2019

A GIS MODEL FOR APIARY SITE SELECTION BASED ON PROXIMITY TO NECTAR SOURCES UTILIZED IN VARIETAL HONEY PRODUCTION ON FORMER MINE SITES IN APPALACHIA

Douglass W. Potter
University of Kentucky, dougpotter7@gmail.com
Digital Object Identifier: https://doi.org/10.13023/etd.2019.098

Recommended Citation

Potter, Douglass W., "A GIS MODEL FOR APIARY SITE SELECTION BASED ON PROXIMITY TO NECTAR SOURCES UTILIZED IN VARIETAL HONEY PRODUCTION ON FORMER MINE SITES IN APPALACHIA" (2019). *Theses and Dissertations--Forestry and Natural Resources*. 46.
https://uknowledge.uky.edu/forestry_etds/46

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Douglass W. Potter, Student

Dr. Jian Yang, Major Professor

Dr. Steven J. Price, Director of Graduate Studies

A GIS MODEL FOR APIARY SITE SELECTION BASED ON PROXIMITY TO NECTAR SOURCES UTILIZED IN VARIETAL HONEY PRODUCTION ON FORMER MINE SITES IN APPALACHIA

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in Forest and Natural Resource
Sciences in the College of Agriculture, Food and Environment
at the University of Kentucky

By

Douglass W Potter

Lexington, Kentucky

Co- Directors: Dr. Jian Yang, Professor of Forest Landscape Ecology

and Dr. Chris Barton, Professor of Forest Hydrology

Lexington, Kentucky

2019

ABSTRACT OF THESIS

A GIS MODEL FOR APIARY SITE SELECTION BASED ON PROXIMITY TO NECTAR SOURCES UTILIZED IN VARIETAL HONEY PRODUCTION ON FORMER MINE SITES IN APPALACHIA

Beekeepers in Appalachia market varietal honeys derived from particular species of deciduous trees; however, finding places in a mountainous landscape to locate new beeyards is difficult. Site selection is hindered by the high up-front costs of negotiating access to remote areas with limited knowledge of the available forage. Remotely sensed data and species distribution modeling (SDM) of trees important to beekeepers could aid in locating apiary sites at the landscape scale. The objectives of this study are *i)* using publicly available forest inventory data, to model the spatial distribution of three native tree species that are important to honey producers in eastern Kentucky: American Basswood, Sourwood and Tulip Poplar, and to assess the accuracy of the models, *ii)* to incorporate a method for discounting the value of a nectar resource as a function of distance based on an energetic model of honeybee foraging, and *iii)* to provide an example by ranking potential apiary locations around the perimeter of a mine site in the study area based on their proximity to probable species habitat using a GIS model.

Logistic regression models were trained using presence-absence records from 1,059 USFS Forest Inventory and Analysis (FIA) sub-plots distributed throughout a 9,000 km^2 portion of the Kentucky River watershed. The models were evaluated by applying them to a separate dataset, 950 forest inventory sub-plots distributed over a 40.5 km^2 research forest maintained by the University of Kentucky. Weights derived from an energic model of honeybee foraging were then applied to the probabilities of tree species occurrence predicted by the SDM. As an example, 24 potential apiary locations around the perimeter of a reclaimed mine site were selected and then ranked according to a site suitability index. Three tributary areas corresponding to different honeybee flight ranges were considered: 500m, 700m, and 1,200m. Results confirm that rankings are dependent on the foraging range considered, suggesting that the number of colonies at an apiary location would be an important factor to consider when choosing a site. However, the methodology makes assumptions that are only anecdotally supported, notably *i)* that colonies will forage preferentially at the target species when it is in bloom and, *ii)* that foragers will exhaust resources closest to the hive first, regardless of patch size. Additional study of how bees deplete the nectar resources surrounding an apiary is needed to verify the usefulness of SDM in site selection for varietal honey production.

KEYWORDS: precision apiculture, apiary site selection, habitat suitability modeling, species distribution modeling, *Tilia* americana, *Apis* melifera, forest inventory analysis.

Doug Potter

February 22, 2019

A GIS MODEL FOR APIARY SITE SELECTION BASED ON PROXIMITY TO NECTAR SOURCES UTILIZED IN VARIETAL HONEY PRODUCTION ON FORMER MINE SITES IN APPALACHIA

By
Douglass W Potter

Dr. Jian Yang
Co-Director of Thesis

Dr. Chris Barton
Co-Director of Thesis

Dr. Steven J. Price
Director of Graduate Studies

4/30/2019
Date

ACKNOWLEDGMENTS

I would like to thank my wife, Tammy Horn, who started Coal Country Beeworks and has shared her love of bees with me. Also, my parents, Genie and John Potter, as well as the members of my committee: Dr. Yang, Dr. Barton and Dr. Rittschof.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

The locations of apiaries (yards where beehives are kept) are critically linked to the success of any beekeeping operation. Among smaller scale beekeepers, it is considered axiomatic that setting will largely determine a hive’s productivity (Sponsler & Johnson, 2015). However, little is known about the site-selection process employed by smaller scale beekeepers. In Appalachia, it is likely that most beekeepers do not engage in any formal site suitability analysis but rely instead on trial and error to assess the quality of sites (personal communication with Kentucky State Apiarist). Central Appalachia is a highly dissected and heterogenous landscape in which empirical approach to selecting apiary sites is challenging – trial and error may be the only viable way to assess forage on such complicated landscapes. However, spatial analysis of honeybee forage in other geographical regions has shown a correlation between remotely sensed landcover and honey production (Kirkpatrick, 2015). Accurately assessing the local forage is especially critical when attempting to collect a mono-floral or “varietal” honey (Campbell & Fearns, 2018) which is derived primarily from the nectar of a single species (Bryant & Jones, 2001). Beekeepers in Appalachia have met with some success marketing varietal honeys derived from particular species of melliferous trees (Mattise, 2014); however, site selection may be hindered by the challenge of negotiating access to remote sites with limited knowledge of the available forage. Modeling the spatial distribution of varietal honey trees can help select among potential apiary sites at the landscape scale. A map would allow beekeepers to remotely evaluate the concentration of nectar resources at multiple locations before attempting to negotiate access and ground-truthing the sites.

Honeybees have a sophisticated system for allocating resources whereby scouting foragers search and locate areas of relatively intense nectar flow (Thenius, et al., 2006).

Upon returning to the hive, scouts communicate this information to other foragers via 'dances' on the face of the comb, indicating the direction and distance to the nectar source with respect to the sun's position (von Frisch, 1967). The vigor and number of dances serve as a recruiting mechanism that concentrate or reallocate the field force based on current conditions (Seeley, 1986). This allows foragers to target the most productive nectar patches on the landscape rather than indiscriminately consuming from the flowering plants they encounter close to the hive (Seeley, 1994).

Honeybees forage in a range that varies according to nectar availability (Beekman & Ratnieks, 2000). Colonies are adept at ramping up collection when conditions are favorable and curtailing activity to conserve energy when conditions deteriorate (Danner et al., 2016). Because of their ability to optimize foraging strategies (Stabentheiner & Kovac, 2016) and to store these resources in the form of honey, colonies in close proximity to major nectar sources will fully exploit them by recruiting reserve foragers (Anderson, 2001). If present within flight range, colonies will invest heavily in the most energetically profitable nectar sources that may only be available during the relatively short anthesis of particular plant species.

Commercial beekeepers take advantage of these windows of opportunity, called "nectar flows" in the industry (Wainselboim et al., 2002). By placing hives in the right place at the right time, the beekeeper helps create the conditions whereby the colony collects nectar over and above its annual caloric requirements. It is this surplus that makes up the honey crop. Aside from the husbandry associated with general hive health, a beekeeper whose primary objective is honey production will manage colonies such that they achieve maximum forager population during the few weeks in the year when nectar is available in copious quantities (Van Engelsdorp & Meixner, 2010). Colonies managed primarily for pollination contracts, queen production, or other hive products (wax, pollen, royal jelly) may employ different management techniques, but for the stationary honey producer, location largely determines output (Pilati & Prestamburgo, 2016). An explicit representation of how individual nectar sources are distributed on the landscape could help beekeepers place apiaries on sites that have highest potential for honey production or steer them toward the sites most appropriate for a specific varietal honey.

Researchers have previously analyzed honeybee forage on the landscape using a Geographic Information System (GIS). A GIS model was employed to analyze the suitability of floral and commercial crop resources for honey production in the Prairie Pothole Region of North Dakota (Kirkpatrick, 2015). The impact of land-use change was highlighted in appraising the suitability of established sites used for honey production in the Northern Great Plains (Otto et al., 2016). In Ohio, researchers found that honey harvest was accurately predicted by the landscape composition surrounding existing apiaries (Sponsler & Johnson, 2015). These studies were primarily concerned with accounting for the probable forage types encountered in a mixed agricultural setting and did not attempt to model the distribution of individual tree species.

Australian researchers have created a web-based application for beekeepers that helps show the phenology of several species of eucalyptus (Arundel et al., 2016). In Western Australia, satellite data was used to detect the flowering of the Red Gum tree (Campbell & Fearns, 2018). However, no such work has been done in the Appalachians or on any forest-based resources in North America. The species rich ‘mixed-mesophytic’ forest (Braun, 1951) and highly dissected landscape make the use of visual indices derived from satellites challenging for species level identification in Eastern Kentucky.

Mapping the distribution of an individual tree species is often undertaken at a larger spatial scale. USGS researchers have created a dataset estimating the basal area per hectare of all major tree species in the Eastern United States (Wilson et al., 2012); however, the coarse spatial resolution and high standard errors indicate that they are not intended for small area estimation and would not be accurate at a scale that is useful to beekeepers. Some researchers have mapped species at the individual tree crown level with greater than 90% accuracy (Lee, et al. 2016), but there is a correlation between resolution and cost. Work in the field to date has utilized hyperspectral imagery (Dalponte et al., 2011; Ghiyamat et al., 2015; Ferreira et al., 2016) and/or Light Detection and Ranging (LIDAR) data (Engler et al., 2013; Asner & Feret, 2012) to identify a single tree species. These kinds of remotely sensed data are cost prohibitive for beekeepers and unavailable in our study area. Some research (Henderson et al., 2014; Gao et al., 2015) has indicated that a nominal level of accuracy can still be achieved with the use of moderate resolution remotely sensed data.

Several native tree species are of interest to beekeepers in Appalachia including Basswood (*Tilia* americana), Sourwood (*Oxydendrum* arboretum) and the Tulip Poplar (*Liriodendron* tulipifera). In the U.S., there is substantial variation in domestic honey pricing (NASS, 2017). Wholesale prices listed in the January 2017 National Honey Report for domestic unprocessed honey ranged from a high of $2.50/*lb* for Basswood honey in New York State to a low of $1.55/*lb* for Buckwheat honey in Washington State (USDA, 2017a). While Basswood honey does not always fetch the highest price, in the four months it was available in 2017 it averaged a 54% price premium over the lowest priced honeys (USDA, 2017b). Varietal honeys generally command a premium in the marketplace over mixes or honeys made primarily from clovers.

The need for sustainable alternatives to extractive industries in Eastern Kentucky is often cited (Holtkamp & Weaver, 2018). The region has experienced a precipitous drop in both coal prices and employment (Klesta, 2016). It has been estimated that from 1985 through 2015 approximately 2,900 km^2 of land have been disturbed by mining in Central Appalachia. Adding in the sites known before 1985 and mining has cumulatively affected around 5,900 km^2 in the region (Pericak et al., 2018). The detrimental effects of mountaintop removal and valley fill mining on the terrain have been dramatic (Wickham, et al., 2013). Native vegetation does not readily recolonize the highly compacted and unweathered landscape left by mining. In addition, mining has affected the physical and chemical alteration of most streams in the area (Bernhardt & Palmer, 2011). It is difficult to quantify the overall social and ecological damage that has resulted from surface mining, but ecosystem services have been broadly curtailed (Zipper, et al., 2011). One effect of mountaintop removal has been the proliferation of denuded plateaus fracturing a landscape of otherwise mature second growth forest. These former mine sites often include a network of roads designed to carry heavy loads. As such, they present a somewhat unique opportunity for apiculture (Angel & Christensen, 1976). Locations around the perimeter of former mine sites may be in close proximity to forest-based nectar resources while also being adjacent to unmown grasslands that provide diverse assemblages of herbaceous forage (Horn et al., 2017). Forest-based beekeeping can make some use of these sites and would work well alongside a developed timber industry, providing supplementary income that is both ecologically sustainable and sorely needed.

