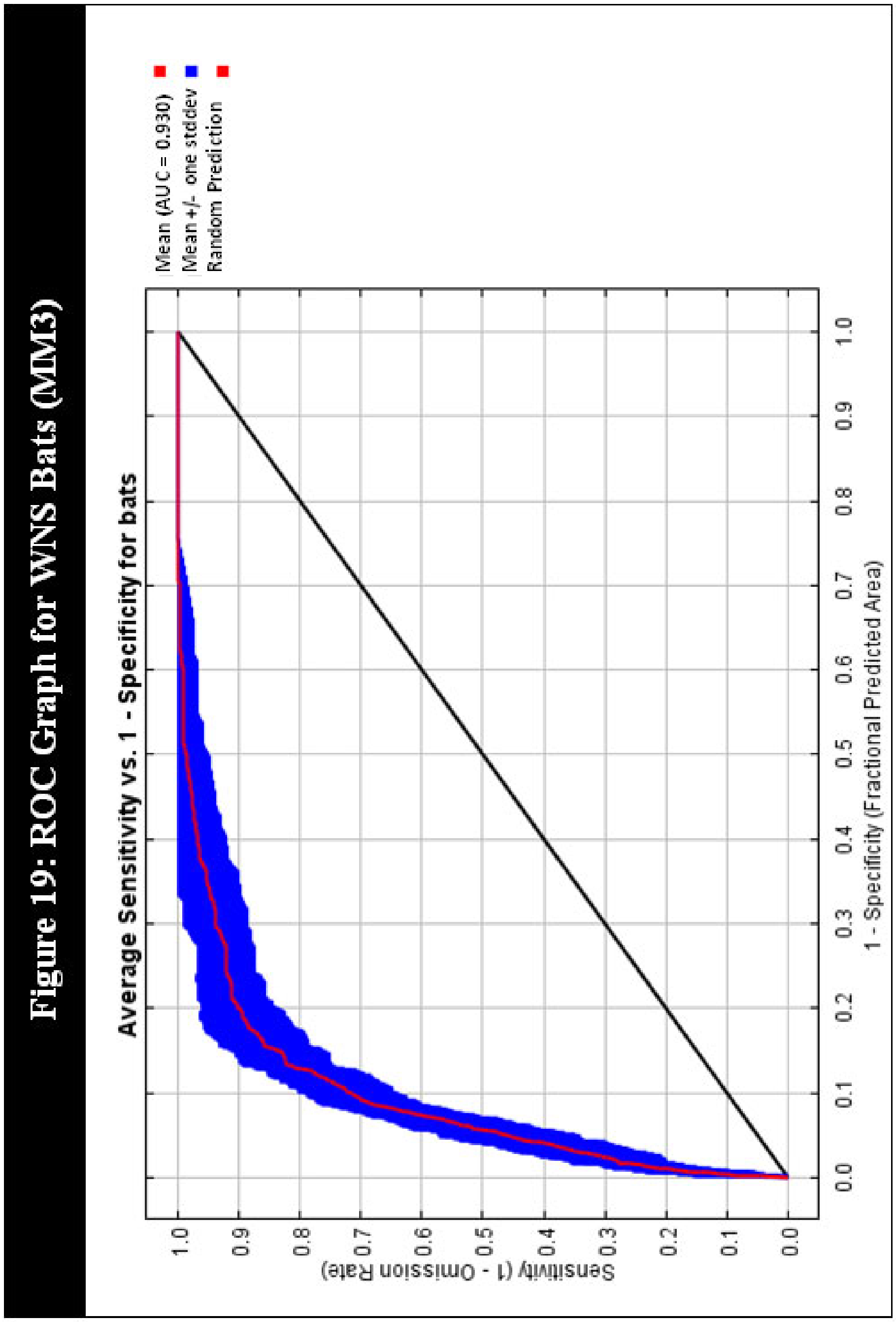


Figure 19: ROC Graph for WNS Bats (MM3)

predictor variables (via case-control study) will determine if there is an actual significant difference between WNS-positive and WNS-negative variable associations, or if this model is simply good at predicting general hibernacula locations.

II. Case-Control Results

A. *Statistical Assumptions*

All continuous variables (Slope, Annual Temperature Range, and Growing-Degree Days) met the required independence and normality assumptions for using a two-sample t-test. See **Table 8** for normality evaluations. The third t-test assumption, equal sample variances, was only met by Annual Temperature Range and Growing-Degree Days (see **Table 9**). Transformations to mediate the non-equivalent variances within Slope data included: reciprocal, square root, logarithmic, and exponential. None of these provided improvement, so Slope statistical testing was performed with both parametric and nonparametric statistical terms[5].

The Chi-Square test statistic was used to evaluate associations between WNS status and Land-Cover. All Chi-Square test assumptions[6] were met by the categorical Land-Cover variable (table not shown) (Agresti 1990, Yates 1999). Assumptions were evaluated using Minitab statistical software, Student Version, Release 14 (Minitab Inc. 2003). Minitab software was also utilized for the logistic regression analysis.

B. *Statistical Test Results*

Statistical tests for Slope, Growing-Degree Days, Annual Temperature Range, and Land-Cover (including test type, hypotheses, significance level, test statistic values, 95% confidence intervals, and results) can be viewed in **Table 10** below.

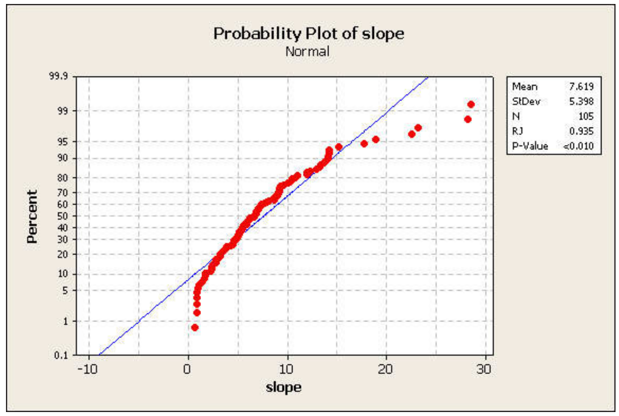
---

[5] Studies of the t-test have shown that the sample variance assumption can be violated to an amazing degree without substantial effects on the results (http://www.fammed.ouhsc.edu/tutor/tstds.htm).

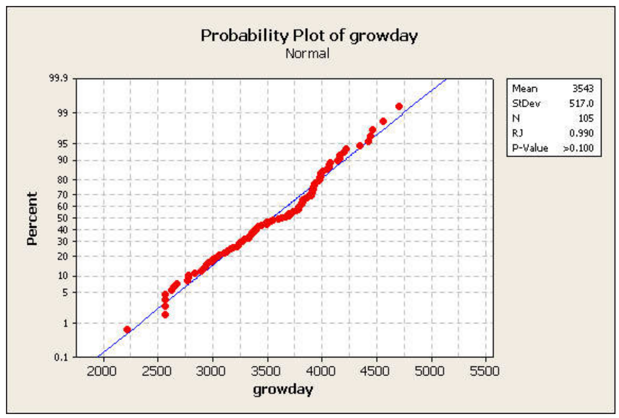
[6] Chi-Square test assumptions include: 1) Each data point must contribute data to only one cell, 2) Each observation is independent of all the others, 3) No more than 20% of the expected values are less than five, and 4) All expected frequencies should be 10 or greater.

