2017

PREDICTING HABITAT DISTRIBUTION FOR FIVE RARE PLANT SPECIES WITHIN THE BLACKFOOT SWAN LANDSCAPE RESTORATION PROJECT

Annalisa Suzan Ingegno

2017

PREDICTING HABITAT DISTRIBUTION FOR FIVE RARE PLANT SPECIES WITHIN THE BLACKFOOT SWAN LANDSCAPE RESTORATION PROJECT

Annalisa S. Ingegno

PREDICTING HABITAT DISTRIBUTION FOR FIVE RARE PLANT SPECIES WITHIN THE BLACKFOOT SWAN LANDSCAPE RESTORATION PROJECT

By

Annalisa Suzan Ingegno

B.A. Biology, University of Montana, Missoula, MT 2013

Thesis

Presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Geography

The University of Montana
Missoula, MT

Spring 2017

Approved by:

Dr. Scott Whittenburg, Dean of the Graduate School
Graduate School

Dr. Anna Klene, Chair
Department of Geography

Dr. Faith Ann Heinsch, Member
Department of Geography, Affiliate

Dr. Ragan Callaway, Outside Member
Department of Biology

Ingegno, Annalisa, M.S., Spring 2017 Geography

Predicting Habitat for Five Rare Plant Species Within the Blackfoot Swan Landscape Restoration Project

Chairperson: Dr. Anna Klene

This study predicted rare-plant habitat at the landscape scale. Using the Maximum Entropy (MaxEnt) algorithm, relationships to and between each environmental variable were quantified for five species in the Blackfoot Swan Landscape Restoration Project (BSLRP) study area in western Montana. This project is part of a greater vegetation assessment for BSLRP that utilizes remotely sensed products for planning and management purposes. The five rare plant species studied in this analysis were common camas (*Camassia quamash*), clustered lady's-slipper (*Cypripedium fasciculatum*), western pearlflower (*Heterocodon rariflorum*), Howell's gumweed (*Grindelia howellii*), and crested shieldfern (*Dryopteris cristata*).

Rare plant models typically do not address dispersal mechanisms in conceptual design. This analysis built dispersal mechanisms into model design by buffering the project area based upon dispersal potential. Plant population data is typically stored as polygons in state and federal databases. This data is usually condensed into a single point before entry into modeling algorithms. This analysis addressed this issue by proportionately placing multiple points inside the polygons. In addition, this analysis considered the effects of using different regularization parameter values in MaxEnt and how it affected model performance. For one species, the efficacy of including LiDAR-derived canopy cover to enhance discrimination of understory communities and its effect on improving model performance was examined. Accuracy assessments were used to better understand predictions and statistical relationships between environmental variables. Lastly, predicted habitat maps were overlaid to identify areas of high probability habitat for multiple species across the project area.

The field surveys identified thirteen new populations of plants. Overall accuracy for the predictions ranged from 26 to 69%. Comparison between predictions based upon centroids versus distributed points yielded an improved AUC and reduced standard deviation. The LiDAR data defined a narrower niche and had an improved AUC and lower standard deviation but did not lower assessed accuracy for crested shieldfern. Species distribution model studies are one way for resource managers to identify potential habitat for rare species before going into the field. They can use this information to prioritize field surveys and inform management decisions. Also, SDM studies provide information on species' environmental associations and can be used to further understand species ecology.

DEDICATION

To Coco.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Anna Klene for her mentorship through times of darkness and light. Your advice has not fallen upon deaf ears! To the love of my life, Dr. William Chad Stasey- your love and support have been a beacon of light in my life. I never could have done this without you.

Thank you to Dr. Ray Callaway and Dr. Faith Ann Heinsch for their assistance and encouragement during my thesis.

I would also like to thank the US Forest Service and the Blackfoot Swan Landscape Restoration Project for this opportunity. Chip Fisher was a pivotal force in acquiring the funding and support necessary to make this happen. Botanists Chantelle Delay, Karen Stockman, and Justina Fedorchuck provided important insight in plant phenology and habitat.

To Heidi Fleury, whose excitement for learning knows no bounds. Thank you for going above and beyond the call of duty to collect data in rain, snow, and around the den of the glorious beaver.

To my friends and family that helped me through the rigors of thesis design, modeling, and writing – thank you! To Kristin Swindle for providing levity to the situation. Thanks to all my wonderful friends who played music with me including Mark and Laurie Chase. And of course to my Dad, Michele, Patrick, and Grandma for their support.

Thank you to Steve Brown and the R1 Geospatial Group for keeping me employed throughout my thesis. I will not soon forget the kindness I have been shown.

I would like to thank the multitude of others who provided me with advice, both scholarly and otherwise.

Lastly, Gus, June, and Bria were always good for a couple of licks but not much else. And of course, Coco. You will not be forgotten.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The future of biological diversity is uncertain. The current rate of extinction on Earth is 1,000 times greater than the historic background rate (IUCN 2017). In Montana alone, 11% of all vascular plant species are considered at risk of extirpation (Montana Natural Heritage Program (MTNHP) 2017). This risk can be mitigated through management of habitat and related environmental resources (Guisan and Zimmerman 2000). However, management relies on conservation and environmental planning to decide realistic courses of action, and that is contingent on understanding species distribution across the landscape (Guisan and Zimmerman 2000, Vaughan and Ormerod 2003).

Plants are designated as sensitive if populations are threatened with extinction due to decreasing numbers or loss of habitat (USFS 2005). The United States Forest Service (USFS) oversees management of endangered, threatened, and sensitive species on national forest land, including the 18% of Montana that is designated as national forest. Management of these sensitive species focuses on maintaining viable populations and mitigating that risk for species listing under the US Endangered Species Act (USFS 2005).