The overarching objective of the thesis is to explore a methodology and develop a spatial modeling framework for apiary site selection based on proximity to nectar sources utilized in varietal honey production. The specific objectives of the study are:

i) Using publicly available forest inventory data, to model the spatial distribution of three native tree species that are important to honey producers in eastern Kentucky: American Basswood, Sourwood and Tulip Poplar.

ii) To incorporate a method for discounting the value of a nectar resource as a function of distance based on an energetic model of honeybee foraging, and

iii) To provide an example by ranking potential apiary locations around the perimeter of a mine site in the study area based on their proximity to probable species habitat using a GIS model.

The methodology presented is not meant to be a comprehensive accounting of all available forage on a landscape, but rather may serve as a guide to further investigation by attempting to quantify a single nectar resource on a particular landscape in a way that is reproducible and extendable. The modeling proceeds under the assumption that a diversity of pollen and nectar sources exist in quantities sufficient to provide for annual colony health and that honey production is the beekeeper's primary goal. It can serve as a guide for smaller scale beekeepers who specialize in varietal honeys and aid in providing a sustainable non-timber forest product on lands that have been heavily impacted by surface mining in Appalachia.

CHAPTER TWO: METHODS

2.1. Species Distribution Modeling of Three Tree Species

Presence-absence records of Basswood, Sourwood and Tulip Poplar were compiled from two sets of forest inventory plots maintained by different entities. One dataset was used to train a species distribution model which estimates the probability of occurrence at a location and the other was used to evaluate the model.

Data Collection: FIA_{train} and CFI_{eval}

The US Forest Service maintains a forest inventory and analysis program in order to monitor the status of forests across North America (USFS-FIA, 2017). A portion of the Kentucky River watershed spanning approximately 9,000 km^2 in eastern Kentucky contained 271 FIA plots, which contain 4 subplots each (see figure 2.1.1 for layout). Some plots lacked information on one or more subplots, resulting in n=1059 subplots that were used to train the species distribution model (hereafter, "FIA_{train} dataset").

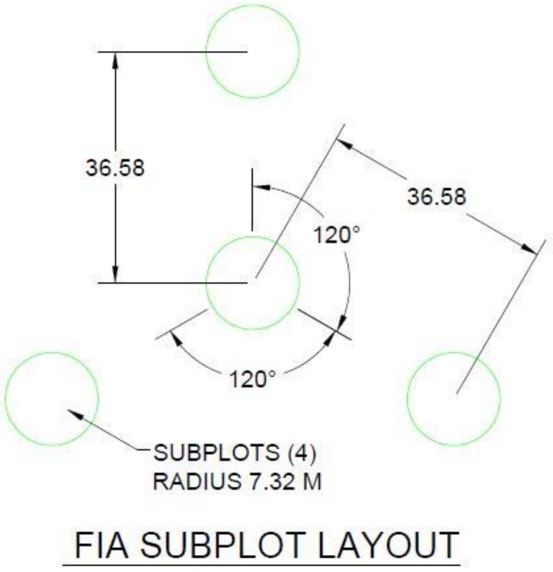

Figure 2.1.1 USFS Forest Inventory Analysis plot, showing subplot layout.

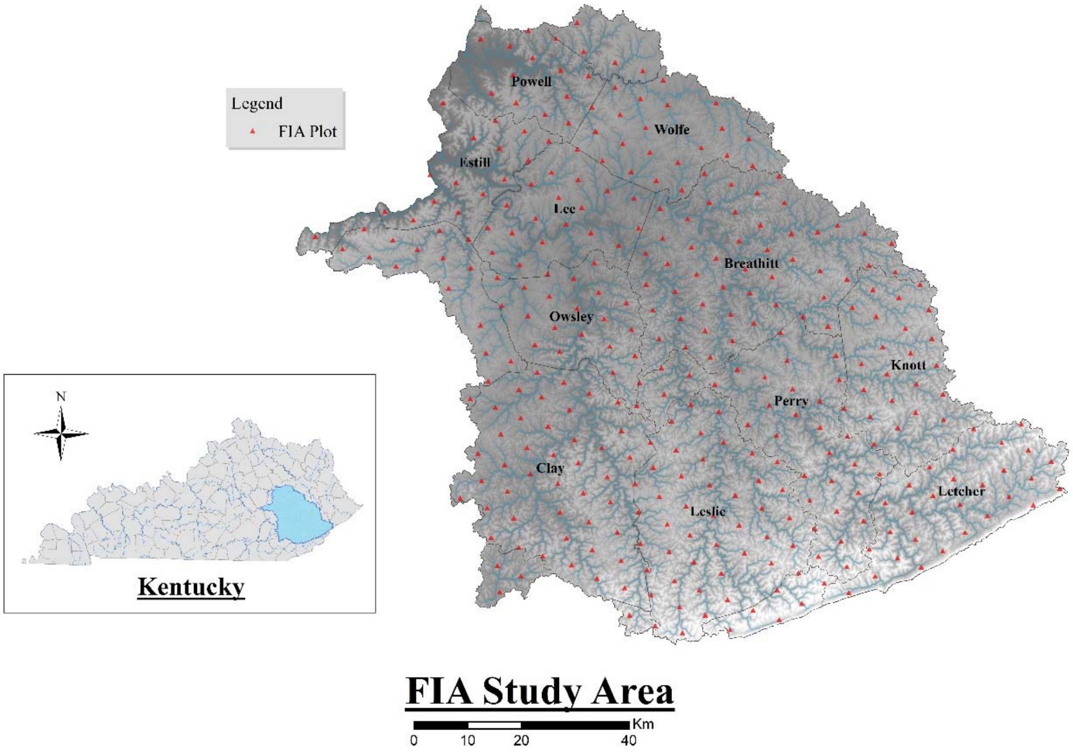

Figure 2.1.2 The FIA study area, a portion of the Kentucky River watershed showing the approximate location of Forest Inventory and Analysis plots maintained by the US Forest Service. The area is approximately 9,000 km^2 (2.2 million acres).

The exact locations of FIA plots are not publicly available. Figure 2.1.2 shows a map of approximate plot locations. Following protocols set forth by the Forest Service, the investigator visited the USDA Forest Service Southern Research Station in Knoxville, TN and provided raster layers of all candidate variables at 10-meter resolution to the FIA officer. Topographic variables and vegetative indices derived from Sentinel-2 satellite data were extracted from a network terminal at the Southern Research Station maintained for that purpose. The locations shown on maps and figures in this thesis are graphic approximations, but the true subplot coordinates were used to extract the FIA$_{train}$ data and train the model.

Nested inside the FIA study site is Robinson Forest, a research forest preserved by the University of Kentucky that spans approximately 4,000 ha across Knott and Breathitt counties in Kentucky, USA. The University maintains 271 continuous forest inventory (CFI) plots that were last surveyed in 2014. Prior to modeling and in coordination with the forest manager, 33 sites that had been recently logged or damaged by fire were removed from consideration. This left 238 plots that best represent the mature second growth forest in the area. The circular area of the CFI plot as sampled in the field was 416 m^2 and the location of trees within the plot was noted in the field. Each presence/absence point was assigned to one of four 104 m^2 quadrants and these were treated as subplots. Two quadrants that crossed open water were discarded, leaving n=950 subplots (hereafter, the "CFI_{eval} dataset"). See

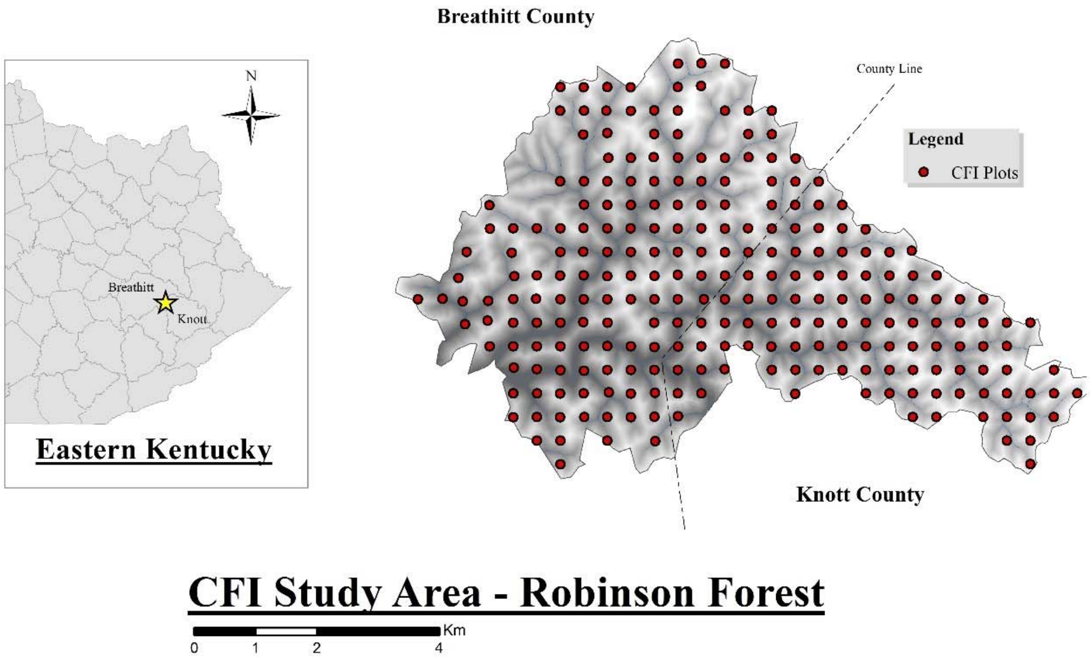

Figure 2.1.3 The CFI study area, a research forest spanning parts of Breathitt and Lee counties in Eastern Kentucky showing the locations of forest inventory plots maintained by the University of Kentucky. The area is approximately 40.5 km^2 (10,000 acres).

figure 2.1.3 for a map of plot locations. The $\mathrm{CFI}_{\mathrm{eval}}$ data was used to test for general concordance with probabilities generated from the broader $\mathrm{FIA}_{\mathrm{train}}$ model. The quadrant clustering and the 4 m^2 discrepancy between the size of the subplot and the 10m x 10m pixel size of the predicted probabilities were ignored for the purposes of evaluating the model.

Modeling and Variable Selection:

Logistic regression is an extensively used generalized linear model that employs a link function to model a binary dependent variable such as the presence or absence of a species. Using a logit link specifies the general form of the model as:

$$\ln\left(\frac{P(Y)}{1-P(Y)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \epsilon \qquad \text{(eqn. 2.1.1)}$$

where $P(Y)$ is the probability of the occurrence of event Y (i.e. the presence of the tree species of interest at a location), X_i $(i = 1, \ldots, k)$ are the independent variables, β_i are the regression coefficients, and ϵ is an error term (Ovaskainen, et al. 2016).

Model analyses were performed in R 3.5.1 (R Core Team, 2018); generalized linear models were fit using the “lme4” package (Bates et al., 2015); generalized estimating equations modeling utilized “geepack” (Hojsgaard et al., 2016); packages “raster” (Hijmans et al., 2015), “pROC” (Robin et al., 2011) and “ResourceSelection” (Solymos & Lele, 2015) were also used throughout the process.

Three measures were considered during variable selection and model comparison: AIC, AUC and plots derived from Hosmer-Lemeshow testing. The Akaike information criterion (AIC) is widely used as a measure for comparing models fit to the same dataset (Akaike, 1974). It is defined as

$$AIC = 2k - 2\ln(L_{max}) \qquad \text{(eqn. 2.1.2)}$$

where k is the number of parameters, and L_{max} is the maximum value of the likelihood function of the model. Forward stepwise variable selection using AIC was utilized to compare different models of the same dataset.

Area Under the receiver operating characteristic Curve (AUC) is a measure of model accuracy (Fielding & Bell, 1997). The receiver operator curve shows how the model's true positive rate (sensitivity) is related to the false positive rate (1-specificity) as the probability threshold used to classify a species presence varies.

Hosmer-Lemeshow testing is used alongside logistic regression to compare the expected presence rate predicted by the model to the presence rate observed in the data (Hosmer & Lemeshow, 2013). It is used to show how well calibrated a model is with the data in different regions of predicted probability. As a single statistic, it is defined as:

$$H = \sum_{q=1}^{g} \frac{(Observed\ P - Expected\ P)^2}{(Expected\ P)} + \frac{(Observed\ A - Expected\ A)^2}{(Expected\ A)} \qquad \text{(eqn. 2.1.3)}$$

where g is the number of divisions used to segment the data, P is the number of species presences and A is the number of species absences.

Initial investigation of covariates was conducted primarily on the CFI_{eval} dataset as the smaller area and readily available plot coordinates allowed for more flexible investigation of the predictive power of multiple variables. A comprehensive table of the candidate variables and their sources is included as an appendix A. The species showed significant response to three topographic variables, denoted here as: Tp_1, Tp_2, and Tp_3. These three variables are used in the modeling to represent the effects of aspect, elevation, and slope respectively.

