



2010-08-27

# Predicting the Distribution of Air Pollution Sensitive Lichens Using Habitat Niche Modeling

Gajendra Shrestha

*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Biology Commons](#)

---

## BYU ScholarsArchive Citation

Shrestha, Gajendra, "Predicting the Distribution of Air Pollution Sensitive Lichens Using Habitat Niche Modeling" (2010). *All Theses and Dissertations*. 2595.

<https://scholarsarchive.byu.edu/etd/2595>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

Predicting the Distribution of Air Pollution Sensitive Lichens Using  
Habitat Niche Modeling

Gajendra Shrestha

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Larry L. St. Clair, Chair  
Steven L. Petersen  
Samuel B. St. Clair

Department of Biology  
Brigham Young University

December 2010

Copyright © 2010 Gajendra Shrestha

All Rights Reserved

## ABSTRACT

### Predicting the Distribution of Air Pollution Sensitive Lichens Using Habitat Niche Modeling

Gajendra Shrestha

Department of Biology

Master of Science

*Usnea hirta* and *Xanthoparmelia cumberlandia* are commonly used as bio-monitors of air quality. In order to more accurately and efficiently determine the distribution of these two sensitive indicator species, we have developed a probabilistic distribution map as a function of 9 macroclimatic and topographic variables for the White River National Forest, Colorado using Non-Parametric Multiplicative Regression (NPMR) analysis. Furthermore, we also developed a logistic regression (LR) model for *X. cumberlandia* in order to evaluate the strengths and limitations of the NPMR model.

The best model for *U. hirta* included four variables – solar radiation, average monthly precipitation, average monthly minimum and maximum temperature ( $\log \beta = 3.68$ ). The presence rate for *U. hirta* based on field validated test sites was 45.5%, 65.4%, and 70.4% for low, medium, and high probability areas, respectively. The best model for *X. cumberlandia* generated by both NPMR and LR involved the same variables - solar radiation, average monthly maximum temperature, average monthly precipitation, and elevation as the best predictor variables ( $\log \beta = 5.10$ ). The occurrence rate for *X. cumberlandia* using the NPMR model was 32%, 44.4%, and 20% for the low, medium, and high probability areas respectively while the LR model had 26%, 50%, and 38% for low, medium and high probability areas respectively. Although the LR model predicted a smaller high probability area compared to the NPMR model there was substantial overlap between the two.

The *U. hirta* model performed better than the *X. cumberlandia* model. The reduced performance of our model especially for *X. cumberlandia* may be due in part to the absence of field measured data in the development of the model. Our study also suggested that the northeast and western part of the forest should be preferentially considered for establishing future air quality bio-monitoring reference sites. Finally, in the future a well defined sampling design with sufficient sampling sites, field measured predictor variables, and microclimatic data should be used in the development of predictive models.

Keywords: habitat niche model, *Usnea hirta*, *Xanthoparmelia cumberlandia*, NPMR, White River National Forest, GIS

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Larry L. St. Clair, for allowing me freedom to explore my own ideas and same time giving appropriate suggestions to complete my research. I am also thankful to my committee members Dr. Steven L. Petersen and Dr. Samuel B. St. Clair for their constant guidance during my research and manuscript preparation. I also appreciate my colleagues Steven D. Leavitt and Murari Paudel for their help during my research and field works. My sincere thanks also go to Monte L. Bean Life Science Museum for providing research facilities.

I especially thank my wife and son for their constant support and encouragement without which I am not able to complete my research.

## TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iii
Table of contents.....	iv
List of Tables .....	v
List of Figures.....	vi
Introduction.....	1
Materials and Method .....	7
Study Sites .....	7
Data Collection .....	9
Response variables .....	9
Predictor variables .....	10
Data Preparation.....	12
Analysis.....	12
Modeling Strategy.....	13
NPMR.....	13
Logistic Regression .....	14
Model Validation .....	16
Results.....	16
NPMR model for <i>Usnea hirta</i> .....	16
NPMR model for <i>Xanthoparmelia cumberlandia</i> .....	18
Logistic Regression model for <i>Xanthoparmelia cumberlandia</i> .....	20
Discussion.....	21
Literature Cited.....	27

## LIST OF TABLES

Table 1: List of Candidate Models.....	15
Table 2: Probability Model selected as best modeling approach for <i>Usnea hirta</i> with different average neighborhood sizes .....	17
Table 3: Probability Model selected as best modeling approach for <i>Xanthoparmelia cumberlandia</i> different average neighborhood sizes .....	19
Table 4: Candidate models with weighted AICc values.....	21
Table 5: Logistic regression models with parameter estimates .....	21

## LIST OF FIGURES

Figure 1: Distribution map of <i>U. hirta</i> for the study area with field verification sites.....	33
Figure 2: Distribution map of <i>U. hirta</i> for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence.....	34
Figure 3: NPMR Model of the distribution map of <i>X.cumberlandia</i> for the study area with field verification sites.....	35
Figure 4: NPMR model of the distribution map of <i>X. cumberlandia</i> for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence.....	36
Figure 5: Logistic Regression Model of the distribution map of <i>X. cumberlandia</i> for the study area with field verification sites.....	37
Figure 6: Logistic Regression model of the distribution map of <i>X. cumberlandia</i> for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence respectively.....	38

## INTRODUCTION

The lichenized condition is a symbiotic association consisting of a fungus (the mycobiont) and an alga and/or a cyanobacterium (the photobiont). These symbiotic systems are common within the ascomycota and are distributed from tropical rainforests to the Arctic tundra. Lichens can be important components in some ecosystems due to the variety of important ecological roles they play. For example, cyanolichens (lichens which contain a cyanobacterium as the photobiont) are important nitrogen fixers in nitrogen-poor systems (Hobara, *et al.*, 2006). Lichens also play a significant role in the recycling of mineral nutrients (Pike, 1978; Shrestha & Baniya, 2006) as well as serving as a significant food source for many animals such as caribou (Purvis, 2000), snails (Baur *et al.*, 1994), and flying fox (Rosentreter *et al.*, 1997).

According to Ferry, *et al.* (1973) lichens are the most thoroughly studied bio-indicators of air quality. The exceptional capacity of lichens to accumulate atmospheric outwash has made them particularly useful as bio-monitors of air quality. Unlike vascular plants, lichens depend directly on atmospheric sources of water and nutrients as they lack roots and a vascular system as well as stomata and a protective cuticle layer. Therefore, air pollution-rich aerosols are effectively absorbed over the entire surface of the lichen (Nash, 2008). Furthermore, lichens are perennial but because they lack deciduous parts air pollutants readily accumulate in lichen thallus and remains there. In the last 50 years researchers have repeatedly demonstrated the feasibility and versatility of using lichens as air quality bio-monitors ((Brodo, 1961; Fuga, *et al.*, 2008; Gilbert, 1973; Jozwiak, 2009; Loppi, 1996; Nobel, *et al.*, 2008; St. Clair, *et al.* 2002a St. Clair, *et al.*, 2002b).

Historically, the Intermountain region of the western United States has supported a wide range of mineral and fossil fuel extraction and processing industries. However, due to increased



pollution-related concerns, some of these activities have been reduced. In recent years, the USDA Forest Service has been required by federal law to monitor forest health. In order to more efficiently and effectively meet this responsibility, prior knowledge of the distribution of pollution sensitive bio-monitors such as lichens has been essential to documenting various forest health issues. Effective and efficient selection of air quality bio-monitoring reference sites with good evidence supporting the occurrence of potential air pollution sensitive indicator species is an important component of any air quality bio-monitoring program. However, accurate and consistent information about the distribution and occurrence of sensitive indicator species is often problematic. Typically, lichen air quality bio-monitoring reference sites are established with very little, if any, prior information about the occurrence of sensitive indicator species. Launching detailed floristic study before establishing bio-monitoring sites is not the most efficient way to document the occurrence of indicator species because of the time and costs involved. One potential solution for overcoming this problem is to develop predictive habitat models that generate potential spatial distributions and probable locations of air pollution sensitive lichen species. Habitat model describes how variation in species performance relates to predictor variables and defines important factors underlying a species' distribution (McCune, 2006). Although there are some limitations with species distribution modeling, this is an important method for generating and testing hypotheses about landscape-scale spatial distributions of species. Habitat modeling provides land managers and ecologists with important information about species' distribution and abundance, thus supporting effective management decisions for rare or threatened species (McCune, 2006). Several parametric as well as non-parametric methods have been used extensively with vascular plants (Petersen & Stringham, 2008; Vogiatzakis & Griffiths, 2006) and animals (Berg, *et al.*, 2004; Bonn & Schroder, 2001; Quinlan, *et al.*, 2004). Although application of prediction modeling is relatively limited in

lichens, due in part to very specific micro-environmental niches, and sensitivity to small scale environmental variation, there is growing interest in developing distribution models of lichen. Waser, *et al.* (2004) used remote sensing data and multiple regression models to develop a model for predicting lichen species richness. Martinez, *et al.*, (2006) evaluated potential distributions for eleven threatened lichen species in Spain and modeled habitat suitability maps for each species with potentially promising results. Bolliger, *et al.*, (2007) modeled the spatial distribution of six epiphytic lichen species in Switzerland as a function of various forest types and climatic variables. Similarly, Bergamini, *et al.*, (2007) modeled the richness of microlichens by evaluating different sets of predictors including richness of the macrolichen flora, along with various field and climatic variables. By using spatial niche modeling, Radies, *et al.* (2009) was able to effectively predict the diversity of canopy lichens in the inland temperate rainforest ecosystem of British Columbia as a function of different environmental variables.

According to Austin (2002), most predictive models require *a priori* assumptions about species response to environmental variables. Therefore, in our study we used Non-parametric Multiplicative Regression Analysis (NPMR), a niche-based habitat modeling program developed by McCune (2006), to develop distribution prediction models for two lichens, *U. hirta* and *X. cumberlandia*, in the White River National Forest in western Colorado. NPMR analyzes environmental gradients, or predictor variables, against locations with known occurrence of the species of interest by using kernel functions to weigh observations multiplicatively, rather than using an additive approach typical of many models (Yost, 2008). NPMR has already been successfully used to predict distribution patterns for both plants (Casazza, *et al.*, 2008; Fenton & Bergeron, 2008; Yost, 2008) and animals (Grundel & Pavlovic, 2007; Kohler, 2007) from various locations. The simplicity of NPMR and its capacity for applying models to a GIS

framework has been very useful in developing probability maps of species occurrence which can then be used to develop predictive maps for species locations across large landscapes.

The use of NPMR in modeling various functions for lichens is also well documented in literature. Berryman & McCune (2006) modeled biomass of epiphytic lichens in the central Cascade Mountains of western Oregon in relation to topography, stand structure, and lichen community composition. Similarly, Ellis *et al.*, (2007) predicted the response of *Lecanora populicola* to climate change scenarios. They modeled both present-day and future distributions for *L. populicola*. Their results showed that there is increased likelihood of occurrence of *L. populicola* in central and north-eastern Scotland, with possible westwards range expansion. Binder & Ellis (2008) examined the response of a lichen species, *Vulpicida pinastri*, to a range of climatic variables using predictive modeling including a projected model of response for *V. pinastri*. Using NPMR Cristofolini, *et al.*, (2008) investigated the response of epiphytic lichens to air pollution, against the background of other ecological predictors in a pre-alpine heterogeneous area.