For land management project planning on USFS lands, managers need to collect specific, detailed information on species within a project area in order to prepare the requisite legal documents, such as those required by the National Environmental Policy Act (NEPA) (1969). This is done through field surveys of proposed project areas. One technique is to field survey the majority of a proposed project area for rare plants (Elzinga et al.1998). Under this methodology, the majority of a project area is surveyed, regardless of whether or not it has been identified as potential habitat beforehand. This is ideal for smaller project areas with high potential for rare plants. While this methodology provides the greatest coverage, it is both time and labor intensive. Targeted sampling is another practice which identifies suitable habitat for threatened species before field assessments are performed (Elzinga et al.1998). Ideally, botanists select only the most highly probable habitat for survey by the field technicians, thus decreasing the amount of time and money spent surveying the area while still achieving an effective sample. Technicians then field survey the potential habitat by walking through proposed units to collect information on species occurrence. While targeted sampling can decrease potential habit to be surveyed, field surveys are still labor intensive, expensive, and often yield poor detection of target species (Elzinga et al.1998).

Geospatial Information Systems (GIS) make it possible to manage landscapes more effectively using remotely sensed and other spatially explicit data (Franklin 2009). These data are used in species distribution models (SDM) to extrapolate relationships between species and environmental factors across the landscape of interest (Guisan and Zimmerman 2000, Elith and Leathwick 2009, Franklin 2009). SDM outputs represent areas of potential habitat across the landscape and can be used to target field surveys. Targeting areas of high probability habitat improves field survey efficiency and rare plant detection prior to project implementation. SDMs can also test how management scenarios may affect species or landscape-level species composition in the future (Guisan and Zimmerman 2000, Elith and Leathwick 2009, Franklin 2009). SDMs can sometimes answer ecological questions about species of interest by evaluating the significance of environmental factors influencing distribution.

This study was designed to predict rare plant habitat at the landscape scale using fine-resolution remotely sensed products, which had not previously been done in this area which has a paucity of knowledge on rare species. These five species are considered at risk of extirpation or are culturally significant to Native American tribes in this area. Modeled species included common camas (*Camassia quamash*), clustered lady's-slipper (*Cypripedium fasciculatum*), western pearlflower (*Heterocodon rariflorum*), Howell's gumweed (*Grindelia howellii*), and crested shieldfern (*Dryopteris cristata*). Using the Maximum Entropy model (MaxEnt), relationships to each environmental variable were quantified for five species within the Blackfoot Swan Landscape Restoration Project (BSLRP). A LiDAR-derived environmental variable was tested for its efficacy in improving model accuracy. This provided some insight into ecological relationships between these rare plant species and their environment. This analysis improves upon methodologies typically followed for predicting rare plant habitat by addressing ecological issues surrounding dispersal and limitations of polygon population data. This project is part of a greater vegetation assessment for BSLRP that utilizes models constructed from remotely sensed raster products for planning and management purposes. It can also serve as a resource to assess risk, prioritize work, and provide additional information on species distribution and ecology for project planning.

CHAPTER 2: BACKGROUND

Species Distribution Models

One type of environmental model, species distribution models (SDMs), predicts potential habitat for a species using observation and environmental data (Guisan and Zimmerman 2000). Typically, in the case of rare plant species, an SDM estimates plant distribution across a study area using known plant locations (observation data) and environmental data such as precipitation. The necessary information to make informed management decisions for rare plant planning is typically limited because the spatial distribution of populations across the landscape is not well understood (Gogol-Prokurat 2011). SDMs can produce habitat suitability maps for rare species to help resource managers prioritize field surveys. They also allow us to test hypotheses on which environmental variables have the greatest effect on species distribution. This information can be useful in biogeographical studies and can provide insight species ecology (Elith & Leathwick 2009). SDMs can also simulate the impacts of various management strategies on plant distribution, as in this study.

In the USFS, sensitive species are defined as those with decreasing population size or those for which reduction in habitat would threaten their survival or distribution (USFS 2005). Management of these species focuses on maintaining population viability to avoid the possibility of an Endangered Species Act (ESA) listing by the US Fish and Wildlife Service (USFS 2005). Sensitive species are selected by the Regional Forester in each USFS Region using information from the Fish and Wildlife Service, state databases, and other relevant sources (USFS 2005). The Montana Natural Heritage Program is a state agency that maintains population information on plants and animals found within all lands in the state of Montana. Their focus is on data collection for species of concern, or plants and animals that are at risk of being extirpated from the state.

MaxEnt

The species distribution model employed in this analysis was Maximum Entropy (MaxEnt). Using species occurrences and environmental data, MaxEnt produces a relative likelihood of observation, or habitat suitability, for a species. MaxEnt is a presence only model, meaning absence data is not required for modeling exercises (Phillips et al. 2006). Instead, MaxEnt randomly samples environmental variable points from the study area to create background data, which is used as a proxy for absence data (Phillips et al. 2006, Phillips et al. 2009, Elith et al. 2011). MaxEnt constructs a statistical distribution of environmental variables, including mean and deviation, found at occurrence and at background points (Elith et al. 2011). Entropy is maximized, and statistically demonstrated, by the output prediction whose distribution is most similar to that of the occurrence points (Phillips et al. 2006, Elith 2011). High probability areas are those with environmental values most similar to the mean environmental values found at occurrence points.

Occurrence Data

Much of the data on rare plant occurrences comes from herbarium records and state databases. Although a rich data source, herbarium data suffers from high locational uncertainty and biased collection records (Franklin 2009). SDM input data is assumed have been randomly collected from across the landscape (Phillips 2005, Kramer-Schadt 2013, Fourcade 2014). Despite this programmatic assumption, most occurrences used in SDMs were collected opportunistically with no predefined sampling framework (Elith 2011). Accuracy of model outputs is usually more dependent on locational certainty of occurrences than the number of occurrences (Engler et al. 2004). Thus, it would be optimal to clean input data by removing occurrences with high locational uncertainty and ensuring taxonomy of occurrence data when possible.

BSLRP Study Area

In the Rocky Mountains, disturbances such as fire, disease, and/or insect outbreaks are principal determinants of landscape heterogeneity, with fire the most important (Bassman 2015). In 2015, over 10 million acres (4,046,860 hectares) burned in the US, with associated suppression costs surpassing $2,000,000,000 (NIFC 2017). State and federal agencies such as the USFS manage responses to wildfires a cost to other agency objectives. Currently, more than 50% of the USFS budget is spent on fire management, while in the 1990s it comprised around 15% (USDA Forest Service 2015).