Tp_1 is a transformation of aspect and can also be described as 'deviation from a bearing' (Jenness, 2007). For modeling, aspect was transformed and treated as a single directional variable with values ranging from 180 facing Southwest and diminishing equally in both directions to zero facing Northeast. Tp_1 was calculated iteratively in three steps beginning with an aspect raster consisting of continuous clockwise degree measurements with 0° as facing map North:

$r_0 = Aspect$ (eqns. 2.1.4)

$r_1 = (r_0 - Bearing)$

$r_2 = if\ r_1 < 0$, then $(r_1 + 360)$; otherwise, r_1

$r_3 = if\ r_2 > 180, then\ (360 - r_2); otherwise, r_2$

where, *Bearing* is the desired direction for zero and r_1, r_2 and r_3 are successive iterations of aspect.

In the course of variable selection, this ‘directional’ transformation was modeled alone in 5° increments to determine which direction would best inform the model. A criticism can be made that doing so runs the risk of ‘tuning’ the data and overfitting to an individual data sample, but the variable showed the same response in both the FIA_{train} and CFI_{eval} datasets. The graph in Figure 2.1.4 shows how the AUC of a logistic regression model using only Tp_1 varies as the transformed aspect is rotated in 5° increments from geographic North in a clockwise fashion. A bearing of 45° was used in calculating the values for Tp_1 for the modeling of all three species.

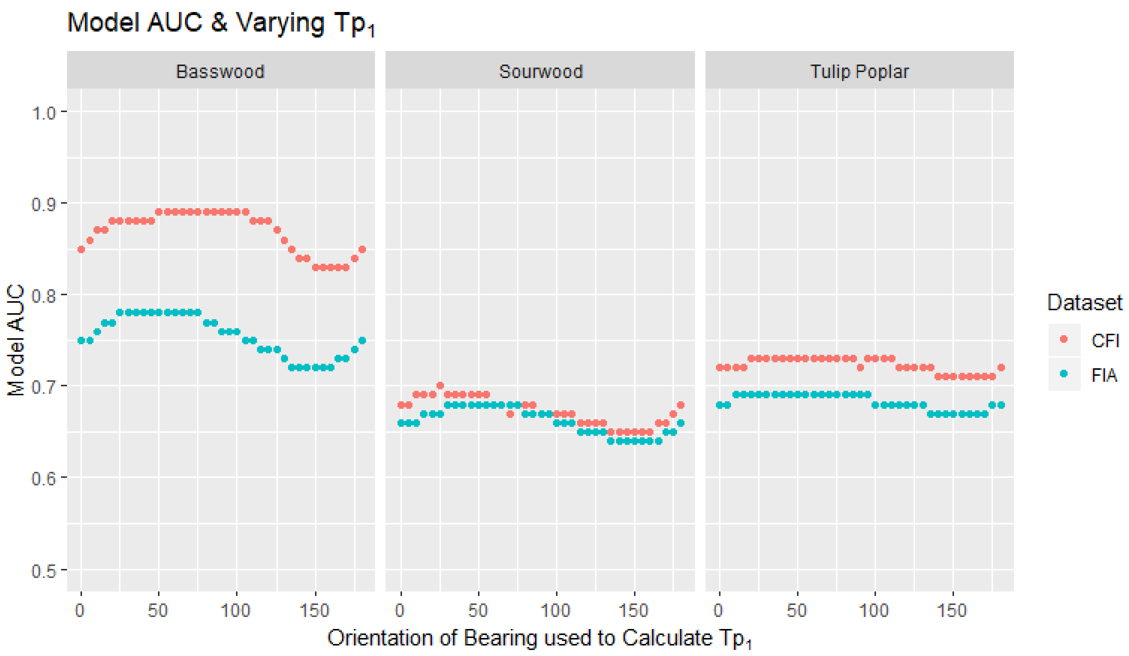

Figure 2.1.4 Graphs showing area under the receiver operator curve (AUC) for logistic regression using only Tp_1 as the bearing used to transform this ‘directional’ variable is rotated in 5□ increments through 180□.

Variable Tp_2 is a measure of relative topographic position. It is often used in conjunction with other measures to create various forms of Topographic Position Index (TPI), but it is used here in the form:

$$Tp_2 = \frac{Elev_{point} - MIN_{neighborhood}}{MAX_{neighborhood} - MIN_{neighborhood}} \qquad \text{(eqn. 2.1.5)}$$

where, $Elev_{point}$ is the elevation of the point, $MIN_{neighborhood}$ is the minimum elevation value present in some square focal area surrounding the point, and $MAX_{neighborhood}$ is the maximum elevation value present in that area.

Tp_2 was similarly modeled individually over varying focal ranges to gauge which would best inform the model in this landscape. The graph in Figure 2.1.5 shows how the AUC of logistic regression modeling using only Tp_2 varies as the focal range used to calculate TPI changes in 5-pixel increments. A value of 25 pixels was used to calculate Tp_2 for modeling of all three species.

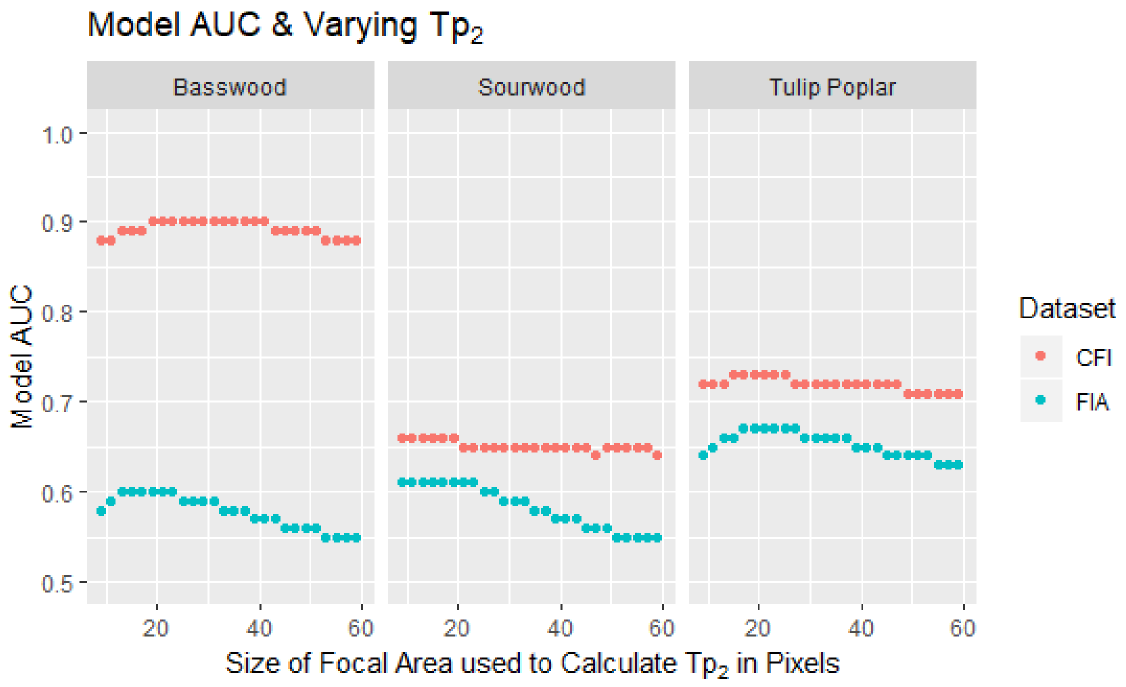

Figure 2.1.5 Graphs showing area under the receiver operator curve (AUC) for logistic regression using only Tp_2. The focal window used to calculate Tp_2 variable used in the regression changes in 5-pixel increments.

Variable Tp_3 has no transformation and is simply topographic slope in degrees derived from the USGS DEM in R using SDMtools.

2.2. Method of Discounting Nectar Resource

Broadly speaking, “currencies” for determining the value of nectar resources to a honeybee colony fall into two categories: net energy efficiency ((benefit – cost)/cost) and the net rate of energy intake ((benefit – cost)/time) (Becher et al., 2013). Models that focus on resource selection often settle on net energy efficiency as the appropriate currency. These models are primarily concerned with navigating a complex set of information feedback mechanisms that affect resource selection and utilization over a range of colony dynamics (Schmickl, & Crailsheim, 2007). In contrast, a varietal honey producer is less concerned with *which* resource will be utilized than *how much* of a given resource is available. They are seeking nectar sources which, from experience, are previously known to be preferentially selected over other forage in the area at that time. Honey producers deliver hives at peak population to sites over the period of time that a specific resource is in bloom (Pilati & Prestamburgo, 2016). For instance, in the case of Basswood, this period may be as short as ten to fourteen days (Anderson, 1976). Commercial beekeepers are ultimately concerned with the quantity of honey stored by a colony over and above its annual energetic requirements as this is what makes up the honey crop. The net rate of energy intake is what is used here.

By first determining the net energy on a per-trip basis and then accounting for the number of possible trips given the time per trip, the net rate of energy intake as a function of resource distance from the colony can be estimated. This function will then be used to discount the probabilities obtained from the species distribution model. Assumptions and notation for the energetic based foraging equations are modified from Baveco et al, 2016.

A Rate of Net Energy Intake in Joules/second (RNEI $J\ s^{-1}$) at the hive can be determined by dividing the Net Energy Intake of a single forager trip in Joules (NEI_{trip} J) by the time in seconds that a single forager trip takes (t_{trip} s):
$RNEI = NEI_{trip} / t_{trip}$ (eqn. 2.2.1)

The Net Energy Intake of a trip can be determined by subtracting the Energy Expended from the Energy Intake at the hive:
NEI_{trip} = Energy Intake (EI_{trip}) - Energy Expended (EE_{trip}) (eqn. 2.2.2)

Energy Intake is related to the mass of the Nectar Load (N_{load}), the fraction of sugars in the nectar (F_s), and the Energy of sugar (E_{sugar}):
$EI_{trip} = (N_{load}) (F_s) (E_{sugar})$ (eqn. 2.2.3)

The Energy Expended for a trip can be broken down into components:

$$EE_{trip} = EE_{transit} + EE_{source} \qquad \text{(eqn. 2.2.4)}$$

Energy expended at the hive between trips is taken as zero because without flight, no energy above baseline metabolic rate is expended ($EE_{hive} \approx 0$).

The Energy Expended flying to and from the nectar source is:

$$EE_{transit} = 2\left(\frac{D}{v}\right)\bar{e} \qquad \text{(eqn. 2.2.5)}$$

where, D is the distance to source (m), v is the flight velocity of the forager (m s^{-1}), and ē is the average energy expended in flight (J s^{-1}).

Average energy expended in flight is used in both the transit and source costs as the forager departs the hive unloaded, fills at a roughly constant rate and returns the same distance loaded with nectar:

$$\bar{e} = \frac{e_u + e_l}{2} \qquad \text{(eqn. 2.2.6)}$$

where e_u is the energy of flight unloaded and e_l is the energy of flight loaded.

The energy cost at the nectar resource are proportional to the Collection Rate of nectar:

$$EE_{source} = \left(\frac{N_{load}}{CR_n}\right)\bar{e} \qquad \text{(eqn. 2.2.7)}$$

The time per trip can also be broken down into components:

$$t_{trip} = t_{hive} + t_{transit} + t_{source} \qquad \text{(eqn. 2.2.8)}$$

Time in transit is:

$$t_{transit} = 2\left(\frac{D}{v}\right) \qquad \text{(eqn. 2.2.9)}$$

where, D is the distance to source (m), v is the flight velocity of the forager (m s^{-1}).

Time elapsed at the nectar source is:

$$t_{source} = \frac{N_{load}}{CR_n} \qquad \text{(eqn. 2.2.10)}$$

where N_{load} is the nectar load (mg) and CR_n is the collection rate of nectar (mg s^{-1}).