As a part of this study, we have developed predictive models using standard statistical modeling methods including logistic regression (LR) in order to evaluate the strengths and limitations of NPMR. Logistic regression predicts the probable occurrence of an event by fitting data to a logistic curve. This method represents one kind of generalized linear model used for binomial regression. Logistic regression models have been used effectively with biological systems for predicting species responses as a function of different environmental variables. Several studies attempting to predict spatial distribution patterns and suitable habitats for lichen species using logistic regression methods have been published (Ask & Nilsson, 2004; Bergamini,

et al., 2007; Berglund, *et al.*, 2009; Bjelland, 2003; Bolliger, *et al.*, 2007; McCune, *et al.*, 2003; Radies, *et al.*, 2009).

In this study, we have developed prediction models to specifically evaluate the distribution patterns of air pollution sensitive lichens using data generated in connection with air quality bio-monitoring programs established throughout the Intermountain West. According to Thuiller, *et al.*, (2004), climate is a major determinant of species distribution patterns. Key variables for the distribution of epiphytic lichens include various abiotic parameters like altitude Dietrich & Scheidegger (1997) as well as climatic variables like precipitation, maximum and minimum temperature (Ellis & Coppins, 2006). Werth, *et al.*, (2005) reported that there is a strong relationship between epiphytic macrolichens and various macroclimatic factors. Thus, the variables used to make predictive models in our study included geographical and macroclimatic variables.

*Usnea hirta* and *Xanthoparmelia cumberlandia* both belong to the lichen family Parmeliaceae. *U. hirta* is fruticose lichen, with a branched thallus and a single point of attachment. This species has a worldwide distribution (Clerc, 1997) and in North America is distributed from the Canadian boreal forest along the Rocky Mountain corridor to Mexico (Thomson, 1984). In the Intermountain West this species is commonly found on the bark and wood of various conifer species and less frequently on rocks. *X. cumberlandia* is large foliose lichen; attached to the exposed surfaces of various rock substrates. This species has also been reported from various locations in Europe and North and South America. It is one of the most common and widespread foliose lichens in North America (Brodo, *et al.*, 2001). This species is prevalent on exposed to somewhat shaded rock outcrops and boulders throughout the Intermountain West. It has also been occasionally reported from soil and/or pebbles in coastal areas (Nash, *et al.*, 2004). Although both *U. hirta* and *X. cumberlandia* are widely distributed

throughout the Intermountain region, it is important to be able to more precisely establish their location in order to more effectively establish air quality bio-monitoring programs and baselines. *Usnea hirta* and *X. cumberlandia* have been used extensively as sensitive indicator species in air pollution bio-monitoring programs throughout the world. A study carried out by Carreras, *et al.*, (1998) involved transplant studies with *Usnea* sp. to determine industrial and traffic pollution levels in the city of Córdoba, Argentina. Garty, *et al.*, (1997) in North Finland showed that *U. hirta* showed a significant capacity for accumulating K, Ca, Na, Mg, Fe, Zn, and Mn. Likewise, van Herk, *et al.*, (2003) reported that there was a strong negative correlation with nitrogen accumulation and the distribution of *U. hirta*. This study showed that the occurrence of *U. hirta* decreased with concentrations as low as  $0.3 \text{ mg N l}^{-1}$  in precipitation at 25 European ICP-IM monitoring sites. St. Clair, *et al.*, (2002a) studied the influence of growth form and substrate on elemental accumulation by using four different lichens, *U. amblyoclada*, *U. hirta*, *Flavaoparmelia caperata* and *Flavopunctelia flaventior*. Their results showed that growth form was a significant factor affecting pollutant element accumulation patterns in lichens. St. Clair, *et al.*, (2002b) reported data for the western Intermountain region documenting a significant relationship between concentrations of copper in ambient air samples and the thalli of foliose and fruticose lichens including *U. hirta* and *X. cumberlandia*. They also documented that foliose lichens are more efficient accumulators of airborne copper than fruticose lichens. Similarly, *X. cumberlandia* has been reported as sensitive indicator of sulfur dioxide in various studies (St. Clair, 2000; St. Clair *et al.*, 2000; St. Clair and Anderson, 1997).

The overall goal of our research has been to develop predictive distribution maps of pollution sensitive lichen species for the Intermountain West. In our study we specifically developed predictive distribution maps for *U. hirta* and *X. cumberlandia* using NPMR methods which we then compared to results generated using a predictive model developed using logistic

regression. In order to test our model we have identified several potential air quality bio-monitoring sites in the White River National Forest, Colorado. Ultimately, our intent has been to generate a reliable predictive tool for use by land managers and lichenologists to more effectively and efficiently locate air quality bio-monitoring reference sites containing air pollution sensitive indicator species.

## **MATERIALS AND METHODS**

### **STUDY SITES**

Research sites for this study are located in three adjacent Intermountain states - Colorado, Utah, and New Mexico. All four study sites (San Juan –Rio Grande National Forest, Colorado, Dinosaur National Monument, Utah-Colorado, Manti La Sal National Forest, Utah, and the Gila Wilderness Area in the Gila National Forest, New Mexico) have unique vegetation patterns, complemented by significant variation in geological formations along with significant summer monsoonal precipitation; resulting in a wide variety of habitat types which ultimately supports diverse and complex lichen communities.

The San Juan-Rio Grande National Forest is located in the southwest corner of Colorado with the San Juan portion of the forest on the western slope of the Continental Divide while the Rio Grande portion occurs on the eastern slope of the Divide. Portions of the San Juan-Rio Grande National Forest occur on the Colorado Plateau which is biogeographically very diverse – consisting of extensive and complex Paleozoic and Mesozoic rocks. Elevation ranges from about 1500 m. to 4400 m. The mountainous areas of southwest Colorado have short, cool summers and long, severe winters with mean annual precipitation ranging between 20 to 50 cm. The Forest has diverse vascular plant communities including, alpine tundra, riparian, Pine-Spruce-Fir,

Aspen, pinyon-juniper woodland, scrub oak, and sage-steppe-saltbush communities (Shrestha and St. Clair, 2009).

Dinosaur National Monument straddles the border of Utah and Colorado along the northern edge of the Colorado Plateau and the southeastern flank of the Uinta Mountains. Average annual precipitation is about 25 – 40 cm including both rainfall and snowfall. Elevation in the monument ranges between 1440 and 2500 m. The monument supports a diverse plant community with Douglas fir, quaking Aspen, and ponderosa pine dominating upper elevation sites while middle elevation sites support pinyon pine, Utah juniper, ephedra, and big sagebrush. Lower elevation sites are dominated by various desert shrub species including sagebrush, greasewood, and shadscale while the riparian community along the Green and Yampa rivers includes cottonwoods, boxelder Maple, willows, river birch, and snowberry.

The Manti-La Sal National Forest consists of four distinct land areas, two located in central Utah and two in the southeast corner of the state one east of Moab and the other west of Monticello. The two central Utah areas lie within the geographical region generally referred to as the Wasatch Plateau, which is part of the much larger Colorado Plateau region; while the other two areas are located in the Paradox Fold Belt of the Colorado Plateau region. Average annual precipitation for the forest ranges between 15 – 35 cm and elevation ranges between 1500 and 3900 m. Habitat types range from mesic riparian communities along canyon streams to drier upland communities including desert shrubland to pinyon-juniper woodland to mixed mountain brush to subalpine mixed conifer forests to alpine tundra.

The Gila Wilderness Area is located in the volcanic highlands of southwestern New Mexico. The wilderness area is dominated by two large calderas, the Bursum Caldera located in the western portion of the wilderness and Gila Cliff Dwellings Caldera in the eastern half of the wilderness. Average annual precipitation ranges between 10 to 30 cm and elevation ranges

between 1300 and 3300 m. Lower elevation locations in the wilderness are dominated by desert scrub (*Artemisia*, and *Yucca* spp) transitioning to pinyon-juniper woodland with increasing elevation. Mid-elevation sites are dominated by ponderosa pine transitioning to mixed conifer forests upslope. Riparian communities include several deciduous species including maple, ash, cottonwood, and Arizona walnut.

Validation of the model was undertaken in the west portion of White River National Forest, Colorado. The White River National Forest, occupies approximately 733,000 acres situated on the White River Plateau. Elevation of the forest ranges between 1780 and 3770 m with an average annual precipitation of 15 – 40 cm. The middle portion of the forest is dominated by subalpine and upper montane aspen-conifer forest with extensive alpine tundra dominating upper elevation areas. The lichen flora and general habitat diversity of the White River National Forest provides an ideal system for testing our ecological niche model based on predictor variables used during calibration of model.

## **DATA COLLECTION**

### **Response variables**

All together 72 air quality bio-monitoring reference sites; 32 in San Juan-Rio Grande National Forest, 18 in Dinosaur National Monument, 15 in Manti-La Sal National Forest, and 7 in Gila Wilderness Area were established by Larry St. Clair and colleagues between 1991 – 1999 as part of a comprehensive regional air quality bio-monitoring program . A total of 316 lichen species in 108 genera were reported from San Juan- Rio Grande National Forest (Shrestha & St Clair, 2009). In addition, 164 lichen species in 55 genera were reported for Dinosaur National Monument; with 143 species in 48 genera from the Manti-La Sal National Forest, and 142 species in 67 genera from the Gila Wilderness Area (Shrestha and St. Clair, in prep). For this study we modeled the distribution of two air pollution sensitive lichen species based on presence



or absence (not recorded) of the two study taxa, *U. hirta* and *X. cumberlandia*, at each of the 72 bio-monitoring reference sites.

### **Predictor variables**

Our predictor variables are GIS derived variables prepared using the ArcGIS 9.3 (Environmental Systems Research Institute) software package. A total of 9 predictor variables: geographical variables – elevation, aspect, and slope; macroclimatic variables –average monthly rainfall, average monthly minimum temperature, average monthly maximum temperature, integrated moisture index, and solar radiation; and land use were prepared for four general study locations.

Several studies (Dietrich & Scheidegger, 1997; Galvich *et al.*, 2005; Petersen and Stringham, 2008; Yost, 2008) have used elevation as an important variable for predicting distribution of both plants and lichens. In this study, elevation data of 1 Arc Second (30 m) for all the study areas was acquired from the National Elevation Dataset (NED), United States Geological Survey (USGS). Elevation of reference plots in the four general study locations ranged between 1490 and 3860 m.

Using the Spatial Analyst tool in ArcGIS 9.3 software (ESRI), the NED was filled to remove pits and peaks in the dataset that may cause small imperfections in the data. Raster data sets for aspect and slope were generated from the filled NED by using Spatial analyst tools in ArcGIS. Values for the slope of the reference plots ranged from 0 to 38°. Aspect was reclassified into four classes – North (315° – 45°), East (45° - 135°), South (135° – 225°), and West (225° – 315°).