Table 8: Testing the Normality Assumption

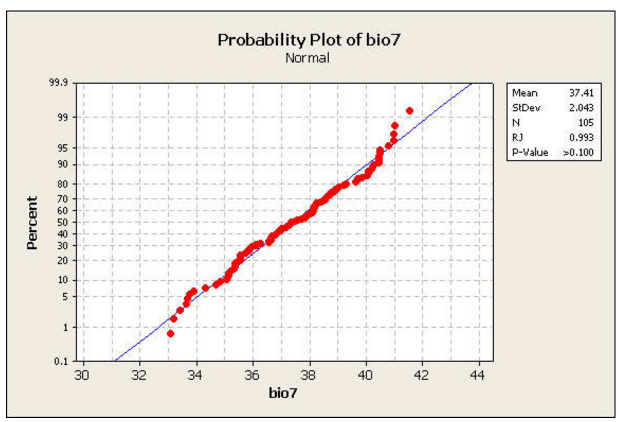
The Ryan-Joiner normality test (C.L. = 95%) for Slope has a p-value of <0.010. Under the hypotheses ‘H_0: Normal’ and ‘H_1: Non-normal’, this p-value is significant, indicating that the data follow a non-normal distribution.



The Ryan-Joiner normality test (C.L. = 95%) for Growdays has a p-value of >0.100. Under the hypotheses ‘H_0: Normal’ and ‘H_1: Non-normal’, this p-value is not significant, indicating that the data follow a normal distribution.



The Ryan-Joiner normality test (C.L. = 95%) for Annual Temperature Range has a p-value of >0.100. Under the hypotheses ‘H_0: Normal’ and ‘H_1: Non-normal’, this p-value is not significant, indicating that the data follow a normal distribution.



Note: C.L. = Confidence Level

Table 9: Testing for Equal Sample Variance

	Slope	Growdays	Annual Temp. Range
Hypotheses	H_0: No difference between sample variances H_1: Difference between sample variances	H_0: No difference between sample variances H_1: Difference between sample variances	H_0: No difference between sample variances H_1: Difference between sample variances
Critical Value $\alpha = 0.05$	F-critical = ~1.62	F-critical = ~1.62	F-critical = ~1.62
Test Statistic	F-test = $\frac{\text{larger sample variance}}{\text{smaller sample variance}} = \frac{6.360^2\ (df = 30)}{4.924^2\ (df = 73)} = 1.668$	F-test = $\frac{\text{larger sample variance}}{\text{smaller sample variance}} = \frac{497.1^2\ (df = 30)}{476.2^2\ (df = 73)} = 1.090$	F-test = $\frac{\text{larger sample variance}}{\text{smaller sample variance}} = \frac{2.063^2\ (df = 30)}{1.748^2\ (df = 73)} = 1.393$
Results $\alpha = 0.05$	1.668 > ~1.62; Reject H_0 (there is a difference between sample variances)	1.090< ~1.62; Fail to reject H_0 (there is no difference between sample variances)	1.393 < ~1.62; Fail to reject H_0 (there is no difference between sample variances)

Note: df = degrees of freedom, ~ = approximately

Table 10: Statistical Significance Testing

Predictor Variable	**Slope**			**Growdays**	**Annual Temp. Range**
Statistical Test	*Parametric*	*Nonparametric*		Two-Sample t-test	Two-Sample t-test
	Two-Sample t-test	Mann-Whitney U test	Kolmogorov-Smirnov two-sample test		
Hypotheses	H_0: Population means are equal H_1: Population means are not equal	H_0: Population medians are equal H_1: Population medians are not equal	H_0: Probability distributions are equal H_1: Probability distributions are not equal	H_0: Population means are equal H_1: Population means are not equal	H_0: Population means are equal H_1: Population means are not equal
Significance Level	$\alpha = 0.05$	$\alpha = 0.05$	$\alpha = 0.05$	$\alpha = 0.05$	$\alpha = 0.05$
Test Statistic And Associated Outputs of Interest	T = -1.15 P-value = 0.257 Mean Diff. = -1.466 95% C.I. for Diff. = (-4.039, 1.106)	U = 1275.0 P-value = 0.372 Mean Diff. = -0.695 95% C.I. for Diff. = (-2.625, 1.078)	D = 0.1486 P-value = 0.684 Case Mean = 7.186 Control Mean = 8.652 Case 95% C.I. (6.05, 8.33) Co. 95% C.I. (6.32, 10.98)	T = -3.99 P-value < 0.001 Mean Diff. = -418.936 95% C.I. for Diff. = (-629.560, -208.312)	T = 4.63 P-value < 0.001 Mean Diff. = 1.955 95% C.I. for Diff. = (1.106, 2.805)
Results	0.257 > 0.05; The two sample means are not significantly different.	0.372 > 0.05; The two samples do not have significantly different medians.	0.684 > 0.05; The two samples do not have significantly different probability distributions.	0.000 << 0.05; The two sample means are significantly different.	0.000 << 0.05; The two sample means are significantly different.

Note: C.I. = Confidence Interval, Diff. = Difference, Co. = Control *i. Slope*

The Mann-Whitney U test, the Kolmogorov-Smirnov two-sample test, and the two-sample t-test were computed with Minitab, Student Version, Release 14 (Minitab n.d.) and the following online calculation website: http://www.physics.csbsju.edu/stats/KS-test.n.plot_form.html. They hypothesized that two population medians were equivalent, two population distributions were equivalent, and two population means were equivalent, respectively. Test results for the Mann-Whitney U test included a mean difference value of -0.70, a 95% confidence interval for this difference of (-2.63, 1.08), a test statistic of U = 1275.0, and a P-value of 0.37. Results from the Kolmogorov-Smirnov two-sample test included a case mean value of 7.19, a control mean value of 8.65, a case 95% confidence interval of (6.05, 8.33), a control 95% confidence interval of (6.32, 10.98), a test statistic of D = 0.15, and a P-value of 0.68. Test results regarding the two-sample t-test for unequal variances included a mean difference value of -1.47, a 95% confidence interval for this difference of (-4.04, 1.11), a test statistic of T = 1.15, and a P-value of 0.26. Using $\alpha=0.05$ for all tests, cases and controls did not have significantly different medians, population distributions, or means.

ii. Growing-Degree Days

At the 95% confidence level, Growing-Degree Days resulted in a statistically significant difference between mean values among cases and controls. T-test results included a T-value of -3.99, a P-value of <0.001, a mean difference of -418.94, and a 95% confidence interval for this difference of (-629.56, -208.31).

iii. Annual Temperature Range

Using the 95% confidence level, Annual Temperature Range also found statistically significant results between mean values among cases and controls. T-test results included a T-value of 4.63, a P-value of <0.001, a mean difference of 1.96, and a 95% confidence interval for this difference of (1.11, 2.81).

iv. Land-Cover

All NLCD land-cover datasets have the potential to contain a total of 29 categorical classification options (categories and descriptions can be found in **Appendix 2**) (USGS Multi 2010). Classifications are coded with numerical values that represent each category. Within the case-control dataset, 11 different categorical codes were reported (11, 21, 23, 24, 31, 41, 42, 43, 81, 90 and 95). Among these, category 41 was most frequent, and category 31 had the highest mean response with WNS (as found in the individual response curve for land-cover).

In order to validate Chi-Square results, categories had to be combined such that cell frequencies were not too low. Two different combined-category Chi-Square tests were of interest: 1) Land-Cover category 41 vs. all others, and 2) Land-Cover category 31 vs. all others. Chi-Square results for Land-cover category 41 vs. all others proved no association between the variables at $\alpha=0.05$ (Pearson's Chi-Square = 0.64, P-Value = 0.43). Chi-Square analysis for Land-cover category 31 vs. all others resulted in low cell frequencies, so Land-cover category 23 (second highest mean response variable) was removed from 'all others' and grouped with 31. Cell frequencies were now high enough, but test results showed no association between cases and controls at $\alpha=0.05$ (Pearson's Chi-Square = 0.2.199, Fisher's Exact test[7] P-Value = 0.318).

v. Logistic Regression

Logistic regression results for the four-variable model supported the case-control study results by showing that both Slope and Land-Cover variables were not statistically significant, and both Growing-Degree Days and Annual Temperature Range were statistically significant.

A second regression model with three variables (the two statistically significant variables plus their interaction term) did not show evidence of effect modification between Growing-Degree Days and Annual Temperature Range.