Table 2.2.1: Fixed energetic coefficients for honeybee:

Coefficient	Symbol	Units	Value	Source
Nectar load (capacity)	N_{load}	mg	32.5	Winston (1987)
Velocity of forager	v	$m\ s^{-1}$	4.17	de Vries & Biesmeijer (1998)
Energetic content nectar	e_{sugars}	$J\ mg^{-1}$	17.2	Seeley (1985)
Flight cost (loaded)	e_l	$J\ s^{-1}$	0.075	Seeley (1986)
Flight cost (unloaded)	e_u	$J\ s^{-1}$	0.037	Seeley (1986)
Average Flight Cost	$\bar{e}$	$J\ s^{-1}$	0.056	
Time unloading in hive	t_{hive}	s	300	Seely, Camazine & Sneyd (1991)

Substituting the fixed values from Table 1 into eqns. (2.2.1-10) yields:

$$\text{RNEI} = \frac{559 * F_s - 0.0269 * D - \frac{1.82}{CR_n}}{300 + 0.48 * D + \left(\frac{35.5}{CR_n}\right)} \qquad \text{(eqn. 2.2.11)}$$

The collection rate and sugar content of different nectars are dependent on the plant species and Basswood is used as an example here:

Table 2.2.2: Resource-specific coefficients for Basswood:

Coefficient	Symbol	Units	Value	Source
Fraction sugars	F_s	$mg\ mg^{-1}$	0.28	Anderson (1976)
Collection rate nectar	CR_n	$m\ s^{-1}$	0.32[a]	Baveco (2016)

[a] Collection Rate of nectar is derived from the attack rate of the foragers, density of flowers and the amount of nectar per flower. It is unknown for Basswood and the mean of the calculated rate for Oilseed Rape (0.42 mg s-1) and White Clover (0.21 mg s-1) is used as a proxy here. The true collection rate is likely higher.

Substituting the resource specific values from Table 2.2.2 into eq. (11) for Basswood yields:

$$\mathrm{RNEI(D)} = \left(\frac{362}{D+856}\right) - 0.056 \qquad \text{(eqn. 2.2.12)}$$

where D is the distance between the hive and the resource in meters.

This gives us the rate of net benefit to the hive in joules per forager-second of a Basswood tree as a function of distance (figure 2.2.1).

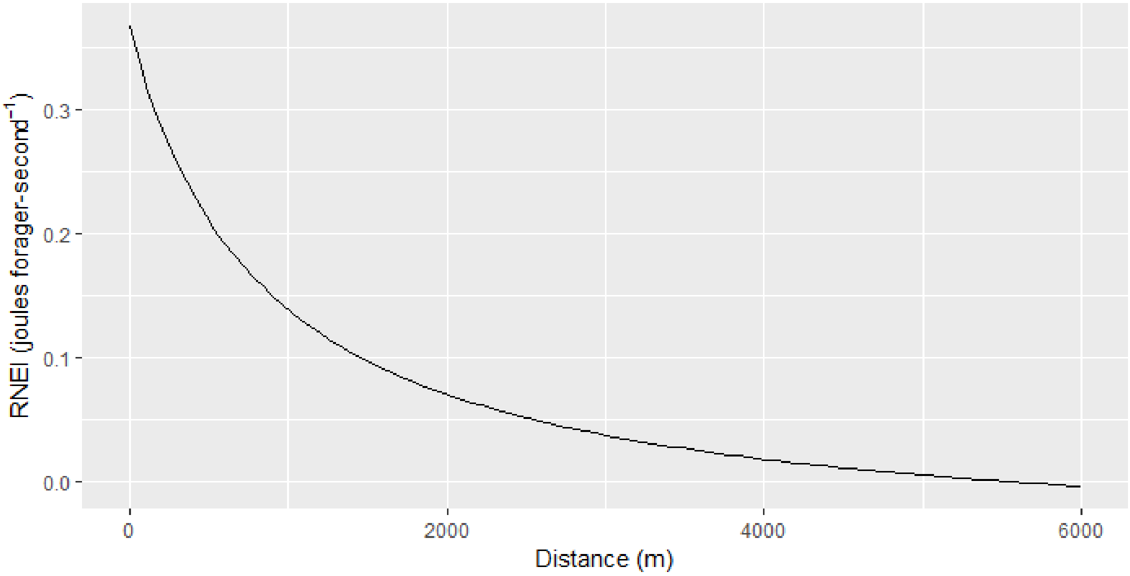

Figure 2.2.1 Graph of nectar value-distance curve for Basswood, showing the Rate of Nectar Energy Intake as a function of Distance: RNEI(D)

Finally, the nectar value-distance function is used to assign RNEI values to the pixels within a given foraging range based on their distance from a point of interest (e.g., a potential colony location). As distance from the hive increases, the potential value of nectar decreases. These values are then used as weights – that is, they are multiplied by the probabilities that the tree species is present generated by the SDM.

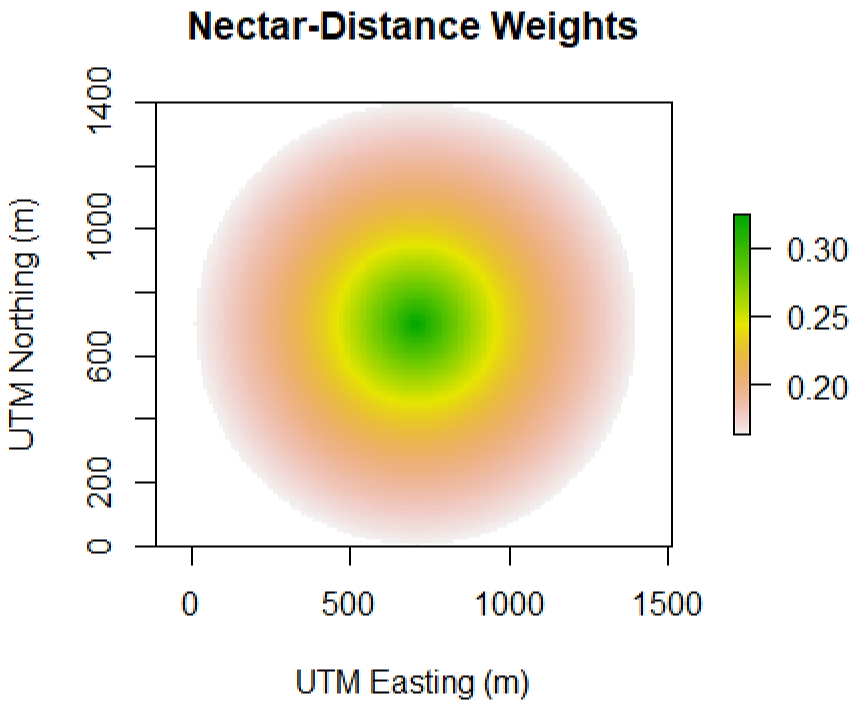

Figure 2.2.2 Shows the nectar value-distance weights for a 700 meter flight range as a raster image.

2.3. Ranking Sites - an Example using Basswood

A reclaimed former mine site adjacent to University owned property was chosen as an example to analyze potential apiary sites with respect to their proximity to Basswood; however, the same process could be applied on any of the approximately 200 mine sites greater than 50 hectares that are located in the watershed or to any of the three tree species modeled. The requirements for consideration as potential apiary site in this project were threefold. First, an area of approximately 20m x 30m (a minimum of six 10-m pixels) must be relatively flat, having a slope of less than 3 degrees; however, to avoid being inundated by storm water, flat areas must not be located in a drainage basin. Secondly, to avoid infestation by an invasive pest of the honeybee called Small Hive Beetle (*Aethina* tumida), sites must be located in the open, at least 10 m from a closed canopy tree line. Thirdly, sites must be accessible by a road capable of supporting a medium size flatbed truck to deliver and maintain the hives (Class 3 gross weight vehicle rating, 10,000 – 14,000 lbs.).

In this example, potential sites were identified manually; however, ArcGIS tools could be used to automate the process. This required overlaying three GIS layers: data from the National Agricultural Imagery Program (USDA-NAIP, 2016), a low slope layer, and a topographic contour layer. NAIP is relatively high resolution (60cm) imagery acquired by the USDA during the growing season every two years. In this context, it allows the user to track the road access to potential sites. The low slope layer identifies pixels with less than 3° slope (identified in blue in Figure 2.3.1). The contour layer is necessary to distinguish between routes that appear to be accessible roads but are in fact drainage rip-rap or terracing. Only areas around the perimeter of the mine site were considered.

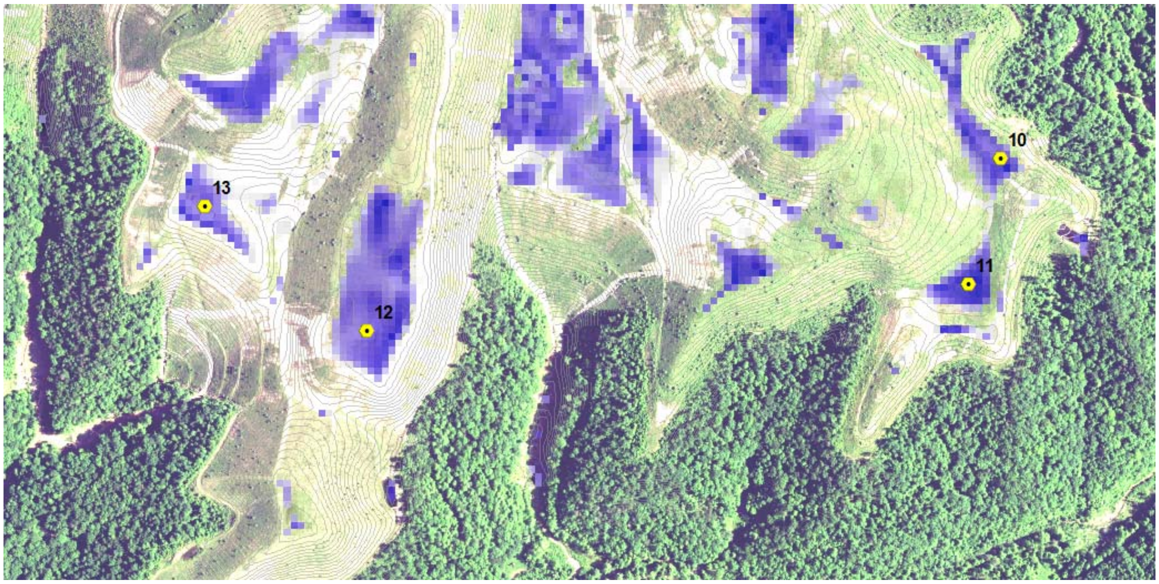

Figure 2.3.1 Selecting potential apiary sites manually: an on-screen process of selecting potential sites (numbered yellow hexagons) based on slope and road access criteria. Blue pixels indicate areas with less than 3° slope and grey contour lines are every 2m.

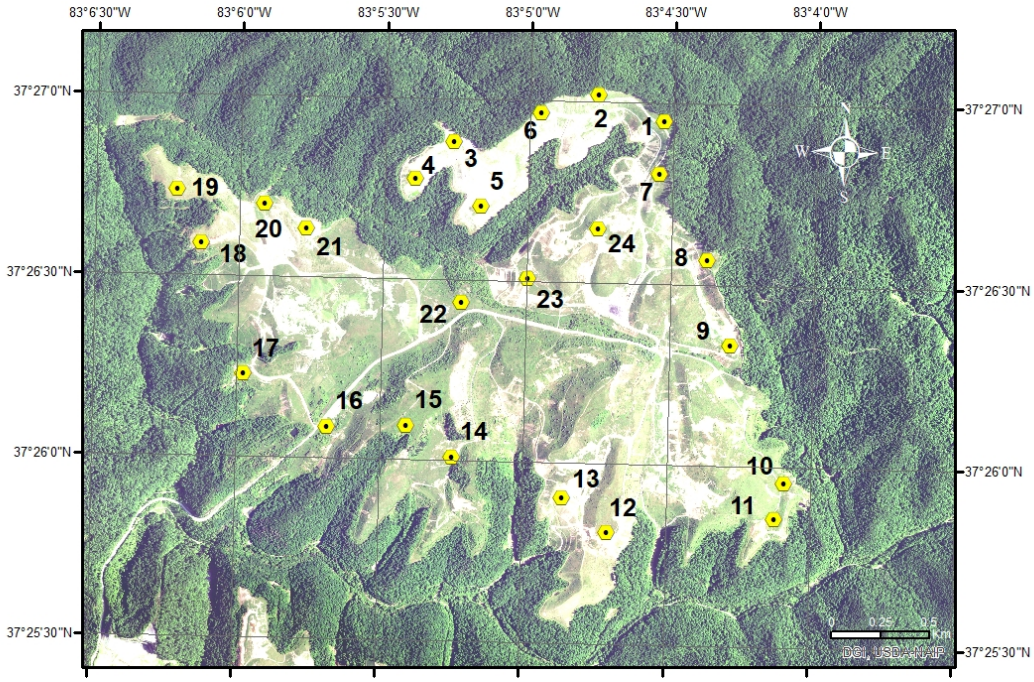

Figure 2.3.2 Example mine site: locations of 24 potential apiary locations.