Beginning in the mid-19th century, naturally occurring wildfires were suppressed when possible (Bassman 2015). Over time, suppression of the natural fire regime has resulted in overgrown forests with much longer fire return intervals (Keane 2002, Finney and Cohen 2003, Bassman 2015). Decades of fire suppression precipitated undesirable forest conditions, increasing management costs. In 2009, the Collaborative Forest Landscape Restoration Program (CFLRP) was passed by Congress, which provided funding for collaborative, landscape-scale forest management (USDA Forest Service 2017a). Its goals include reducing wildfire risk, improving forest health, and creating jobs. The Blackfoot Swan Landscape Restoration Project (BSLRP) is a landscape-level project funded under CFLRP in western Montana. The goal of

BSLRP is to restore a natural fire regime to a landscape that has the potential to support more resilient and diverse ecosystems (Bassman 2015).

BSLRP encompasses 1.2-million acres (4,856,228 hectares) in the Mission Ranges of the Rocky Mountains in western Montana (Figure 1). Approximately 70% of the project area is publicly-owned land and consequently, over twenty local, state, and federal collaborative groups have been involved in development and implementation of project goals (USDA Forest Service 2017b). It spans three ranger districts across three National Forests (NF): Lincoln Ranger District on the Helena – Lewis & Clark NF, Seeley Lake Ranger District on the Lolo NF, and Swan Lake Ranger District on the Flathead NF. The project is bordered by some of Montana's most remote mountainous areas including the Bob Marshall Wilderness Complex in the East, the Mission Mountain range in the West, and the Garnet Mountain Range in the South.

Research Questions

There is little information on the distribution of rare plant species within BSLRP. Information on rare plant distribution is necessary for resource managers to complete project planning. This analysis sought to address this need using remotely sensed imagery products and the MaxEnt algorithm to predict rare-plant habitat for five rare plant species in BSLRP and answering the following questions:

(1) How well can we predict habitat for rare plants in BSLRP?

(2) Can we improve SDM modeling for rare plant species by addressing:

(a) the dispersal potential of each species,

(b) occurrence data stored as polygons,

(c) if LiDAR data improves model performance, and

(d) over-prediction of rare plant habitat.

CHAPTER 3: METHODS

Computational modeling was a desirable and pragmatic option for identifying habitat within the BSLRP project area. Botanists on each National Forest included in the BSLRP project provided knowledge on potential plant habitat that was incorporated into models when possible. Botanists were included in almost every aspect of the modeling process, from design to field assessment, and their participation was an important aspect of the process because they will use the results to complete impact statements.

Preliminary analysis included compiling sensitive species located within the project area from the Natural Resource Manager (NRM) Threatened and Endangered Species (TESP) database. Botanists on all Forests in BSLRP met with the regional botanist to discuss potential species, needs, as well as a generalized modeling approach in February 2016. At this meeting, a list of all potential species as well as two potential modeling methods were given to the botanists by the author.

Species Descriptions

Five species were selected by botanists for modeling: common camas (*Camassia quamash)*, clustered lady's-slipper (*Cypripedium fasciculatum)*, western pearlflower (*Heterocodon rariflorum)*, Howell's gumweed (*Grindelia howellii)*, and crested shieldfern (*Dryopteris cristata*; Table 1, Figure 2).

Camassia quamash (Frederick Traugott Pursh) Edward Lee Greene: Common camas

Common camas is a bulb perennial in the lily family. Growing up to 60 cm tall, it has slender basal leaves and a thin leafless stalk that leads to a dark blue to purple inflorescence of six tepaled (petals for monocots) flowers (Lesica et al. 2012, MTNHP 2017a). After flowering, the capsule fruits mature to produce dozens of tiny black seeds. Common camas is found in moist meadows from valley bottoms to subalpine areas below 3,300 m (USDA Plants Guide 2006). It prefers vernally moist soils that dry out later in the year. Cover is moderate to fairly open, with little competition from shrubs or weeds (USDA Plants Guide 2006). Common camas is found throughout western Montana and is not considered rare or threatened by state or federal agencies. It is a sensitive species, however, because of its cultural significance to Native American tribes in the area (Lackschewitz 1991).

Common camas is a perennial that reproduces via seed and bulbs (USDA 2006). Seeds must be cold stratified to germinate but are not adapted for long distance dispersal (Emery 1988). Their secondary reproductive mechanism, bulbs, also have a relatively low dispersal potential. Literature suggests that they can be dispersed short distances by gophers (Watson 1988, Beckwith 2004). Because there are two mechanisms for dispersal, this species has a higher combined dispersal potential than from either single mechanism and was given a large dispersal buffer of 2.5 miles outside the project area.

Cypripedium fasciculatum Albert Kellog ex Sereno Watson: Clustered lady's-slipper

Clustered lady's-slipper is a rhizomatous perennial in the orchid family. Growing up to 20 cm tall, it has a slender hairy stem with two oval opposite leaves (MTNHP 2017b). The inflorescence contains 2 to 4 purple-brown flowers, each with a single sac. Each flower produces a capsule with thousands of tiny seeds that are easily dispersed by the wind, and thus it has a high dispersal potential. It was given a dispersal buffer of 2.5 miles outside the project area. (Seevers and Lang 1998). Clustered lady's-slipper is found in dry to moist shaded habitats with a dense (60 – 100%) overstory consisting of Douglas-fir, grand fir, subalpine fir and shrubs

(MTNHP 2017b). Temperature and precipitation are thought to be the most important determinants affecting orchid distribution (Cornell 1978). This species is hypothesized to need a mycorrhizae, or fungal symbiont, to derive nutrients from the soil and thus has a fragile root system (Arditti 1967, Wells 1981, Doherty 1997, Seevers and Lang 1998). This species is ranked as sensitive by the USFS Region 1 and is designated as a species of concern by the MTNHP. Threats include timber harvest activities, but this species' listing is due to its fragile relationship with soil mycorrhizae (MTNHP 2017b).

