PREDICTING SITE SUITABILITY FOR KUDZU (*PUERARIA MONTANA*) IN THE GREAT LAKES BASIN AND SURROUNDING REGION

by

Ashley Marie Kovach-Hammons

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Biological Science at Purdue Fort Wayne

Fort Wayne, Indiana

May 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Jordan M. Marshall, Chair
Department of Biological Sciences

Dr. Mark Jordan
Department of Biological Sciences

Dr. Bruce Kingsbury
Department of Biological Sciences

Approved by:
Dr. Jordan M. Marshall

Dedicated in loving memory to Mary C. D'Angelo and Ursula J. Crisp, who taught me strength, kindness, and above all, to enjoy life's journey. I would not be who I am today without you.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my committee chair, Dr. Jordan Marshall, for his guidance and patience. You were the first professor to take me on a field survey and teach me Indiana tree identification, and since then, you've taught me so much more in the classroom and beyond. I would also like to thank my committee members, Dr. Mark Jordan and Dr. Bruce Kingsbury for their insight and support of my thesis. Furthermore, I thank for Ken Cote (Indiana Department of Natural Resources), Rich Gardner (Ohio Department of Natural Resources), Abbie Bristol (SW x SW Corner Cooperative Invasive Species Management Area), and Fields Ratliff (Northwest Michigan Invasive Species Network) for their communication and help in finding field survey sites and helping me organize my field environmental surveys for this project. I would also like to thank IndianaView for funding my field survey travel through their IndianaView Consortium Grant.

A special thanks goes to my sisters, Carly and Kelsey, my grandmother, Mommoo, my partner, Conner, and my best friend, Ashley, for your company and help with my field surveys. I would also like to thank my friend, Austin, for his support, encouragement, and grammar expertise. Thank you to Bria and Elle for being my friends since the day we met and for all of the lunch breaks and advice. You both have made me feel supported and welcome since the day we all met. I would like to thank everyone in the Department of Biological Sciences, staff and fellow graduate students, for your encouragement and kindness over the past two years. Lastly, I would like to express my utmost gratitude for my partner, Conner, my family, and my friends for your continuous support and encouragement throughout all of my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Kudzu (*Pueraria montana*) is an invasive woody vine widespread throughout the southeastern United States, with recent studies predicting that its habitat will expand northward. New occurrences and recent studies using climatic parameters suggest that the Midwestern region of the United States is at the greatest risk of kudzu invasion. As there have already been 25 reports of kudzu within the Great Lakes basin, and no previous landscape models exist for the basin, I developed probability models from existing spatial data (land cover, hydrology, geology, annual precipitation, elevation, aspect, and known kudzu locations) using generalized additive, bioclimate envelope, and maximum entropy methods. I further expanded each model to include the basin and a 2.25-degree buffer in order to include 193 reported kudzu sites. For each predictive model, I determined the area under the curve (AUC) for a receiver operating characteristic curve (ROC) comparing false positive and false negative rates. I performed field surveys at eight known sites of kudzu presence in Michigan, Indiana, and Ohio. Each presence site was paired with a control (known kudzu absence site). I collected environmental data including canopy cover, volumetric soil moisture, soil pH, litter depth, midstory species diversity and diameter at breast height (DBH), and overstory basal area. Each environmental measure was compared between kudzu presence and control survey sites as well as between in-basin and out-of-basin survey sites using a two-way ANOVA. Maximum entropy models produced the highest AUC in both the basin and buffer models during model development. These models showed that urban and disturbed habitats resulted in the greatest probability of potential habitat for kudzu. I found no statistically significant differences in environmental characteristics between kudzu absent and presence sites or between in- and out-of-basin sites, suggesting kudzu might be dispersal-limited rather than limited by environmental characteristics. Continuing existing management and further monitoring of kudzu spread is likely necessary to limit further introduction and to mitigate spread of kudzu within the Great Lakes region.

CHAPTER 1. INTRODUCTION

1.1 Invasive Species in the United States

Over the second half of the 20th century, the rate of biodiversity loss has increased, correlating with human activity, land use changes, and invasive species (Millennium Ecosystem Assessment 2005). Current species extinction rates are dramatically higher than the historical background rate (i.e., 1000x greater than background rates), and as this diversity is lost, nearly all ecosystems are affected (De Vos et al. 2015, Millennium Ecosystem Assessment 2005). However, such estimations have been questioned due to empirical evidence not matching theoretical simulation, especially when investigating broad taxonomic groups (e.g., Stork 2009). Additionally, there are disagreements regarding the potential overestimation and underestimation of extinction rates using the same techniques of quantification (e.g., species-area relationships; Fattorini & Borges 2012, He & Hubbell 2011). Regardless, natural biodiversity levels are crucial for ecosystem processes and services, and can impact the well-being, health, and security of humans through these processes and services as well (Millennium Ecosystem Assessment 2005). To reduce this rate of biodiversity loss and to preserve native flora and fauna, it is important to monitor threats to biodiversity, like invasive species and habitat fragmentation, and to invoke management plans to preserve native species. Long term consequences of previous invasions and fragmentation persist well after management intervention (Vellend et al. 2006). In my thesis, I model habitat suitability for the invasive vine kudzu within the Great Lakes basin of the United States. Such modeling for widespread invasive species can provide insight to management to combat these invasive species.

There are over 15,000 non-native or introduced species within the United States, with over 8,000 in the conterminous United States (Simpson et al. 2021). Improvements in technology lead to more efficient travel and transportation mechanisms, which ultimately leads to increases in species introductions and subsequent survival (Simpson et al. 2021). Non-native species that establish and spread beyond the point of introduction and pose economic or ecological impacts as a result of establishment in an introduced region are labeled as invasive species (Simpson & Eyler 2018). Of the non-native species introduced to the United States, approximately 3,882 of these are recognized as invasive species, which by their label definition can further threaten native biodiversity (Simpson et al. 2021). A commonly cited and well accepted estimate indicated that invasive species were the primary risk for approximately 42% of threatened and endangered species (Pimentel et al. 2005).

Non-native species can be transferred across long distances via vectors (e.g., ship-ballast or imported materials) (Pimentel et al. 2005). Those non-native species that survive transportation are then introduced into novel areas (Lodge et al. 2006). Some of these species will not survive once they are introduced due to ecological and climatic constraints, but if they can establish (i.e., survive and reproduce), they have the potential to affect local communities. Furthermore, if a species can spread, they can pose a larger impact on communities within the introduced range, subsequently posing economic and ecological costs (Lodge et al. 2006). Thus, it is important to mitigate every step of the invasion process, where possible, through management to prevent the introduction, establishment, and spread of potentially invasive species (Lodge et al. 2006, Simpson & Eyler 2018).

1.2 Costs of Invasion

Invasive species cost upwards of $120 billion per year to manage, with upwards of $33 billion allocated towards invasive plant management (Pimentel et al. 2005). These costs may increase year-by-year as additional invasive plant species are introduced, compounding with the established populations spreading. In addition to these management costs, invasive species also impact agricultural crops, forests, and other natural ecosystems, which results in damages and lost yield, totaling approximately $24 billion in losses per year (Pimentel et al. 2005). As these estimates are nearing two decades in age, they are likely becoming underestimates of the true costs. Furthermore, invasive species can impact existing communities by out-competing native species, ultimately displacing those native species (Pimentel et al. 2005). Such ecological costs are far more complicated to estimate than economic costs with various algorithms proposed that may be highly species and ecosystem dependent (e.g., Parker et al. 1999, Ricciardi 2003). Finally, invasive species can also impact ecosystem productivity, ecosystem mitigate of natural disasters and habitat degradation, and even human health through novel pathogen introductions (Lodge et al. 2006).

One of the most prevalent invasive plant species in the United States is kudzu (*Pueraria montana* [Lour.] Merr. var. *lobata* [Willd.] Maesen & S.M. Almeida ex Sanjappa & Predeep, Fabaceae), a woody vine native to eastern Asia (Lindgren et al. 2013). Kudzu alone is estimated to cost upwards of $336 million each year in losses, primarily due to its toll on forest productivity, with additional costs of approximately $81 per hectare per year for management (Boyette et al. 2002). Beyond just monetary implications, kudzu poses ecological costs as well. Kudzu can cause losses in forest productivity, and it can also pose threats to local species diversity by reducing light availability resulting in shade-induced mortality in other vegetative species (Britton et al. 2002, Forseth & Innis 2004). Because kudzu is a woody vine, it can climb existing vegetation and bridge tree canopies together, which exacerbates the effects of windthrow and fire (Forseth & Innis 2004, Munger 2002).

1.3 Kudzu Biology

Kudzu is a perennial woody vine within the legume family (Fabaceae), and is characterized by its trifoliate leaves, with one trilobate leaflet in the center and two slightly smaller bilobate lateral leaflets (Lindgren et al. 2013, Mitich 2000). The vine is able to grow up to 30 cm per day, and up to 30 m per stem (Lindgren et al. 2013, Mitich 2000). This fast-growing nature allows it to overtop, shade out, and kill native vegetation, making it an aggressive competitor (Britton et al. 2002; Forseth & Innis 2004). Kudzu produces fragrant panicle inflorescence containing reddish-purple, pea flower-like florets in late summer to early fall (Lindgren et al. 2013, Mitich 2000). Kudzu fruits are flattened, long pods containing up to 10 oval- to kidney-shaped seeds (Lindgren et al. 2013, Mitich 2000).

Kudzu has an extensive system of tuberous roots, which can act as a nutrient storage system. These roots can reach up to 3 meters deep and new roots can develop at nodes of existing vines (Abramovitz 1983, Forseth & Innis 2004). This hardy root system allows new kudzu stems to sprout each spring, adding to the difficulty of extermination kudzu as new stems can sprout from roots left after treatment (Guertin et al. 2008, Lindgren et al. 2013). Additionally, kudzu exhibits a large root surface area, allowing it to extract nutrients that may be harder for other plants to obtain in the soil (Abramovitz 1983). Like many legumes, kudzu associates with nitrogen-fixing bacteria, which may provide an advantage over other non-native species in nutrient-poor or disturbed environments (Hickman 2010, Lindgren et al. 2013). *Bradyrhizobium* spp. and *Mesorhizobium* spp. bacteria were isolated from kudzu in Korea (Kown et al 2005), though it appears unclear if these genera form symbioses with kudzu globally or if they are site-specific. Additionally, this nitrogen-fixation results in kudzu having high foliar N content resulting in nitrogen-rich litter, which can alter biogeochemical processes, like litter decomposition rates (Hickman & Lerdau 2014, Lindgren et al. 2013). Because of its high foliar N content and symbiosis with nitrogen-fixing bacteria, kudzu may increase available nitrogen and alter nitrification and denitrification rates, but this seems to also depend on other species in the community and location (Coiner 2012, Forseth & Innis 2004, Hickman et al. 2010, Hickman & Lerdau 2014).

Kudzu primarily reproduces asexually by growing new crowns and roots at the nodes of existing vines, with a clonal rate of approximately 80% within a site (Abramovitz 1983, Bentley & Mauricio 2016, Winberry & Jones 1973). Kudzu's high genetic variation, due to its multiple introductions, and phenotypic plasticity likely allow for kudzu's success (Bentley & Mauricio 2016). Because kudzu does not produce many seeds within its invasive range, little is known about the seed dispersal and germination of this species (Aurambout & Endress 2018). Typically, there is low seed viability and seeds exhibit a physical dormancy (Lindgren et al. 2013). This low investment into sexual reproduction may allow kudzu to invest more in growth, which further expands its abilities to be a strong competitor in introduced environments (Abramovitz 1983).

1.4 Kudzu Native Distribution

Kudzu is native to eastern Asia, ranging across southern China, the Korean Peninsula, and Japan. In its native range, kudzu prefers hot summers with long growing seasons and mild winters. Additionally, kudzu prefers high precipitation and open areas with high light availability (Foreseth and Teramura 1987). Kudzu plants are dormant throughout winter after dropping foliage following the first frost. Though kudzu thrives on forest edge habitat and deep loamy soil, it is able to grow in a variety of habitats, including within forests and across multiple soil types (Li et al. 2011). In eastern Asia, kudzu is susceptible to many predators, including a variety of insect stem, leaf, and root feeders. These were originally proposed as potential biological control agents (Britton et al. 2001). However, many are not host specific and feed readily on soybean (*Glycine max*), which is confamilial with kudzu (Frye et al. 2007, Zhang et al. 2012). Some mammals may also consume kudzu foliage, though due to kudzu's extensive root system, these mammals are not able to control kudzu populations (Li et al. 2011). Kudzu has been a cultivated crop in eastern Asia, with roots and foliage used in fibers and textiles, food, and medicine. Harvesting kudzu, especially kudzu roots, has allowed for kudzu control within its native range (Li et al. 2011).

1.5 Kudzu United States (Invasive) Distribution

Initially introduced as an ornamental into North America in 1876 as part of the Japanese gardens in the World's Fair, kudzu has since had multiple intentional introductions with intended uses of ornamental, fodder, and for erosion control (Brown 2008, Guertin et al. 2008, Pieters 1932, Winberry & Jones 1973). Through much of the early 20th century, kudzu was actively promoted with research targeting successful propagation, especially for soil conservation (Myers et al. 1938). Interestingly, Pieters (1932) explicitly stated that there is no danger of kudzu becoming a pest, while admitting the growth can smother other vegetation. These introductions were primarily within the southern United States, where most of the kudzu within the United States is still found today (Abramovitz 1983, EDDMapS 2019). Since these introductions, kudzu has established and spread to over three million hectares within the United States, with reports in 32 states (Callen and Miller 2015, EDDMapS 2019). Kudzu is widely established throughout the southeastern United States, but populations of kudzu have also established along the east coast and as far northwest as Portland, Oregon (EDDMapS 2019).