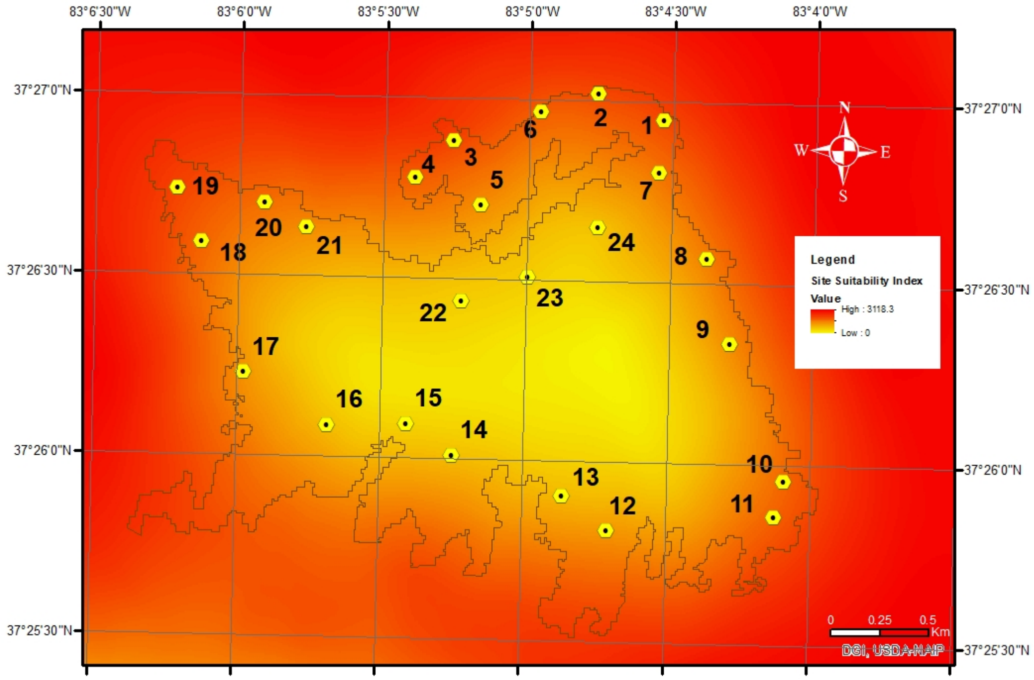

Figure 2.3.3 Example mine site: Site Suitability Index (SSI).

In the example, this process identified 24 potential apiary locations around the perimeter of the mine site (see figure 2.3.2 for site map). A raster of Site Suitability Index (SSI) was derived by multiplying the probability of a species presence by the rate of nectar energy intake as a function of distance at each pixel within a flight radius of 700 m and summing the products in that tributary area:

$$SSI = \sum RNEI(D) * P(Y) \qquad \text{(eqn. 2.3.1)}$$

where RNEI(D) is the Rate of Nectar Energy Intake as a function of Distance (the 'nectar-distance weights') and P(Y) is the probability that one or more Basswood stems is present within that pixel. Point values at potential locations were extracted and the locations ranked according to their SSI reflecting their weighted proximity to Basswood. A flow-chart of the modeling and GIS procedures used to rank potential apiary locations is shown in Figure 2.3.4.

The probabilities generated by the model are based on an FIA plot which is 168 m^2 while the pixel size of the rasters used in the GIS (and the approximate size of the CFI plots) is 100 m^2. The probabilities were treated as a Poisson point-process and 'scaled' for use as probability over the smaller raster cells area:

$$P(Y)_{raster} = 1 - [1 - P(Y)_{model}]^{\left(\frac{100}{168}\right)} \qquad \text{(eqn. 2.3.2)}$$

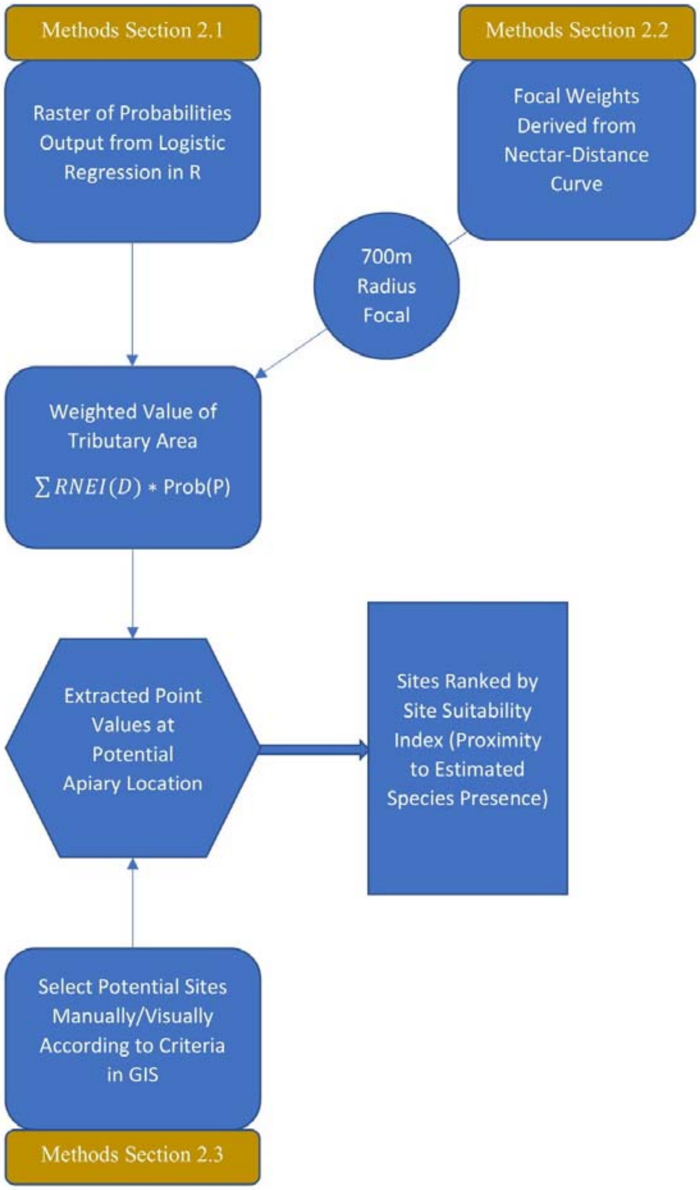

Figure 2.3.4 Flowchart of methods used to rank potential apiary sites.

CHAPTER THREE: RESULTS

3.1 Species Distribution Modeling Results:

This study models the distribution of three deciduous species using topographic variables that are both publicly available and easily derived. To examine the potential effects of subplot clustering (see Figure 2.1.1), the FIA data was modeled separately utilizing Generalized Estimated Equations (GEE) using plot as a grouping factor and an exchangeable correlation structure. The resulting coefficients varied somewhat but after comparing the AUC and Brier score (Table 3.1.1) developed from the traditional GLM, it was concluded that the quasi-clustering of the FIA subplots had a negligible effect on probabilities generated from this dataset.

Table 3.1.1: AUC, Prevalence and Brier scoring to compare models that considered possible clustering effects (GEE) and those did not (GLM).

Model-Species	AUC (FIA$_{train}$)	Prev (FIA$_{train}$)	Brier (FIA$_{train}$)	AUC (CFI$_{eval}$)	Prev (CFI$_{eval}$)	Brier (CFI$_{eval}$)
GLM-Basswood	0.7753	6.2%	0.05456	0.8780	3.9%	0.03321
GEE-Basswood	0.7749		0.05453	0.8800		0.03298
GLM-Sourwood	0.6668	22.5%	0.16549	0.6269	10.5%	0.09464
GEE-Sourwood	0.6687		0.16557	0.6220		0.09607
GLM-Poplar	0.7213	36.5%	0.20023	0.7296	18.3%	0.14053
GEE-Poplar	0.7212		0.20027	0.7300		0.14020

The final model coefficients are taken from traditional GLM logistic regression with a binomial logit link function. The coefficients, standard errors, and p-values for the final models derived using the FIA$_{train}$ dataset are given in Table 3.1.2. Among these three topographic variables, aspect and relative elevation exerted significant effects on the distribution of all three species, but slope was identified as a significant predictor for Basswood only.

Table 3.1.2: Coefficients of logistic regression models from a sample of n=1059 subplots in the FIAtrain dataset.

Species	β_0(INT)	SE β_0	P β_0	$\beta_1 Tp_1$	SE $\beta_1 Tp_1$	P $\beta_1 Tp_1$	$\beta_2 Tp_2$	SE $\beta_2 Tp_2$	P $\beta_2 Tp_2$	$\beta_3 Tp_3$	SE $\beta_3 Tp_3$	P $\beta_3 Tp_3$
Basswood	-3.218879	0.6495	< 0.01	-0.015648	0.003069	< 0.01	-0.040801	0.01594	< 0.05	0.067563	0.021245	< 0.01
Sourwood	-2.090492	0.190064	< 0.01	0.008072	0.001585	< 0.01	0.029749	0.006439	< 0.01			
Tulip Poplar	0.177784	0.16722	0.29	-0.008556	0.001513	< 0.01	-0.058913	0.007375	< 0.01			

Plots showing how the three species models predicted on the $\text{FIA}_{\text{train}}$ dataset (the dataset used to fit the model) as well as how it predicted on the independent CFI_{eval} dataset are show in Figures 3.1.1-3. These plots take the probability of a species presence predicted by the SDM and, arranging them ordinally, compares them to the actual presences and absences in the data. If the model is performing well, the ratio of presences (green dots) to absences (red dots) should increase as the predicted probabilities increase. Another way to visualize the performance of the model are plots derived from the Hosmer-Lemeshow statistic, shown in Figure 3.1.4-6. These take the same ordered probabilities and divide them into a number of equal groups (bins) and sum the probabilities in each bin. This is considered the expected number of presences in that bin which can then be plotted against the actual presences.

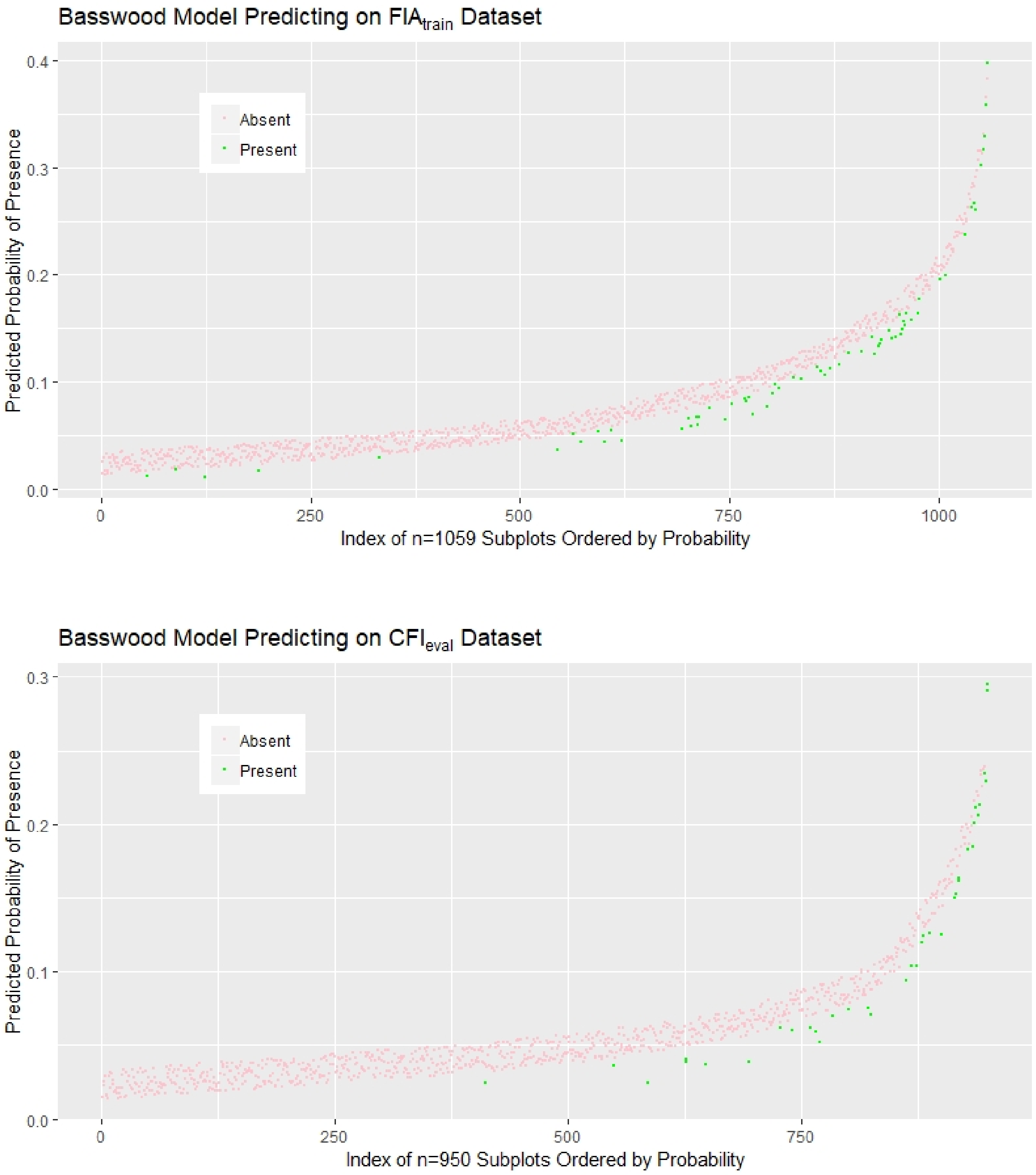

Figure 3.1.1 The Basswood model predicting on the FIA$_{train}$ and CFI$_{eval}$ datasets.