Solar radiation is another important variable in predicting species response. According to Boucher & Stone, (1992) solar radiation influences lichen abundance. Similarly, Bjelland (2003) found that radiation is an important predictor for describing floristic variation in lichen

communities. Solar radiation raster data were also generated from NED data by using Spatial analysis tools in ArcGIS; default values were adopted for all parameters. According to Coxson and Stevenson (2007) maximum growth for lichens occurs between March and October so solar radiation ( $W/m^2$ ) layers for the month of March to October for the year of 1995 were generated and averaged for each of the four study areas. We also tried to average the solar radiation values over the entire study period (1991 – 1999). However, due to computational problem, taking several hours to generate the raster layer, as well as relatively constant values for each year we opted to simply calculate solar radiation value for the year of 1995 as representative of entire study period.

Moisture availability is also an important predictor variable because lichen obtains their water directly from atmosphere. Boucher & Stone, (1992) considered moisture availability as one of the most influential environmental variables affecting lichen abundance. Iverson, *et al.*, (1997) developed a raster based image called Integrated Moisture Index (IMI) for estimating relative moisture of a landscape. To produce IMI images three topographic features, hillshade, flow accumulation, and curvature were derived from the NED data in ArcGIS 9.3 using spatial analyst tools. Hillshade is an index of the drying effects of differential solar radiation due to variation in slope angle, aspect and position, and shading from adjacent cells. Curvature measures convexity and concavity across landscape. Flow accumulation is a count of the number of cells that would contribute water downslope to an evaluation cell (Yost, 2006). All three raster layers were rescaled to 0 to 100 range and finally combined to a single layer using the equation in the raster calculator function of Spatial Analyst in ArcGIS:

$$IMI = (\text{hillshade} \times 0.5) + (\text{curvature} \times 0.15) + (\text{flow accumulation} \times 0.35)$$

Integrated Moisture Index ranged from 0 to 100; with 0 indicating no moisture accumulation and retention and 100 indicating the highest moisture accumulation and retention.

Land use data was also obtained from the United States Geological Survey (USGS) website.

Average monthly maximum temperature, average monthly minimum temperature, and average monthly precipitation data at 1-km resolution for the months of March through October of the climatological period 1971 – 2000 were obtained from Parameter-Elevation Regressions on Independent Slopes Model (PRISM; PRISM Climate Group, Oregon State University 2008). All raster datasets used in the analysis were re-sampled to a 30-m resolution. For this study, the average value of monthly minimum temperature and maximum temperature, and precipitation for the period of March to October was used for all analyses as lichen attains maximum growth during this period (Coxson and Stevenson, 2007).

### **DATA PREPARATION**

GIS layers representing each study area were generated and the GPS coordinates for all sites with their response variables (presence and absence of *U. hirta* and *X. cumberlandia*) were plotted in their respective layer. A 55 meter circular buffer was created around each reference point to roughly represent about 1 hectare of sample area for the response variables. Then predictor variable statistics for each response variable location was calculated using the Zonal Statistics function in ArcGIS 9.3.

### **ANALYSIS**

Non-Parametric multiplicative regression analysis (NPMR) was used to analyze and develop predictive models using the Hyperniche v1 software developed by (McCune & Mefford, 2004). According to Yost (2008) NPMR parsimoniously produces probability of occurrence for a given species by modeling response of a species multiplicatively to the complex interactions among several ecological factors. NPMR, like linear regression, seeks to establish relationships between a response variable and one or more predictors, but it differs from linear regression in

that it seeks an optimal fit to the data without reference to a specific model form (McCune, 2006). The model obtained from NPMR has advantages over other parameterized modeling techniques such as linear regression because in parametric regression it is difficult to fit the complex responses resulting from multiple factors and their interaction (Yost, 2006). NPMR uses a kernel weighting function to select the best predictor variables based on their relationship to species occurrence or abundance. Habitat modeling using NPMR generally has two phases, calibration and application. In the calibration phase, empirical data on species abundance or species presence-absence are used to evaluate the model's ability to estimate abundance or likelihood of occurrence from the independent variables. So, in the calibration phase, we determine the predictor variables that should be used, select smoothing parameters (tolerance) for quantitative variables, and rank each model base on descriptive statistics. In the application phase, we predict species occurrence or abundance in unknown sites based on the values of the predictor selected in calibration phase (Davis, 2009; McCune, 2006).

## **MODELING STRATEGY**

### **NPMR**

Predictive response models were constructed using a local mean (LM) with Gaussian weighting. This is a stepwise free search function which seeks a range of models with different combination of predictors. Model quality is assessed by the descriptive statistic  $\log \beta$ , which is the log likelihood ratio for two competing models (McCune, 2006). According to Yost (2008)  $\log \beta$  is calculated as the  $\log_{10}$  of the ratio between the developing NPMR model, or “posterior” model and the naïve or prior model. This prior model is simply the average frequency of the species in the full data set. A variable was retained only when its inclusion increased the  $\log \beta$  value by at least 5% over the next largest model. To compare whether the default minimum average neighborhood size (3.6 which is 5% of total sampling units) was appropriate for the

study we also developed another model with a neighborhood size of 10 (14% of the sample size). The average neighborhood size is simply the sum of neighborhood sizes for all plots divided by the total number of plots; this value decreases with the increase in the number of variables Yost (2006). After the best model was selected, it was then fine tuned, which involved adjusting the continuous variable tolerances by 1% of the variable's value range (Yost, 2008). The model with the highest log  $\beta$  value is considered the best model. According to Yost (2008), models with log  $\beta$  values higher than 2.0 provide decisive evidence against the naïve model. We also evaluated our model on the basis of percentage improvement over the naïve model. The percent improvement is calculated dividing the sum of the number of plots with improvements by the total number of plots. Here improvement represents those plots having estimates higher than naïve model when species are present in that plot and lower than the naïve model when species are absent (Yost, 2008). The Model was further evaluated with a Monte Carlo permutation test, by comparing the estimation of the response variable done by the selected models with an average estimation calculation by N random permutations among the data sets.

By using the GIS function in Hyperniche, predictive estimation maps were created for the models with the highest log  $\beta$  values with a minimum neighborhood size of 1.

### **Logistic Regression**

In order to compare the capabilities of NPMR, models were also developed using standard parametric statistical modeling tools. We used logistic regression with logit link function for modeling distribution of *X. cumberlandia*. We tested ten combinations of predictor variables that served as a plausible explanatory hypothesis for the distribution of *X. cumberlandia* in the White River National Forest (Table 1).

Table 1: List of Candidate Models

Model #	Candidate Models
Model 1	Slope + Elevation + IMI +Land use+ Aspect +Tem (min) +Tem (max) + Precipitation +Solar Radiation
Model 2	Precipitation + Elevation+ Tem (max)
Model 3	Precipitation +Solar Radiation +Elevation +Tem (max)
Model 4	IMI + Elevation + Slope + Aspect
Model 5	Precipitation + Tem (min) + Tem (max) + Solar Radiation
Model 6	Slope + Aspect + Land use
Model 7	Precipitation + Tem (max) + Tem (min)
Model 8	Elevation +IMI + Tem (max) +Tem (min)
Model 9	Solar Radiation + IMI + Tem (max) +Tem (min)
Model 10	Elevation + Land use + Solar Radiation + Precipitation

IMI = Integrated Moisture Index; Tem (min) = Average monthly minimum temperature; Tem (max) = Average monthly maximum temperature

The most parsimonious logistic regression model was selected by using Akaike’s Information Criterion with a correction for small sample size (AIC<sub>c</sub>). We calculated the AIC<sub>c</sub> weights (AIC<sub>c</sub>w) and interpreted this value as the approximate probability that the model with the largest value was the most parsimonious of the set (Johnson & Omland, 2004). Since our results received support from more than one model we did model averaging to find the best coefficient for each variable. Model averaging reduces model selection bias and accounts for model selection uncertainty (Johnson & Omland, 2004). We used statistical program R to perform logistic regression analysis.

The probability map was generated in ArcGIS by using the raster calculator function based on the following equation used to calculate the probability of species occurrence:

$$\frac{1}{(1 + (Exp (-(\beta_0 + X_1\beta_1 \dots \dots \dots X_n\beta_n))))}$$

Where  $\beta_0$  = intercept  
 $B_{(1\dots n)}$  = predictor variables  
 $X_{(1\dots n)}$  = coefficient of variables

Some of the predictor variables used in the analysis showed strong correlation with other variables. The problem of collinearity is serious if our main focus is on the regression coefficient

used to identify important and specific combinations of the variables (Yost, 2006). But when the study goal is to simply predict a response variable out of a set of predictor variables, the problem of multicollinearity is not a problem and the predictions will still be accurate (Motulsky, 2002). As we are only concerned with predicting response variables as a function of a set of predictor variables, we negated the problem of multicollinearity.

## **MODEL VALIDATION**

The models generated in Hyperniche and Logistic Regression were tested for the western portion of White River National Forest, Colorado. The model generated a predictive map with a range of values related to the probability of occurrence of *U. hirta* and *X. cumberlandia*. The values were first normalized in the range of 0-1 and then reclassified into ten probability regions using the natural breaks function in ArcMap. The 10 probability regions were again classified in three regions – a lower value range of 0 – 0.35, a middle range of 0.35 – 0.65 and a higher range of 0.65 to 1.00. Ninety points were randomly selected for the validation of model for each species. All accessible field validation points were visited and at each point we searched approximately 100 m. radius area for the presence of both species. Finally, the percentage of sites with *U. hirta*, and *X. cumberlandia* were calculated for each probabilities class.

## **RESULTS**

### **NPMR MODEL FOR *USNEA HIRTA***

*U. hirta* was reported for 30 of the 72 total sites. It was widely distributed in the San Juan-Rio Grande National Forest and the Gila Wilderness Area but poorly represented in Dinosaur National Monument and the Manti La Sal National Forest.

An optimum non-parametric multiplicative model having the highest log  $\beta$  value (3.62) was selected from 1282 competing models generated in a stepwise free search when default average neighborhood size of 3.6 was set (Table 2). The best model had 4 variables – average

monthly minimum temperature, average monthly maximum temperature, average monthly precipitation, and average monthly solar radiation. The best model was then fine tuned which increased the  $\log \beta$  value to 3.68. Among the four best predictors, average monthly maximum temperature proved to be the most important predictor in this model (sensitivity 0.71).