Kudzu preferential climate had previously been defined as mild winters and hot summers (Lindgren et al. 2013), and kudzu was believed to have a cold threshold of -20 °C (Coiner et al. 2018). However, recent studies suggest that this threshold may vary due to kudzu acclimation to cold and therefore the invasive plant may survive further north than previously expected (Coiner et al. 2018). This is of particular concern as kudzu has already been reported in multiple locations in Michigan, some as far north as Benzie County, MI (EDDMapS 2019). As kudzu may be able to withstand colder temperatures than initially thought, management may be necessary to control these populations. For example, Benzie County has experienced minimum temperatures at or below -20 °C in 56 of the last 120 years, as well as 13 of the last 30 years (NOAA 2022).

1.6 Habitat Modeling for Invasive Species

Though species distribution models (SDM) were first used in the 1920s, they have become increasingly popular within the past few decades as useful tools to understand changes in individual species ranges and biodiversity patterns, as well as to predict invasive species distributions (Guisan & Thuiller 2005). These models generally include presence-absence, presence, or abundance data for a species which is then fit with additional data, including environmental characteristics at species locations. Models are built to fit the existing data and then used to produce spatial predictions of abundance or probabilities of occurrence for the distributions. These models can then be evaluated based on the accuracy of these predictions (Guisan & Thuiller 2005). There are many approaches to species distribution modeling, but climatic approaches are more common today with the heightened concern for how climate change can influence future distributions and for conservation of species of concern (Guisan & Thuiller 2005).

Common models for invasive species modeling include bioclimate envelope (BioClim) and maximum entropy models (MaxEnt), though generalized linear (GLM) and generalized additive models (GAM) were historically common in literature (Gastón & García-Viñas 2011, Guisan et al. 2002, Guisan & Thuiller 2005). GLMs and GAMs are primarily based in logistic regression, with GAMs being able to account for more non-linear relationships and inclusion of categorical data (Guisan et al. 2002). BioClims, on the other hand, correlate species locations and climate variables at those locations. While limits do exist for BioClims, many of the concerns are often associated with overfitting, multicollinearity, or spatial autocorrelation between climate variables (Heikkinen et al. 2006). Additionally, inclusion of non-climate predictor variables (i.e., not limiting predictors to only climate variables) may lead to improvements in model appropriateness (Harris et al. 2014). MaxEnt models have become increasingly popular in species distribution modeling as they include machine-learning concepts and can account for more complicated parameters and interactions between those parameters (Elith et al. 2006, Gastón & García-Viñas 2011). In fact, Baldwin et al. (2009) and Elith et al. (2006) compared

MaxEnt models to the well-established GLM, GAM, and BioClim techniques and both studies found that MaxEnt outperformed these other model types. MaxEnt was found to be particularly advantageous with limited presence locations and with more complicated parameters (Baldwin et al. 2009, Elith et al. 2006).

Species distribution models have previously been used to predict kudzu distributions across the United States. Bradley et al. (2010) used bioclimatic envelope and maximum entropy modeling techniques with future climate data predictions to determine invasion risk for kudzu, privet (*Ligustrum sinense, L. vulgare*), and cogongrass (*Imperata cylindrica*) in Eastern United States. Furthermore, Callen and Miller (2015) also used maximum entropy modeling techniques based on climate variables to compare kudzu's native range with suitable habitat within the United States. While these studies focused on the United States at the continent scale, few studies have sought to model kudzu distribution at the landscape scale (e.g., the Great Lakes basin). Two recent studies simulating kudzu spread at this scale in Oklahoma, using a Monte Carlo simulation, and in Illinois, using techniques similar to a general model of biological invasion (Arambout & Endress 2018, Harron et al. 2020). Additionally, Liang et al. (2020) incorporated novel Lidar data to predict kudzu distribution over Knox County, TN, using a random forest model. Despite much of the Midwestern United States being subject to high risk for kudzu invasion, there appears to be no studies that consider kudzu at the landscape level within the Great Lakes basin, (Bradley et al. 2010, Callen & Miller 2015).

1.7 Kudzu within the Great Lakes Basin

Kudzu has already been reported within at least twelve counties within the Great Lakes basin across Illinois, Indiana, Michigan, and Ohio, (EDDMaps 2019, Figure 1). Callen and Miller (2015) found that the invasive niche of kudzu resembles its native niche in Asia, but kudzu has not yet inhabited all suitable niche space in North America. This includes the Midwestern region of the United States, which is at highest risk for kudzu invasion due to its proximity to the invasion front from the southeastern region, where kudzu populations are widespread and well-established (Bradley et al. 2010, Callen & Miller 2015). Previous studies on kudzu invasive site habitability have focused on climate factors (Bradley et al. 2010, Callen & Miller 2015), but other factors may be necessary to determine the susceptibility of specific sites.

These factors may include the soil type and conditions, hydrology, dominant vegetation types, herbivory, anthropogenic disturbances, and reproductive strategies. While studies have previously shown kudzu distribution in the United States using species distribution modeling techniques, few provide insight to landscape-scale dynamics, which is necessary for management decisions.

1.8 Objectives

The objectives of my thesis were (1) to predict kudzu habitat suitability within the US portion of the Great Lakes basin, within the Midwest region of the United States, by comparing bioclimate envelope, generalized additive, and maximum entropy modeling techniques, (2) to determine if kudzu site characteristics are comparable between locations within the Great Lakes basin and outside of the Great Lakes basin, and (3) to test the hypothesis that habitat suitability can be predicted for such a wide-ranging invasive species like kudzu.

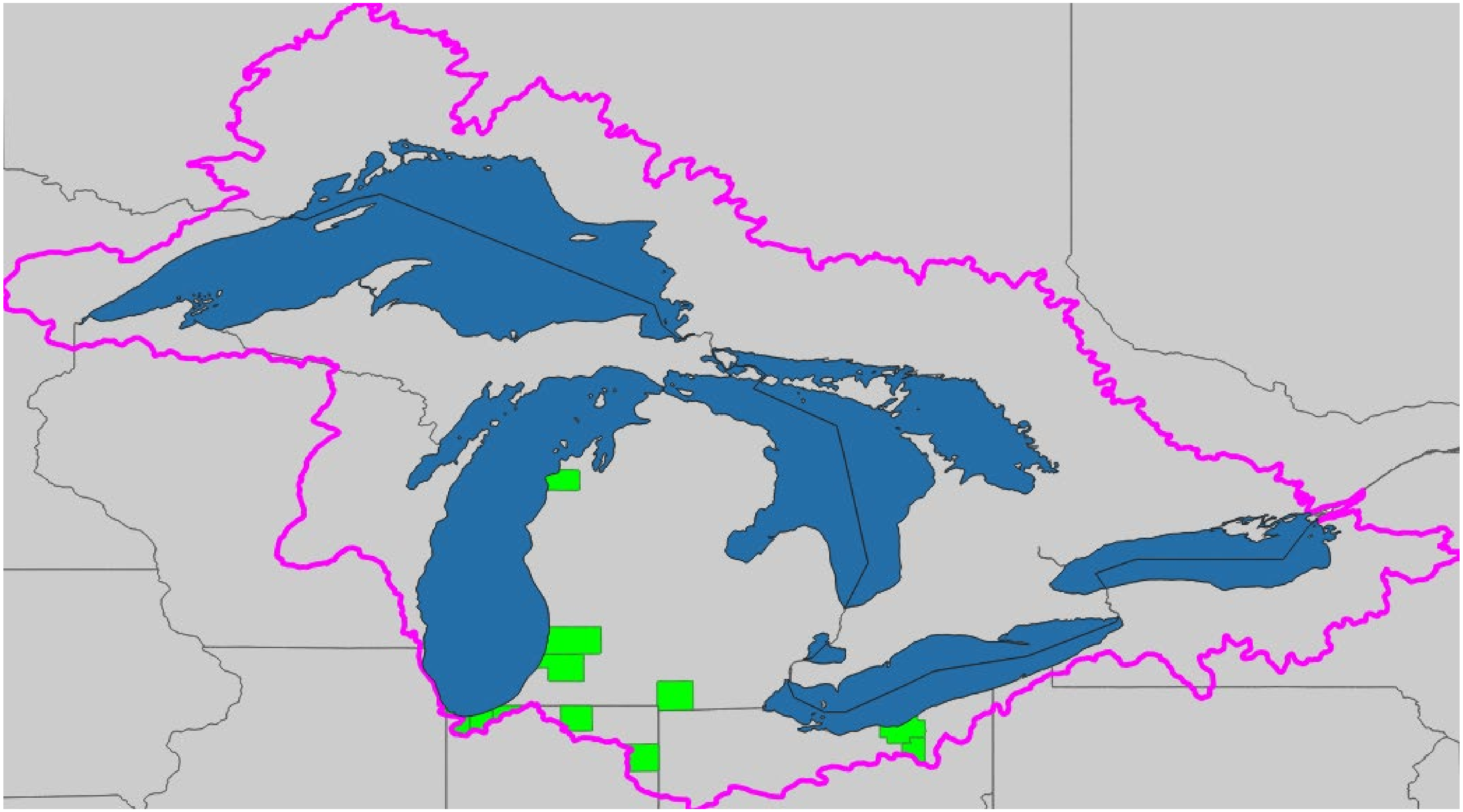

Figure 1: Counties with reported kudzu locations (highlighted green) within the Great Lakes basin boundary (solid pink line) (EDDMapS 2020).

CHAPTER 2. METHODS

2.1 Predicting Habitat Suitability within the Basin

Models were developed to predict probability of habitat suitability within the United States portion of the Great Lakes basin based on climatic and physical environmental variables using geographic data layers. Independent variables for model development included total annual precipitation (mm), mean annual temperature (ºC), land cover, geology (parent material), hydrology, elevation (m), and known kudzu locations (Appendix A, EDDMapS 2020). Aspect (degrees) and slope (degrees) were derived from digital elevation models. I omitted mean annual temperature due to its collinearity with total annual precipitation and likewise omitted slope due to its collinearity with aspect. The resulting models included the mean annual temperature, land cover, geology, hydrology, elevation, and aspect layers. These raster layers were resampled in QGIS (2020) to standardize resolution to 30 m x 30 m pixels across the Great Lakes basin. Environmental layer values were extracted for 25 known kudzu locations within the Great Lakes basin identified via the Early Detection & Distribution Mapping System (EDDMapS 2020) and the Midwest Invasive Species Information Network (MISIN 2020; Figure 2). The environmental values at these 25 locations were used to develop predictive probability models (bioclimate envelope, generalized additive, and maximum entropy models) using R (R Core Team 2020). Subsequent models were used to compare predicted probabilities for kudzu site suitability. These probabilities were then used to identify areas of the Great Lakes region at most risk of invasion based on available suitable habitat (i.e., those areas with a higher predicted probability of kudzu site suitability are at risk of invasion).

2.2 Expanding Site Suitability Predictions Using Expanded Models

Due to the limited availability of kudzu points within the Great Lakes basin (n = 25), I made additional predictive models that included a buffer of 2.25 degrees (geographic coordinate system = North American Datum 1983; approximately 250 km) around the Great Lakes basin. This buffer size was chosen such that it did not extend too far outside of the Great Lakes basin, so that models could still represent this landscape, as well as to include Brown County, IN, where my out-of-basin field sites were located, allowing for comparison between expanded models and kudzu presence and absence locations from field surveys. Extending the boundary from the Great Lakes basin to the new buffer increased the known kudzu locations to 193 (EDDMapS 2020). Independent environmental layers were the same as the basin models (mean annual precipitation [mm], land cover, geology, hydrology, elevation [m], and aspect [degrees]). Like the basin models, I developed bioclimate envelope, generalized additive, and maximum entropy models using R and used these to compare predicted probabilities for kudzu site suitability. These probabilities were used to determine the areas at risk of kudzu invasion within the Great Lakes region and expanded area.

2.3 Field Site Vegetation and Environmental Surveys

I collected site-specific environmental data at 8 sites in Indiana, Michigan, and Ohio during June-August 2021 in order to assess any differences in vegetative or environmental characteristics between in- and out-of-basin sites as well as between sites with known kudzu presence and known kudzu absence (Figure 3). Sites were selected based on current or previous kudzu infestation records from the Indiana Department of Natural Resources, Ohio Department of Natural Resources, and Midwest Invasive Species Information Network. Four sites were located within the Great Lakes basin in Ohio (n = 1) and Michigan (n = 3), with the remaining sites were located outside of the basin in Brown County, Indiana (n = 4).

For each site with current or previous kudzu infestation, I established one 5 m x 5 m plot to survey midstory and overstory vegetative strata. I also measured canopy cover (percent), volumetric soil moisture (percent), and soil pH as environmental characteristics. Additionally, I measured litter depth (cm) at each plot corner and the plot center. The diameter at breast height (DBH, 1.3 m above soil surface; cm) for each midstory stem was measured and identified to species. DBH was converted to basal area (m^2/ha) for comparison and analysis. The basal area (m^2/ha) was also measured for overstory species using a 10-factor angle gauge, and each overstory species was identified to species. Each kudzu-infested site was paired with a neighboring site without known kudzu infestation at least 40 m away that consisted of similar vegetative structure (i.e., similar forest stands) for a total of 16 field survey sites (8 known infested sites, 8 known non-infested sites). Environmental and vegetation data was collected at the neighboring, non-infested sites to serve as a control.