[7] Fisher's exact test provides a more accurate p-value for low cell counts and is provided in this study when Minitab software output results show it is appropriate (Agresti 1990, Yates 1999).

CHAPTER 4: DISCUSSION

I. Findings

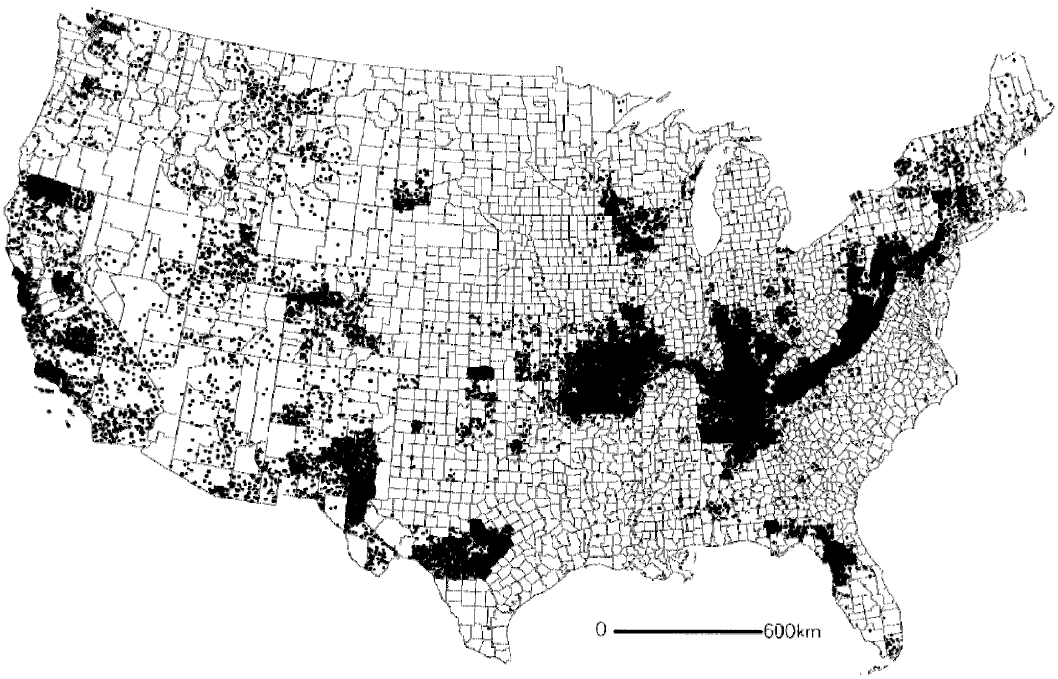
As described in the section on multicollinearity, there were several instances in which one variable took the place of a group of highly correlated variables. Due to this, the interpretation of one variable may include various ecological explanations.

A. *Slope and Land-Cover Variables*

Based on the case-control analyses, Slope and Land-Cover variables were not statistically significantly associated with bat WNS status. Neither variable was highly correlated with any other variables, so the lack of a significant association found via case-control analysis is only applicable to these variables alone. Due to Maxent's high ranking of these variables as predictors, Slope and Land-Cover are most likely associated with the general location of bat hibernacula, but not WNS-positive bat hibernacula specifically. To visually examine this, three figures are presented in **Figure 20 (a.-c.)** below. This figure is given mainly because literature that addresses the possible influences of landscape structure on spatial patterns in bat communities is critically lacking (Jaberg and Guisan 2001). One study performed by Jaberg and Guisan showed that cave locations generally coincide with areas of hilly or mountainous terrain (2001), but additional information is sparse. In lieu of adequate literature, **Figure 20** presents **(a.)** a dot-density map of U.S. caves per county, **(b.)** a U.S. terrain map, and **(c.)** a U.S. land-cover map. These three figures can be used to visually compare locations of U.S. caves with slope and land-cover patterns. Based

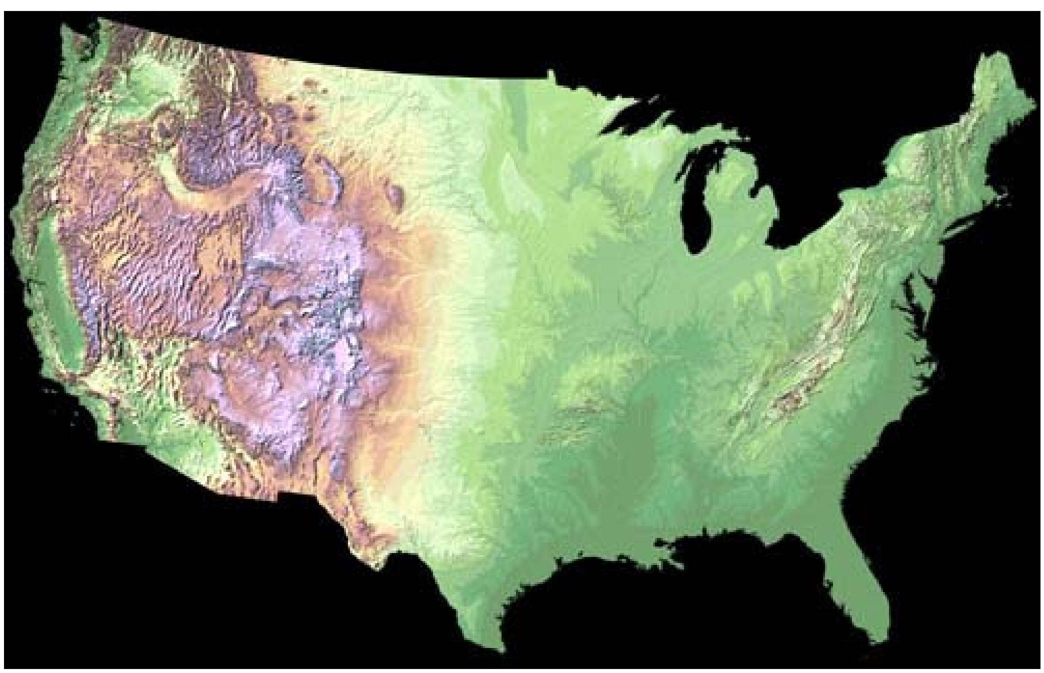
Figure 20 (a.-c.): Visual Comparisons Between Number of Caves, Terrain, and Land-Cover in the Contiguous United States

a. Dot density map of the number of caves per county in the contiguous United States.



(Source: Culver et al. 1999)

b. United States terrain map.



(Source: http://egsc.usgs.gov/isb/pubs/factsheets/fs10602.graphics/usa.jpg)

c. United States Land-Cover Map

(Source: http://www.anglonautes.com/hist_map_flag/us_maps_land_cover_1.jpg)

on these figures, it seems that caves may overlap with certain slope values (specifically, mountainous terrain) and certain vegetation classifications (specifically, more vegetative areas). Maxent's ranking of Slope and Land-Cover as top predictor variables shows that they have some relationship with the geographic location of bat hibernacula, but since the case-control statistical results were not significant it is unlikely that they have a specific relationship with the WNS positive/negative status of the hibernacula. Logistic regression results supported this conclusion.

B. <u>*Growing-Degree Days and Annual Temperature Range*</u>

According to two-sample t-test case-control study results, both Growing-Degree Days and Annual Temperature Range variables resulted in significantly different mean values among cases and controls (two-sided test; statistical association's direction is unknown). Basic descriptive statistics of the case-control dataset showed that Growing-Degree Days cases had a lower mean value than controls, and Annual Temperature Range cases had a higher mean value than controls (refer to page 40). Maxent's ranking of these two variables as top predictors suggests that there is some form of association between them and the WNS site locales. Additionally, the case-control study analyses indicate that the relationship may actually be linked to the infection status of a site and not just bat hibernacula locales.