Table 2: Probability Model selected as best modeling approach for *Usnea hirta* with different average neighborhood sizes

Number of Variables	Log $\beta$	Variables and Tolerances			
Average Neighborhood Size = 3.6					
4	3.62	5.20-Tmin	1.85-Tmax	1.16-Precip	20813.30-SR
Fine Tuned	3.68	4.68-Tmin	2.04-Tmax	1.08-Precip	19425.75-SR
Average Neighborhood Size = 10					
4	3.23	9.75-Tmin	2.77-Tmax	1.16-Precip	17344.42-SR
Fine Tuned	3.25	9.23-Tmin	2.59-Tmax	1.16-Precip	19425.75-SR

Tmin = Average Monthly minimum Temperature; Tmax = Average Monthly Maximum Temperature; Precip = Average Monthly Precipitation; SR = Average Monthly Solar Radiation

For the 30 sites where *U. hirta* occurred, 20 probability estimates showed improvements in excess of the naïve model. Similarly, for the plots where *U. hirta* was absent, 29 had estimates lower than the naïve model. Overall the improvement was 68.1%. Our best model further resulted statistically significant ( $p < 0.01$ ) when evaluated with a Monte Carlo permutation test (1000 randomized runs).

The same four variables – average monthly minimum temperature, average monthly maximum temperature, average monthly precipitation, and average monthly solar radiation were selected for the best model ( $\log \beta = 3.23$ ) with the average neighborhood size value set at 10 (Table 2). In this model, average monthly precipitation was relatively more important than the other predictor variables (Sensitivity = 0.64). When the model was fine tuned there was only a 0.02 unit increase in the  $\log \beta$  value (3.25). Of the 30 sites where *U. hirta* occurred 21 response



variables sites showed improvement over the naïve model while 29 sites out of the remaining 42 sites where *U. hirta* did not occur showed improvements. The overall improvement percentage over the naïve model was 69.4%. This model was also statistically significant ( $p < 0.05$ ) when evaluated with a Monte Carlo permutation test (1000 randomized runs).

Although the model with an average neighborhood size of 10 had the higher overall improvement over the model with average neighborhood size of 3.6 we selected a model with average neighborhood size of 3.6 for generating the predictive maps because of a higher log  $\beta$  value.

The predictive map for *U. hirta* in White River National Forest is shown in Fig. 1. The map identifies 96% of the forest as above a minimum neighborhood size of 1.

Out of the possible 90 points we visited 64 (11 low, 26 medium and 27 high). In the low probability area *U. hirta* was present at 5 sites (45.5%); with *U. hirta* specimens present in 17 (65.4%) and 19 (70.4%) of the medium and high probability sites respectively (Fig 2).

#### **NPMR MODEL FOR *XANTHOPARMELIA CUMBERLANDIA***

*X. cumberlandia* was reported for 36 of 72 total sites and similar to *U. hirta* was more common in the San Juan – Rio Grande National Forest and Gila Wilderness Area and much less common in the other two study areas.

For *X. cumberlandia* the best model was selected from 1003 competitive models when the default average neighborhood value was set at 3.6 with a log  $\beta$  value of 4.79. This model identified elevation, average monthly maximum temperature, average monthly precipitation, and average monthly solar radiation as the best predictor variables (Table 3). After fine tuning, the log  $\beta$  value increased to 5.1. Among the variables, average monthly precipitation (sensitivity = 2.84) was relatively more important than the other variables.

Table 3: Probability Model selected as best modeling approach for *Xanthoparmelia cumberlandia* different average neighborhood sizes

Number of Variables	Log $\beta$	Variables and Tolerances			
Average Neighborhood Size = 3.6					
4	4.79	1063.35-Ele	3.70-Tmax	0.39-Precip	13875.54-SR
Fine Tune	5.10	1157.87-Ele	3.70-Tmax	0.31-Precip	16650.64-SR
Average Neighborhood Size = 10					
2	2.90	1654.1- Ele	0.39-Precip		
Fine Tune	2.92	1559.59-Ele	0.39-Precip		

Ele = Elevation; Tmax = Average Monthly Maximum Temperature; Precip = Average Monthly Precipitation; SR = Average Monthly Solar Radiation

In this model, 5 sites were excluded from the model as they had smaller average neighborhood sizes. Of the 32 sites, where *X. cumberlandia* was present, 18 showed improvement over the naïve model; among the 35 sites where *X. cumberlandia* was absent 31 sites showed improvement. The overall improvement was 73.1% (49 out of 67). Our best model was statistically significant ( $p < 0.01$ ) when evaluated with a Monte Carlo permutation test (1000 randomized runs).

The model with an average neighborhood size of 10 resulted in best model with 2 predictor variables – elevation and average monthly rainfall. The best model was derived from 329 candidate models (Table 3). The model had a log  $\beta$  value of 2.90 which increased to 2.92 after fine tuning. Average monthly rainfall (sensitivity = 2.71) was relatively more important than elevation (sensitivity = 0.06). Out of the 36 sites where *X. cumberlandia* occurred 25 sites showed improvement. Likewise 25 out of 36 sites where *X. cumberlandia* was absent showed improvement. Overall improvement was 69.4%. Our best model showed statistical significance at  $p = 0.05$  when evaluated with a Monte Carlo permutation test with 1000 randomized runs. The *X. cumberlandia* model with the default average neighborhood size (3.6) had a higher log  $\beta$  value

and better improvement percentage than the model with an average neighborhood size of 10. The stronger model was selected for generating the predictive map.

The predictive map for *X. cumberlandia* in the White River National Forest is shown in Fig 3. Ten percent of the area was not mapped due to a probability value lower than the minimum neighborhood size of 1. Out of 90 total points 63 (25 Low, 18 Medium and 20 High) were accessible. In the low probability area *X. cumberlandia* was recorded from 8 sites (32%); while *X. cumberlandia* was reported for 8 sites (44.4%) in medium probability area and 4 sites (20%) in the high probability area (Fig 4).

#### **LOGISTIC REGRESSION MODEL FOR *XANTHOPARMELIA CUMBERLANDIA***

The best logistic regression model was selected on the basis of the AICc value. Of the ten candidate models two models (Numbers 7 and 8) explained about 91% of variation (Table 4). Consequently these two models were averaged to identify the best coefficient for each predictor variable. The logistic regression model predicted average monthly precipitation, elevation, solar radiation, and average monthly maximum temperature as the best predictors. The best coefficients for each of the predictor variable are reported in Table 5.

The GIS predictive map generated using the logistic regression model identified a large portion of the White River National Forest as being low probability for occurrence of *X. cumberlandia* (Fig 5). Out of the 90 points we were able to visit 67 (53 Low, 6 Medium and 8 High). We found *X. cumberlandia* at 14 (26.4%) out of 53 sites in the low probability area while *X. cumberlandia* specimens were identified at 3 sites in the medium probability area (50.0%) and 3 sites (38%) in the high probability area (Fig 6).

Table 4: Candidate models with weighted AICc values

S.N.	Model #	Model	AIC	Parameter	AICc	ΔAIC	AICc (Wi)
1.	7	Xancum~Precip+Elevation+Tmax	55.5	4	56.09	0	0.689
2.	8	Xancum~Precip+SR+Elevation+Tmax	57.5	5	58.4	2.31	0.217
3.	9	Xancum~Elevation+IMI+Tmax+Tmin	59.19	5	60.1	4	0.093
4.	10	Xancum~Precip+Tmin+Tmax+SR	72.45	5	73.36	17.27	0.000
5.	1	Xancum~Slope+Elevation+IMI+luse+aspect+Tmin+Tmax+Precip+SR	70.84	10	74.45	18.36	7.12E-05
6.	5	Xancum~Precip+Tmax+Tmin	81.49	4	82.09	25.99	1.57E-06
7.	2	Xancum~SR+IMI+Tmax+Tmin	84.51	5	85.42	29.33	2.95E-07
8.	6	Xancum~Elevation+luse+SR+Precip	95.22	5	96.13	40.04	1.40E-09
9.	4	Xancum~IMI+Elevation+Slope+Aspect	103.9	5	104.81	48.72	1.82E-11
10.	3	Xancum~Slope+Aspect+luse	108.67	4	109.27	53.17	1.96E-12

Tmax = Average Monthly Maximum Temperature; Tmin = Average Monthly Minimum Temperature; Precip = Average Monthly Precipitation; SR = Solar Radiation; IMI = Integrated Moisture Index; luse = Land use

Table 5: Logistic regression models with parameter estimates

Predictor Variables	Parameter estimates
Intercept	-85.093
Elevation	0.016
Average Monthly Maximum Temperature	2.36
Average Monthly Precipitation	0.51
Solar Radiation	0.0000015

## DISCUSSION

We developed a habitat niche model for predicting the distribution of two air pollution sensitive lichens, *Usnea hirta* and *Xanthoparmelia cumberlandia* in the White River National Forest, Colorado. Comparatively, the *U. hirta* model performed significantly better than the *X. cumberlandia* model. The *U. hirta* model showed that a large percentage of the forest occurred within the high probability area. In addition, field collections documented that *U. hirta* also occurred at a higher frequency than predicted in the low and medium probability areas of the forest. Likewise the model for *X. cumberlandia* performed better in the low and medium probability areas than expected; however, its performance in the high probability area was much

poorer than predicted based on both NPMR and LR models. *U. hirta* was almost exclusively recorded from various conifers such as, ponderosa pine, Douglas fir, white fir, and engelmann spruce, while *X. cumberlandia* was found on a variety of rock types but never on dolomite limestone – a dominant rock type in the forest.

The model predicted that the high probability elevation range for *U. hirta* was between 2400 and 3000 meters and in the field we found that sites within this range consistently demonstrated abundant quantities of *U. hirta*. We were also able to show that *U. hirta* occurred with much lower frequency above and below this high probability elevation range; a logical conclusion in light of the general distribution of the subalpine coniferous forest.

For *X. cumberlandia*, both models (NPMR and LR) predicted a similar high probability elevation range – 2300 meters to 3600 meters. Although we observed *X. cumberlandia* within this elevation range the actual occurrence was much lower than expected. We speculate that the low performance of the models for *X. cumberlandia* may be due in part to the lack of suitable rock substrates in the high probability area. Studies have shown that the availability of appropriate substrates is an important determinant in the distribution patterns of lichen species (Pharo and Beattie, 2002). More specifically, according to Benedict and Nash (1990) substrate characteristics directly affect the distribution of *Xanthoparmelia* spp. During the field work phase of this study we found that most of the high probability areas were dominated by coniferous forests or open subalpine and alpine meadows with few to no rock outcrops. Furthermore, most of the limited rock outcrops were predominantly limestone; generally not a particularly suitable substrate for *Xanthoparmelia* spp. This pattern suggests that geologic substrate is most likely an important factor in accurately predicting the distribution of this species. Future models for this species should include geological substrate as an essential predictor variable.

The NPMR model identified average monthly precipitation, average monthly minimum temperature, average monthly maximum temperature, and solar radiation as the major predictor variables out of 9 geographical and macroclimatic variables governing the distribution of *U. hirta*. Similarly, for *X. cumberlandia* average monthly rainfall, elevation, average monthly maximum temperature, and solar radiation were the best predictor variables. Our study aligned closely with the results of two other studies (Dietrich & Scheidegger 1997 and Ellis & Coppins 2006) both of which reported that altitude and climatic variables (rainfall and temperature) were major drivers controlling the distribution of lichen species. Werth et al. (2005) has also demonstrated a strong relationship between the species composition of epiphytic macrolichen communities and various macroclimatic factors. In addition, studies conducted by Bolliger *et al.* (2007) and Radies *et al.* (2009) have documented that climatic variables similar to our study are important for predicting lichen species distribution patterns.