2.4 Data Analysis

All analyses were conducted in R (2020). Bioclimate envelope models were developed using the *dismo* package (Hijmans et al. 2021). Default options were used in the *bioclim* function to produce bioclimate envelope models. Generalized additive models were developed using the *mgcv* package (Wood 2011). Five-thousand background points were randomly selected and used as absence points for the basin GAM, and 10,000 background points were randomly selected for the absence points of the expanded GAM. Options within the *gam* function for generalized additive models included the restricted maximum likelihood (REML) method using a binomial family distribution. I used a smoothing spline fit for total annual precipitation, elevation, and aspect. Maximum entropy modes were developed using the *dismo* package as an interface with the Maxent software by Phillips et al. (2017). Options within the *maxent* function included land cover, geology, and hydrology as factors, and jackknifing. I used presence and absence data collected from survey sites to assess model accuracy by determining the percentage of true positives, true negatives, false positives, and false negatives for each model. These values were used to determine sensitivity (true positive rate) and specificity (true negative rate) as well as to determine the area under the curve (AUC) for a receiver operating characteristic (ROC) curve comparing false positive and true positive rates for each model.

From the field site vegetation data, I calculated Shannon's diversity index for each field survey site as a measure of species diversity. Mean canopy cover, overstory basal area, soil moisture, soil pH, litter depth, midstory species richness, and midstory species diversity were compared between infested (presence) and non-infested (absence) survey sites as well as between in-basin and out-of-basin survey sites through two-way ANOVA. Tukey's HSD was used as a post-hoc pairwise comparison between infested and non-infested sites and between in- and out-of-basin sites.

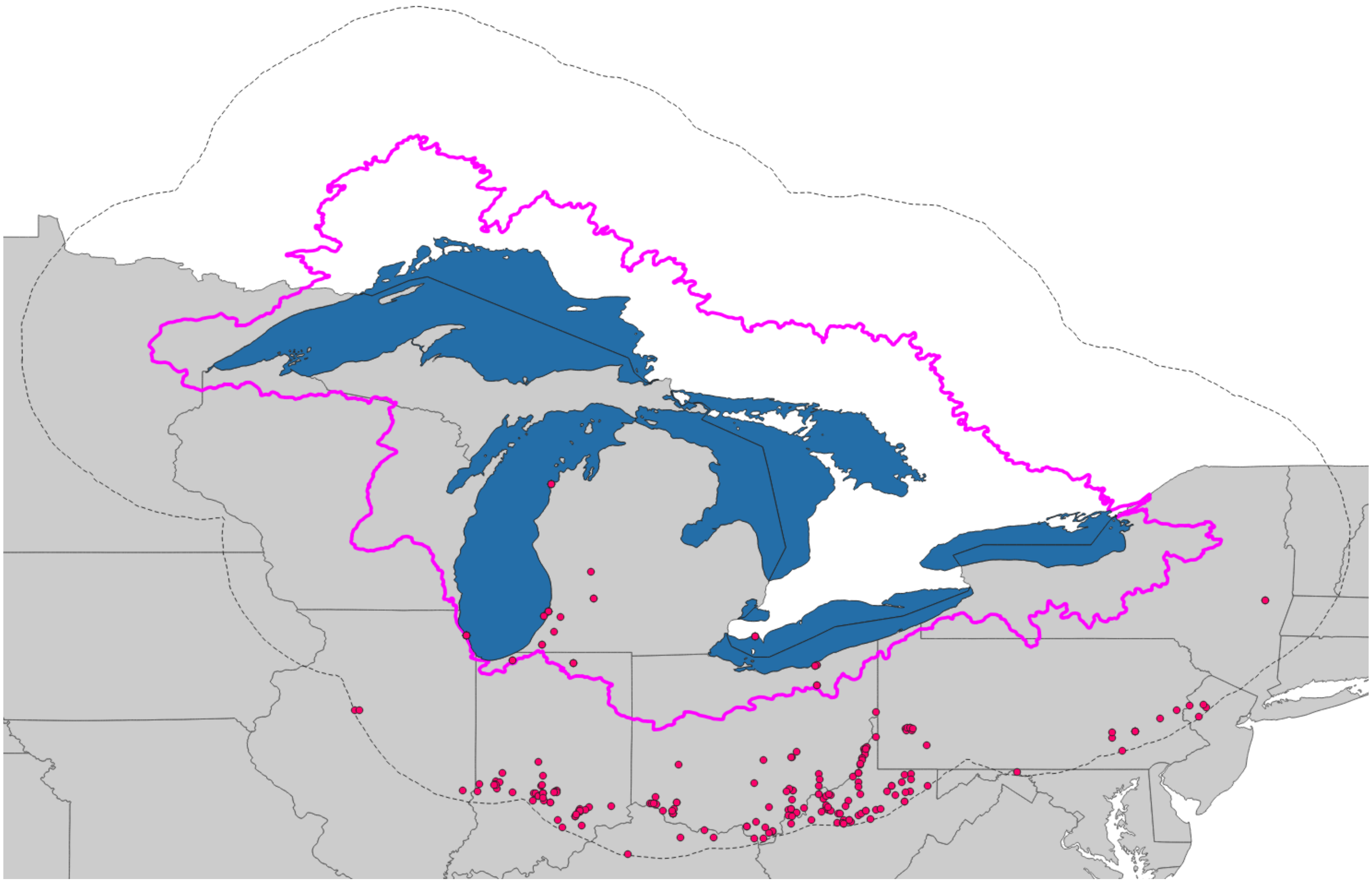

Figure 2: Reported kudzu locations (red dots) throughout the Great Lakes basin (solid pink line) and expanded area (dashed line; expanded with a 2.25-degree buffer including the basin; coordinate reference system = North American Datum 1983) (EDDMapS 2020).

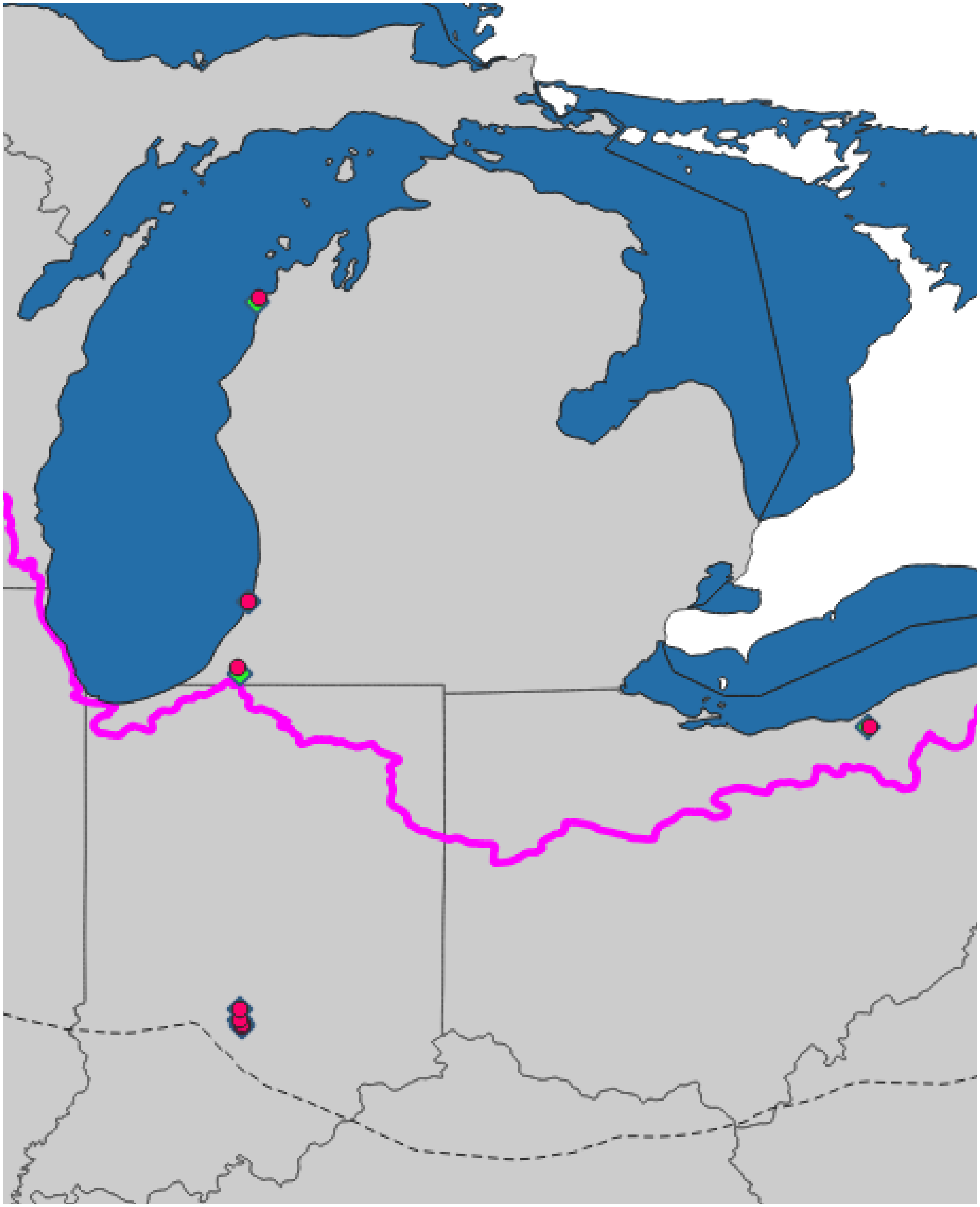

Figure 3: Survey locations of known kudzu sites (red dots) and control sites (green diamonds) throughout the Great Lakes basin (solid pink line) and expanded area (dashed line; expanded with a 2.25-degree buffer including the basin; coordinate reference system = North American Datum 1983) (EDDMapS 2020, MISIN 2020).

CHAPTER 3. RESULTS

3.1 Predicting Habitat Suitability

Based on the independent variables from 25 reported kudzu locations within the Great Lakes basin, I was able to produce a bioclimate envelope (BioClim) model, a generalized additive model (GAM), and a maximum entropy (MaxEnt) model within the basin (Figures 4-9). Additionally, expanding the basin with a 2.25-degree buffer, I was able to produce expanded probability models using 193 locations with generalized additive, bioclimate envelope, and maximum entropy approaches (Figures 10-15). All models contained total annual precipitation (mm), land cover, geology, hydrology, elevation (m), and aspect (degrees) as independent variables. The two model areas (basin and expanded) covered approximately 405,000 km^2 and 1,300,000 km^2, respectively (Appendix B).

Basic model evaluations for GAM and MaxEnt provide information about the individual models. In the GAM versions of the basin and expanded models, the spline smoothed independent variables were all retained as the effective degrees of freedom (edf) did not approach the basic model dimensions (k') (Table 3). Additionally, edf for all spline smoothed variables were indicative of linear relationships (edf $\geq$ 1). In the MaxEnt models, the land cover raster layer contributed the most in the basin (70%) and the buffer (40%) models (Table 4). The other variables contributed smaller percentage explanations of the model variance, compared to land cover. Within the basin MaxEnt model, land cover and elevation contributed 87% of the model explanation. The expanded area buffer MaxEnt model had 84% explanation contributed by land cover, geology, and elevation (Table 4).

I defined arbitrary thresholds of low probability, medium probability, and high probability of kudzu habitat suitability (inclusive of the upper limit) as 0.00-0.33, 0.33-0.66, and 0.66-1.00, respectively, in order to better understand kudzu habitat suitability predictions. The BioClim model and GAM had less than 1% of pixels in the medium and high threshold categories within the basin and the expanded area (Table 1; Figures 4, 5, 10, and 11). Conversely, the maximum entropy models in the basin and expanded area had 11% and 18%, respectively, of pixels in the medium and high categories (Table 1; Figures 9 and 15). The BioClim model had the smallest developmental area under the ROC curve (AUC), and the

MaxEnt produced the greatest probability. However, all basin models developmental AUC values were within a 0.059 range, and each of these AUC values were greater than 0.89 (Table 2).

In the expanded models, including the basin and the 2.25-degree buffer around the basin, my BioClim, GAM, and MaxEnt models predicted 0.09%, 0.0%, and 6.10% of 30 m x 30 m pixels to have a high (greater than 0.66) probability of kudzu habitat suitability, respectively (Table 1; Figures 13, 14, and 15). Similar to the basin models, the BioClim model had the smallest AUC value, and MaxEnt had the greatest. Each of the expanded models had AUCs of greater than 0.85, across a range of 0.057 (Table 2).

Table 1: Percentage of 30 m x 30 m pixels for bioclimate envelope model (BioClim), generalized additive model (GAM), and maximum entropy model (MaxEnt) for within the Great Lakes basin (Basin) as well as the basin and a surrounding 2.25-degree buffer (geographic coordinate system = North American Datum 1983) with low (0.00-0.33), medium (0.33-0.66), and high (0.66-1.00) probability for kudzu habitat suitability (inclusive of the upper limit). Independent variables for all models included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS). Basin models also included reported kudzu sites from the Midwest Invasive Species Information Network (MISIN).

Area	Model	Low %	Med %	High %
Basin	BioClim	99.58	0.40	0.02
	GAM	99.16	0.63	0.21
	MaxEnt	88.47	6.39	5.14
Expanded Area	BioClim	99.00	0.91	0.09
	GAM	99.20	0.80	0.00
	MaxEnt	81.90	12.00	6.10

Table 2: Developmental area under the receiver operator characteristic (ROC) curve values (AUC) for bioclimate envelope model (BioClim), generalized additive model (GAM), and maximum entropy model (MaxEnt) for within the Great Lakes basin (Basin) as well as the basin and a surrounding 2.25-degree buffer (geographic coordinate system = North American Datum 1983).

Area	Model	AUC
Basin	BioClim	0.894
	GAM	0.945
	MaxEnt	0.953
Expanded Area	BioClim	0.852
	GAM	0.874
	MaxEnt	0.909

Table 3: GAM results for the basis dimension (k') and effective degrees of freedom (edf) for continuous independent variables within the basin and expanded area GAM models.