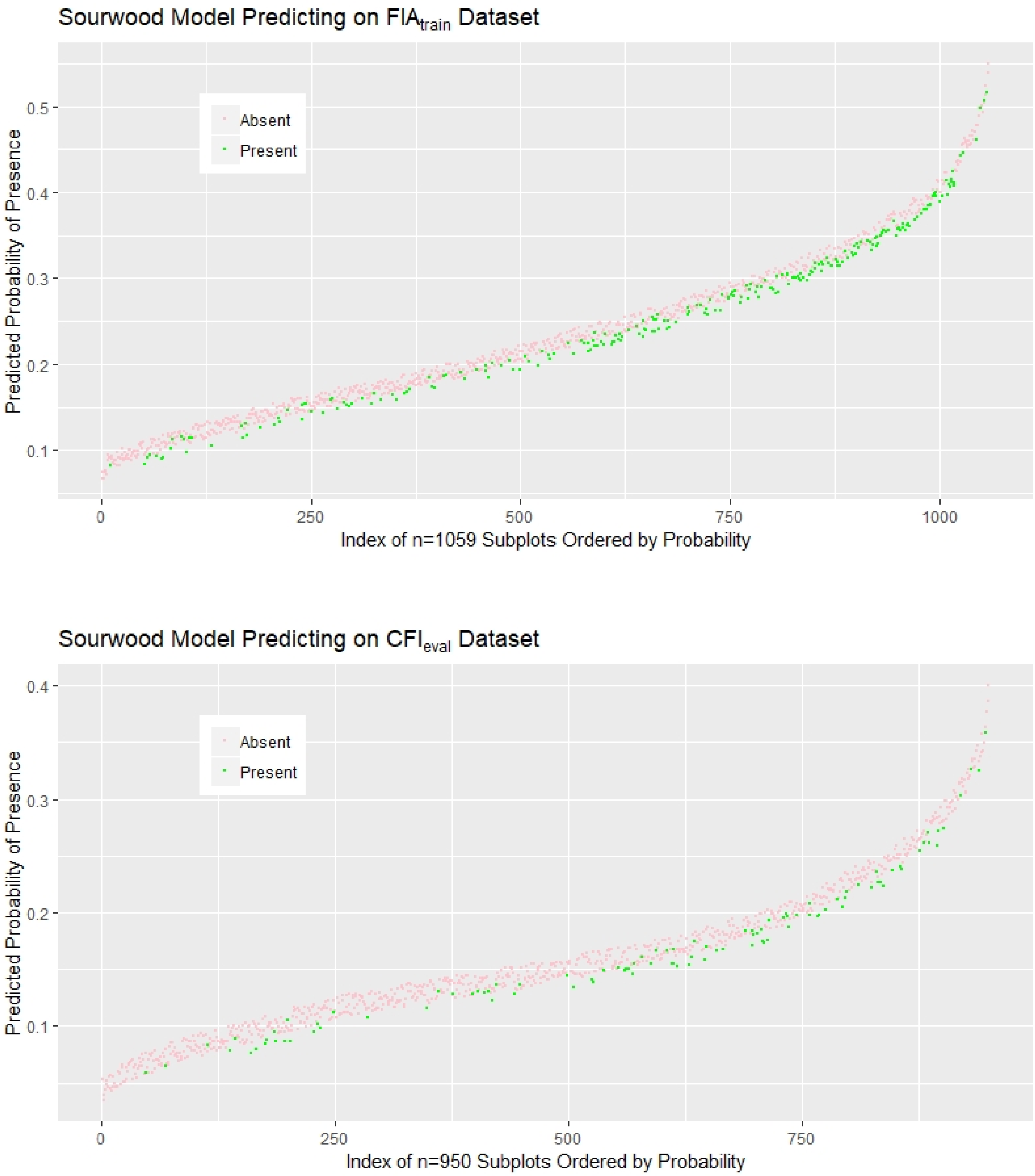

Figure 3.1.2 The Sourwood model predicting on the FIA$_{train}$ and CFI$_{eval}$ datasets.

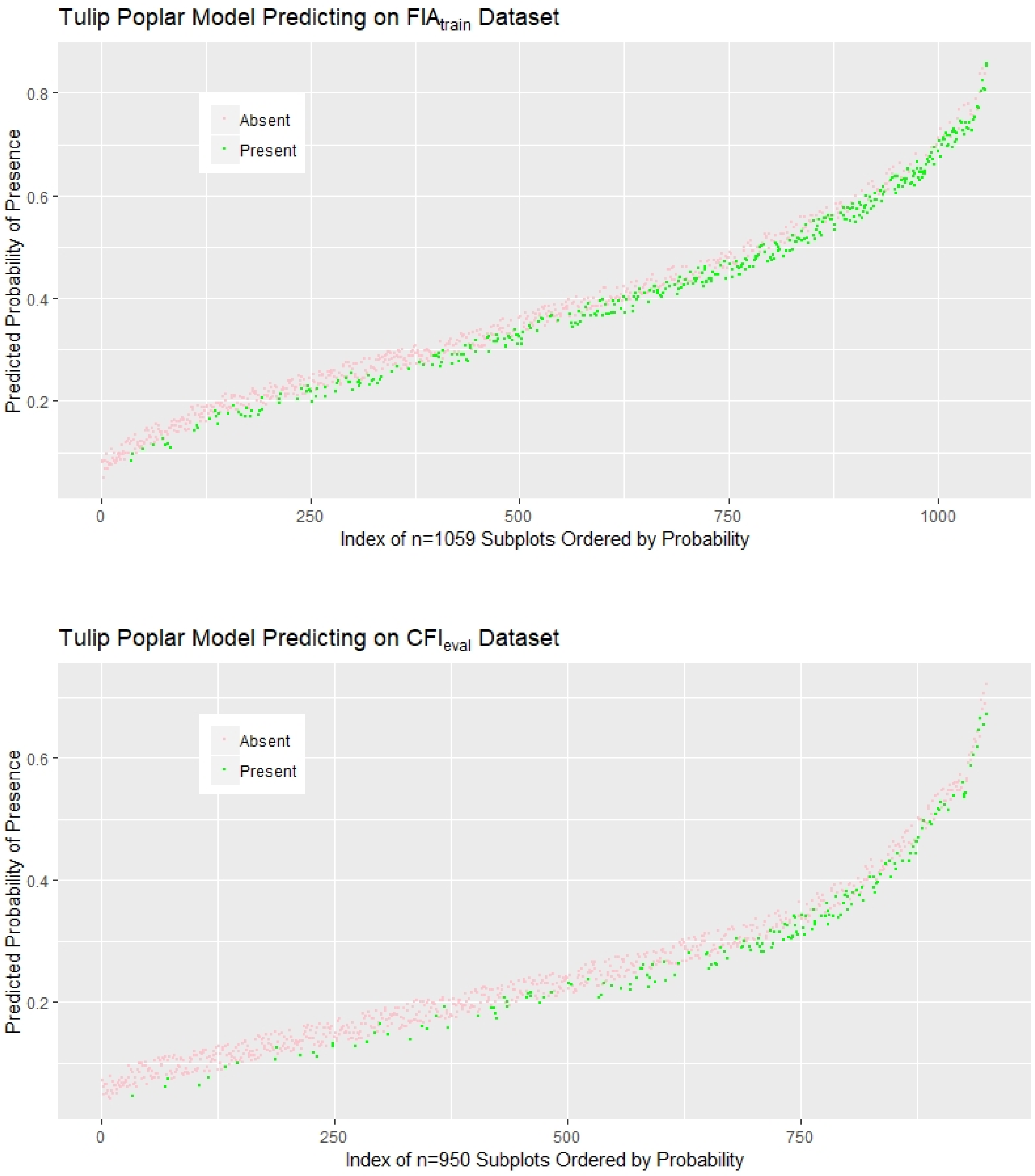

Figure 3.1.2 The Tulip Poplar model predicting on the FIA$_{train}$ and CFI$_{eval}$ datasets.

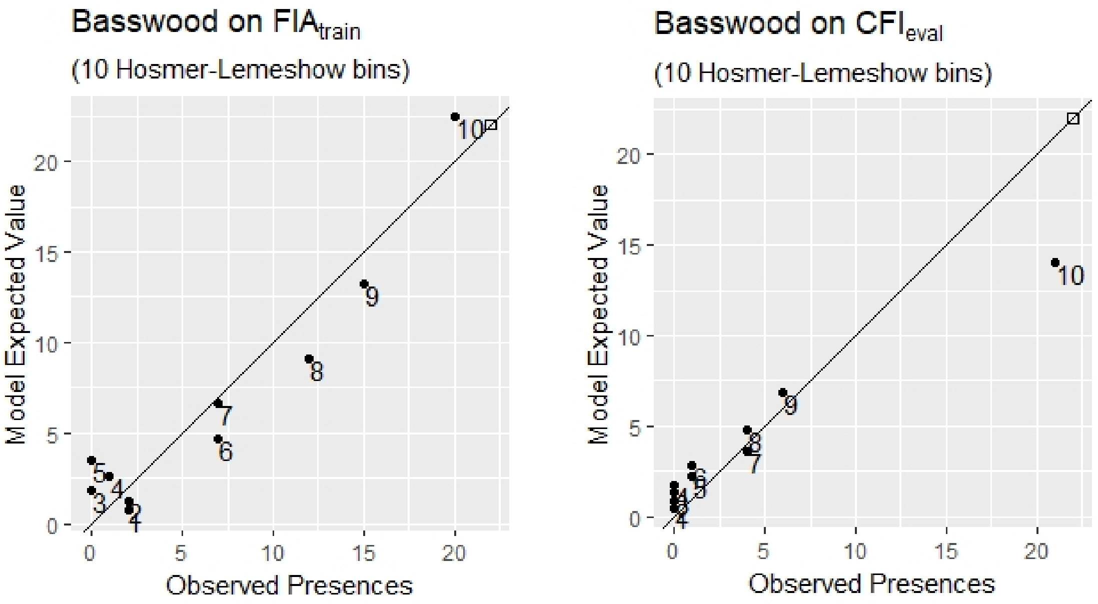

Figure 3.1.4. Basswood Hosmer-Lemeshow plots showing the model predicting on both datasets with 10 sub-divisions of the ordered plots.

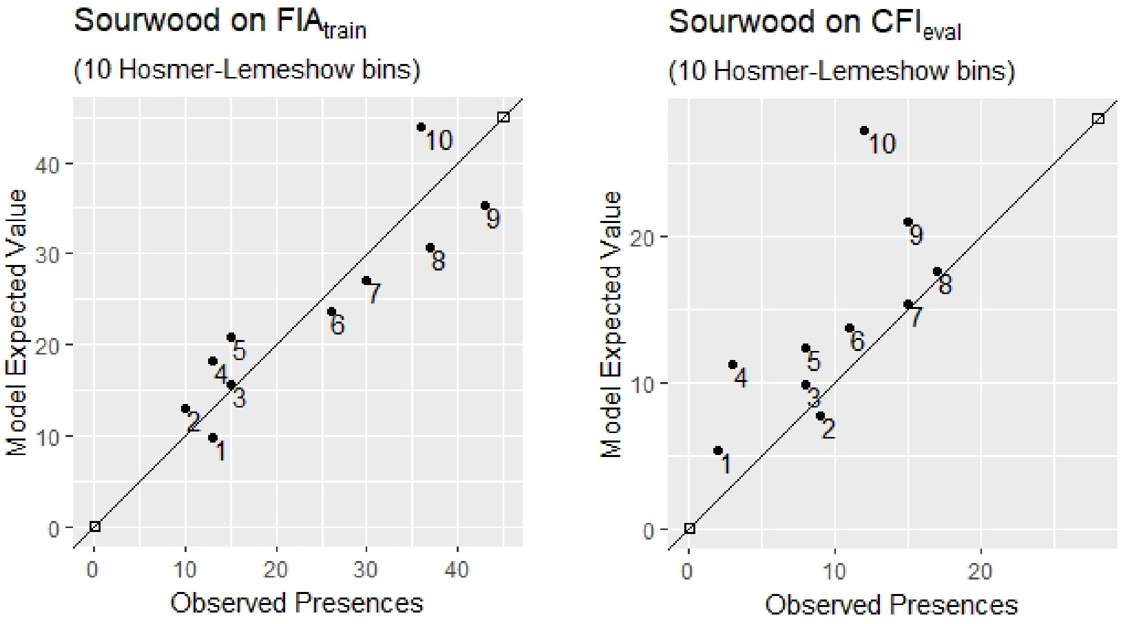

Figure 3.1.5. Sourwood Hosmer-Lemeshow plots showing the model predicting on both datasets with 10 sub-divisions of the ordered plots.