Heterocodon rariflorum Thomas Nuttall: Western Pearl-flower

Western pearl-flower is an annual in the Campanula family. It is the only species found in the *Heterocodon* genus. Growing up to 30 cm tall, it can have one to many thin stems with small, serrated circular leaves along the stem. Single flowers are found in leaf axils and are cuplike with blue to purple petals (Morin 2017). After flowering it produces a capsule filled with hundreds of small seeds (MTNHP 2017c). The dispersal potential for this species is relatively low since these seeds are gravity dispersed over a short distance (MTNHP 2017c). Because of its limited dispersal potential, western pearlflower was given a dispersal buffer of 1.0 miles outside the project area. It is typically found on wet soils, including ephemeral mossy seeps, rocks, and swales from foothills to montane areas. It is ranked sensitive by USFS Region 1 and is ranked as a species of concern by the MTNHP due to weed encroachment on populations (MTNHP 2017c).

Grindelia howellii Steyermark: Howell's Gumweed

Howell's gumweed is a perennial in the Aster family. Growing up to 90 cm tall, Howell's gumweed has gland tipped hairs along its stem, which has a woody base (NatureServe 2015, MTNHP 2017d). Its ray flowers are yellow, with hooked, sticky involucre bracts (NatureServe 2015). It produces several dozen seeds per flowering head; the sticky flowering head attaches easily to things that brush against it. Thus, long distance dispersal for this species is possible; a 2.5 mile dispersal buffer was drawn around the project area for this species. This species is found only in Idaho and Montana, where it is restricted to Missoula and Powell counties (NatureServe 2015). It is a ruderal species that prefers disturbance and is often found along roadsides. Habitat includes the foothills, valleys, and grasslands at elevations lower than 1,700 meters (NatureServe 2015). Howell's gumweed is ranked sensitive by USFS Region 1 and is categorized as a species of concern by MTNHP due to weed encroachment on populations (MTNHP 2017d).

Dryopteris cristata (Linnaeus) Asa Gray: Crested shieldfern

Crested shieldfern is a spore-bearing perennial fern. Each frond grows up to 60 cm long and is elliptical with pairs of staggered leaflets along the stem (eFloras 2008, MTNHP 2017e). Fertile leaves with sori, or spore containing structures, die back in the fall leaving sterile fronds that persist through winter. Sori are covered in a "u" shaped indusium, or sheath. It produces thousands of spores that are dispersed by wind and gravity. Spores are specifically designed to withstand harsh conditions and are known to disperse far distances (Peck et al. 1990). Thus, the dispersal potential for this species is high; Table 1 shows it was given a dispersal buffer of 2.5 miles outside the project area. It is typically found in montane areas along the edges of swamps, wetlands, and heavily moist soils. Crested shieldfern typically occurs with Alder (*Alnus sp.)* in shallow standing water. Preferred areas include bottomlands near meadows, rivers, inlet streams, and wetlands. It is ranked as a sensitive species by USFS Region 1 and is considered a species of concern by the MTNHP because of few known occurrences across the state (MTNHP 2017e).

Data Sources

Three data sources were utilized in the modeling effort: Natural Resource Manager (NRM) TESP, Pacific Northwest Herbaria (PNWH), and MTNHP. The NRM TESP is a USFS database of rare plants found on federal lands. It is a robust database that includes pertinent species information such as canopy cover, associated species, and threats to the population. PNWH is an online consortium of herbaria located across the Northwestern US. Herbarium data can be a rich as they span a longer time period than federal databases. However, older records typically have high locational uncertainty, or no geographic information whatsoever, making them essentially unusable for modeling. Lastly, MTNHP maintains a statewide database of plant occurrence data collected by various agencies. Their database is inclusive of most herbarium data and includes occurrence data from the NRM database but will occasionally contain additional records.

Species data were collected from all three sources as both point and polygon data. Multipart polygon observations were broken into single polygons. Redundant observations were removed using ArcMap 10.2.2 (ESRI 2013). When observation data for a single record spanned multiple databases, care was taken to retain all information from each database. Herbarium specimen data were used to improve the locational certainty of records when possible. The locational uncertainty for each feature was assessed, and those observations with uncertainty greater than 1,000 m were removed.

Model Design

Dispersal Buffer

Plants differ from animals because they are sessile and cannot move across the landscape. However, plants have evolved a variety of unique ways to disperse across their environment. For example, some perennial plants have underground bulbs that split to produce offspring called bulblets. Other plants produce winged seeds that permit them to disperse miles from their parent plant. The five species included in this analysis have a broad spectrum of reproductive strategies including seeds, bulbs, and spores.

Typically, SDMs fail to address dispersal in their conceptual design. To address dispersal potential in this model, data on species dispersal mechanisms such as dispersal type and dispersal method were gathered (Table 1). A qualitative assessment was made for each species based on our understanding of its dispersal mechanisms.

Gravity dispersed species received the smaller value – the seed simply falls to the ground instead of traveling away from the parent. For species with two dispersal mechanisms, the longer distance was used, as these species had a greater potential to disperse farther from their parent. Wind dispersed species had a higher potential to disperse much farther from their parent plant than gravity-dispersed plants. Similarly, spores were given the longer dispersal qualitative value, since they are known to travel far distances (Peck et al. 1990). The dispersal values of 1.0 and 2.5 miles outside of the project area we given based on the recommendation of project botanists.

Polygon-to-Points

MaxEnt requires occurrence data to be input as point locations. Because population size and extent are used by resource managers in project analysis, rare-plant data are typically collected as, and represented by, polygons when stored in databases. A single polygon can potentially encompass a range of environmental values that are suitable habitat for that species. Condensing a population polygon into a single point is the most common practice when entering polygon occurrence data into MaxEnt. However, this method of representing the data loses some of the ecological information, since it usually attributes only a single set of environmental values to that population. As seen in Figure 3, the actual population might be represented by a gradient of environmental values, such as elevation. For example, a single plant population may span an area that varies in elevation from 1,250 to 1,450 m. The centroid method would use one value, but to address this, the point-to-polygon method proportionately assigns points to polygon areas. This allows for a range of environmental values to be represented for each population in the model, potentially producing a more accurate output.