Area	Variable	k'	edf
Basin	Total Annual Precipitation	9.00	2.34
	Aspect	9.00	2.76
	Elevation	9.00	1.00
Expanded Area	Total Annual Precipitation	9.00	1.55
	Aspect	9.00	4.60
	Elevation	9.00	3.99

Table 4: MaxEnt model results for percent contribution and permutation importance for each independent variable within the basin and expanded area MaxEnt models.

Area	Variable	Percent Contribution	Permutation Importance
Basin	Land Cover	69.6	43.4
	Elevation	17.5	50.1
	Total Annual Precipitation	5.3	0.6
	Aspect	3.8	1.4
	Hydrology	3.5	4.0
	Geology	0.3	0.5
Expanded Area	Land Cover	40.4	37.9
	Geology	23.9	10.8
	Elevation	20.2	33.2
	Aspect	7.3	12.7
	Hydrology	6.2	3.3
	Total Annual Precipitation	2.0	2.1

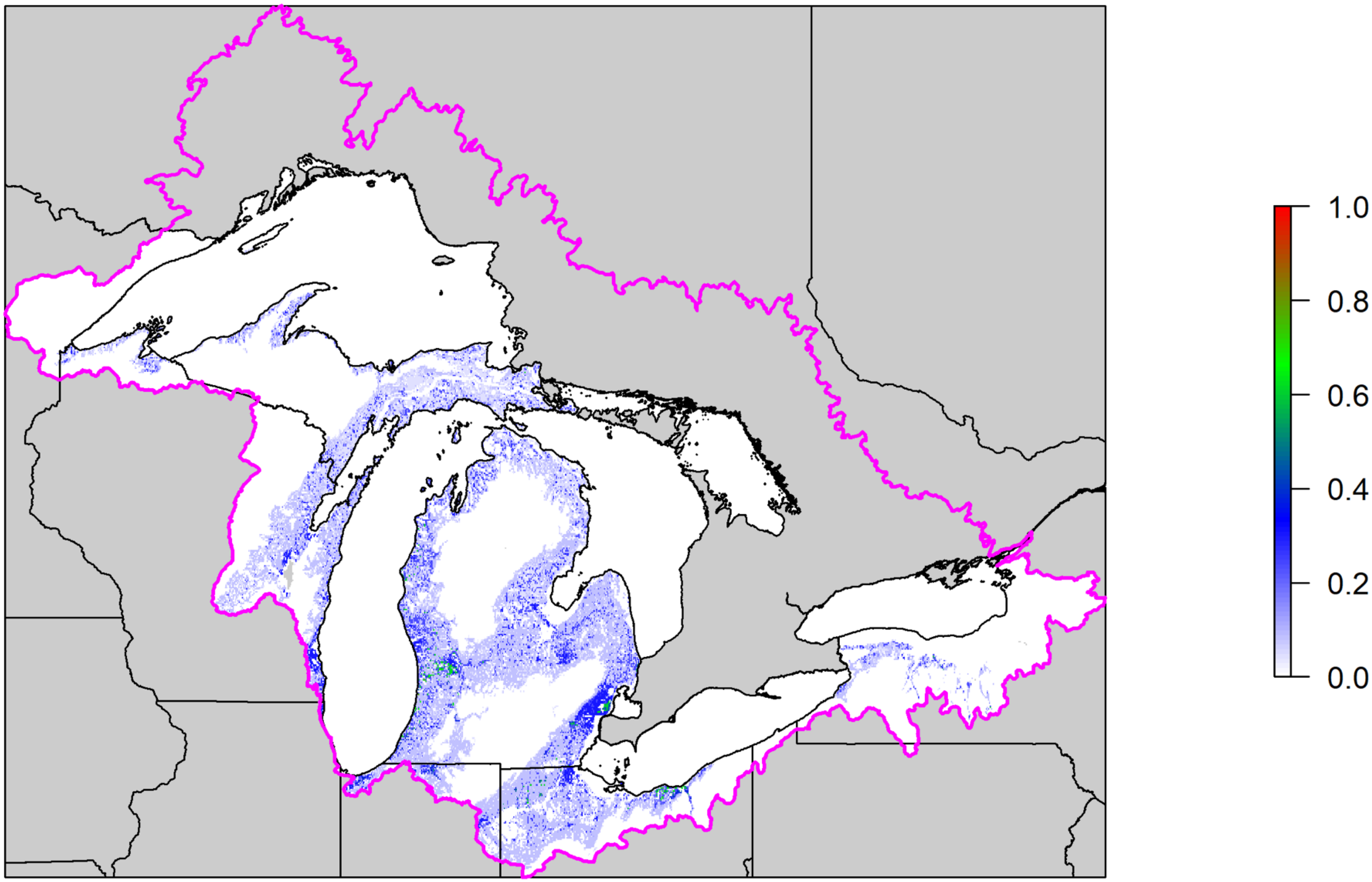

Figure 4: Bioclimate envelope (BioClim) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN)

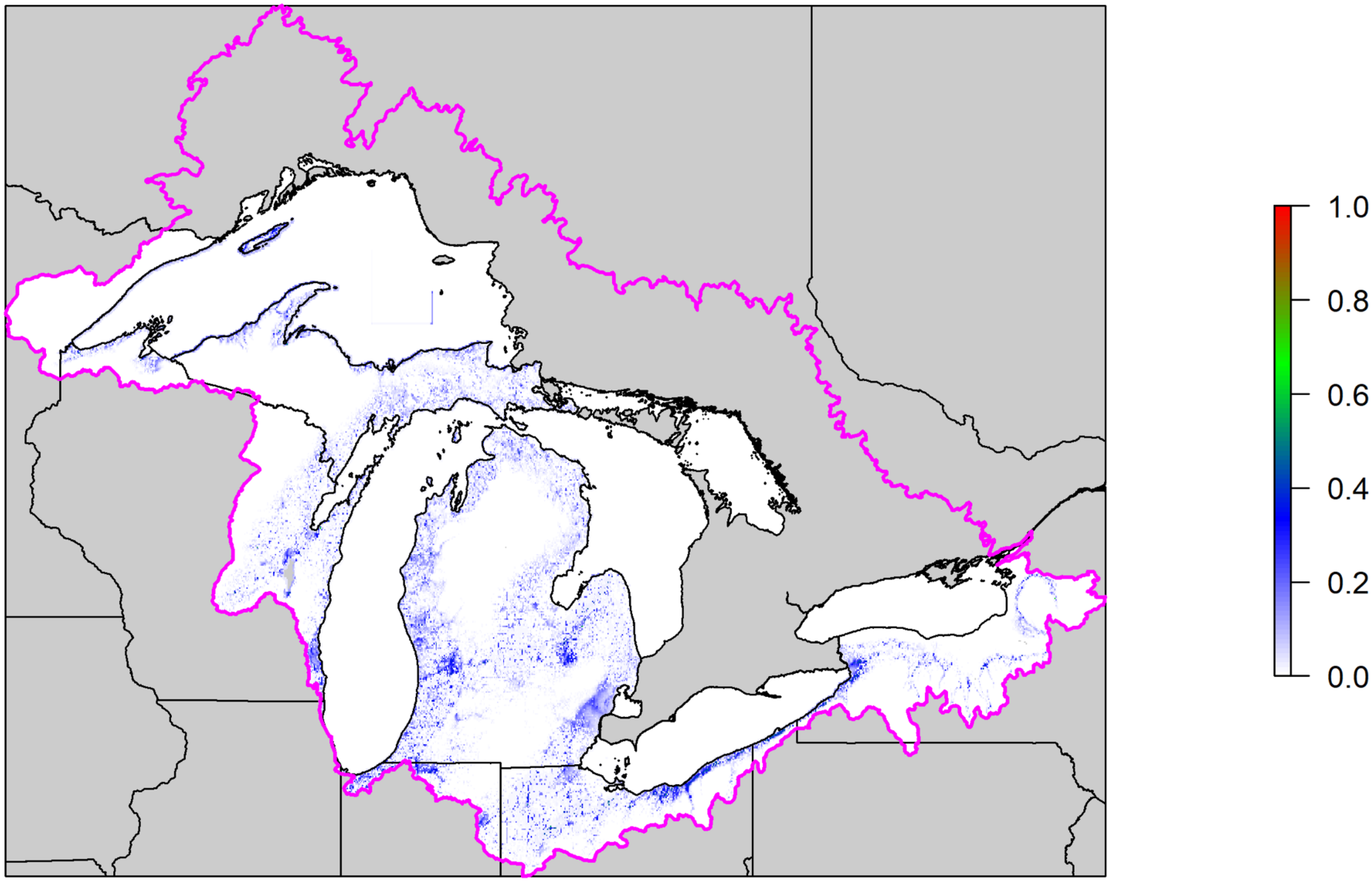

Figure 5: Generalized additive model (GAM) depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN).

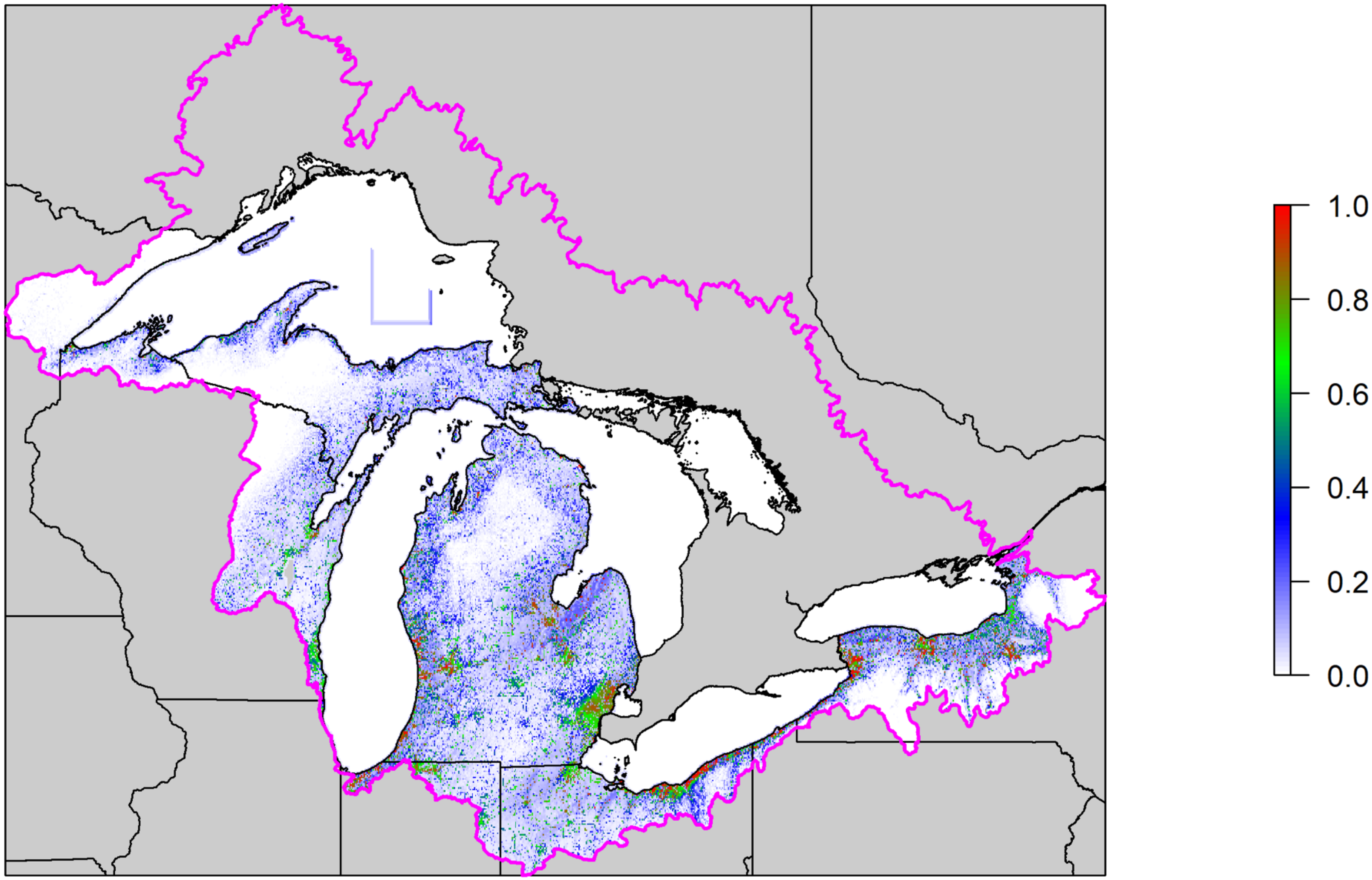

Figure 6: Maximum entropy (MaxEnt) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN).

Figure 7: Bioclimate envelope (BioClim) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN).

Figure 8: Generalized additive model (GAM) depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN).