Logistic regression results supported these t-test findings for Growing-Degree Days and Annual Temperature Range variables when all four predictors were included in the model.

i. Variable Correlations

Growing-Degree Days was highly correlated with Humidity, Number of Frostdays, Annual Mean Temperature (BIO-1), Maximum Temperature of Warmest Month (BIO-5), Minimum Temperature of Coldest Month (BIO-6), Mean Temperature of Warmest Quarter (BIO-10), and Mean Temperature of Coldest Quarter (BIO-11). Annual Temperature Range was only correlated with one variable, Temperature Seasonality (BIO-4). Individual response curves for each correlated variable would be needed in order to analyze the specific relationship each variable might have with bat WNS, but these cannot be viewed since the terms were never actually added to the Maxent model. Regardless, it is important to consider these correlated variables when interpreting results.

In general, temperature-related correlation factors seem to dominate both variables. Even the predictor variables themselves (Growing-Degree Days and Annual Temperature Range) are highly temperature related, suggesting that temperature may play a large role in the spread of WNS-related *Geomyces destructans*. As previously mentioned, psychrophilic fungi (such as *Geomyces destructans*) are known to thrive best in cool temperatures (around 5-10ºC) (Boyles and Willis 2009), but the causal pathways are unknown (Hajek 1994; Blehert *et al.* 2009). Discovery of these seasonally-specific temperature-related variables provides insight to potential annual climatic variations that may foster WNS bat hibernacula infection and may lay the foundation for causal pathway discoveries. Two benefits of identifying these variables within this study include: 1) narrowing down the range of possible factors influencing WNS occurrence, and 2) providing a starting-point for future work on WNS environmental determinants.

II. Study Strengths and Limitations

A. *Limitations*

This study's sample size was probably one of its largest limitations. Due to the recent appearance of WNS, current data and information on site locations is difficult to find. Many northeastern U.S. states have known infected sites, but many of these sites are not even recorded as geographic coordinate locations, making large sample sizes and power within a study more difficult ascertain. Regardless, the ecological problems surrounding WNS occurrence must be tackled immediately.

Another limitation included several state agencies' hesitation with officially declaring certain hibernacula sites negative due to their close proximity to infected sites, but these were still defined as 'negative' within the study. The effect of this limitation would most likely bias results towards the null. Other state agencies expressed a similar hesitancy regarding long spans of time without site inspection, which would also most likely bias results towards the null. The number and percentage of WNS-positive and WNS-negative hibernacula points retrieved from each state can be viewed in **Table 11**.

Several other types of limitations have been introduced to this study due to biasing effects. Cases and controls were chosen based on what was available, not what was ideal, so selection bias may have been introduced. For example, the association between ‘Growdays’ and ‘WNS status’ may differ between those that were selected and those that theoretically could have been selected. There is no way of knowing whether “theoretically” selected data would have had a similar association as “actually” selected data, so the biasing effect direction remains undetermined.

Sources contributing to this selection bias might include: number of field inspectors per state, number of field inspections per year, accessibility of site locations, and proximity of sites to state wildlife agency offices. Each of these factors may potentially bias the geographic location at which sites were collected, and therefore represent, or fail to represent, the environmental characteristics found at all sites. For example, if New York only had 2 field inspectors for the entire state (i.e. not enough to adequately assess hibernacula for the entire state), they might only get to examine sites within a certain radius of their employment office. Therefore, certain sites outside of this radius may be missed. Spatially speaking, the geographic characteristics found at sites within this radius may or may not be similar to the geographic characteristics of those that were missed.

Table 11: Percentage of Positive and Negative Data Points in Each State			
State	**% WNS-Positive Sites (# in state/total)**	**% WNS-Negative Sites (# in state/total)**	**Raw Difference Between Positive & Negative Percentages**
Connecticut	**4.05%** (3/74)	**6.45%** (2/31)	**2.40**
Maine	**N/A**	**6.45%** (2/31)	**6.45**
Massachusetts	**9.46%** (7/74)	**3.23%** (1/31)	**6.23**
New Hampshire	**6.76%** (5/74)	**0/31** (0.0%)	**6.76**
New Jersey	**4.05%** (3/74)	**N/A**	**4.05**
New York	**37.84%** (28/74)	**9.68%** (3/31)	**28.16**
Pennsylvania	**10.81%** (8/74)	**N/A**	**10.81**
Vermont	**14.86%** (11/74)	**6.45%** (2/31)	**8.41**
Virginia	**6.76%** (5/74)	**25.81%** (8/31)	**19.05**
West Virginia	**5.41%** (4/74)	**41.94%** (13/31)	**36.53**
TOTAL	100% (74/74)	100% (31/31)	0

N/A = Not Available

Other types of biasing effects within the study include: 1) GPS instrument error or measurement inaccuracies, 2) variation in accuracy among point data (some dead-on, some estimated, some deliberately inaccurate for privacy purposes), 3) miss-reporting or failure to report infected and non-infected sites, and 4) data transformations and resolution resampling error.

Additionally, MODIS data used for Maxent modeling purposes was only available for the years 2006-2007, but WNS data points were collected over a period of four years (2006, 2007, 2008, and 2009). Ideally, averaged MODIS data from the entire WNS collection year range would be obtained to produce the most accurate results, but values have most likely not changed considerably throughout that time period. Similarly, the Daymet climate dataset represented data averaged over the years 1980-1997. Again, current data would have been preferred, but averaged values over a period of 17 years are unlikely to be considerably different to averaged values for 2006-2009.

Lastly, it is important to remember that these results are not abstractable to other bat species or geographic regions.

B. <u>*Strengths*</u>

Strengths of this study include the use of all available 2009 WNS data, inclusion of MODIS' most recent and best resolution satellite imagery capturing biomass/vegetation indices, incorporation of a new modeling technique (Maxent modeling; proven very effective for various types of ecological niche modeling and only requires presence locale data), utilization of a large number of environmental predictor variables for Maxent's modeling procedures (totaling 58), and it's representativeness of the study's sample population (i.e., representativeness is most likely good because both case and control groups contained hibernacula locations that housed all six bat species and both groups were selected from the same geographic area). In addition to these, a case-control analysis of the top four predictor variables was performed (allowing for a more in-depth look at possible associations between the top predictor variables and WNS status), an output map of continuous WNS prediction spanning the study area was created, and identical transformation procedures (by state) were used on all WNS points (preventing over exaggeration of the effect estimate).

Manual detection of spatially correlated hibernacula points replaced the more-often used spatial auto-correlation technique in order to reduce the model's inaccuracy (Phillips 2008). Most field survey efforts tend to be strongly spatially biased (typically due to proximity and accessibility) and are therefore also strongly spatially auto-correlated (Phillips et al. 2009). High spatial correlation introduces model error because environmental data gets collected at random, but hibernacula site data does not. With a small sample size, automatic detection and deletion of spatially correlated data is likely to remove too many points. Veloz comments that automated spatial correlation for presence-only niche modeling can wrongly inflate measures of accuracy and significantly reduce your sample size (2009). Manually selecting spatially autocorrelated points helped reduce the presence of spatial bias, while leaving enough data for extractable study results.

Prediction of bat hibernacula distribution based on environmental input factors has never been attempted before in published material. This makes focusing the prediction of bat hibernacula on a specific pathogen a major mile-stone in field of bat ecology, and a major strength of this study.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

Using exploratory approaches, this study aimed to identify possible environmental factors linked to the occurrence of mortality-causing bat white-nose syndrome in the northeastern United States. Study results indicated that Humidity, Frostdays, Growdays, and various measures of temperature were significantly associated with the occurrence of white-nose syndrome in bats.