Polygon population areas were calculated in ArcMap and binned into size classes based on model resolution (30 m × 30 m). Populations were binned, by area, into one cell or less, which represented a total area of 900 m^2. The second bin contained populations with areas between one and four pixels, or between 900 and 3,600 m^2. The third bin contained all populations with areas between four and nine pixels, or with areas between 3,600 and 8,100 m^2. Random points were then assigned to each size class by squaring the bin number. For the first bin just one point was assigned, for the second bin four points were used, and for the third bin nine points were randomly placed inside that population area. Table 2 displays the number of random points created from each species' samples using the polygon-to-point method. The variation in the number of points for each species is due to differences in population areas.

Environmental Variables

There is substantial literature on selecting environmental variables for use in plant modeling exercises. Before selecting environmental variables, it is important to develop a conceptual model for environmental variables affecting plant distribution (Franklin 2009). The "primary environmental regime" framework states that plant distribution is controlled by light energy, heat, water, and nutrients (Mackey 1993). Another framework suggests that distribution is directed by plant response to water, light energy, nutrients, and the sum of heat energy (Guisan and Zimmerman 2000). Franklin (1995) suggests that geology, climate, and topography are commanding factors for the primary environmental regime, which in turn determines distribution. Because plants are the primary producers in the food chain, Franklin (2009) suggests constructing a generalized suite of environmental variables that can be used to explain distribution.

The primary issue when selecting environmental variables for SDM modeling is the availability and scale of data across the area of interest (Goodchild 1996). Interactions between plants and their environment is exceedingly complex. Thus constructing a suite of relevant environmental variables for modeling can be difficult. In addition, statistical principles suggest that reducing model complexity by incorporating fewer environmental variables, especially when faced with small sample sizes, generally produces better models (Gogol-Prokurat 2011, Warren and Seifert 2011). Research suggests that because it is rare to have environmental data that describes actual distribution, SDMs should use whatever appropriate data can be acquired, with the knowledge that there will be some uncertainty introduced to the model (Barry and Elith 2006).

MaxEnt is capable of ranking variables by statistical contribution to the overall model. This feature can be used to test which variables are best suited for a model. This capability was not fully employed in this analysis, but variable rankings were used to confirm ecological relevance of each output. Correlation between environmental variables can result in erroneous rankings of variables by importance in MaxEnt (Phillips et al. 2006), and uncorrelated variables are preferable. For this study, a suite of ecologically relevant environmental variables were selected and utilized for all models. This suite of variables was chosen based not only on their ecological relevance but also on their availability and usage in similar modeling studies. Previous rare-plant modeling by Nock (2008) and Burbach (2011) for the USFS in Region 1 presented a generalized framework for environmental variable selection that was followed in this analysis.

Seven environmental variables were selected according to their biological and ecological relevance for plants in the study area. Table 3 describes the seven variables utilized in this analysis as well as the data resolution and source. These included climatic (PRISM precipitation; PRISM 2015) and topographic (solar insolation, elevation, and topographic wetness index) variables which are useful in plant SDM for areas with high topographic variation (Zimmerman and Kienast 1999).

Elevation is a topographic variable commonly used in plant SDM studies because it affects temperature and precipitation which relate to water and heat availability in the primary environmental regime (Moore et al. 1991, Raven 2009), which then affects distribution. A 30 m Digital Elevation Model (DEM) was constructed by Chip Fisher the GIS specialist on BSLRP from the 30 m National Elevation Dataset (NED) from 2005.

Vegetation, or land-cover classification, schemes are a description of the ecosystem as a whole (Kochler 1973). They are used in plant SDMs to capture a species' association with specific habitat types. Vegetation classifications are frequently used in SDMs because they are widely available at a usable scale (Leyequien et al 2007). VMap is a raster vegetation classification geospatial database developed by USFS Region 1 Geospatial Group (Barber et al. 2012). Its primary purpose is to provide information on current vegetation conditions for USFS Region 1. Remotely sensed data were classified into vegetation units that describe lifeform, tree dominance type, tree canopy cover class, and tree diameter for each Forest. The VMap product utilized here describes a single vegetation species with greater than 60% cover within a stand. VMap was resampled from 10 ×10 m resolution to 30 × 30 m resolution (Barber et al. 2012).

Solar insolation is the amount of solar energy hitting an area, which also affects the amount of heat that area receives. The amount of incoming solar radiation is an important determinant of plant community composition, density, and distribution (Satterlund and Means 1978). Accurate calculation of solar insolation is computationally involved, requiring knowledge of cloud cover, latitude, and reflectivity for the area of interest on a specific date. For this analysis, solar insolation was estimated in ArcMap using the Spatial Analyst extension. It was used to calculate a raster layer to estimate the maximum amount of solar energy over a specified period hitting the surface of each pixel in KWh/m^2. This calculation, based solely upon geometry derived from latitude and elevation, does not include meteorological or surficial observations such as cloud or canopy cover.

Although climatic conditions influence species distribution at coarser scales, edaphic variables, or those related to soil conditions, can influence distribution at fine scales (Zhang 2013). Soils contain moisture and nutrients necessary for plant growth and thus are an important consideration in modeling. In the primary environmental regime, nutrients are a key factor determining distribution (Mackey 1993). Soil and related geology affect the accessibility of both nutrients and water (Franklin 2009). For example, many rare species are restricted to microhabitats that contain a certain soil type. Soil is composed of both organic and non-organic substances derived from parent material. Breakdown of parent material such as weathering of rocks, provides minerals and structure for soil (Jenny 1941). Thus, geology is an important determinant of soil type and is used frequently in plant SDM studies (Beauregard and de Blois 2014). The Northern Rockies geology layer is managed by the USGS and USFS and is used for project analysis and research dealing with geologic resources (Zientek et al. 2005). This database was constructed from 1:100,000 to 1:250,000 scale maps across Montana, Idaho, and Washington. It is symbolized according to the map unit, which contains a single type of rock of the same age.

Topography is an important determinant of surface hydrology. These hydrological conditions in turn, influence soil characteristics including moisture and pH (Sorensen 2006). Soil moisture is frequently cited as an important determinant of plant distribution (Giesler et al. 1998). The Topographic Wetness Index (TWI) was created as a proxy for measuring soil moisture conditions (Sorensen 2006). It is calculated using flow accumulation and slope upstream derived from a DEM. The calculation for TWI is:

$$TWI = \ln \frac{FlowAccumulation * Pixel\ Area}{\tan \frac{Slope * \pi}{180}}$$

(Beven and Kirkby 1979, Vendrusculo and Kaleita 2013).