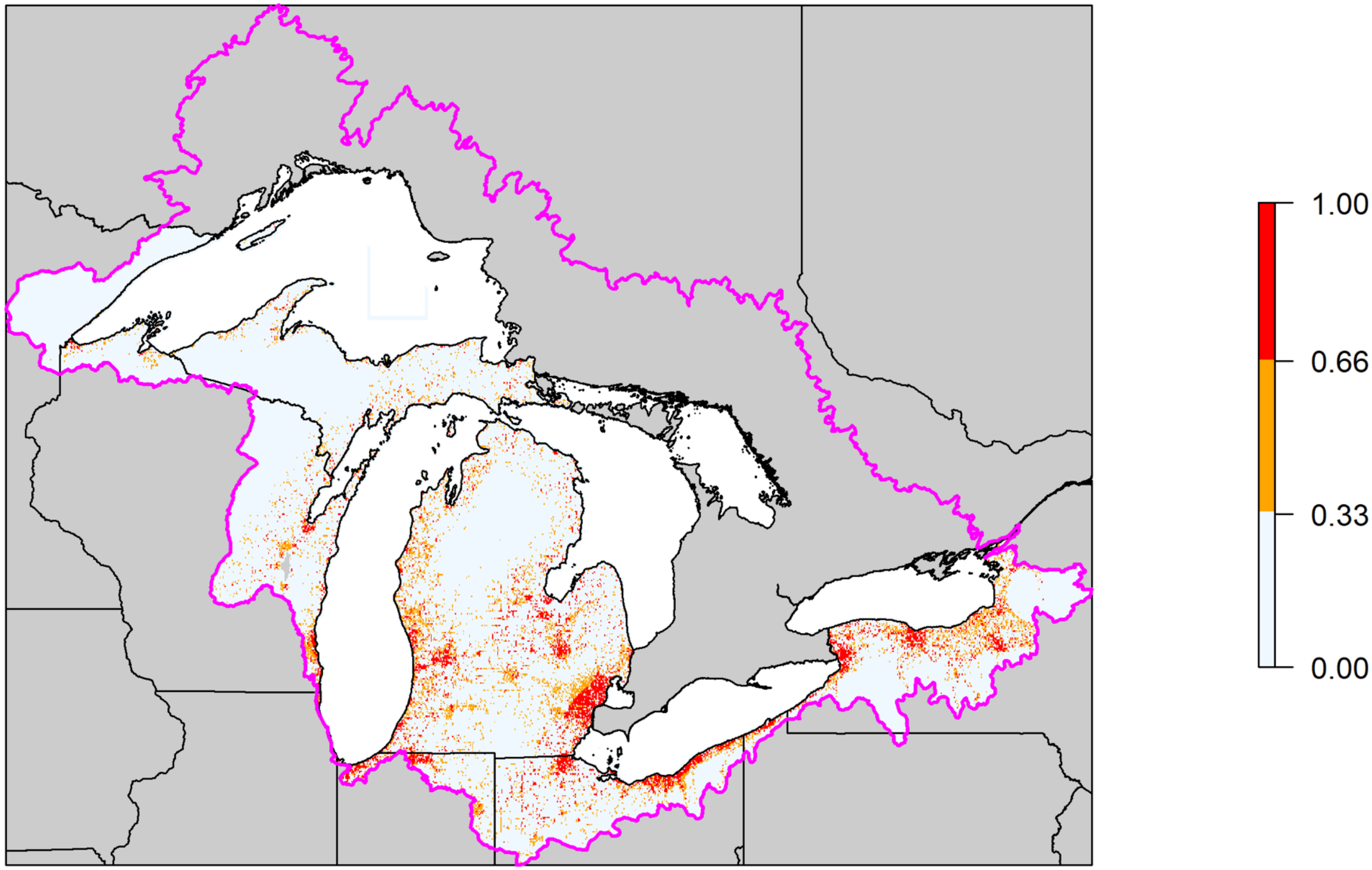

Figure 9: Maximum entropy (MaxEnt) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line). Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS) and the Midwest Invasive Species Information Network (MISIN).

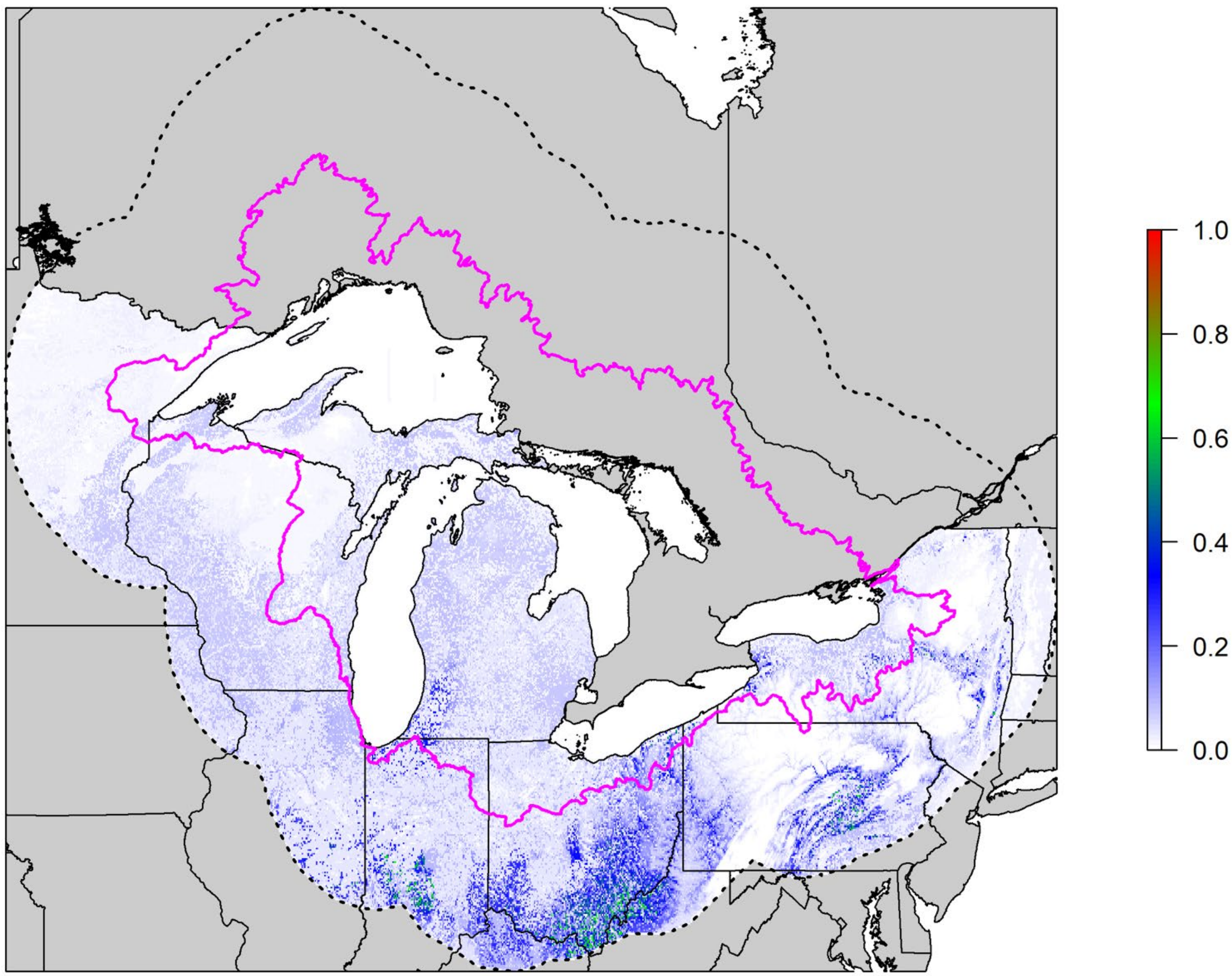

Figure 10: Bioclimate envelope (BioClim) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

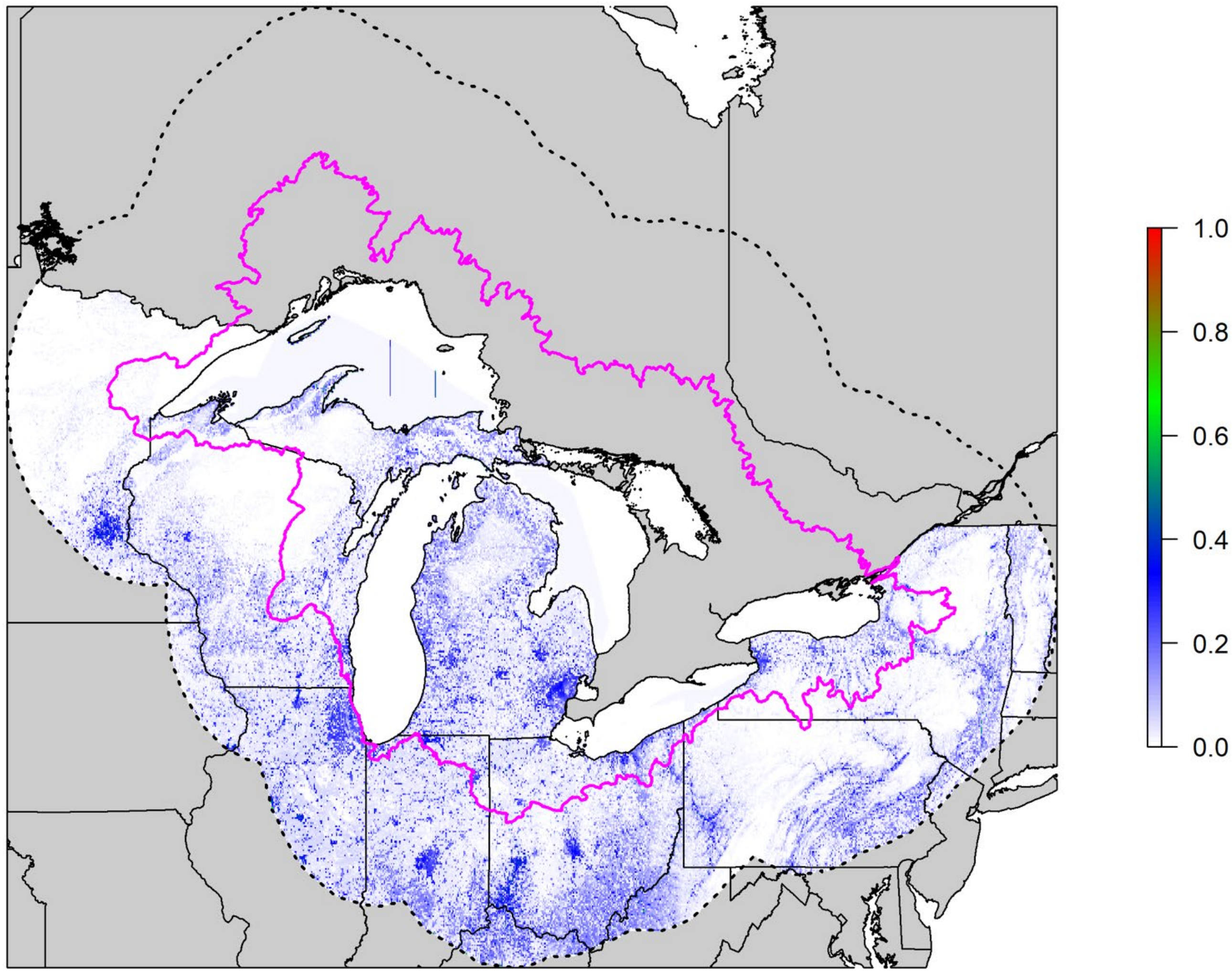

Figure 11: Generalized additive model (GAM) depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

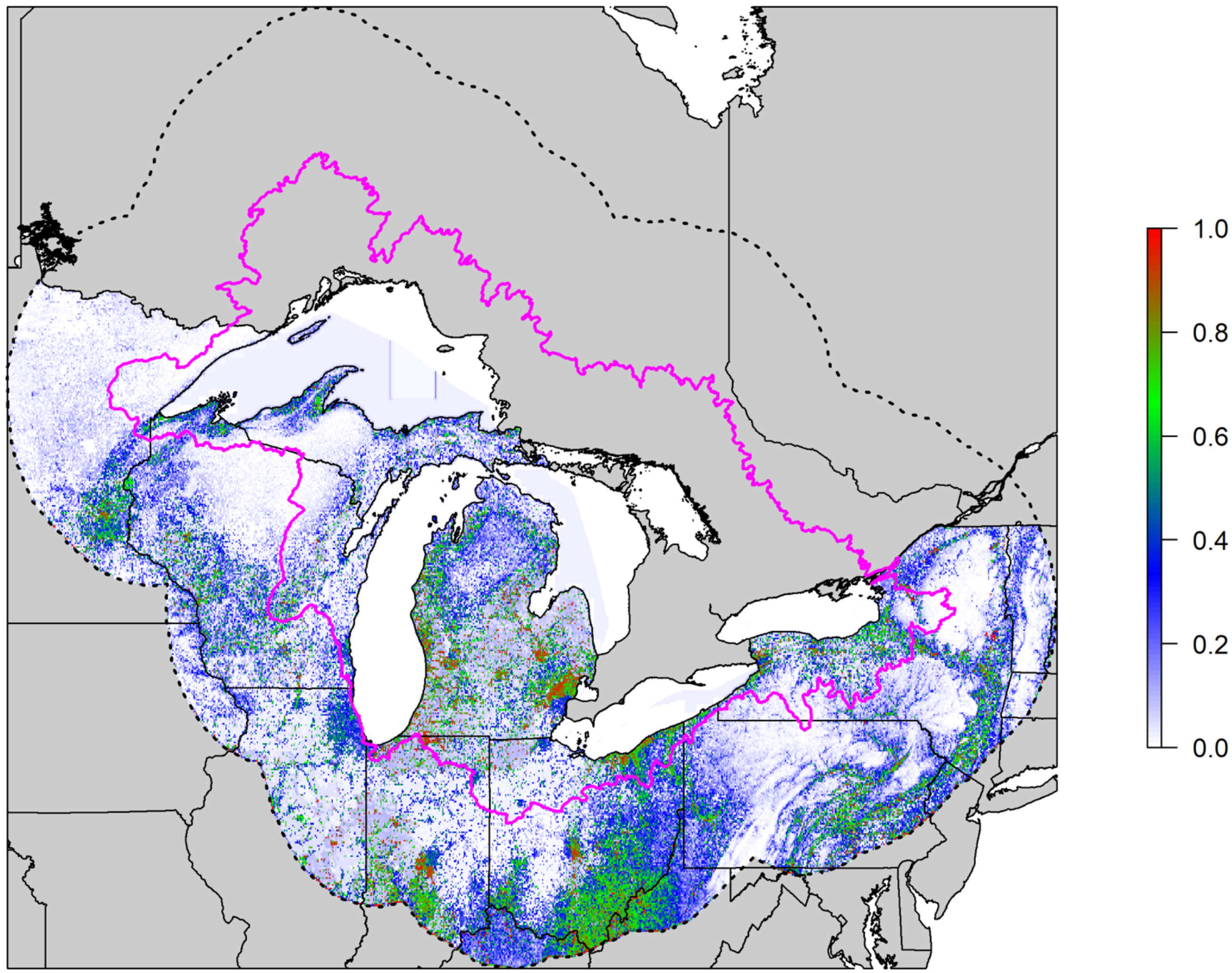

Figure 12: Maximum entropy (MaxEnt) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