It is highly likely that the pathogenic fungus *Geomyces destructans* is a factor in causing bat mortality (Blehert *et al.* 2009, Boyles and Willis 2009; USGS White 2009), so measures of association between *Geomyces destructans* and the significant variables found in this study may be an ideal direction for future WNS research. Additionally, future studies could aim for the following: 1) improve the resolution at which Maxent modeling was performed (requires more precise data), 2) include recent infected hibernacula data in the Maxent model (data that was not available at the start of this study), 3) consider adding other variables to the model (such as ownership, i.e. private, government, or public land, as a measure of human visitation and the role of humans in pathogenic spread), 4) obtain interior cave measurements for more accurate exposure assessments, 5) include a Karst topography layer (cave-like landscape formed by drainage eroding away limestone), 6) increase the number of control points used in the case-control analysis, 7) expand the study area to broaden prediction range, and 8) try changing the testing and training data percentages within Maxent once more data is available.

Implications of variable correlation associations with the WNS-related fungal pathogen could be extremely relevant for future WNS prediction and prevention, but at this point it is difficult to determine the exact role that the co-factors may have with *Geomyces destructans*' ability to cause bat mortality. Individual response curves for each correlated variable were not given in the Maxent results. Future research could aim to investigate these individual response curves in order to determine the range of exact values in which WNS occurrence is most likely to occur.

As previously stated, WNS occurrence is predicted in several areas of southern Maine (avoiding it's eastern-most coasts), scattered areas of Pennsylvania and Ohio, mildly around eastern Kentucky and eastern Tennessee, along the Ohio and Michigan border, within parts of east-central Indiana, and possibly even stretching up to the northern tip of Michigan's 'glove'. These are the recommended target areas for future WNS-prevention efforts.

Out of the originating 58 variables, study results suggest that cave locations may be best predicted by Slope and Land-Cover geographic characteristics (Maxent results), and that infected cave locations are most impacted by Growing-Degree Days and Annual Temperature Range (including correlated variables) (case-control results).

The cold-loving *Geomyces destructans* fungus has very recently been found in the following previously unaffected areas: Canada, Tennessee, Missouri (officially released May 14th, 2010), and Oklahoma (officially released May 19, 2010) (Cryan 2010, Griffin 2010, National Park Service 2010). *Geomyces destructans* has also been found in certain

European countries, including France, but does not seem to present the same mortality rate that has been dominating infected bats in the United States (Puechmaille 2010).

Surveillance soil testing for *Geomyces destructans* outside of the northeastern U.S. has not yet begun; only visual signs of the disease have been noted (Cryan 2010). As the potential for this fungal pathogen to spread increases over time, appropriate and timely steps must be taken to narrow down causative and influencing factors. The ability of environmental variables to impact disease emergence and dispersion patterns *must* be considered when approaching causal research regarding new pathogens. Future goals for combating bat WNS include understanding the life cycle of this fungus and, ultimately, preventing the decimation of susceptible bat species worldwide. Use of this study's results are intended as a starting point for the determination of environmental factors associated with bat WNS. Ideally, ecological field specialists and researchers will continue this environmental investigation via future research for more detailed information regarding environmental associations with bat white-nose syndrome.

REFERENCES

Agresti, A. (1990). Categorical Data Analysis. New York: Wiley.

Arc Macro Language – AML (2009). ArcInfo 7x Programming. GIS Lounge and DM Geographics, LLC. Retrieved October 8, 2009 from http://gislounge.com/arc-macro-language-aml/

Bahls, L.L. (2009). Potential habitat distribution for the freshwater diatom Didymosphenia geminate in the continental US. *Frontiers in Ecology and the Environment*, 7, doi: 10.1890/080054.

Benito, B.M., M.M. Martinez-Ortega, L.M. Munoz, J. Lorite, J. Penas (2009). Assessing extinction-risk of endangered plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses. *Biological Conservation*, doi: 10.1007/s10531-009-9604-8.

Beyer, H.L. (2004). Hawth's Analysis Tools for ArcGIS. Retrieved March 2, 2010 from http://www.spatialecology.com/htools/overview.php

Blehert, D.S., A.C. Hicks, M. Behr, et al. (2009). Bat White-Nose Syndrome: An Emerging Fungal Pathogen? *Science*, 323, 227.

Boyles, J.G. and C.K.R. Willis (2009). Could Localized Warm Areas Inside Cold Caves Reduce Mortality of Hibernating Bats Affected by White-Nose Syndrome? *Frontiers in Ecology and the Environment*, doi: 10.1890/080187.

Cockrum, E.L. (1956). Homing, Movement, and Longevity of Bats. *Journal of Mammalogy*, 37, 48-57.

Courtin, F.; W.B. Stone; G Risatti; K. Gilbert; H.J. Van Kruiningen. Pathologic Findings and Liver Elements in Hibernating Bats With White-Nose Syndrome. *Veterinary Pathology*, Published online 28 January 2010, DOI: 10.1177/0300985809358614

Cryan, Paul (2010). White-Nose Syndrome Threatens the Survival of Hibernating Bats in North America. United States Geological Survey (USGS). Retrieved May 25, 2010 from http://www.fort.usgs.gov/wns/

Culver, D.C., H.H. Hobbs III, M.C. Christman, L.L. Master (1999). Distribution Map of Caves and Cave Animals in the United States. *Journal of Cave and Karst Studies* 61(3), 139-140.

Daszak, P., A.A. Cunningham, and A.D. Hyatt (2001). Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Tropica*, 78(2), 103-116.

Daymet U.S. Data Center (n.d.). Daily Surface Weather and Climatalogical Summaries (DAYMET). Retrieved February 21, 2010 from http://www.daymet.org/

Elith, J., Catherine H. Graham, Robert P. Anderson, *et al.* (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129-151.

Environmental Systems Research Institute (ESRI) (n.d.). GIS Software, Version 9.2, Redlands, California, USA. http://www.esri.com/products/index.html.

Environmental Systems Research Institute (ESRI) (n.d.). How Euclidean Functions Work. ArcGIS 9.2 Desktop Help. Retrieved March 1, 2010 from http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Euclidean%20functions%20work

Environmental Systems Research Institute (ESRI) (2008). How Nibble Works. ArcGIS 9.2 Desktop Help. Retrieved February 21, 2010 from http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Nibble%20works

Fenton, M.B. (1969). Summer Activity of *Myotis lucifugus* (Chiroptera: Vespertilionidae) in Ontario and Quebec. *Canadian Journal of Zoology*, 47, 597-602.

Fenton, M.B. (1970). Population Studies of *Myotis lucifugus* (Chiroptera: Vespertilionidae) at Hibernacula in Ontario. *Life Science Miscellaneous Publications*, Royal Ontario Museum, 77, 1-34.

Findley, James S. (1993). Bats: A Community Perspective. Great Britain: The Press Syndicate of the University of Cambridge.

Fry, J.A., Coan, M.J., Homer, C.G., Meyer, D.K., and Wickham, J.D., 2009, Completion of the National Land Cover Database (NLCD) 1992–2001 Land Cover Change Retrofit product: U.S. Geological Survey Open-File Report 2008–1379. Retrieved March 2, 2010 from http://pubs.usgs.gov/of/2008/1379/

Gargas, A.; M.T. Trest; M. Christensen; T.J. Volk; and D.S. Blehert (2009). Geomyces destructans sp. nov. associated with bat white-nose syndrome. *Mycotaxon*, 108, 147-154.

GeoGratis (2009). Natural Resources Canada, Earth Sciences Sector. Retrieved February 21, 2010 from http://geogratis.cgdi.gc.ca/geogratis/en/index.html.

Graham, Catherine H., and Robert J. Hijmans (2006). A comparison of methods for mapping species ranges and species richness. *Global ecology and Biogeography*, 15: 578-587.

Griffin, Andrew W. (2010). Woodward County bat may have 'white-nose syndrome'. Oklahoma Watchdog. Retrieved May 25, 2010 from http://oklahoma.watchdog.org/766/woodward-county-bat-may-have-white-nose-syndrome/

Guisan, A. and U. Hofer (2003). Predicting reptile distributions at the mesoscale: relation to climate and topography. *Journal of Biogeography*, 30, 1233-1243.

Guisan, A., S.B. Weiss, and A.D. Weiss (1999). GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, 143, 107-122.