Nearly every SDM at the regional scale utilizes some form of climate data, such as mean annual precipitation, for predicting habitat (Franklin 2009). Precipitation was the only climatic data used in the analysis. Thirty-year (1971 – 2000) mean annual precipitation data were downloaded from the PRISM climate group at 800 × 800 m resolution (PRISM 2016). The 30-yr average climate normal dataset is the most commonly used and peer-reviewed of all PRISM data (PRISM 2016). This layer was resampled using bilinear interpolation to a 30 × 30 m raster.

LiDAR

Light Detection and Ranging (LiDAR) is a highly accurate remote-sensing technique that uses an active sensor to construct a 3 dimensional (x, y, z) point cloud of the area of interest (Collis 1966). Laser pulses emitted from the sensor is reflected off forest canopy objects and returned to the sensor. The time delay is used to calculate the heights of various feature on the landscape (Collis 1966, Gatziolis and Andersen 2008). A single LiDAR pulse can potentially have up to five returns, each representing a different height object on the landscape. For instance in a heavily forested area, the first return could represent tree tops, second return understory trees, third return shrubs, and the last return might have reached the ground surface. LiDAR data is utilized for highly accurate topographic mapping operations, investigating vegetation structure, and constructing other forest derivatives (Lefsky et al. 2002).

LiDAR-derived canopy metrics are used to improve fuel models, provide information on vegetation structure for wildlife management, and create highly accurate elevation models, among others (Lefsky et al. 2002, Andersen et al. 2005). Forest canopy affects the availability of light and water for the understory and is therefore highly influential on understory community composition (Jennings et al. 1999). Forest canopy is typically calculated using canopy cover, which is the percentage of the ground obscured by leaves and limbs (Jennings et al. 1999). Canopy-cover raster layers were derived from LiDAR data collected between 2013 and 2015 for 28% of the project area at a resolution of 8 points/m^2. The raw LiDAR data were processed by

USFS contractor Red Castle Resources in Salt Lake City, Utah to produce canopy cover using FUSION software.

Variable Correlations

The principle of parsimony states that an observed event is best explained by the simplest terms (Epstein 1984, Vandekerckhove et al. 2014). In modeling, we try to achieve the most parsimonious model by balancing the tradeoff between model complexity and accuracy. Exceedingly complex models may have many potentially correlated variable terms and thus provide little information to enhance the model, while greatly enhancing the potential error that is often multiplicative (Vandekerckhove et al. 2014). Thus, variable correlations were calculated using Microsoft Excel for the first five models using the first six variables (elevation, solar insolation, precipitation, geology, land cover, and TWI). No correlations were found between them (Table 4). Separate variable correlations were also calculated for the seven variables used in the crested shieldfern model, which included LiDAR-derived canopy cover. No variable correlations were found in this model as seen in Table 5. Note that the correlation values between similar variables vary between these models due to the different extent of the crested shieldfern model with LiDAR-derived canopy cover.

MaxEnt

MaxEnt was run as a standalone Graphical User Interface (GUI) accepting occurrence and environmental data to produce a distribution prediction. As mentioned before, all environmental data must use the same projection, cell size, and extent. All environmental variable data were resampled to 30 × 30 m resolution in ArcMap with a projection of NAD83

UTM 12 N in ASC format. Occurrence data were entered as a single CSV file of latitude and longitude points. The following settings were utilized for runs based on literature review of similar studies: 25 random test percentage (25% of the presence data was withheld as test data to evaluate model performance), 10,000 maximum points (number of points selected as background data), 15 replicates (number of times the model is run; these are averaged to produce the final output), subsample (how test data is withheld; 75% of the data is used to train the model, 25% is used to test the model), and 5,000 max iterations (number of times the model is allowed to iterate before converging). Output maps were classified into moderate and high probability habitat using Jenks natural breaks (Table 6).

Regularization Parameter

The primary objective of modeling is to mathematically determine the relationship between sample (observation) data and variables of interest so greater conclusions can be made at the population level. Overfitting occurs when a model is fit too closely to the sample data by using too many variables. Instead of describing a population-level relationship, it also includes the relationship between noise in the sample data and model variables (Babyak 2004). While overfitting was not specifically evaluated in this analysis, it was addressed in model design using the regularization parameter in MaxEnt.

There are several parameters available to customize the output of a model run for an individual species. The regularization parameter allows the user to change how the output's distribution is fit to the observation data (Phillips 2005). The default regularization parameter is 1.0, and can be scaled to give a result that includes more area (it is less tuned to the observation data) or to give an output that is more closely attuned to the observation data. Values smaller than 1.0 mean the output distribution is fit more closely to the observation data; thus a smaller area will be defined as suitable. Decreasing the regularization parameter may result in overfitting (Phillips 2005).

Initial runs of MaxEnt with the regularization parameter set at 1.0 predicted large swaths of habitat across the landscape and appeared to over-predict habitat. For this analysis, maps displaying smaller amounts of high probability habitat were more highly ranked by the author because most rare species have a small probability of occurring across the landscape (Franklin 2009). Each species model was run with regularization parameters of both 1.0 or 0.01 and the outputs from both runs were compared.

Initial and final model results were assessed using the Area Under the Curve (AUC) provided as part of MaxEnt's statistical output. AUC is derived from the Receiver Operating Characteristic (ROC). The ROC is a threshold-independent measure of model performance where the proportion of true positives is plotted against the proportion of false positives (Fielding and Bell 1997). The AUC is the area found under this curve and is used to rate the performance of MaxEnt runs. AUC ranges from 0 to 1.0, with 0.5 representing random performance and values above 0.7 distinguishing well performing models (Hanley and McNeil 1982, Bradley 1997). Models whose output agreed more closely with the observation data, had higher AUC values, and whose omission rates were closer to the predicted omission rate were selected. The accuracy of a model was assessed by comparing field collected data with model results. These results were then summarized as likely habitat and overall accuracy.