In comparing both models (NPMR and LR), we found that they both identified the same predictor variables – elevation, average monthly maximum temperature, average monthly rainfall, and solar radiation as the best variables for predicting lichen distribution. Among the four variables the NPMR model identified average monthly rainfall was the most sensitive predictor variable; however, it was not significant with the LR model at  $p = 0.05$ . Generally, the LR model predicted a smaller high probability area; however, most of the high probability areas overlapped for both models within the 2,300 to 3,600 m elevation range. McCune et al. (2003) and Yost (2008) showed that the NPMR model is generally stronger than the LR model; however, our results, especially for *X. cumberlandia*, showed no significant distinction in the performance of the two models.

Having used data from various habitat types across three Intermountain Area states for the calibration phase of our model suggests the possibility of effectively applying our model across significant portions of the Intermountain western United States.

In part the limitations of our model may be related to misidentification of species. We assumed that the *Usnea* and *Xanthoparmelia* spp. used in developing our model had been correctly identified; however, recent molecular studies of various lichen genera suggest that species delimitation is far more complex than previously thought. This issue is particularly manifest in the large foliose genus *Xanthoparmelia* where species concepts based on morphology and chemistry become profoundly problematic when compared to the molecular data (Leavitt, 2010). Therefore, it may be that what we have been calling *X. cumberlandia* may in reality represent not a single species but multiple species. This loss of taxonomic resolution may have resulted in a general failure to effectively identify appropriate tolerances for predictor variables during the calibration phase of model development; thus resulting in lower performance for both the NPMR and LR models. Likewise, there may have been similar issues in distinguishing between various *Usnea* species based on conflicts between the application of traditional taxonomic characters and molecular data. While field testing our model we encountered what appears to be *U. lapponica* at some sites. As both *U. hirta* and *U. lapponica* have somewhat similar habitat preferences and are both sensitive to various air pollutants we recorded all sites with *U. lapponica* as positive records for *U. hirta*.

Another important limitation with our study is related to the structure of the database used to develop the model. The original data set was not collected with the intent to model lichen species distribution; rather it was collected as a part of a floristic and air quality bio-monitoring survey. Therefore, the variables used in the original database for predicting lichen species occurrence were based on GIS derived data. Designing a clear sampling procedure is an

important aspect of modeling species distribution – because selection of explanatory variables as well as the prediction accuracy of the model generally depend heavily on the sampling strategy for the dependent variable(s) (Edwards *et al.* 2006). This study compared two classification models: 1) a design approach (with probability-based sampling efforts which support design-based inferences) and 2) a purposive approach (with a non-probability, or purposive, framework, where biologists used knowledge of lichen life histories to search for and “sample” lichen occurrence). Edwards’ results showed that there was significantly lower prediction accuracy for the purposive tree models when compared to the design model. Our model has certainly performed better than expected but would likely have been much more effective if our sampling design had been based specifically on modeling lichen species distribution.

Field measured variables like total tree basal area, number of tree species colonized by lichens, proportion of trees including deciduous trees, canopy cover, vegetation characteristics etc. are also key variables for effectively modeling the distribution of lichen species. The studies of Hyvärinen *et al.* (1992) and Gustafsson *et al.* (2004) have emphasized the importance of field measured variables in accurately determining the distribution of lichen species. Similarly, Bergamini *et al.* (2007) found that models based on field measured variables performed better in predicting microlichen richness than models based on climatic variables. Our model would certainly have performed better if we had included field measured variables in our predictor list along with macroclimatic and geographical variables.

The sample size used to develop predictive models for species occurrence greatly affects the effectiveness of the model (Stockwell and Peterson, 2002). Ideal sample size can be determined statistically using the data/predictor ratio. For binary responses, the data/predictor ratio is the number of observations in the least represented category (presences or absences) divided by the number of predictors in the model (McCune, 2006) and according to Harrell *et al.*



(1996) for binary data a minimum ratio should be 10. Our study was based on a smaller than ideal sample size; therefore, by increasing the number of sampling sites our model would likely have performed better.

Another potential reason for the reduced performance of our model may have been due to fact that we used climatic variables rather than field measured data. According to Bergamini *et al.* (2007) depending only on climatic variables as predictors lowers the performance of the model because of potential mismatches between real climatic and modeled variables; a condition further complicated by the fact that lichens are inherently more sensitive to microclimatic conditions. Furthermore, we predicted the distribution of species beyond the range of the climatic data used to calibrate our model and according to Thuiller (2004) model application becomes more uncertain outside the climatic range of the calibrated data set.

In conclusion, our predictive model provided some reasonably valuable information about the distribution of *U. hirta* and *X. cumberlandia* although it performed at a lower level than expected especially for *X. cumberlandia*. Our model also indicated that more emphasis should be placed on the northeast and western part of the White River National Forest when establishing air quality bio-monitoring reference sites using *U. hirta* as the sensitive indicator species. However, when considering *X. cumberlandia* as a potential sensitive indicator species emphasis would more profitably be placed on the northeast quadrant of the forest. In the future when developing models for predicting the distribution of lichen species we should have a clear sampling design with sufficient sampling sites involving field measured predictor variables as well as microclimatic data.

## LITERATURE CITED

- Ask, P. & Nilsson, S. G. (2004). Stand characteristics in color-infrared aerial photographs as indicators of epiphytic lichens. *Biodiversity and Conservation*, 13: 529-542
- Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. *Ecological Modeling*, 157: 101-118.
- Baur, A., Baur, B., & Froberg, L. (1994). Herbivory on Calicolous lichens: different food preferences and growth rates in two co-existing land snails. *Oecologia*, 98: 313-319.
- Benedict, J. B., & Nash, T. H. III (1990). Radial growth and habit selection by morphologically similar chemotypes of *Xanthoparmelia*. *The Bryologist*, 93: 319-327.
- Berg, A., Gardenfors, U., & von Proschqitz, T. (2004). Logistic regression models for predicting occurrence of terrestrial molluscs in southern Sweden - importance of environmental data quality and model complexity. *Ecography*, 27: 83-93.
- Bergamini, A., Stofer, S., Bolliger, J., & Scheidegger, C. (2007). Evaluating macrolichens and environmental variables as predictors of the diversity of epiphytic microlichens. *The Lichenologist*, 39: 475-489.
- Berglund, H., O'Hara, R. B., & Jonsson, B. G. (2009). Quantifying Habitat Requirements of Tree-Living Species in Fragmented Boreal Forests with Bayesian Methods. *Conservation Biology*, 23: 1127-1137.
- Berryman, S., & McCune, B. (2006). Estimating epiphytic macrolichen biomass from topography, stand structure and lichen community data. *Journal of Vegetation Science*, 17: 157-170.
- Binder, M. D., & Ellis, C. J. (2008). Conservation of the rare British lichen *Vulpicida pinastri*: changing climate, habitat loss and strategies for mitigation. *The Lichenologist*, 40: 63-79.
- Bjelland, T. (2003). The influence of environmental factors on the spatial distribution of saxicolous lichens in a Norwegian coastal community. *Journal of Vegetation Science*, 14: 525-534.
- Bolliger, J., Bergamini, A., Stofer, S., Kienast, F., & Scheidegger, C. (2007). Predicting the potential spatial distributions of epiphytic lichen species at the landscape scale. *The Lichenologist*, 39: 279-291.
- Bonn, A., & Schroder, B. (2001). Habitat models and their transfer for single and multi-species groups: a case study of carabids in an alluvial forest. *Ecography*, 24: 483-496.
- Boucher, V. L., & Stone, D. F. (1992). Epiphytic lichen biomass. In Carroll G. C. & D. T. Wicklow (Eds.), *The fungal community. Its organization and role in the ecosystem*. 2nd ed., pp. 583-599, Marcell Dekker, New York.
- Brodo, I. M. (1961). Transplant experiments with corticolous lichens using a new technique. *Ecology*, 42: 838-841.

- Brodo, I. M., Sharnoff, S. D., & Sharnoff, S. (2001). *Lichens of North America*: Yale University Press/New Haven and London.
- Carreras, H. A., Gudiño, G. L., & Pignata, M. L. (1998). Comparative biomonitoring of atmospheric quality in five zones of Córdoba city (Argentina) employing the transplanted lichen *Usnea* sp.. *Environmental Pollution*, 103: 317-325.
- Casazza, G., Zappa, E., Mariotti, M. G., Médail, F., & Minuto, L. (2008). Ecological and historical factors affecting distribution pattern and richness of endemic plant species: the case of the Maritime and Ligurian Alps hotspot. *Diversity and Distributions*, 14: 47-58.
- Clerc, P. (1997). Notes on the genus *Usnea* Dill. ex Adanson. *The Lichenologist*, 29: 209-215.
- Coxson, D. S. & Stevenson, S. K. (2007). Growth rate responses of *Lobaria pulmonaria* to canopy structure in even-aged and old-growth cedar-hemlock forests of central-interior British Columbia, Canada. *Forest Ecology and Management*, 242: 5-16.
- Cristofolini, F., Giordani, P., Gottardini, E., & Modenesi, P. (2008). The response of epiphytic lichens to air pollution and subsets of ecological predictors: A case study from the Italian Prealps. *Environmental Pollution*, 151: 308-317.
- Davis, D. B. (2009). *Predictive Modeling of Sulfur Flower Buckwheat (Eriogonum umbellatum Torrey) Using Non-Parametric Multiplicative Regression Analysis*. Master Thesis, Department of Plant and Wildlife Science, Brigham Young University, Provo, UT.
- Dietrich, M., & Scheidegger, C. (1997). Frequency, diversity and ecological strategies of epiphytic lichens in the Swiss Central Plateau and the Pre-Alps. *The Lichenologist*, 29: 237-258.
- Edwards, T. C. Jr., Cutler D. R., Zimmermann, N. E., Geiser L., & Moisen, G. G. (2006). Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling*, 199: 132-141.
- Ellis, C. J., & Coppins, B. J. (2006). Contrasting functional traits maintain lichen epiphyte diversity in response to climate and autogenic succession. *Journal of Biogeography*, 33: 1643-1656.
- Ellis, C. J., Coppins B. J. & Dawson, T. P. (2007). Predicted response of the lichen epiphyte *Lecanora populicola* to climate change scenarios in a clean-air region of Northern Britain. *Biological Conservation*, 135: 396-404
- Fenton, N. J., & Bergeron, Y. (2008). Does time or habitat make old-growth forests species rich? Bryophyte richness in boreal *Picea mariana* forests. *Biological Conservation*, 141: 1389-1399.
- Ferry, B. W., Baddeley, M. S., & Hawksworth, D. L. (1973). *Air Pollution and Lichen*: The Athlone Press, London.