Hajek, A.E., and R.J. Leger (1994). Fungal Pathogens and Insect Hosts. Annual Review of Entomology, 39, 293-322. Retrieved February 18, 2009 from http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.en.39.010194.001453

Haneisen, R. (2009). Deadly fungus could wipe out bats. The Metrowest Daily News. Retrieved March 26, 2010 from http://www.metrowestdailynews.com/news/x187210749/Deadly-fungus-could-wipe-out-bats

Hansen, M., R.S. DeFries, J.R.G. Townshend, M. Carroll, C. Dimiceli, R.A. Sohlberg (2003). "Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm," Earth Interactions. 7(10): 1-15.

Hernandez, P.A., C.H. Graham, L.L. Master, D.L. Albert (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29: 773-785.

Hijmans, R. J. (2006). MkBCvarsAML version 2.3. www.worldclim.org/mkBCvars.aml. Viewed March 22, 2009.

Hosmer, D.W. and S. Lemeshow (2000). Applied Logistic Regression. Wiley Series in Probability and Statistics. Canada: Wiley-Interscience Publication; 2nd edition, pg. 162.

Humphrey, S.R. and J.B. Cope (1976). Population Ecology of the Little Brown Bat, *Myotis lucifugus*, in Indiana and South-Central Kentucky. *American Society of Mammalogists*, Stillwater, OK, pp. 81.

Huq, Anwarul and R.R. Colwell (1996). Environmental Factors Associated with Emergence of Disease with Special Reference to Cholera. *Eastern Mediterranean Health Journal*, 2(1), 37-45.

Irwin, H. (2004). Modeling Fire Behavior on the Landscape of the Southern Appalachians. Southern Appalachian Forest Coalition. Retrieved March 1, 2009 from http://proceedings.esri.com/library/userconf/proc04/docs/pap1538.pdf

Jaberg, C. and A. Guisan (2001). Modelling the distribution of bats in relation to landscape structure in a temperate mountain environment. *Journal of Applied* Ecology, 38(6), 1169-1181.

Jones, G., D.S. Jacobs, T.H. Kunz, M.R. Willig, P.A. Racey (2009). Carpe noctem: the importance of bats as bioindicators. *Endangered Species Research*, 8, 93-115.

Jonsson, P. and L. Eklundh (2004). TIMESAT - a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30, 833-845.

Kumar, S., Sarah A. Spaulding, Thomas J. Stohlgren, Karl A. Hermann, Travis S. Schmidt, Loren L. Bahls (2009). Potential habitat distribution for the freshwater diatom Didymosphenia geminate in the continental US. *Frontiers in Ecology and the Environment*, 7, doi: 10.1890/080054.

Kumar, S., T.J. Stohlgren, and G.W. Chong (2006). Spatial Heterogeneity Influences Native and Nonnative Plant Species Richness. *Ecology*, 87(12), 3186-3199.

Kumar, S., and T.J. Stohlgren (2009). Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. *Journal of Ecology and Natural Environment*, 1(4): 94-98.

Kunz, T.H. (Ed.) (1982). Ecology of Bats. New York, NY: Plenum Publishing Corporation.

Kunz, T.H.; C.L. Frank; E.P. Widmaier; and J.D. Reichard (2008). Death by Starvation: An Hypothesis-Based Investigation of White-Nose Syndrome in the Little Brown Myotis (Myotis lucifugus): A Collaborative Proposal Submitted to the U.S. Fish and Wildlife Service. Retrieved February 17, 2010 from https://secure.caves.org/WNS/Kunz%20Proposal.htm

Lamb, J.M., T.M.C. Ralph, S.M. Goodman, et al. (2008). Phylogeography and predicted distribution of African-Arabian and Malagasy populations of giant mastiff bats, *Otomops* spp. (Chiroptera: Molossidae). *Acta Chiropterologica*, 10(1), 21-40.

Li, G. and Q. Weng (2005). Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering & Remote Sensing*, 71(8), 947-958.

Maxent Software version 3.1. Retrieved September 28, 2009 from www.cs.princeton.edu/~schapire/maxent/

Meteyer, C.U., E.L. Buckles, D.S. Blehert, A.C. Hicks, D.E. Green, V. Shearn-Bochsler, N.J. Thomas, A. Gargas, M.J. Behr (2009). Histopathologic Criteria to Confirm White-Nose Syndrome in Bats. *Journal of Veterinary Diagnostic Investigation*, 21, 411-414.

McMichael, A.J. (2004). Environmental and social influences on emerging infectious diseases: past, present and future. *Philosophical Transactions of The Royal Society*, 359(1447), 1049-1058.

Minitab Inc. (2003). Student Version of MiniTab: Release 14. Statistical Software for Education.

Moore, I.D., R.B. Grayson and A.R. Ladson (1991). Digital Terrain Modelling: A Review of Hydrological, Geomorphological and Biological Applications. *Hydrological Processes An International Journal*, January - March, 3-30.

Morisette, J.T.; A.D. Richardson; A.K. Knapp; J.I. Fisher; E.A. Graham; J. Abatzoglou; B.E. Wilson; D.D. Breshears; G.M. Henebry; J.M. Hanes; and L. Liang (2009). Tracking the rhythm of the seasons in the face of global change: phonological research in the 21st century. *Frontiers in Ecology and the Environment*, 7(5), 253-260.

Morita, R.Y. (1975). Psychrophilic Bacteria. *American Society for Microbiology*, 39(2), 144-167.

Nagler, P; A. Jetton; J. Fleming; et al. (2007). Evapotranspiration in a cottonwood (*Poplus fremontii*) restoration plantation estimated by sap flow and remote sensing methods. *Agricultural and Forest Meteorology*, 144, 95-110.

NASA (National Aeronautics and Space Administration) (2010). MODIS for NACP, Goddard Space Flight Center. Retrieved February 8, 2010, from http://accweb.nascom.nasa.gov/

National Park Service (2010). U.S. Department of the Interior. *Ozark Scenic National Riverways*, New Release. Retrieved May 25, 2010 from http://www.nps.gov/ozar/parknews/upload/OzarkNSR-Cave-Closures-due-to-WNS.pdf

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman (1996). Applied linear statistical models: regression, analysis of variance, and experimental designs. Fourth edition. Irwin, Chicago, Illinois, USA.

Nix, H. A. (1986). A biogeographic analysis of Australian elapid snakes. In: Longmore R (Ed). Australian flora and fauna series 8. Canberra, Australia: Australian Government Publishing Service.

Nuckols, J.R., M.H. Ward, and L. Jarup (2004). Using Geographic Information Systems for Exposure Assessment in Environmental Epidemiology Studies. *Environmental Health Perspectives*, 112(9), 1007-1015.

Ortega-Huerta, M.A., and A.T. Peterson (2008). Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. Revista Mexicana De Biodiversidad 79: 205-216.

Papes, M., and P. Gaubert (2007). Modeling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity and Distributions*, 13: 890-902.

Pearson, R.G. (2007). Species' Distribution Modeling for Conservation Educators and Practitioners. *Synthesis*. American Museum of Natural History. Available at http://ncep.amnh.org

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

Phillips, S. J. (n.d.). A Brief Tutorial on Maxent. Maxent Software for Species Habitat Modeling. Retrieved September 28, 2009 from www.cs.princeton.edu/~schapire/maxent

Phillips, S.J. (2008). Transferability, sample selection bias and background data in presence-only modeling: a response to Peterson et al. (2007). *Ecography*, 31, 272-278.

Phillips, S.J., M. Dudik, J. Elith, C.H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier (2009). Samples selection bias and presence-only distribution models: implication for background and pseudo-absence data. *Ecological Application*, 19(1), 181-197.

Phillips, S.J., M. Dudik, and R.E. Schapire (2004). A Maximum Entropy Approach to Species Distribution Modeling. *ACM International Conference Proceeding Series,* 69, 83-90.