Accuracy Assessment

All models were assessed for accuracy in the field using transects located across the project area. Ten transects were selected for each model, for a total of sixty transects. Each one was 210 meters from a random point placed in a cell found at least 60 m from a road to reduce disturbance effects. Additionally, transects were only placed in moderate and high probability habitat to maximize probability of finding new populations.

Transects were constructed in ArcMap using knowledge of species site requirements and 2015 National Agricultural Imagery Program (NAIP) imagery. A series of steps was created for selecting transect locations for each species. First, species descriptions were reviewed and notable land features associated with each species were documented. Next, NAIP imagery for known locations was reviewed to gain an image of potential habitat and habitat description lists were adjusted accordingly. Transects had to also be relatively easy to access by motor vehicle, bicycle, or on foot. Therefore, transects were preferentially placed in areas that were close to roads or trails. Additionally, previously treated areas such as those with obvious signs of timber harvest or fire were avoided. After transects were chosen, maps were sent out to project botanists, and feedback was used to modify transect locations.

Heidi Fleury was hired as a field technician to assist in surveying field transects. The field technician was trained in data collection protocols as well as vegetation sampling for the project. Trimble GPS units were utilized to perform field samples and collect data. Transects were systematically walked and intensely visually surveyed by both Ms. Fleury and the author. Data on habitat, associated species, overstory, invasive species, disturbance, as well as photos were recorded for each transect (data collection form is shown in Appendix 1). Transects were visually determined to be likely habitat or not in the field based on previous visitation of known occurrence sites and species descriptions.

Species Overlay Map

The predicted habitats for all species were overlaid to show areas of congruence between species. The predicted habitat maps for each species were first reclassified, with moderate and high probability habitat represented as 1. The sum of all rasters was computed to show areas of congruence for one, two, three, and four species (Figure 11).

CHAPTER 4: RESULTS AND DISCUSSION

Results will be presented for the main models first and then the adjustments. Overall, thirteen previously unknown populations of rare plants were found in the project area during the field validation. Table 7 summarizes the results from the predictive model as well as those from the field accuracy assessment for the five main species models and LiDAR comparison.

Dispersal Buffers

The dispersal buffer for each species was either 1.0 or 2.5 miles (Table 1) based on their dispersal potential. Incorporating dispersal buffers resulted in output maps with predicted habitat outside the proposed BSLRP project area. These maps proved advantageous as there was high probability habitat for crested shieldfern located close to the northern boundary of the project. Transects placed in habitat located outside the project boundary did not uncover additional populations. Resulting output maps were not clipped to the final project area and can be seen in Figures 5 – 10. Each species' results will be discussed separately below.

Species Results

Camassia quamash (Frederick Traugott Pursh) Edward Lee Greene: Common camas

Surveys for common camas were performed on June 8 – 11 and June 15, 2016 (Table 2). For common camas 22,169 acres (8,972 hectares) was classified as moderate or high probability habitat. A total of eleven new populations, 6 acres (2 hectares), of new common camas populations were found. Of these new populations, 2 acres (0.8 hectares) were located in areas predicted as moderate or high probability habitat by MaxEnt (Figure 18).

The mean AUC for common camas was 0.968 (± 0.007; Table 7). This model, like all others in this analysis, had an AUC value above 0.7 and therefore was considered 'high performing.' To be comparable, models must have the same extent (Phillips et al. 2006). Due to differences in dispersal buffer size and project area, this model is only directly comparable to the clustered lady's-slipper, Howell's gumweed, and crested shieldfern models. Differences between AUC and standard deviation values between these four models are presented in Table 7. Out of these four models, the common camas model performed the best; it had both a high AUC value and a low standard deviation. Although the crested shieldfern model had a higher AUC value, it had a much larger standard deviation value.

Figure 5 is a map of high and moderate probability habitat for common camas. Elevation had the largest effect on predicted habitat, explaining 52% of the variability. In the modeled area, the probability of presence for this species was highest between 1,000 – 1,500 m, and presence dropped sharply above 1,500 m. According to the USDA Plants Guide (2006), this species is typically found at elevations below 3,300 m. As seen in the predicted habitat map for common camas, much of the predicted habitat is in and around the lower elevation valley bottoms where there are many moist meadows. This provides further support to the accuracy and ecological relevance of this model. In addition, much of the common camas habitat predicted by the model is in the Seeley Lake Ranger District (RD). Very little habitat for this species was predicted in the Lincoln RD but it is unclear why the model did not predict habitat in that area.

Cypripedium fasciculatum Albert Kellog ex Sereno Watson: Clustered lady's-slipper

Surveys for clustered lady's-slipper took place on June 13 and June 15 – 17, 2016 (Table 2). For clustered lady's-slipper 12,452 acres (5,039 hectares) was classified as moderate or high probability habitat. Within transects, 53% of the habitat was classified as likely by field accuracy assessment. No new populations of this species were found during field surveys.

The mean AUC for this model (Table 7) was 0.928 (± 0.068). This model, like all others in the analysis, had an AUC above 0.7 is classified as 'high performing.' Due to differences in dispersal buffer size and project area, this model is only directly comparable to the common camas, Howell's gumweed, and crested shieldfern models. Of these four models, this model had the lowest AUC value and the highest standard deviation. This is likely a result of the small sample size (4) and small number of random points (18) constructed during the polygon-to-points process. Although MaxEnt has been shown to perform well with small sample sizes (< 20) (Phillips et al. 2006, Wisz et al. 2008) the extremely small sample size used in this model can produce questionable results. In addition, the regularization parameter utilized in this model was 0.01, meaning it was may have been over-fit to the data, though it is almost unavoidable with a sample size this small..

Geology was ranked as having the highest percent contribution to the overall model. As seen in Figure 6, predicted habitat for clustered lady's-slipper was centered around valley bottoms primarily in the Swan Lake RD and Seeley Lake RD in the project area. Once again, it is unclear why so little habitat was predicted in the Lincoln RD.