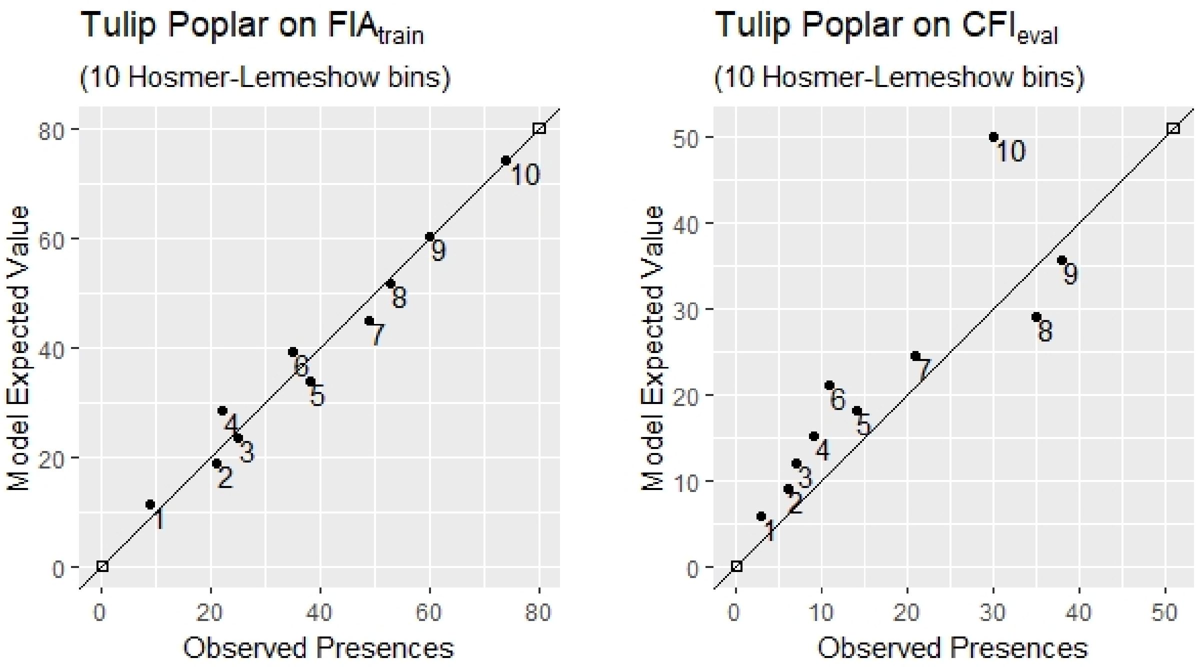

Figure 3.1.6. Tulip Poplar Hosmer-Lemeshow plots showing the model predicting on both datasets with 10 sub-divisions of the ordered plots.

3.2 Value-Distance Weighting and Site Rankings:

The probabilities derived in the Basswood species distribution model were used alongside distance weights based on an energetic model of honeybee foraging to specify a site suitability index and rank potential apiary locations at an example mine site. The ranking and corresponding site index for the 24 selected sites are given in Table 3.2.1.

Table 3.2.1. Example Site Rankings using 500 m, 700 m, and 1.2 km as tributary areas. The top 5 sites in each tributary area are highlighted in green and the next 5 sites are highlighted in blue:

Site No.	500 m Tributary SSI	500 m Rank	700 m Tributary SSI	700 m Rank	1.2 km Tributary SSI	1.2 km Rank
7	1,214	1	2,255	1	5,491	1
2	1,190	2	2,179	4	5,081	3
5	1,189	3	1,942	5	3,763	12
1	1,172	4	2,181	3	5,161	2
4	1,169	5	2,229	2	4,342	9
3	1,113	6	1,931	6	4,815	5
23	939	7	1,879	7	4,692	6
10	888	8	1,865	8	4,835	4
21	886	9	1,778	9	4,557	7
6	830	10	1,636	11	4,313	10
13	830	11	1,490	12	3,753	13
9	798	12	1,712	10	4,516	8
8	771	13	1,332	14	3,052	17
17	710	14	1,233	16	3,723	14
11	692	15	1,353	13	3,810	11
16	682	16	1,305	15	3,437	15
24	512	17	1,153	17	3,186	16
22	421	18	779	18	2,501	20
12	356	19	641	21	2,963	18
20	309	20	696	20	2,139	22
15	301	21	492	23	1,564	24
14	265	22	502	22	1,846	23
19	264	23	719	19	2,855	19
18	139	24	405	24	2,172	21

The rankings are dependent on the tributary area used to calculate them. As the foraging range of the bees is unknown, several estimated ranges are considered. Comparing sites would necessarily involve a ‘fuzzy’ comparison of the rankings over several field-realistic foraging ranges.

CHAPTER FOUR: DISCUSSION AND CONCLUSION

4.1 Limitations of the SDM and Quality of the training Data:

Initial investigation of covariates included a number of spectral indexes derived from 10-meter Sentinel-2 data (Copernicus ESA, 2017). Also, following the procedures of Sadeghi et al. (2017), an optical trapezoidal model (OPTRAM) was parametrized for several dates to produce a wetness index. Attempts to achieve significant separation using a variety of modeling techniques were unsuccessful with bands and indices extracted from the Sentinel rasters. This may indicate the limitations of remotely sensed data to achieve species level identification at 10-meter resolution in a complex and highly dissected landscape, a problem of co-registration with the inventory plots or a limitation of the modeler's ability and the time required to investigate alternate dates. Maximum entropy modeling (MaxEnt) was not explored as the true absence data was considered reliable.

From previous work, co-registration errors were known to be problematic in the CFI data and only 35 of the plots had been geolocated with an accuracy +/-2.0 m. The remaining plots could be as much +/-10.0 meters out, meaning that they could only reliably be attributed to any of (9) pixels. In the FIA dataset, the Forest Service indicated that the range of horizontal control was likely +/-5 to 15 m (personal communication with USFS SRS). However, the researcher was unable to verify sites due to the opaque nature of the plot's true coordinates. Variables were extracted from the 'unmasked' coordinates at the Southern Research Station, but it was not possible to utilize the coordinates after that. This proved problematic on several fronts and it is recommended that other researchers who make use of this invaluable dataset take the time necessary to arrange a research agreement with the USFS to allow them access to the unmasked plots coordinates at their University.

4.2 Assumptions and Drawbacks of the Foraging Model:

This approach to apiary site selection, a frequentist statistical model to locate an individual species and a nectar-distance curve to value it, makes several assumptions that may not be valid. Firstly, it assumes that colonies will forage preferentially at the target species when it is in bloom. The continued existence and marketing of varietal honey suggest that, at least for Basswood and Sourwood, it is a reasonable assumption to make. Secondly, the approach implies that foragers will exhaust resources closer to the hive before expending the energy to forage further afield, regardless of the 'size' of the resource. There is ample research to suggest that this is generally *not* the case (Couvillon et al., 2015; Beekman et al., 2004). Forage selection is a complex and varying arrangement of opportunity and feedback that may be too complicated to model accurately at the landscape scale.

The results in the example show that rankings are dependent on the foraging range considered, implying that each potential site should have some optimal number of colonies that would maximize honey production. Field studies with apiaries of varying size could help determine the number of colonies that would effectively saturate a given location. Without knowing how many colonies the landscape can support, it leaves open the question of whether an unknown number of colonies would be more likely to gather nectar from a small patch close to the colony or a larger patch that is farther away from the colony. It would likely depend not only on the number of colonies emanating from that apiary but also on the presence of other apiaries within flight range.

Finally, the impact on native pollinators of introducing numerous honeybee colonies is unknown. Recent work suggests that over 40% of insect species worldwide are threatened with extinction (Sánchez-Bayo & Wyckhuys, 2019). Increased pressure on native bees in areas that previously served as a refuge for threatened or endangered species of insect pollinators is a concern.

4.3 Conclusions:

A better understanding of nectar resources on the landscape and methods to quantify their value to honeybees is important to beekeepers and any step in that direction is generally welcomed. With the caveats noted above, the modeling approach show here can distinguish in a rudimentary way between sites that are likely to be better for a particular varietal honey and those sites that should be avoided. However, the plethora of equations, numbers and data in this thesis should not be mistaken for precision by a commercial beekeeper. More study is needed before commercial application.

The term ‘precision agriculture’ refers broadly to the integration advanced technologies into the practice of agriculture (Lee, et al., 2010). Innovative uses of technology have lagged in apiculture. The techniques, equipment and business models used in beekeeping have not changed as quickly as those used in other agricultural sectors (Zacepins, et al, 2014). This study attempted to demonstrate that the use of remotely sensed data and computer modeling could be useful to the practice of apiculture as well.

APPENDIX A: CANDIDATE VARIABLES

Table A.1 Candidate Variables derived from Sentinel Satellite images:

Band Index	Source	Dates of Imagery
SAVI[1]	https://www.sentinel-hub.com	6/9/2016; 3/16/2017; 10/17/2017; 11/26/2017; 12/16/2017; 1/25/2018
MCARI[2]	https://www.sentinel-hub.com	6/9/2016; 3/16/2017; 10/17/2017; 11/26/2017; 12/16/2017; 1/25/2018
EVI[3]	https://www.sentinel-hub.com	6/9/2016; 3/16/2017; 10/17/2017; 11/26/2017; 12/16/2017; 1/25/2018
GNDVI[4]	https://www.sentinel-hub.com	6/9/2016; 3/16/2017; 10/17/2017; 11/26/2017; 12/16/2017; 1/25/2018
SARVI[5]	https://www.sentinel-hub.com	6/9/2016; 3/16/2017; 10/17/2017; 11/26/2017; 12/16/2017; 1/25/2018

[1] SAVI – Soil Adjusted Vegetation Index;
Sentinel Band combination:
[(B08 - B04) / (B08 + B04 + 0.428)]

[2] MCARI – Modified Chlorophyll Absorption Reflectance Index;
Sentinel Band combination:
[1.2 * (2.5 * (B08 - B04) - 1.3 * (B08 - B03))]

[3] EVI - Enhanced Vegetation Index;
Sentinel Band combination:
[(2.5 * (B08 - B04)) / (B08 + (6 * B04) - (7.5 * B02) + 1)]

[4] GNDVI - Normalized Difference Vegetation Index – Green;
Sentinel Band combination:
[B03* (B08 - B04) / (B08 + B04)]

[5] SARVI - Soil and Atmospherically Resistant Vegetation Index;
Sentinel Band combination:
[(1.0 + 0.487) * (B08 - (0.740 - 0.735 * (0.560 - 0.740))) /
(B08 + -(0.740 - 0.735 * (0.560 - 0.740)) + 0.487)]

Table A.2 Topo-climatic Candidate Variables:

Variable Abrev.	Name	Derived from / Source	Notes
Tp1	Aspect transformation	USGS 10M DEM	(see eqns. 2.1.4)
Tp2	Topographic Position Index	USGS 10M DEM	(see eqn. 2.1.5)
Tp3	Slope	USGS 10M DEM	
Tp4	Solar Radiance	ArcGIS 10.3	(7 month growing season)
Tp5	Topograpic Wetness Index	ArcGIS 10.3	
Tp6	Soil type	USDA/NRCS SSURGO	(inconsistent across counties)
Tp7	Temp/Precipitation	WorldClim bio-climatic	(B1-B12)

References:

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on, 19*(6), 716-723. doi:10.1109/TAC.1974.1100705

Anderson, C. (2001). The adaptive value of inactive foragers and the scout-recruit system in honey bee (Apis mellifera) colonies. *Behavioral Ecology, 12*(1), 111-119. doi:10.1093/oxfordjournals.beheco.a000372

Anderson, G. J. (1976). The pollination biology of Tilia. *American Journal of Botany, 63*(9), 1203-1212. doi:10.1002/j.1537-2197.1976.tb13205.x

Angel, P. N., & Christensen, C. M. (1976, 1976). *Honey Production on Reclaimed Strip Mine Soil.* Paper presented at the Hill Lands: Proceedings of an International Symposium, West Virginia University, Morgantown.