Phillips, S. J., R.P. Anderson, and R.E. Schapire (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modeling*, 190, 231-259.

Plonsky, M. (2009). Psychological Statistics: Summary of Statistical Tests. Retrieved April 2, 2010 from http://www.uwsp.edu/psych/Stat/indexTests.htm

Puechmaille, S.J., P. Verdeyroux, H. Fuller, M.A. Gouilh, M. Bekaert, and E.C. Teeling (2010). White-Nose Syndrome Fungus (*Geomyces destructans*) in Bat, France. *Emerging Infectious Diseases*, 16(2), 290-293.

Sattler, T., Bontadina, F., Hirzel, A. H., and Arlettaz, R. 2007. Ecological niche modelling of two cryptic bat species calls for a reassessment of their conservation status. *Journal of Applied Ecology*, 44, 1188–1199.

Shore, J.E. and R.W. Johnson (1980). Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy. *IEEE Transactions on Information Theory*, 26(1), 26-37. IEEE Xplore Ditigal Library. Retrieved October 8, 2009 from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1056144&isnumber=22712

Skerratt, L.F., L. Berger, R. Speare, et al. (2007). Spread of chytridiomycosis has caused the rapid globaldecline and extinction of frogs. *EcoHealth,* 4, 125-134.

Stata/SE 10.1 (2007). Statistical Software. StataCorp LP, College Station, TX. Version 10.1 for Windows.

StatSoft Electronic Statistics Textbook (n.d.). Nonparametric Statistics. Retrieved March 23, 2010 from http://www.statsoft.com/textbook/nonparametric-statistics/

Strickland, E. (2008). Bats are Dying from White Nose Mold, But Researchers Don't Know Why. *Discover*. Retrieved March 26, 2010 from http://blogs.discovermagazine.com/80beats/2008/ 10/30/bats-are-dying-from-white-nose-mold-but-researchers-dont-know-why/

Sukumar, N. (2008). Maximum-entropy basis functions. Information-Theoretic Approximation Schemes in Mechanics Computations. University of California, Davis. Retrieved October 8, 2009 from http://dilbert.engr.ucdavis.edu/~suku/maxent

SYSTAT Software, version 13 (2010). Statistical and Graphical Software. Retrieved March 3, 2010 from http://www.systat.com/

Tan, B., J.T. Morisette, R.E. Wolfe, F. Gao, G.A. Ederer; J. Nightingale, and J.A. Pedelty (2008). Vegetation Phenology Metrics Derived From Temporally Smoothed and Gap-Filled MODIS Data. *International Geoscience and Remote Sensing Symposium*, 3, 593-596.

Thomas, D.W., M.B. Fenton, R.M.R. Barclay (1979). Social Behavior of the Little Brown Bat, *Myotis lucifugus*: I. Mating Behavior. *Behavioral Ecology and Sociobiology*, 6(2), 129-136.

Two-Sample T-Tests Using Minitab (2010). Minitab Technical Support Document. Retrieved May 21, 2010 from http://www.minitab.com/support/documentation/Answers/BowerTwoSamplet.pdf

United States Fish and Wildlife Service (USFWS) (2010). Conserving the Nature of America. Retrieved February 21, 2010 from http://www.fws.gov/

United States Geological Survey (USGS) (2009). HYDRO-1k Elevation Derivative Database, Earth Resources Observation and Science (EROS) Center. Retrieved March 1, 2009 from http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro

United Sates Geological Survey – North American Political Boundaries, North America 2006 (2008). Government of Canada, Natural Resources Canada, Canada Centre for Remote Sensing, The Atlas of Canada. Retrieved February 8, 2008, from http://finder.geocommons.com/overlays/2264

United States Geological Survey (USGS) (2009). Investigating White-Nose Syndrome in Bats. Fact Sheet 2009-3058. Retrieved September 28, 2009 from http://www.nwhc.usgs.gov/publications/fact_sheets/pdfs/2009-3058_investigating_wns.pdf

United States Geological Survey (USGS) (2009). White-Nose Syndrome (WNS). *National Wildlife Health Center*. Retrieved February 18, 2009 from http://www.nwhc.usgs.gov/disease_information/whitenose_syndrome/index.jsp

United States Geological Survey (USGS) (2010). Multi-Resolution Land Characteristics Consortium, National Land Cover Database. Retrieved March 10, 2010 from http://www.mrlc.gov/

United States Geological Survey (USGS) (2010). The National Map Seamless Server. Retrieved December 18, 2009 from http://seamless.usgs.gov/

Veilleux, J.P. (2008). Current Status of White-Nose Syndrome in the Northeastern United States. *Bat Research News*, 49, 15-17.

Veloz, S.D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, 36, 2290-2299.

Weier, J. and D. Herring (2010). Measuring Vegetation (NDVI & EVI). National Aeronautics and Space Administration (NASA), Earth Observatory. Retrieved February 10, 2010, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/printall.php

Wilcox, B.A. and D.J. Gubler (2009). Environmental change and new infectious diseases. United Nations Environment Programme. Retrieved July 10, 2010 from http://www.grida.no/publications/et/ep4/page.aspx

Wisz, M.S., R.J. Hijmans, J. Li, A.T. Peterson, C.H. Graham, A. Guisan, NCEAS Predicting Species Distributions Working Group (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14: 763-773.

WNS Science Strategy Group (2008). Questions, observations, hypotheses, predictions, and research needs for addressing effects of white-nose syndrome (WNS) in hibernating bats. Retrieved February 17, 2010 from http://batcon.org/pdfs/WNSMtgRptFinal2.pdf

WorldClim (n.d.). Global Climate Data. Retrieved January 7, 2010 from http://seamless.usgs.gov/

Yates, D., Moore, D., McCabe, G. (1999). The Practice of Statistics (1st Ed.). New York: W.H. Freeman.

Zimmerman, R. (2009). Biologists Struggle to Solve Bat Deaths. *Science*, 324, 1134-1135.

APPENDICIES

Appendix 1: Input Environmental Variables for WNS Maxent Modeling		
Variable ID #	**Environmental Variable**	**Description**
1	[1,2]BIO-1	Annual mean temperature (°C)
2	[1,2]BIO-2	Mean diurnal range in temperature (Mean of monthly (max temp – min temp)) (°C)
3	[1,2]BIO-3	Isothermality (BIO-2 / BIO-7) (*100)
4	[1,2]BIO-4	Temperature seasonality (Standard Deviation*100)
5	[1,2]BIO-5	Maximum temperature of warmest month (°C)
6	[1,2]BIO-6	Minimum temperature of coldest month (°C)
7	[1,2]BIO-7	Temperature annual range (°C) (BIO-5 – BIO-6)
8	[1,2]BIO-8	Mean temperature of wettest quarter (°C)
9	[1,2]BIO-9	Mean temperature of driest quarter (°C)
10	[1,2]BIO-10	Mean temperature of warmest quarter (°C)
11	[1,2]BIO-11	Mean temperature of coldest quarter (°C)
12	[1,2]BIO-12	Mean annual precipitation (cm)
13	[1,2]BIO-13	Precipitation of wettest month (cm)
14	[1,2]BIO-14	Precipitation of driest month (cm)
15	[1,2]BIO-15	Precipitation seasonality (Coefficient of Variation)