- Fuga, A., Saiki, M., Marcelli, M. P., & Saldiva, P. H. N. (2008). Atmospheric pollutants monitoring by analysis of epiphytic lichens. *Proceedings of the 4th International Workshop on Biomonitoring of Atmospheric Pollution (With Emphasis on Trace Elements)*, 151: 334-340.
- Galvich, D. A., Geiser, L. H. & Mikulin, A. G. (2005). Rare epiphytic coastal lichen habitats, modeling, and management in the Pacific Northwest. *The Bryologist*, 108: 377-390.
- Garty, J., Kauppi, M., & Kauppi, A. (1997). The influence of air pollution on the concentration of airborne elements and on the production of stress-ethylene in the lichen *Usnea hirta* (L.) Weber em. Mot. transplanted in urban sites in Oulu, N. Finland. *Archives of Environmental Contamination and Toxicology*, 32: 993-1009.
- Gilbert, O. L. (1973). Lichens and air pollution. In V. H. Ahmadjian (Ed.), *The Lichens*. pp. 443-472, Academic Press, London.
- Grundel, R., & Pavlovic, N. E. (2007). Response of bird species densities to habitat structure and fire history along a midwestern open-forest gradient. *The Condor* 109: 734-749.
- Gustafsson, L., Appelgren, L., Jonsson, F., Nordin, U., Persson, A. & Weslien, J. O. (2004). High occurrence of red-listed bryophytes and lichens in mature managed forests in boreal Sweden. *Basic and Applied Ecology*, 5: 123-129.
- Harrell, F. E. Jr. Lee, K. L., & Mark D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15: 361-387.
- Hyvärinen, M., Halonen, P. & Kauppi, M. (1992). Influence of stand age and structure on the epiphytic lichen vegetation in the middle-boreal forest of Finland. *The Lichenologist*, 24: 165-180.
- Hobara, S., McCalley, C., Koba, K., Giblin, A. E., Weiss, M. S., Gettel, G. M., et al. (2006). Nitrogen fixation in surface soils and vegetation in an arctic tundra watershed. A key source of atmospheric nitrogen. *Arctic Antarctic and Alpine Research*, 38: 363-372.
- Iverson, L. R., Dale M. E., Scott, C. T., & Prasad, A. (1997). A GIS-derived intergrated moisture index to predict forest composition and productivity of Ohio forests (USA). *Landscape Ecology*, 12: 331-348.
- Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution*, 19: 101-108.
- Jozwiak, M. (2009). Influence of cement industry on accumulation of heavy metals in bioindicators. *Ecological Chemistry and Engineering* 16: 323.
- Kohler, G. R. (2007). Predators associated with hemlock woolly adelgid (Hemiptera: Adelgidae) infested western hemlock in the Pacific Northwest. PhD Dissertation, Oregon State University, Corvallis, Oregon.

- Leavitt S. (2010) Assessing traditional morphological- and chemistry- based species circumscriptions in lichenized ascomycetes: character evolution and species delimitation in common western North American lichens. Phd Dissertation, Brigham Young University, Provo, Utah.
- Loppi, S. (1996). Lichens as bioindicators of geothermal air pollution in central Italy. *The Bryologist*, 99: 41-48.
- Martinez, I., Carreno, F., Escudero, A., & Rubio, A. (2006). Are threatened lichen species well-protected in Spain? Effectiveness of a protected area network. *Biological Conservation*, 133: 500-511.
- McCune, B. (2006). Non-parametric habitat models with automatic interactions. *Journal of Vegetation Science*, 17: 819-830.
- McCune, B., Berryman, S. D., Cissel, J. H., & Gitelman, A. I. (2003). Use of a smoother to forecast occurrence of epiphytic lichens under alternative forest management plans. *Ecological Applications*, 13: 1110-1123.
- McCune, B., & Mefford, M. J. (2004). HyperNiche. Nonparametric multiplicative habitat modeling, Version 1, MjM Software, Gleneden Beach, OR, US.
- Motulsky, H. (2002). Multicollinearity in multiple regression. <http://www.graphpad.com/articles/Multicollinearity.htm>
- Nash, T. H. (2008). Lichen sensitivity to air pollution. In T. H. Nash (Ed.), *Lichen Biology*, pp. 299 - 314. Cambridge University Press, New York.
- Nash, T. H., Ryan, B. D., Diederich, P., Gries, C., & Bungartz, F. (2004). *Lichen flora of the Greater Sonoran Desert Region* (Vol. 2): Lichen Unlimited, Arizona State University, Tempe, Arizona.
- Nobel, W., Kostka-Rick, R., & Bartholmess, H. (2008). Assessment of air pollution by traffic-related metals near a motorway using bioindicators. *Gefahrstoffe, Reinhaltung der Luft*, 68: 245.
- Petersen, S. L., & Stringham, T. K. (2008). Development of GIS-based models to predict plant community structure in relation to western juniper establishment. *Forest Ecology and Management*, 256: 981-989.
- Pharo, E. J., and Beattie, A. J. (2002). The Association between substrate variability and bryophyte and lichen diversity in eastern Australian forests. *The Bryologist*, 105:11-26.
- Pike, L. H. (1978). The Importance of Epiphytic Lichens in Mineral Cycling. *The Bryologist*, 81: 247-257.
- Purvis, W. (2000). *Lichens*: Smithsonian Institution Press, Washington D. C. in Association with The Natural History Museum, London.

- Quinlan, K., Moro, D., & Lund, M. (2004). Improving trapping success for rare species by targeting habitat types using remotely sensed data: a case study of the heath mouse (*Pseudomyn shortridgei*) in Western Australia. *Wildlife Research*, 31: 219-227.
- Radies, D., Coxson, D., Johnson, C., & Konwicky, K. (2009). Predicting canopy macrolichen diversity and abundance within old-growth inland temperate rainforests. *Forest Ecology and Management*, 259: 86-97.
- Rosentreter, R., Hayward, G. D., and WicklowHoward, M. (1997). Northern flying squirrel seasonal food habits in the interior conifer forests of Central Idaho, USA. *Northwest Science*, 71: 97-102.
- Shrestha, G., & Baniya, C. B. (2006). *Nutrients Content in Lichens. In: Environment and Plants: Glimpses of Research in South Asia. (eds.): Ecological Society (ECOS), Kathmandu.*
- Shrestha, G., & St Clair, L. L. (2009). The lichen flora of Southwestern Colorado. *Evansia*, 26: 102-123.
- Shrestha, G., & St. Clair, L. L. (In Prep<sup>n</sup>) A comparison of the lichen floras of four Intermountain Area Locations.
- St. Clair, L. L., and Anderson K. T. (1997). Establishment of a lichen air quality bio-monitoring program and baseline for selected sites in the Gila Wilderness Area, Gila National Forest, New Mexico. Final Report. pp 35.
- St. Clair, L. L. (2000). Establishment of a lichen air quality bio-monitoring program and baseline for selected sites in the Manti La Sal National Forest, Utah, Final Report. pp 70.
- St. Clair, L. L., St. Clair, S. B., and Newberry, C. C. (2000). Establishment of a lichen air quality bio-monitoring program and air quality baseline for Dinosaur National Monument. Final Report, pp 56.
- St. Clair, S. B., St. Clair, L. L., Mangelson, N. F., & Weber, D. J. (2002a). Influence of growth form on the accumulation of airborne copper by lichens. *Atmospheric Environment*, 36: 5637-5644.
- St. Clair, S. B., St. Clair, L. L., Weber, D. J., Mangelson, N. F., & Eggett, D. L. (2002b). Element Accumulation Patterns in Foliose and Fruticose Lichens from Rock and Bark Substrates in Arizona. *The Bryologist*, 105: 415-421.
- Stockwell D. R. B. & Peterson A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148: 1-13.
- Thomson, J. W. (1984). *American Arctic Lichens 1. The Macrolichens*: Columbia University Press, NY.
- Thuiller, W. (2004). Patterns and uncertainties of species' range shift under climate change. *Global Change Biology*, 10: 2020-2027.

- Thuiller, W., Araújo, M. B., & Lavorel, S. (2004). Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography*, 31: 353-361.
- van Herk, C. M., Mathijissen-Spiekman, E. A., & de Zwart, D. (2003). Long distance nitrogen air pollution effects on lichens in Europe. *The Lichenologist*, 35: 347-359.
- Vogiatzakis, I. N., & Griffiths, G. H. (2006). A GIS-based empirical model for vegetation prediction in Lefka Ori, Crete. *Plant Ecology*, 184: 311-323.
- Waser, L. T., Stofer, S., Schwarz, M., Kuchler, M., Ivits, E., & Scheidegger, C. (2004). Prediction of biodiversity - regression of lichen species richness on remote sensing data. *Community Ecology*, 5: 121-133.
- Werth, S., Tommervik, H., & Elvebakk, A. (2005). Epiphytic macrolichen communities along regional gradients in northern Norway. *Journal of Vegetation Science*, 16: 199-208.
- Yost, A. (2006). *Probabilistic Modeling of Understory Vegetation Species in a Northeastern Oregon Industrial Forest*. PhD Dissertation, Oregon State University, Corvallis, Oregon.
- Yost, A. C. (2008). Probabilistic modeling and mapping of plant indicator species in a Northeast Oregon industrial forest, USA. *Ecological Indicators*, 8: 46-56.