Arundel, J., Winter, S., Gui, G., & Keatley, M. (2016). A web-based application for beekeepers to visualise patterns of growth in floral resources using MODIS data. *Environmental Modelling and Software, 83*, 116-125. doi:10.1016/j.envsoft.2016.05.010

Baveco, J. M., Focks, A., Belgers, D., van der Steen, J. J. M., Boesten, J. J. T. I., Roessink, I., & Jonsson, M. (2016). An energetics-based honeybee nectar-foraging model used to assess the potential for landscape-level pesticide exposure dilution. *PeerJ, 4*(8). doi:10.7717/peerj.2293

Becher, M. A., Osborne, J. L., Thorbek, P., Kennedy, P. J., & Grimm, V. (2013). Towards a systems approach for understanding honeybee decline: a stocktaking and synthesis of existing models. *Journal of Applied Ecology*, 50: 868-880.

Beekman, M., & Ratnieks, F. L. W. (2000). Long-Range Foraging by the Honey-Bee, Apis mellifera L. *Functional Ecology, 14*(4), 490-496.

Beekman, M., D.J.T. Sumpter, N. Seraphides and F.L.W. Ratnieks 2004. Comparing foraging behaviour of small and large honey-bee colonies by decoding waggle dances made by foragers. *Functional Ecology,* 18: 829-835.

Bernhardt, E. S., & Palmer, M. A. (2011). The environmental costs of mountaintop mining valley fill operations for aquatic ecosystems of the Central Appalachians. *Annals of the New York Academy of Sciences, 1223*(1), 39. doi:10.1111/j.1749-6632.2011.05986.x

Braun, E. L. (1951). Deciduous Forests of Eastern North America. *Soil Science, 71*(2), 155. doi:10.1097/00010694-195102000-00012

Bryant, V. M., & Jones, G. D. (2001). The R-Values of Honey: Pollen Coefficients. *Palynology, 25*, 11-28.

Campbell, T., & Fearns, P. (2018). Honey crop estimation from space: detection of large flowering events in western Australian forests. *The International Archives of the Photogrammetry*, 79-86. doi:10.5194/isprs-archives-XLII-1-79-2018

Copernicus-ESA. (2017). European Space Agency. In: https://sentinel.esa.int/web/sentinel/home.

Couvillon, M.J., F.C.R. Pearce, C. Accleton, K.A. Fensome, S.K.L. Quah, E.L. Taylor and F.L.W. Ratnieks (2015). Honey bee foraging distance depends on month and forage type. *Apidologie* 46: 61-70.

Dalponte, M., Bruzzone, L., & Gianelle, D. (2012). Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment, 123*, 258-270. doi:https://doi.org/10.1016/j.rse.2012.03.013

Danner, N., Molitor, A. M., Schiele, S., Härtel, S., & Steffan□Dewenter, I. (2016). Season and landscape composition affect pollen foraging distances and habitat use of honey bees. *Ecological Applications, 26*(6), 1920-1929. doi:10.1890/15-1840.1

de Vries, H., & Biesmeijer, J. C. (1998). Modelling collective foraging by means of individual behaviour rules in honey-bees. *Behavioral Ecology and Sociobiology, 44*(2), 109-124. doi:10.1007/s002650050522

Douglas, B., Martin, M., Ben, B., & Steve, W. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Engler, R., Waser, L. T., Zimmermann, N. E., Schaub, M., Berdos, S., Ginzler, C., & Psomas, A. (2013). Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *Forest Ecology and Management, 310*(C), 64-73. doi:10.1016/j.foreco.2013.07.059

Ferreira, M. P., Zortea, M., Zanotta, D. C., Shimabukuro, Y. E., & de Souza Filho, C. R. (2016). Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sensing of Environment, 179*, 66-78. doi:https://doi.org/10.1016/j.rse.2016.03.021

Fielding, A. H. a. B., J.F. . (1997). A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. *Environmental Conservation, 24*, 38-49.

Frisch, K. v. (1967). *The dance language and orientation of bees*. Cambridge, Mass.: Cambridge, Mass., Belknap Press of Harvard University Press.

Gao, T., Zhu, J., Zheng, X., Shang, G., Huang, L., & Wu, Sr. (2015). Mapping Spatial Distribution of Larch Plantations from Multi-Seasonal Landsat-8 OLI Imagery and Multi-Scale Textures Using Random Forests. *Remote Sensing, 7*(2), 1702-1720. doi:10.3390/rs70201702

Ghiyamat, A., Shafri, H. Z. M., Mahdiraji, G. A., Ashurov, R., Shariff, A. R. M., & Mansour, S. (2015). Airborne hyperspectral discrimination of tree species with different ages using discrete wavelet transform. *International Journal of Remote Sensing, 36*(1), 318-342. doi:10.1080/01431161.2014.995272

Gregory, P. A., & Jean-Baptiste, F. (2012). Semi-Supervised Methods to Identify Individual Crowns of Lowland Tropical Canopy Species Using Imaging Spectroscopy and LiDAR. *Remote Sensing, 4*(8), 2457-2476. doi:10.3390/rs4082457

Henderson, E. B., Ohmann, J. L., Gregory, M. J., Roberts, H. M., & Zald, H. (2014). Species distribution modelling for plant communities: stacked single species or multivariate modelling approaches? *Applied Vegetation Science, 17*(3), 516-527. doi:10.1111/avsc.12085

Hijmans, R. J. (2015). Raster: Geographic Data Analysis and Modeling. R Package Version 2.4-15. . In: http://CRAN.R-project.org/package=raster.

Holtkamp, C. R., & Weaver, R. C. (2018). Quantifying the relationship between social capital and economic conditions in Appalachia. *Applied Geography, 90*, 175-186. doi:10.1016/j.apgeog.2017.12.010

Horn, T., Angel, P., Zipper, C., Ulyshen, M., French, M., Burger, J., & Adams, M. B. (2017). Re-establishing pollinator habitat on mined lands using the forestry reclamation approach. *Appalachian Regional Reforestation Initiative. Forest Reclamation Advisory No. 14. 7 p.*, 1-7.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398): John Wiley & Sons.

James, W., Petra Bohall, W., Matthew, C. N., William, J., Daniel, D., Glenn, W. S., . . . John, A. (2013). The Overlooked Terrestrial Impacts of Mountaintop Mining. *BioScience, 63*(5), 335-348. doi:10.1525/bio.2013.63.5.7

Jenness, J. (2007). Some Thoughts on Analyzing Topographic Habitat Characteristics. In: http://www.jennessent.com/downloads/topographic_analysis_online.pdf.

Kirkpatrick, E. (2015). Modeling the suitability of landscapes for managed honeybees - a case study in the northern Great Plains. In A. Davis, S. Dumyahn, & M. Henry (Eds.): ProQuest Dissertations Publishing.

Klesta, M. (2016, Spring 2016). Eastern Kentucky: A Region in Flux. *Forefront Federal Reserve Bank of Cleveland, 7,* 18-22.

Lee, J., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., . . . Coomes, D. A. (2016). Individual Tree Species Classification From Airborne Multisensor Imagery Using Robust PCA. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 9*(6), 2554-2567. doi:10.1109/JSTARS.2016.2569408

Lee, W. S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D., & Li, C. (2010). Sensing technologies for precision specialty crop production. *Computers and Electronics in Agriculture, 74*(1), 2-33. doi:10.1016/j.compag.2010.08.005

Lobo, J. M., Jiménez□Valverde, A., & Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography, 17*(2), 145-151. doi:10.1111/j.1466-8238.2007.00358.x

Luciano, P., & Mario, P. (2016). Sequential Relationship between Profitability and Sustainability: The Case of Migratory Beekeeping. *Sustainability, 8*(1), 94. doi:10.3390/su8010094

Mattise, J. (2014). Beekeepers eye Appalachian surface mines for hives. *The San Diego Union-Tribune*.

NASS. (2017). Honey 2017. *National Agricultural Statistics Service (USDA).*

Otto, C. R. V., Roth, C. L., Carlson, B. L., & Smart, M. D. (2016). Land-use change reduces habitat suitability for supporting managed honey bee colonies in the Northern Great Plains. *Proceedings of the National Academy of Sciences of the United States of America, 113*(37), 10430. doi:10.1073/pnas.1603481113

Ovaskainen, O., de Knegt, H. J., & del Mar Delgado, M. (2016). *Quantitative Ecology and Evolutionary Biology: Integrating Models with Data*: Oxford University Press.

Pericak, A., Kroodsma, D., Wasson, M., Ross, M., Clinton, N., Campagna, D., . . . Amos, J. (2018). Mapping the yearly extent of surface coal mining in Central Appalachia using Landsat and Google Earth Engine. *PLoS One, 13*(7), e0197758. doi:10.1371/journal.pone.0197758

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. In (Vol. 12, pp. 77-77).

Sadeghi, M., Babaeian, E., Tuller, M., & Jones, S. B. (2017). The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. *Remote Sensing of Environment, 198*, 52-68. doi:https://doi.org/10.1016/j.rse.2017.05.041

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation, 232*, 8-27. doi:https://doi.org/10.1016/j.biocon.2019.01.020

Schmickl, T., & Crailsheim, K. (2007). HoPoMo: A model of honeybee intracolonial population dynamics and resource management. *Ecological Modelling, 204*(1), 219-245. doi:10.1016/j.ecolmodel.2007.01.001

Seeley, T. (1986). Social foraging by honeybees: how colonies allocate foragers among patches of flowers. *Behavioral Ecology and Sociobiology, 19*(5), 343-354.

doi:10.1007/BF00295707

Seeley, T. (1994). Honey bee foragers as sensory units of their colonies. *Behavioral Ecology and Sociobiology, 34*(1), 51-62. doi:10.1007/BF00175458

Seeley, T., Camazine, S., & Sneyd, J. (1991). Collective decision-making in honey bees: how colonies choose among nectar sources. *Behavioral Ecology and Sociobiology, 28*(4), 277-290. doi:10.1007/BF00175101

Seeley, T. D. (1985). *Honeybee ecology : a study of adaptation in social life*. Princeton, N.J.: Princeton, N.J. : Princeton University Press.

Solymos, P., & Lele, S. R. (2015). Revisiting resource selection probability functions and single-visit methods: Clarification and extensions. *7*(2). doi:10.1111/2041-210X.12432

Søren, H., Ulrich, H., & Jun, Y. (2005). The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software, 15*(1), 1-11. doi:10.18637/jss.v015.i02

Sponsler, D. B., Johnson, R. M., & Jonsson, M. (2015). Honey bee success predicted by landscape composition in Ohio, USA. *PeerJ, 3*(3). doi:10.7717/peerj.838

Stabentheiner, A., & Kovac, H. (2014). Energetic Optimisation of Foraging Honeybees: Flexible Change of Strategies in Response to Environmental Challenges. *PLoS One, 9*(8), e105432. doi:10.1371/journal.pone.0105432

Thenius, R., Schmickl, T., & Crailsheim, K. (2006). Economic optimisation in honeybees: Adaptive behaviour of a superorganism. *From Animals To Animats 9, Proceedings, 4095*, 725-737.

USDA. (2017). National Honey Report. In *Federal Market News Service* (Vol. Number XXXVII): United States Department of Agriculture.

USDA-NAIP. (2016). National Agricultural Imagery Program. In. https://gis.apfo.usda.gov/arcgis/rest/services.

USFS-FIA. (2017). *US Forest Service Forest Inventory and Analysis*. https://www.fia.fs.fed.us/tools-data/index.php.

Vanengelsdorp, D., & Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *Journal of Invertebrate Pathology, 103*(1), S80-S95. doi:10.1016/j.jip.2009.06.011

Wainselboim, A. J., Roces, F., & Farina, W. M. (2002). Honeybees assess changes in nectar flow within a single foraging bout. *Animal Behaviour, 63*(1), 1-6. doi:10.1006/anbe.2001.1879

Wilson, B. T., Lister, A. J., & Riemann, R. I. (2012). A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management, 271*(C), 182-198. doi:10.1016/j.foreco.2012.02.002

Winston, M. L. (1987). *The biology of the honey bee*. Cambridge, Mass.: Cambridge, Mass. : Harvard University Press.

Zacepins, A., Brusbardis, V., Meitalovs, J., & Stalidzans, E. (2015). Challenges in the development of Precision Beekeeping. *Biosystems Engineering, 130*, 60-71. doi:10.1016/j.biosystemseng.2014.12.001

Zipper, C., Burger, J., Skousen, J., Angel, P., Barton, C., Davis, V., & Franklin, J. (2011). Restoring Forests and Associated Ecosystem Services on Appalachian Coal Surface Mines. *Environmental Management, 47*(5), 751-765. doi:10.1007/s00267-011-9670-z

Vita

Doug Potter was born in Mobile, Alabama. He received a Bachelor of Arts in economics from Brown University in 2000.

Douglass W Potter