16	[1,2]BIO-16	Precipitation of wettest quarter (cm)
17	[1,2]BIO-17	Precipitation of driest quarter (cm)
18	[1,2]BIO-18	Precipitation of warmest quarter (cm)
19	[1,2]BIO-19	Precipitation of coldest quarter (cm)
20	[3]NDVI-1	NDVI phenology metric for beginning of season
21	[3]NDVI-2	NDVI phenology metric for end of season
22	[3]NDVI-3	NDVI phenology metric for length of season
23	[3]NDVI-4	NDVI phenology metric for base value (average of NDVI-1 and NDVI-2 minimum values)
24	[3]NDVI-5	NDVI phenology metric for peak time (middle of season)
25	[3]NDVI-6	NDVI phenology metric for peak value (of the fitted function)
26	[3]NDVI-7	NDVI phenology metric for amplitude (difference between peak and base values)
27	[3]NDVI-8	NDVI phenology metric for the left derivative (rate of increase at the beginning of the season)
28	[3]NDVI-9	NDVI phenology metric for the right derivative (rate of increase at the beginning of the season)
29	[3]NDVI-10	NDVI phenology metric for integral over season – absolute (large seasonal integral, from season start to end)
30	[3]NDVI-11	NDVI phenology metric for integral over season – scaled (small seasonal integral, relative to the base value)
31	[3]NDVI-12	NDVI phenology metric for maximum value
32	[3]NDVI-13	NDVI phenology metric minimum value
33	[3]NDVI-14	NDVI phenology metric mean value
34	[3]EVI-1	EVI phenology metric for beginning of season
35	[3]EVI-2	EVI phenology metric for end of season

36	[3]EVI-3	EVI phenology metric for length of season
37	[3]EVI-4	EVI phenology metric for base value (average of NDVI-1 and NDVI-2 minimum values)
38	[3]EVI-5	EVI phenology metric for peak time (middle of season)
39	[3]EVI-6	EVI phenology metric for peak value (of the fitted function)
40	[3]EVI-7	EVI phenology metric for amplitude (difference between peak and base values)
41	[3]EVI-8	EVI phenology metric for the left derivative (rate of increase at the beginning of the season)
42	[3]EVI-9	EVI phenology metric for the right derivative (rate of increase at the beginning of the season)
43	[3]EVI-10	EVI phenology metric for integral over season – absolute (large seasonal integral, from season start to end)
44	[3]EVI-11	EVI phenology metric for integral over season – scaled (small seasonal integral, relative to the base value)
45	[3]EVI-12	EVI phenology metric for maximum value
46	[3]EVI-13	EVI phenology metric minimum value
47	[3]EVI-14	EVI phenology metric mean value
48	[5]NED	National elevation dataset
49	[5]Slope	Slope of NED (in degrees)
50	[5]Aspect	Aspect of NED (in degrees)
51	[5,6]Dist_Water	Distance to water raster grid (water = streams and waterbodies)
52	[7]ComTopInd	Compound topographic index (wetness index)
53	[5]Veg_Types	National land-cover dataset (NLCD) (land-cover types)
54	[1]Grow_days	Growing degree days (degree-days)
55	[1]Frost_days	Frost days (days)

56	[1]FreqPrecip	Frequency of precipitation (number of wet days/total days)
57	[1]Humidity	Humidity (Pa)
58	[1]PrecipSize	Annual precipitation event size (cm per day)

Notes:
BIO = the "bioclim" variable (Nix 1986; www.worldclim.org/bioclim) that was calculated via ARC AML script (Hijmans 2006) and the Daymet climate dataset.
NDVI = Normalized Difference Vegetation Index from MODIS
EVI = Enhanced Vegetation Index from MODIS
MODIS = Moderate-Resolution Imaging Spectroradiometer
CTI = function of slope and upstream contributing area per unit width orthogonal to the flow direction (Irwin 2004)

Data Sources:
[1] Daymet: www.daymet.org/
[2] Worldclim: www.worldclim.org/bioclim
[3] MODIS Vegetation Indices: http://accweb.nascom.nasa.gov/
[4] National Land Cover Dataset (2001): http://www.mrlc.gov/index.php
[5] The National Map Seamless Server; www.seamless.usgs.gov
[6] GeoGratis; http://geogratis.cgdi.gc.ca/geogratis/en/download/northamerica.html
[7] USGS Hydro 1k Elevation Derivative Database; http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro

Appendix 2: NLCD 2001 Land-Cover Class Definitions

11. Open Water - All areas of open water, generally with less than 25% cover of vegetation or soil.

12. Perennial Ice/Snow - All areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.

21. Developed, Open Space - Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes

22. Developed, Low Intensity - Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.

23. Developed, Medium Intensity - Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50-79 percent of the total cover. These areas most commonly include single-family housing units.

24. Developed, High Intensity - Includes highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80 to100 percent of the total cover.

31. Barren Land (Rock/Sand/Clay) - Barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.

32. Unconsolidated Shore* - Unconsolidated material such as silt, sand, or gravel that is subject to inundation and redistribution due to the action of water. Characterized by substrates lacking vegetation except for pioneering plants that become established during brief periods when growing conditions are favorable. Erosion and deposition by waves and currents produce a number of landforms representing this class.

41. Deciduous Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change.

42. Evergreen Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species maintain their leaves all year. Canopy is never without green foliage.

43. Mixed Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75 percent of total tree cover.

51. Dwarf Scrub - Alaska only areas dominated by shrubs less than 20 centimeters tall with shrub canopy typically greater than 20% of total vegetation. This type is often co-associated with grasses, sedges, herbs, and non-vascular vegetation.

52. Shrub/Scrub - Areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.

71. Grassland/Herbaceous - Areas dominated by grammanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.

72. Sedge/Herbaceous - Alaska only areas dominated by sedges and forbs, generally greater than 80% of total vegetation. This type can occur with significant other grasses or other grass like plants, and includes sedge tundra, and sedge tussock tundra.

73. Lichens - Alaska only areas dominated by fruticose or foliose lichens generally greater than 80% of total vegetation.

74. Moss - Alaska only areas dominated by mosses, generally greater than 80% of total vegetation.

81. Pasture/Hay - Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation.

82. Cultivated Crops - Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.

90. Woody Wetlands - Areas where forest or shrubland vegetation accounts for greater than 20 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

91. Palustrine Forested Wetland* -Includes all tidal and non-tidal wetlands dominated by woody vegetation greater than or equal to 5 meters in height and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is below 0.5 percent. Total vegetation coverage is greater than 20 percent.

92. Palustrine Scrub/Shrub Wetland* - Includes all tidal and non-tidal wetlands dominated by woody vegetation less than 5 meters in height, and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is below 0.5 percent. Total vegetation coverage is greater than 20 percent. The species present could be true shrubs, young trees and shrubs or trees that are small or stunted due to environmental conditions.

93. Estuarine Forested Wetland* - Includes all tidal wetlands dominated by woody vegetation greater than or equal to 5 meters in height, and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is equal to or greater than 0.5 percent. Total vegetation coverage is greater than 20 percent.

94. Estuarine Scrub/Shrub Wetland* - Includes all tidal wetlands dominated by woody vegetation less than 5 meters in height, and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is equal to or greater than 0.5 percent. Total vegetation coverage is greater than 20 percent.

95. Emergent Herbaceous Wetlands - Areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

96. Palustrine Emergent Wetland (Persistent)* - Includes all tidal and non-tidal wetlands dominated by persistent emergent vascular plants, emergent mosses or lichens, and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is below 0.5 percent. Plants generally remain standing until the next growing season.

97. Estuarine Emergent Wetland* - Includes all tidal wetlands dominated by erect, rooted, herbaceous hydrophytes (excluding mosses and lichens) and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is equal to or greater than 0.5 percent and that are present for most of the growing season in most years. Perennial plants usually dominate these wetlands.

98. Palustrine Aquatic Bed* - The Palustrine Aquatic Bed class includes tidal and nontidal wetlands and deepwater habitats in which salinity due to ocean-derived salts is below 0.5 percent and which are dominated by plants that grow and form a continuous cover principally on or at the surface of the water. These include algal mats, detached floating mats, and rooted vascular plant assemblages.

99. Estuarine Aquatic Bed* - Includes tidal wetlands and deepwater habitats in which salinity due to ocean-derived salts is equal to or greater than 0.5 percent and which are dominated by plants that grow and form a continuous cover principally on or at the surface of the water. These include algal mats, kelp beds, and rooted vascular plant assemblages.

* Coastal NLCD class only