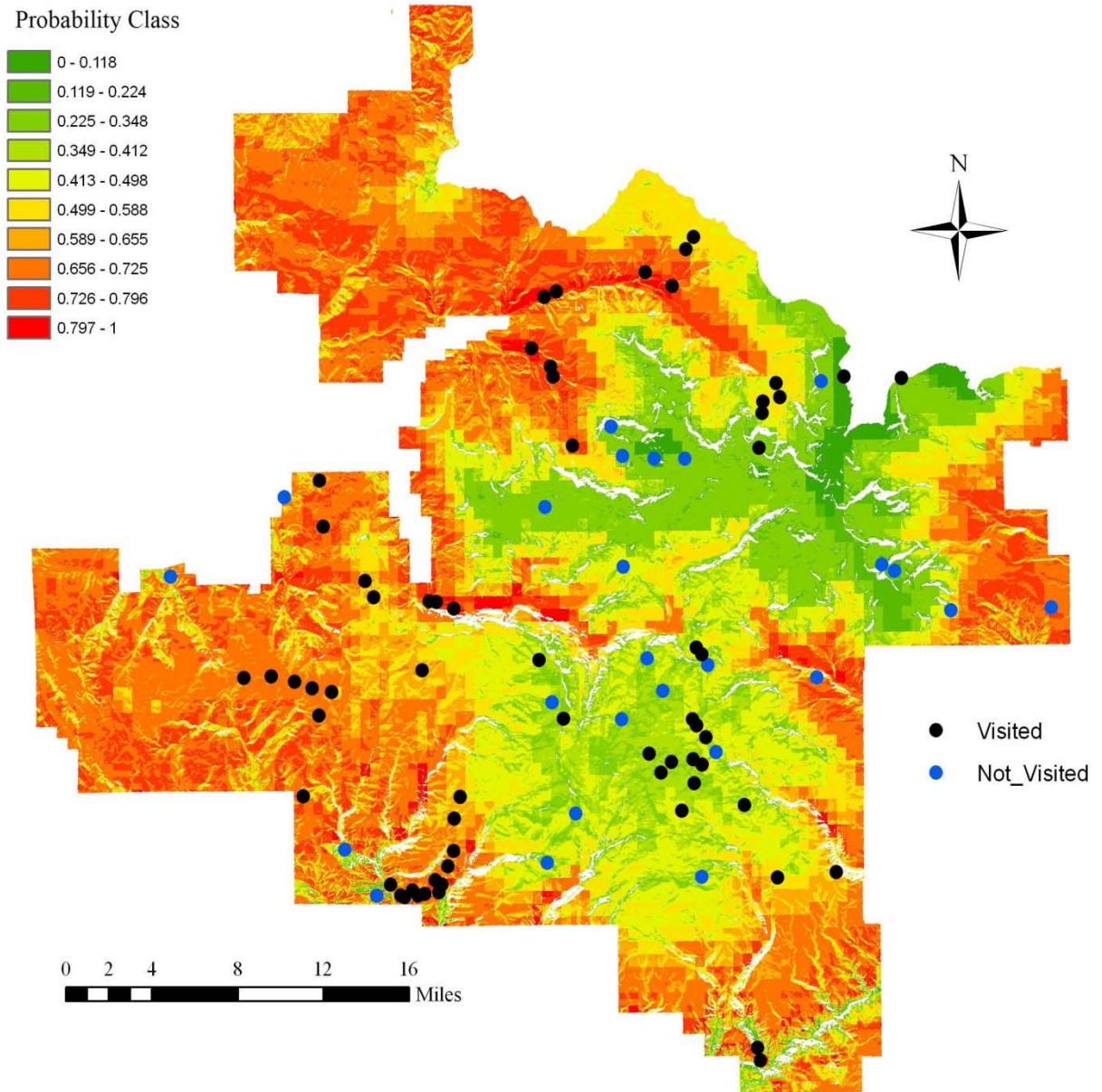


Figure 1: Distribution map of *U. hirta* for the study area with field verification sites.



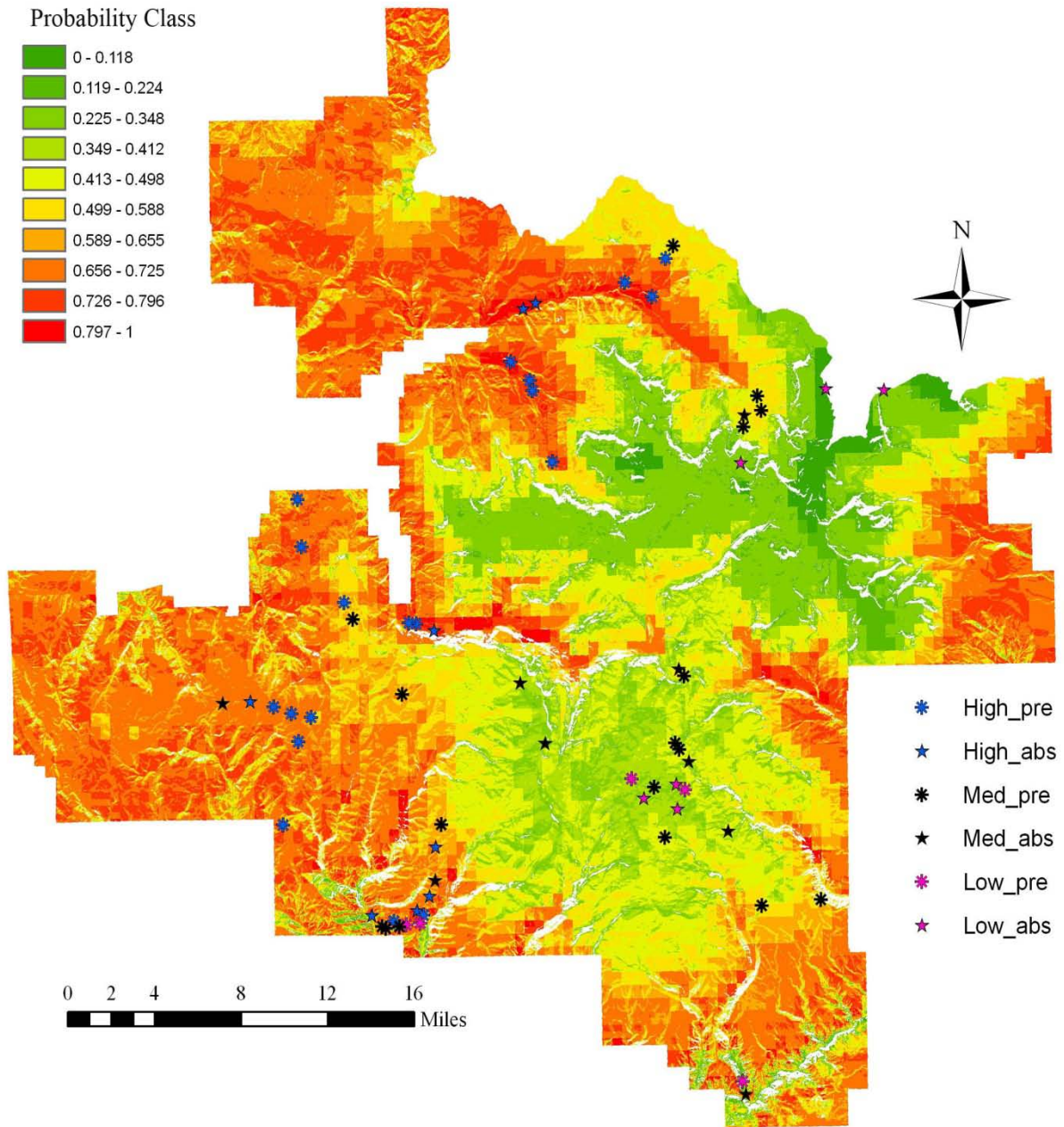


Figure 2: Distribution map of *U. hirta* for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence.

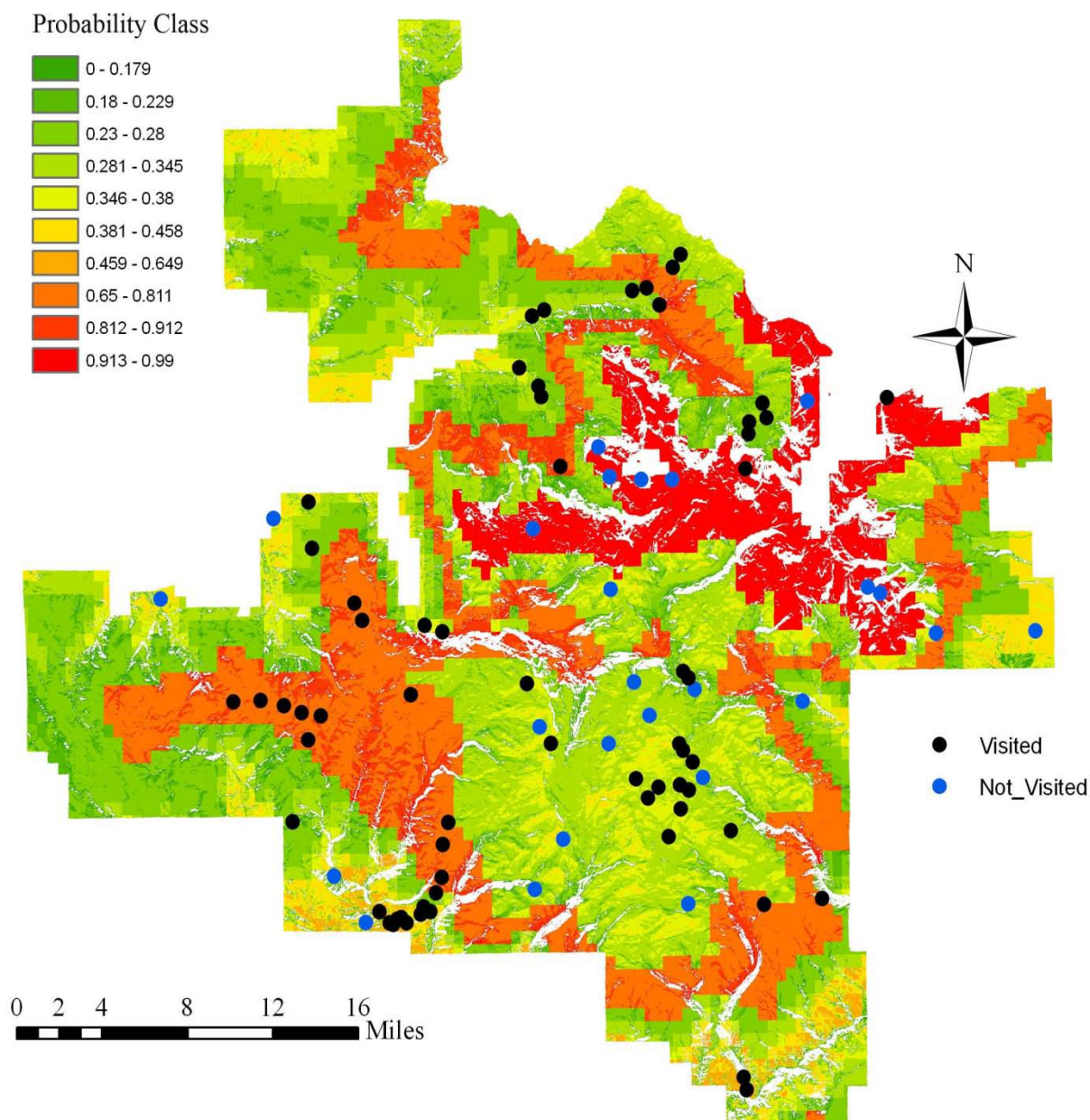


Figure 3: NPMR Model of the distribution map of *X. cumberlandia* for the study area with field verification sites.