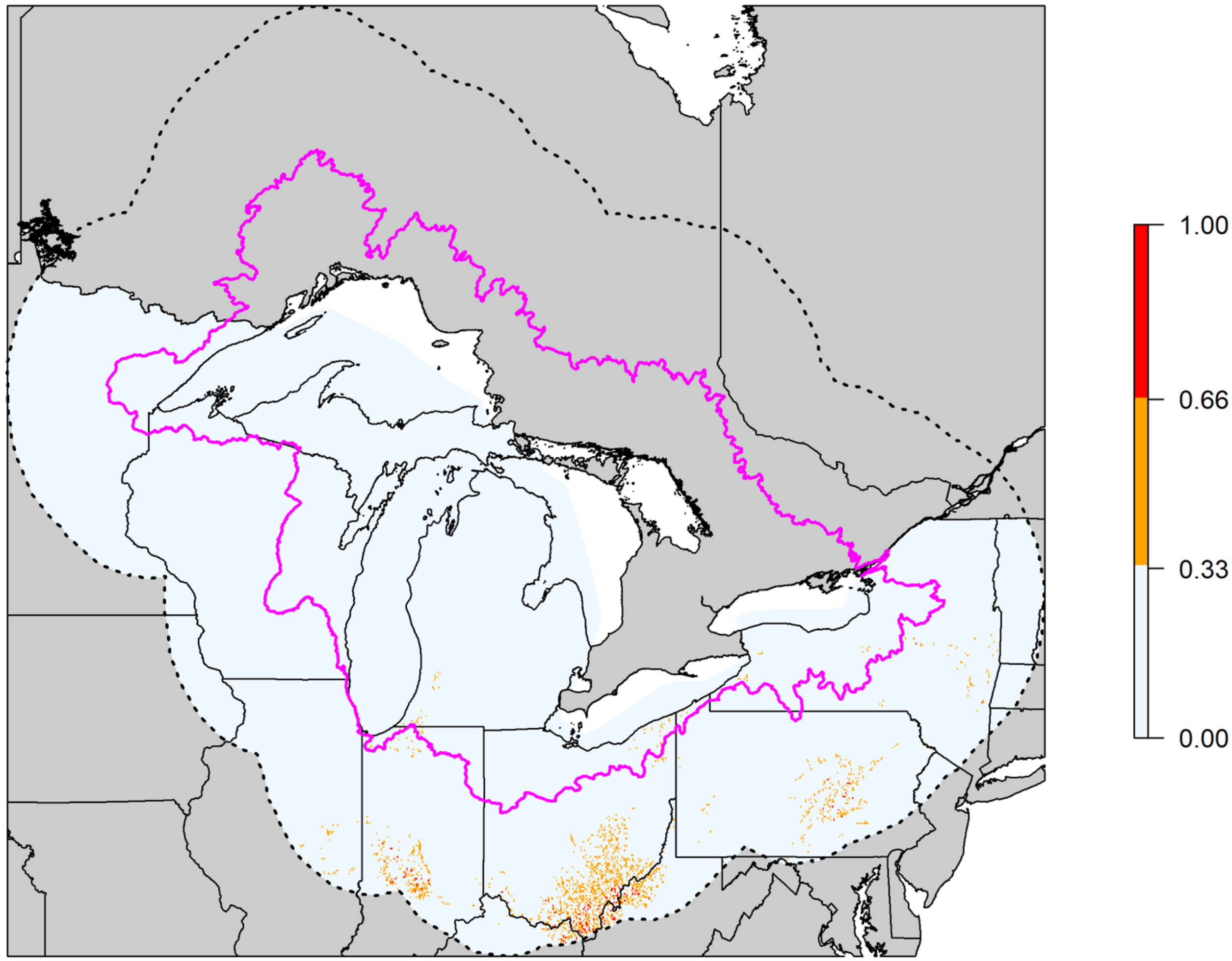

Figure 13: Bioclimate envelope (BioClim) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

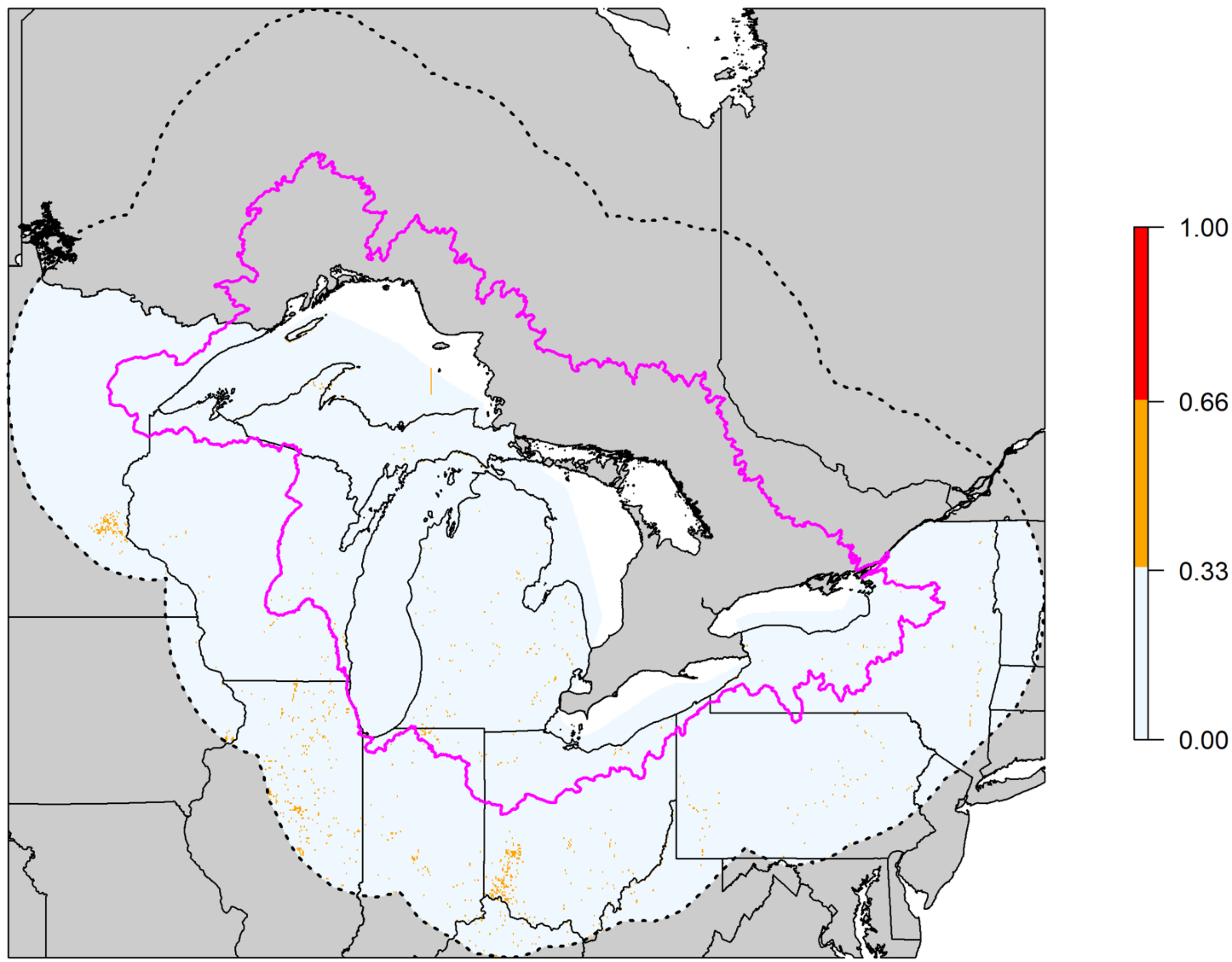

Figure 14: Generalized additive model (GAM) depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

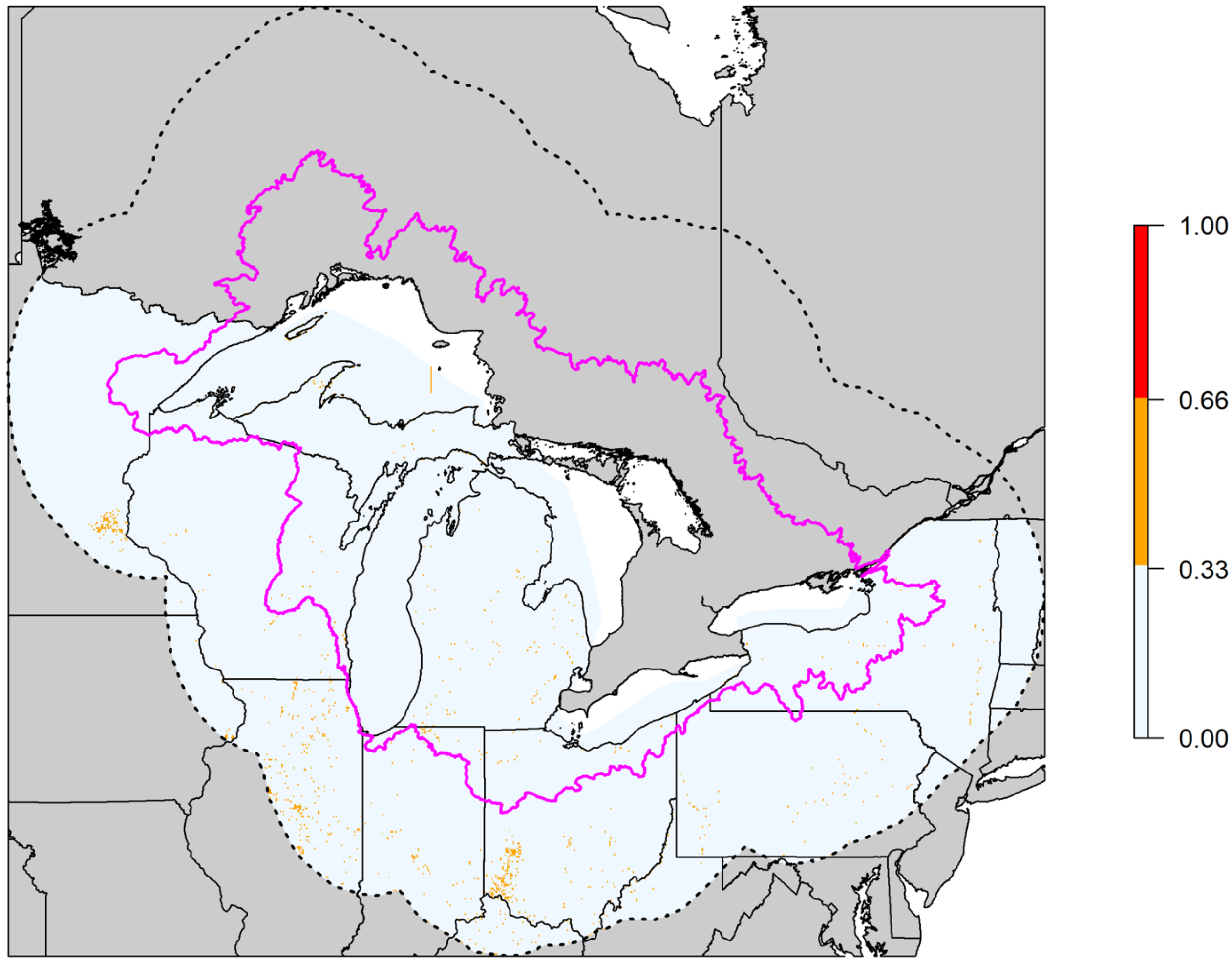

Figure 15: Maximum entropy (MaxEnt) model depicting probability of kudzu habitat suitability probability within 30 m x 30 m pixels within the United States portion of the Great Lakes basin (solid pink line) and a surrounding 2.25-degree (dashed line; coordinate reference system = North American Datum 1983) buffer around the basin. Pixels are defined as low (light blue; 0-0.33, inclusive of upper limit), medium (orange; 0.33-0.66, inclusive of upper limit), or high (red; 0.66-1, inclusive of upper limit) probability. Independent variables included land cover, mean annual precipitation, geology, hydrology, elevation, and aspect layers along with reported kudzu sites from the Early Detection and Distribution Map System (EDDMapS).

3.2 Field Site Vegetation and Environmental Surveys

I was able to visit eight known kudzu sites (four in the basin, four in the buffer). I found no statistically significant differences in local environmental variables (i.e., canopy cover, forest overstory basal area, soil moisture, soil pH, and litter depth) between the kudzu site and the paired site without kudzu using a two-way ANOVA (Table 6). Likewise, I found no statistically significant differences between these environmental characteristics when comparing kudzu presence/absence sites within and outside of the basin. There were also no statistically significant differences in vegetative characteristics (i.e., midstory basal area, midstory richness, midstory diversity, and overstory basal area) between kudzu presence and absence sites nor between inside and outside of the basin (Tables 5 and 6).

Field sites were used for testing the prediction ability of the three spatial models within the basin and in the expanded model area. These field sites were not part of the model development sites. When comparing known kudzu presence and absence sites with model predictions, the maximum entropy model had the highest sensitivity (true positive rate) for both basin and expanded models (0.50 and 0.75, respectively; Table 7). Additionally, the basin and expanded maximum entropy models were the only models to produce a specificity (true negative rate) unequal to one (0.75 and 0.38, respectively; Table 7). The basin generalized additive and bioclimate envelope models predicted all sites (kudzu absent and present) to not have suitable habitat for kudzu (i.e., the predicted probability for each location was less than 0.5), and thus were not able to produce any true positive predictions. Likewise, the expanded generalized additive model predicted all sites to not have suitable habitat for kudzu, and also were not able to produce any true positive predictions. The expanded bioclimate envelope model was able to produce true positive predictions, with sensitivity of 0.25 (Table 7). The BioClim, GAM, and MaxEnt models for the basin all showed alignment between predicted and observed kudzu presence and absence sites (McNemar $X^2 = 0.134$, 0.134, and 0.00, respectively; $p > 0.05$; Table 8). Of the expanded models, only the MaxEnt model showed alignment between predicted and observed kudzu presence and absence sites (McNemar $X^2 = 0.57$; $p > 0.05$; Table 8).

Table 5: Mean canopy cover, soil moisture, soil pH, litter depth, forest basal area, midstory richness, and midstory diversity for kudzu present and absent field sites located inside the Great Lakes basin boundary and outside the boundary.

Location	Kudzu	Canopy Cover (%)	Soil Moisture (%)	Soil pH	Litter Depth (cm)	Forest Basal Area (m^2/ha)	Midstory Richness	Midstory Diversity	Midstory Basal Area (m^2/ha)
Basin	Present	60.9	22.0	7.0	3.5	11.2	2.8	1.5	1.0
	Absent	60.0	33.1	7.1	6.8	10.3	2.8	1.3	4.2
Outside Basin	Present	72.0	23.9	6.9	5.1	13.8	4.0	1.2	3.2
	Absent	82.1	17.3	7.3	5.6	19.2	2.3	1.2	3.0

Table 6: Descriptive statistics for two-Way ANOVA for field survey environmental and vegetative characteristics for kudzu present and absent field sites (Kudzu) located inside the Great Lakes basin boundary and outside the boundary (Location).

Independent Variable	Source of Variation	Df	Mean Square	F-Value	p-value
Canopy Cover	Location	1	1084.7	1.069	0.320
	Kudzu	1	101.8	0.100	0.756
	Interaction	1	130	0.128	0.726
Soil Moisture	Location	1	156.1	0.419	0.531
	Kudzu	1	31	0.083	0.777
	Interaction	1	286.7	0.771	0.399
Soil pH	Location	1	0.00167	0.006	0.940
	Kudzu	1	0.25644	0.895	0.361
	Interaction	1	0.0451	0.157	0.698
Litter Depth	Location	1	0.167	0.020	0.891
	Kudzu	1	13.342	1.558	0.234
	Interaction	1	8.52	0.995	0.337
Overstory Basal Area	Location	1	125.17	0.735	0.407
	Kudzu	1	25.59	0.150	0.704
	Interaction	1	41.95	0.246	0.628
Midstory Species Richness	Location	1	0.944	0.090	0.769
	Kudzu	1	3.572	0.342	0.569
	Interaction	1	3.224	0.308	0.589
Midstory Diversity	Location	1	0.05538	0.183	0.682
	Kudzu	1	0.05064	0.167	0.695
	Interaction	1	0.01816	0.060	0.814
Midstory Basal Area	Location	1	1.693	0.082	0.779
	Kudzu	1	6.62	0.321	0.581
	Interaction	1	14.522	0.704	0.416

Table 7: Sensitivity (true positive rate) and specificity (true negative rate) for bioclimate envelope (BioClim), generalized additive (GAM), and maximum entropy (MaxEnt) models for within the Great Lakes basin (Basin) as well as the basin and a surrounding 2.25-degree buffer (Expanded Area; geographic coordinate system = North American Datum 1983) compared to field data. Eight sites were surveyed within the basin (4 kudzu, 4 control), and 16 sites were surveyed for the expanded area in total (8 kudzu, 8 control).

Area	Model	Sensitivity	Specificity
Basin	BioClim	0.00	1.00
	GAM	0.00	1.00
	MaxEnt	0.50	0.75
Expanded Area	BioClim	0.25	1.00
	GAM	0.00	1.00
	MaxEnt	0.75	0.38

Table 8: McNemar X^2 test results comparing observed (kudzu field survey presence sites) with expected (predicted presence or absence for those sites) for bioclimate envelope (BioClim), generalized additive (GAM), and maximum entropy (MaxEnt) models for within the Great Lakes basin (Basin) as well as the basin and a surrounding 2.25-degree buffer (Expanded Area; geographic coordinate system = North American Datum 1983) compared to field data. Eight sites were surveyed within the basin (4 kudzu, 4 control), and 16 sites were surveyed for the expanded area in total (8 kudzu, 8 control).

Area	Model	McNemar X^2	df	p-value
Basin	BioClim	2.25	1	0.134
	GAM	2.25	1	0.134
	MaxEnt	0.00	1	1.000
Expanded Area	BioClim	4.17	1	0.041
	GAM	6.13	1	0.013
	MaxEnt	0.57	1	0.450

CHAPTER 4. DISCUSSION

4.1 Predicting Habitat Suitability – Great Lakes Basin

I produced three predictive models within the basin, and these models consistently predicted higher kudzu probability near urban areas and areas with high human activity. The bioclimate envelope (BioClim) model predicted the most low-probability (≤ 0.33) pixels out of the three basin models. This model predicted higher probabilities (though still < 0.5) near human-populated and tourist areas, including Holland, Grand Rapids, Detroit, Ludington, and Flint in Michigan, and Toledo and Cleveland in Ohio, agreeing with the predictions from the basin GAM model. These trends are supported by previous studies that have shown that kudzu establishes well in disturbed areas, which are likely to be near human-populated areas (Munger 2002). Additionally, kudzu has had multiple human-mediated introductions as fodder, ornamentals, and for flood control, tying it closely to human populations (Li et al. 2011, Lindgren et al. 2013, Mitich 2000). It is likely that kudzu has stayed close to those areas with higher human populations due to its low seed viability and thus limited long-distance dispersal (Geerts et al. 2016, Munger 2002). Conversely, these higher probabilities could be due to my limited sample size and improved abilities to detect kudzu populations better near areas with higher populations and better invasive species management.

Though the BioClim model did predict the most low-probability pixels, the generalized additive (GAM) model was likely the most conservative model within the Great Lakes basin boundary as the medium- and high-probability (> 0.33) pixels were concentrated along the entire basin boundary. Therefore, the bulk of the basin appears to contain lower probability pixels overall. Like the BioClim model, the GAM model within the basin had higher probabilities near human-populated areas like Grand Rapids, Detroit, and Flint in Michigan as well as Green Bay and Milwaukee in Wisconsin. Additionally, this model predicted higher probabilities along the southern coastlines of Lake Superior and Lake Erie.

The maximum entropy (MaxEnt) model within the basin predicted the most high-probability (> 0.66) pixels for kudzu habitat suitability out of the three basin models. In this model, urban areas tend to have the highest probability of kudzu habitat suitability, with additional high-probability regions along the southern coast of Lake Michigan (including coastal tourist cities like Grand Haven, Holland, South Haven, and Benton Harbor in Michigan, as well as Michigan City, IN and Chicago, IL). The basin MaxEnt model also predicted large areas of high probabilities near Detroit, MI and Milwaukee, WI as well as along the southern coast of Lake Erie, including Cleveland, OH. While these regions agree with the models previously discussed in that urban areas have a higher probability of suitable habitat, the MaxEnt model demonstrates a wider range of high-probability and low-probability areas.

4.2 Predicting Habitat Suitability – Expanded Area

I produced three predictive models across the basin and a 2.25-degree buffer, including 193 known kudzu points. These models tended to produce more pixels with high probability values than the basin models alone, which only had 25 known kudzu points. Expanded models predicted higher probabilities concentrated in urban areas within the northern portions of the landscape, but greater densities of higher probabilities widespread in the southern parts of the study area, where there is already a high-density of existing kudzu populations. The expanded BioClim model showed some higher probabilities (though still < 0.5) within the southwest portion of Michigan, including aforementioned urban areas like Grand Rapids, Holland, Michigan City, IN, and Cleveland, OH. This model also showed substantially more medium probability areas of pixels in southern Indiana and Ohio, along the northern front of kudzu's widespread southeastern distribution where there is a high density of known kudzu points. This suggests that while urban areas may play a role in kudzu habitat suitability, there may be other factors, like kudzu density, driving kudzu habitat suitability in the southern portions of the study area. These ranges of high probabilities are mostly concentrated where there are known kudzu presence points, like the expanded GAM model, suggesting that using this buffer and increasing the number of known points appears to provide more accurate predictions. While these expanded models seem to be more accurate visually, they do provide lower developmental AUC values. This could be due to the increased heterogeneity in the greater area, and thus including other or more refined environmental factors, like soil type or altitude, might help increase these models' accuracy (Mitich 2000).

The generalized additive model was the most conservative model of the expanded models, predicting the most low- probability (≤ 0.33) pixels for habitat suitability. In this model, I did not see the high probability of kudzu habitat suitability along the boundary of the basin nor the boundary of the buffer. This suggests that the basin model was perhaps limited in its ability to predict, which could be due to the limited sample size or differences in drivers within the basin area. This agrees with Liang et al. (2020) who found that sample size may be linked to model accuracy and number of predictions. Like in the basin models, the expanded GAM predicts higher probabilities (though still less than 0.5) near aforementioned urban centers as well as Chicago, IL, Indianapolis, IN, and Minneapolis, MN. The expanded GAM model also predicted medium probabilities (around 0.5) in regions along the southern coast of Lake Superior, Lake Erie, and Lake Ontario as well as the northern coast of Lake Michigan. While some cities reside in these medium probability predictions, these areas are not limited to the population centers of the coast, suggesting coastlines might contain factors that are driving kudzu habitat suitability within this model. There are additional small patches of medium probability pixels in the expanded GAM model within the southern portions of the study area, outside of the Great Lakes basin. This agrees with the substantial number of known kudzu populations within these areas, coming from the northward expansion of kudzu from the Southeastern United States (EDDMapS 2019).

The expanded MaxEnt model predicted the most high-probability (> 0.66) pixels for kudzu habitat suitability out of the three expanded models, similar to the MaxEnt basin model. This model depicted a combination of these high probabilities in both urban areas, like in the expanded GAM model, as well as in the Appalachian Mountain Range and southern portions of the study area, like the expanded BioClim model. Here, the highest concentration of these high probability pixels is within the southeastern portions of Indiana and Ohio, which is where many of the existing kudzu populations are found (EDDMapS 2019). Interestingly, none of the expanded areas showed high probabilities (> 0.66) near Chicago, IL, which is the most populated city (> 2.5 million people) within the study region. This could be due to the limited sample size of known kudzu points around this area, which may limit the models' predictability within this area. However, this does indicate that the models, including MaxEnt, were not limited to only urban and suburban areas for predicting high probabilities of habitat suitability.

4.3 Overall Trends

Overall, the MaxEnt models for both the basin and expanded areas not only predicted the most high-probability pixels for kudzu habitat suitability, but they also had the highest developmental AUC (AUC = 0.953 and 0.909, respectively; Table 2), indicating that these predictive models had the greatest overall fit for the data used. This agrees with previous studies which have also found MaxEnt to be useful models for ecological data (Baldwin 2009, Gastón & García-Viñas 2011, Kaky et al. 2020, Phillips et al. 2006). Additionally, the expanded MaxEnt model had the highest sensitivity (true positive rate) out of all models. This suggests that MaxEnt modelling may be a reasonable candidate to predict areas where kudzu management is most necessary.

In both basin and expanded areas across all models, there was a higher probability of kudzu habitat suitability around populated urban areas. This could be a result of kudzu's ability and preference to invade disturbed areas, like roadsides and forest edges which are more concentrated in these urban areas, combined with kudzu's multiple human-mediated introductions (Lindgren et al. 2016, Mitich 2000, Munger 2002). While higher probabilities of kudzu habitat suitability seem to be mostly concentrated to these urban areas for each the basin models, in the expanded models, kudzu is not as limited to these areas in much of the study area. While kudzu is widespread in the southeastern United States, there are few existing populations in the Midwestern United States, including the northern part of the study area, which may indicate that kudzu spread is limited to existing populations in the north (EDDMapS 2019, MISIN 2021). Additionally, kudzu primarily reproduces through clonal reproduction, which requires existing crowns of kudzu plants (Bentley & Mauricio 2016, Geerts et al. 2016, Lindgren et al. 2013, Munger 2002). The disjunct nature of kudzu records in Michigan (northern most site in Benzie County, ~250 km from the closest location) exemplifies the importance of human transport for kudzu invasion. In the southeastern areas, kudzu populations are more widespread and thus have many existing plants that may increase reproduction rates. Thus, environmental factors, like climate, topography, and hydrology might be larger determining factors here. Despite this discrepancy between the area within the basin and outside of the basin of the study area, I found no differences between in-basin and out-of-basin field environmental measurements, suggesting what is already known about kudzu in the southeast may be used to understand kudzu in the Great Lakes basin. However, further studies with greater sample sizes and a variety of areas outside of the basin should be conducted to determine if existing knowledge on kudzu habitat suitability, like that of the southeastern United States, can be used to accurately predict suitability within the basin. Additionally, kudzu might be limited by temperature in northern areas, though recent studies have found that the vine can grow in colder temperatures than initially believed, suggesting that kudzu may be adapting to these colder climates (Coiner et al. 2018, Lindgren et al. 2013). Because local characteristics did not differ between my kudzu presence and absence locations, kudzu may be dispersal-limited and not habitat-limited. Being dispersal-limited means that not all suitable kudzu habitat will be occupied; however, the introduction of individuals will lead to establishment (Ehrlén & Eriksson 2000). As we continue to study kudzu biology, future studies should also assess whether kudzu habitat suitability can be predicted by these environmental factors or if, alternatively, kudzu is not limited by habitat characteristics due to its generalist nature.