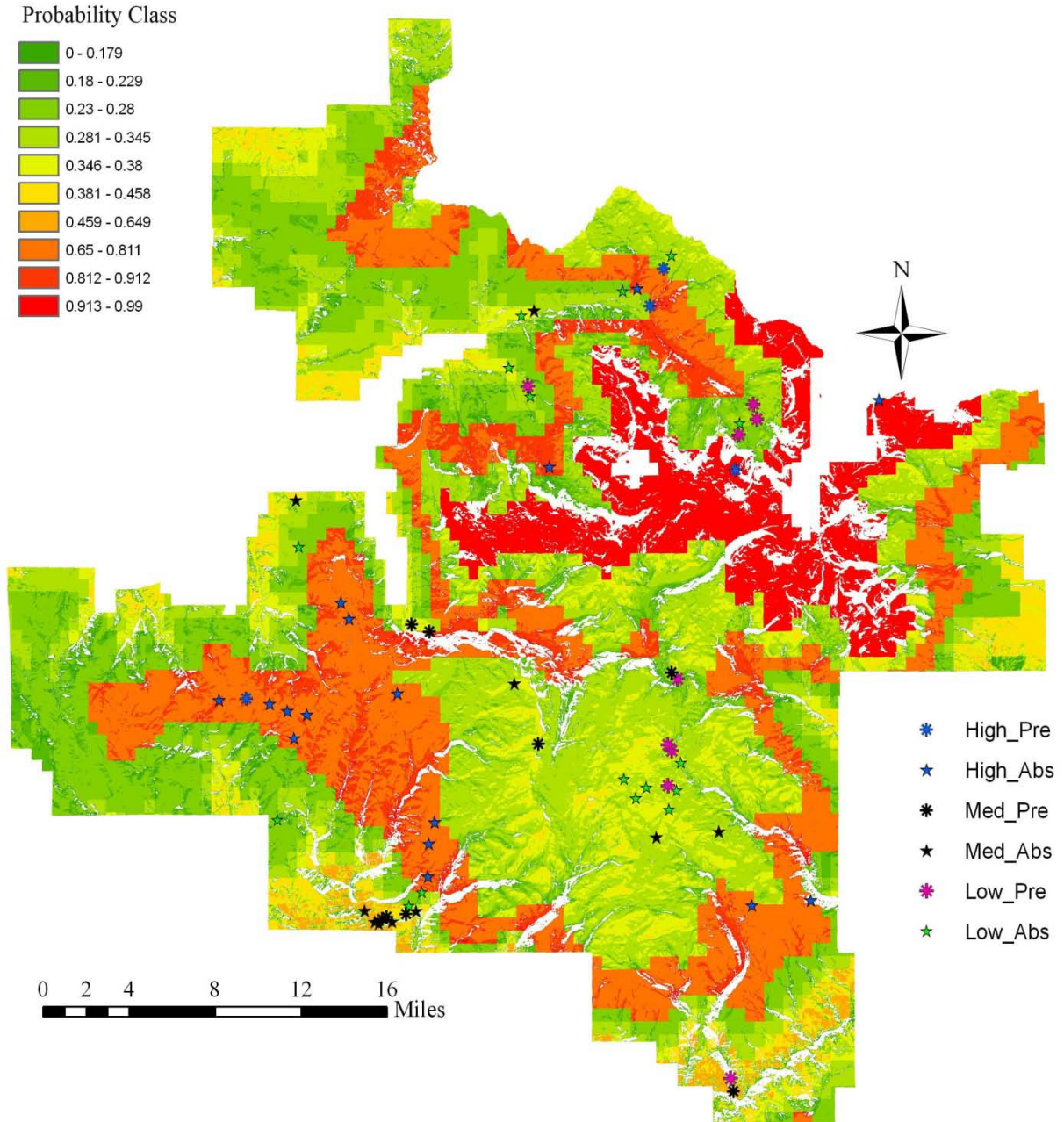


Figure 4: NPMR model of the distribution map of *X. cumberlandia* for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence.

Probability Class

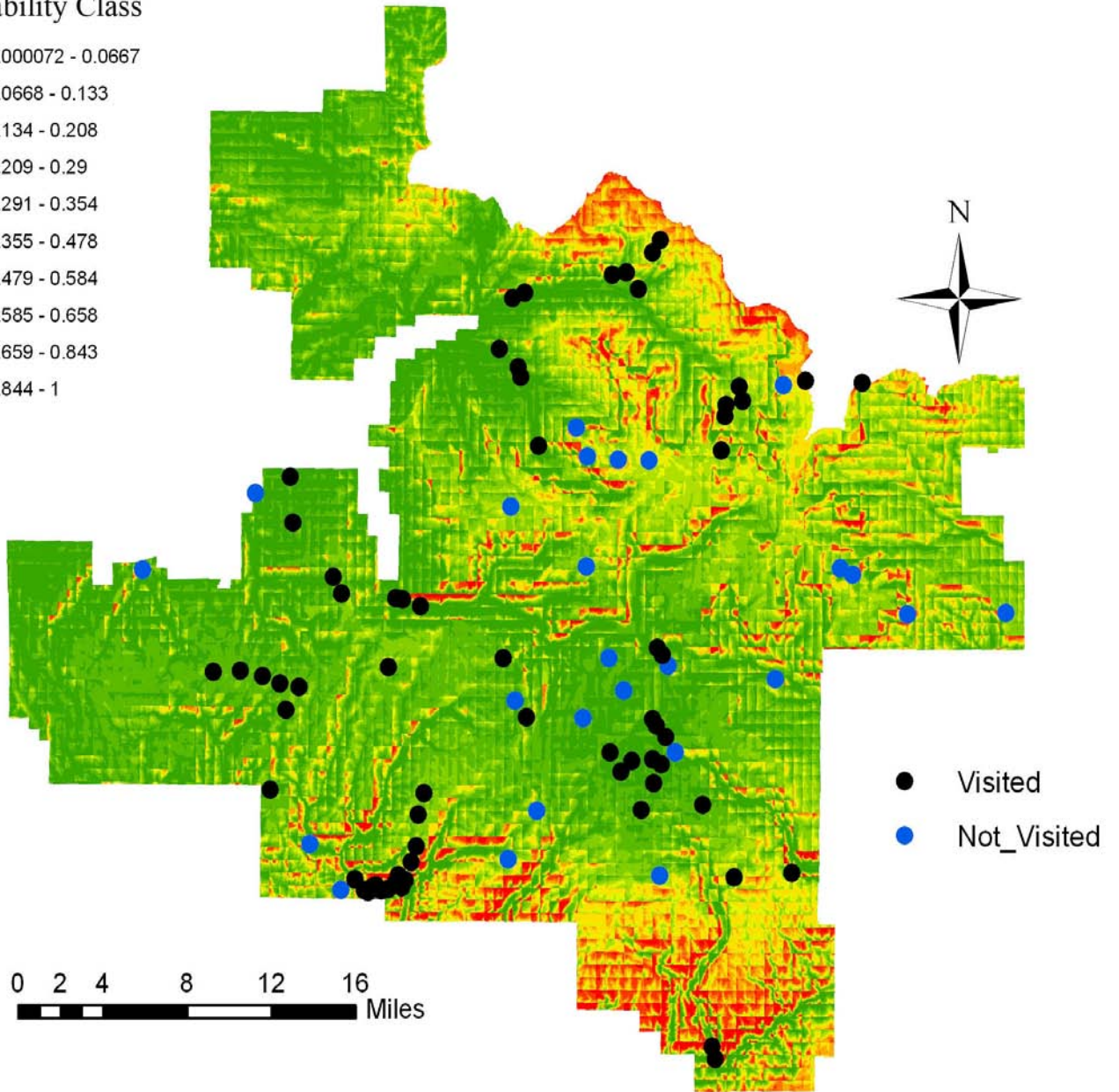


Figure 5: Logistic Regression Model of the distribution map of *X. cumberlandia* for the study area with field verification sites.



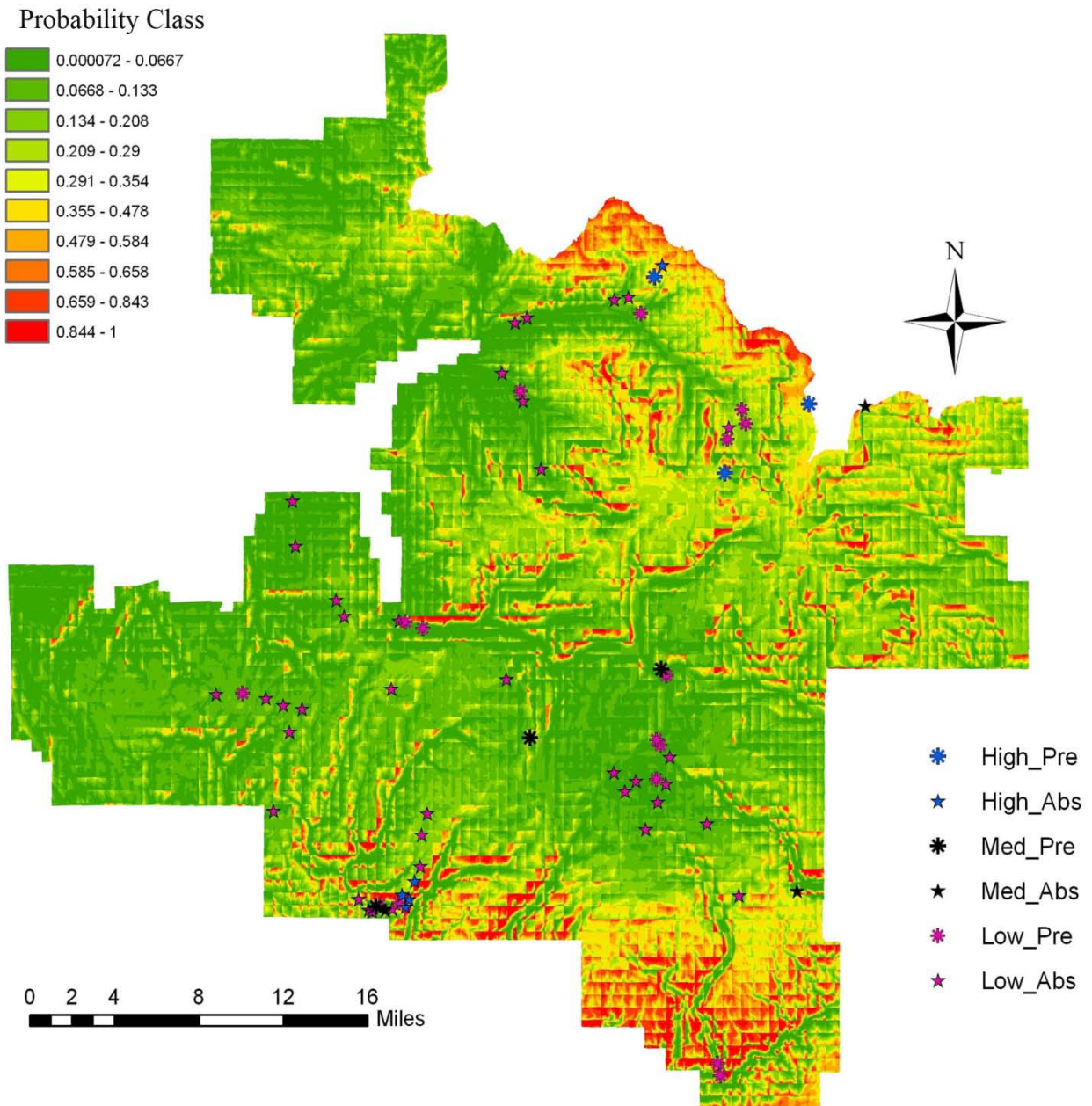


Figure 6: Logistic Regression model of the distribution map of *X. cumberlandia* for the study area showing sites with Presence and Absence of the species. Blue points are those with expected High Probability of occurrence while Black and Pink points represent Medium and Low Probability of occurrence respectively