4.4 Future Directions

These models appear to confirm existing knowledge about kudzu habitat environments and spread, but further refinement is needed in order to use such models at the land management scale. While MaxEnt seems like a promising approach, other environmental factors at finer resolutions might be necessary to fully understand kudzu habitat suitability (Mitich 2000). Additionally, these models only have a resolution of 30 m x 30 m due to the resolution of available data such as land cover and climate variables, which is impractical for land management use. As higher resolution data become available, these predictive models should be refined in order to understand and predict locations of potential kudzu habitat. Additionally, other model approaches may be necessary in order to predict kudzu habitat and spread. Lattice models might be useful to predict the spread of kudzu, as they predict the probability of introduction to areas adjacent to existing populations and kudzu is known to reproduce through existing crowns (Bentley & Mauricio 2016, Lindgren et al. 2013, Munger 2002). Such a model would help predict the probability of invasion itself, rather than habitat suitability alone. Ultimately, these models should also be used in conjunction with future climate predictions to show how the habitat availability for kudzu in the Great Lakes basin might evolve in the face of climate change. Previous studies have shown that kudzu is expanding northward, with the midwestern and northeastern United States at greatest risk for kudzu invasion with respect to climate change (Bradley et al. 2010, Callen & Miller 2015, Jarnevich & Stohlgren 2009, Lindgren et al. 2013). However, such models need to be refined to specific regions, like the Great Lakes basin, and then to the management scale in order to be useful in invasive species management.

REFERENCES

Abramovitz JN (1983). *Pueraria lobata* Willd. (OHWI), Kudzu: Limitations to Sexual Reproduction. MS thesis. University of Maryland, College Park, MD.

Aurambout JP, Endress AG (2018). A model to simulate the spread and management cost of kudzu (*Pueraria montana* var. *lobata*) at landscape scale. Ecological Informatics 43: 146-156.

Baldwin RA (2009). Use of maximum entropy modeling in wildlife research. Entropy 11: 854-866.

Bentley KE, Mauricio R (2016). High degree of clonal reproduction and lack of large-scale geographic patterning mark the introduced range of the invasive vine, kudzu (*Pueraria montana* var. *lobata*), in North America. American Journal of Botany 103(8): 1499-1507.

Boyette CD, Walker HL, Abbas HK (2002). Biological control of kudzu (*Pueraria lobata*) with an isolate of *Myrothecium verrucaria*. Biocontrol Science and Technology 12: 75-82.

Bradley BA, Wilcove DS, Oppenheimer M (2010). Climate change increases risk of plant invasion in the Eastern United States. Biological Invasions 12: 1855-1872.

Britton KO, Orr D, Sun J (2002). Kudzu In: Van Driesche R, Lyon S, Blossey B, Hoddle M, Reardon R (tech. coords.) Biological Control of Invasive Plants in Eastern United States. USDA Service Publication FHTET-2002-04, 413.

Britton KO, Sun J, Orr D, Yang Z, Ping C, Mingyi T, Zi-de J (2001). Potential biocontrol agents for kudzu from China. Pp. 24-27 in Fosbroke SLC, Gottschalk KW (eds.). Proceedings U.S. Department of Agriculture Interagency Research Forum on Gypsy Moth and Other Invasive Species. Forest Service, Northeastern Research Station. General Technical Report NE-285.

Brown KH (2008) Fair Japan: Japanese gardens at American World's Fairs, 1876-1940. SiteLINES: A Journal of Place 4: 13-16.

Callen ST, Miller AJ (2015). Signatures of niche conservatism and niche shift in the North American kudzu (*Pueraria montana*) invasion. Diversity and Distributions 21: 853-863.

Coiner HA (2012). The role of low temperatures in determining the northern range limit of kudzu (*Pueraria montana* var. *lobata*), an invasive vine in North America. PhD dissertation. University of Toronto.

Coiner HA, Hayhoe K, Ziska LH, Van Dorn J, Sage RF (2018). Tolerance of subzero winter cold in kudzu (*Pueraria montana* var. *lobata*). Oecologia 187: 839-849.

De Vos JM, Joppa LN, Gittleman JL, Stephens PR, Pimm SL (2015) Estimating the normal background rate of species extinction. Conservation Biology 29: 452-462.

EDDMapS (2019). Early Detection & Distribution Mapping System. The University of Georgia - Center for Invasive Species and Ecosystem Health. Available online at http://www.eddmaps.org/.

Ehrlén J, Eriksson O (2000). Dispersal limitation and patch occupancy in forest herbs. Ecology 81: 1667-1674.

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson KS, Schchetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006). Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

Fattorini S, Borges PAV (2012). Species-area relationships underestimate extinction rates. Acta Oecologica 40: 27-30.

Forseth Jr. IN, Innis AF (2004). Kudzu (*Pueraria montana*): History, physiology, and ecology combine to make a major ecosystem threat. Critical Reviews in Plant Sciences 23(5): 401-413.

Forseth Jr. IN, Teramura, AH (1987). Field photosynthesis, microclimate and water relations of an exotic temperate liana, *Pueraria lobata*, kudzu. Oecologia 71: 262-267.

Frye MJ, Hough-Goldstein J, Sun J-H (2007). Biology and preliminary host range assessment of two potential kudzu biological control agents. Environmental Entomology 36: 1430-1440.

Gastón A, García-Viñas JI (2011). Modelling species distributions with penalized logistic regressions: A comparison with maximum entropy models. Ecological Modelling 222: 2037-2041.

Geerts S, Mashele BV, Visser V, Wilson JRU (2016). Lack of human-assisted dispersal means *Pueraria montana* var. *lobata* (kudzu vine) could still be eradicated from South Africa. Biological Invasions 18: 3119-3126.

Guertin PJ, Denight ML, Gebhart DL, Nelson L (2008). Invasive species biology, control, and research. Part 1: Kudzu (*Pueraria montana*). Engineer Research and Development Center. US Army Corps of Engineers.

Guisan A, Edwards Jr. TC, Hastie T (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157: 89-100.

Guisan A, Thuiller W (2005). Predicting species distribution: offering more than simple habitat models. Ecology Letters 8: 993-1009.

Harris DB, Gregory SD, Brook BW, Ritchie EG, Croft DB, Coulson G, Fordham DA (2014). The influence of non-climate predictors at local and landscape resolutions depends on the autecology of the species. Austral Ecology 39: 710-721.

Harron P, Joshi O, Edgar CB, Paudel S, Adhikari A (2020). Predicting Kudzu (*Pueraria montana*) spread and its economic impacts in timber industry: A case study from Oklahoma. PLoS ONE 15(3): e0229835.

He F, Hubbell SP (2011). Species-area relationships always overestimate extinction rates from habitat loss. Nature 473: 368-371.

Heikkinen RK, Luoto M, Araujo MB, Virkkala R, Thuiller W, Sykes MT (2006). Methods and uncertainties in bioclimatic envelop modelling under climate change. Progress in Physical Geography 30: 751-777.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2021). Dismo: species distribution modeling. R package version 1.3-5. https://CRAN.R-project.org/package=dismo.

Hickman JE, Wu S, Mickley LJ, Lerdau MT (2010). Kudzu (*Pueraria montana*) invasion doubles emissions of nitric oxide and increases ozone pollution. The Proceedings of the National Academy of Sciences 107: 10115-10119.

Hickman JE, Lerdau MT (2014). Biogeochemical impacts of the northward expansion of kudzu under climate change: the importance of ecological context. Ecosphere 4(10): 121.

Jarnivich CS, Stohlgren TJ (2009). Near term climate projections for invasive species distributions. Biological Invasions 11: 1373-1379.

Kaky E, Nolan V, Alatawi A, Gilbert F (2020). A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: A case study with Egyptian medicinal plants. Ecological Informatics 60: 101150.

Kwon S, Park J, Kim J, Kang J, Cho Y, Lim C, Parker M, Lee G (2005). Phylogenetic analysis of the genera *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* on the basis of 16S rRNA gene and internally transcribed spacer region sequences. International Journal of Systematic and Evolutionary Microbiology 55: 263-270.

Li Z, Dong Q, Albright TP, Guo Q (2011). Natural and human dimensions of a quasi-wild species: the case of kudzu. Biological Invasions 13: 2167-2179.

Liang W, Abidi M, Carrasco L, McNelis J, Tran L, Li Y, Grant J (2020). Mapping vegetation at species level with high-resolution multispectral and lidar data over a large spatial area: A case study with kudzu. Remote Sensing 12(4): 609.

Lindgren CJ, Castro KL, Coiner HA, Nurse RE, Darbyshire (2013). The biology of invasive alien plants in Canada. 12. *Pueraria montana* var. *lobata* (Willd.) Sanjappa & Predeep. Canadian Journal of Plant Science 93: 71-95.

Lodge DM, Williams S, MacIssac HJ, Hayes KR, Leung B, Reichard S, Mack RN, Moyle PB, Smith M, Andow DA, Carlton JT, McMichael A (2006). Biological invasions: Recommendations for US policy and management. Ecological Applications 16: 2035-2054.

Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-being: Biodiversity Synthesis. World Resources Institute, Washington, DC.

MISIN (2021). Midwest Invasive Species Information Network. Michigan State University – Department of Entomology. Available online at http://www.misin.msu.edu.

Mitich LW (2000). Kudzu [*Pueraria lobata* (Willd.) Ohwi] 1. Weed Technology 14(1): 231-235.

Munger GT (2002). *Pueraria montana* var. *lobata*. In: Fire Effects Information System. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. Available online at https://www.fs.fed.us/database/feis/plants/vine/puemonl/all.html.

Myers MC, Bowden RA, Hardisty FE (1938). Stimulation of kudzu cuttings. Science 88: 67.

NOAA (2022). Climate Data Online: Benzie County, MI (FIPS:26019). National Centers for Environmental Information. Available online at https://www.ncdc.noaa.gov/cdo-web/

Parker IM, Simberloff D, Lonsdale WM, Goodell K, Wonham M, Kareiva PM, Williamson MH, Von Holle B, Moyle PB, Byers JE and Goldwasser L (1999). Impact: towards a framework for understanding the ecological effects of invaders. Biological Invasions 1: 3–19.

Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017). Opening the black box: an open-source release of Maxent. Ecography 40:887-893.

Pieters AJ (1932). Kudzu, a forage crop for the Southeast. U.S. Department of Agriculture, Leaflet No. 91.

Pimentel D, Zungia R, Morrison D (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 52: 273-288.

QGIS.org (2020) QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.org

R Development Core Team (2020). R: A language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/.

Ricciardi A (2003) Predicting the impacts of an introduced species from its invasion history: an empirical approach applied to zebra mussel invasions. Freshwater Biology 48: 972-981.

Simpson A, Eyler MC (2018). First comprehensive list of non-native species established in three major regions of the United States: US Geological Survey Open-File Report 2018-1156, https://doi.org/10.3133/ofr20181156.

Simpson A, Turner R, Blake R, Liebhold A, Dorado M (2021). United States Register of Introduced and Invasive Species: US Geological Survey data release, https://doi.org/10.5066/P95XL09Q.

Stork NE (2009). Re-assessing current extinction rates. Biodiversity Conservation 19: 357-371.

Vellend M, Verheyen K, Jacquemyn H, Kolb A, Van Calster H, Peterken G, Hermy M (2006) Extinction debt of forest plants persists for more than a century following habitat fragmentation. Ecology 8: 542-548.

Winberry JJ, Jones DM (1973). Rise and decline of the “miracle vine”: Kudzu in the southern landscape. Southeastern Geographer 13(2): 61-70.

Wood SN (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73:3-36.

Zhang Y, Hanula JL, Horn S (2012). The biology and preliminary host range of *Megacopta cribraria* (Heteroptera: Plataspidae) and its impact on kudzu growth. Environmental Entomology 41: 40-50.

APPENDIX A. ENVIRONMENTAL LAYERS FOR HABITAT SUITABILITY MODELS

Table A1: Environmental layers used in habitat suitability model creation including dataset where each layer was obtained and data type for each layer.

Layer	Dataset	Data Type
Land Cover	USGS National Landcover Database	Categorical; 20 land cover categories
Total Annual Precipitation	Historical Climate Data	Continuous
Mean Annual Temperature	Historical Climate Data	Continuous
Geology	USGS State Geological Map Compilation	Categorical; 66 structure feature categories
Hydrology	USFWS National Wetlands Inventory	Binary; water presence and absence
Elevation	Historical Climate Data	Continuous

APPENDIX B. MIDSTORY SPECIES INFORMANTION

Table B1: Mean number of stems for midstory species in kudzu presence (Kudzu) and absence (Control) sites within the basin and outside of the basin boundary.

	In Basin		In Buffer	
Species	Kudzu	Control	Kudzu	Control
Acer saccharum Marshall	0.00	0.00	1.00	0.75
Carya ovata (Mill.) K. Koch	0.00	0.00	0.00	0.25
Cornus rugosa Lam.	0.75	0.00	0.00	0.00
Fagus grandifolia Ehrh.	0.00	0.00	0.25	0.50
Fraxinus pennsylvanica Marshall	0.50	1.00	2.75	0.50
Impatiens capensis Meerb.	1.25	0.50	0.00	0.00
Juglans nigra L.	0.25	0.00	0.00	0.00
Lonicera x *bella* Zabel [*marrowii* x *tatarica*]	0.00	0.50	0.00	0.00
Lonicera maackii (Rupr.) Herder	0.00	0.50	0.00	0.25
Ostrya virginiana (Mill.) K. Koch	0.00	0.00	0.25	0.00
Prunus serotina Ehrh.	0.00	0.25	0.00	0.00