Cold weather likely contributed to later flowering times because orchids are highly sensitive to temperature and thus this species might not have been in flower when surveyed. The model predicted habitat in several steep and thickly vegetated locations for this species. Thus, transects were difficult and sometimes dangerous to access. In addition, other orchids, including mountain lady's-slipper (*Cypripedium montanum)*, were found in transects when surveying for this species.

Heterocodon rariflorum Thomas Nuttall: Western Pearlflower

As seen in Table 2, surveys for western pearlflower were performed from June 23 – 25, 2016. For western pearlflower, 3,517 acres (1,423 hectares) was classified as moderate or high probability habitat. Within transects, 34% of the habitat was classified as likely by the field technician. No new populations of this species were located during field surveys. Western pearlflower is a small herbaceous plant with a thin stem and few small-cupped flowers that grows in wet soils with thick overstory and can be very difficult to identify, even with intensive survey techniques. In fact, known populations of western pearlflower were unable to be located. This might have been due to it flowering before or after surveys had taken place.

This model had an AUC of 0.978 (± 0.002; Table 7). This model, like all others in the analysis, had an AUC above 0.7 and is classified as 'high performing.' Due to differences in dispersal buffer size, this model is not directly comparable to any other model. As seen in Figure 11, very little habitat was predicted for this species across the project area. However, unlike the common camas and clustered lady's-slipper, this model predicted the most habitat in the Lincoln RD. Predicted habitat for this species seemed to be clustered around lower elevation valley bottoms near streams. Because it contributed 50% to the model, elevation had the largest effect on predicted habitat. Probability of presence for this species was highest from 750 – 1,000 m, becoming negligible at higher elevations. It is typically found from the foothills to montane area.

The western pearlflower model had the lowest overall accuracy (26%) of all models. Within the modeled area, western pearlflower occurs in large patches that may only have a few individuals (MTNHP 2017c). The polygon-to-point method might result in more points being placed in a polygon than individuals that exist in that population, potentially training the model on inaccurate data. Additionally, it was difficult to determine suitable habitat from the aerial imagery when placing transects for this species. Transects were preferentially placed in what appeared to be areas around streams and vernally moist areas in the aerial imagery, but when surveyed these sites were much drier than expected.

Grindelia howellii Steyermark: Howell's Gumweed

Field surveys for Howell's gumweed took place from July 19 – 22, 2016 (Table 2). For Howell's gumweed 16,344 acres (6,614 hectares) was classified as moderate or high probability habitat in the modeled area. Within transects, 31% of habitat was classified as likely by the field technician. No new populations of this species were observed during field surveys. Known populations of this species were in flower during surveys. The mean AUC for this model was 0.931 (± 0.005; Table 7). Due to differences in dispersal buffer size and project area, this model is only directly comparable to the common camas, clustered lady's-slipper, and crested shieldfern models. Of these four models, this model had the lowest standard deviation value.

Figure 8 is a map of high and moderate probability habitat for Howell's gumweed, and shows that this model predicted the most habitat in the Seeley Lake RD. Habitat was centered around lower elevations, and elevation had the highest (43%) contribution to the overall model. The highest probability of occurrence for this species was between 1,200 – 2,000 m and dropped off sharply at higher elevations. This outcome is supported by the literature; this species is typically found at elevations lower than 1,700 meters (NatureServe 2015).

Dryopteris cristata (Linnaeus) Asa Gray: Crested shieldfern without LiDAR

Field surveys for crested shieldfern without LiDAR took place on August 16 – 19, 2016 (Table 2). For crested shieldfern, two types of models were field sampled: one including LiDAR-derived canopy cover variables and one without. For the model without LiDAR 7,340 acres (2,970 hectares) were classified as moderate or high probability habitat in the modeled area. Within transects, 68% of habitat was classified as likely by the field technician. No new populations of this species were found when surveying for this species.

As seen in Table 7, the mean AUC for this model was 0.977 (± 0.021). Due to differences in dispersal buffer size and project area, this model is only directly comparable to the clustered lady's slipper, Howell's gumweed, and common camas models. Differences between AUC and standard deviation values between these four models can be seen in Table 7. Of these four models, the crested shieldfern model without LiDAR had the highest AUC value and second highest standard deviation value.

Figure 9 is a map of high and moderate probability habitat for crested shieldfern. Because it was ranked as contributing 57% to the overall model, elevation had the largest effect on predicted habitat. The highest probability of presence occurred between 900 – 1,300 m, any higher and presence dropped sharply. Habitat for this species was scattered throughout the modeled area with patches in all ranger districts.

Dryopteris cristata (Linnaeus) Asa Gray: Crested shieldfern with LiDAR

The field survey for this species was used to evaluate both the model with and without the LiDAR layer. For the crested shieldfern model with LiDAR-derived canopy cover 2,756 acres (1,115 hectares) were classified as moderate or high probability habitat. Within transects, 61% of habitat was classified as likely by the field technician. Two new populations of crested shieldfern were found when surveying transects for this model, which were in areas which were predicted to be high probability by the model using canopy cover, but were low probability in the model without LiDAR data.

For this model, the mean AUC was 0.990 (± 0.005; Table 7). LiDAR data were only available for 28% of the project area, thus the AUC values for the crested shieldfern cannot be compared against each other (Fielding and Bell 1997; Phillips et al. 2006). The AUC value for the crested shieldfern model without LiDAR was 0.977, while the model with LiDAR and a smaller extent had an AUC value of 0.990. This was unexpected and contradictory to Burbach's (2011) finding that larger extent models have higher AUC values.

Figure 10 is a map of high and moderate probability habitat for crested shieldfern. Elevation had the largest effect on predicted habitat because it was ranked as contributing 51% to the overall model. The highest probability of presence occurred between 900 – 1,500 m and dropped off sharply at higher elevations. This finding is similar to the crested shieldfern model without LiDAR. For both crested shieldfern models, elevation was ranked as having the highest contribution to the model. This too is contradictory to Burbach's (2011) finding of variable rankings at different extents. Burbach (2011) found that variable contributions varied between model extents, but this supports elevation as being key for this species in this area.

Surveys for this species were centered around lower elevation ephemeral wetlands. However, many of the sites surveyed were either too dry or infested with the invasive canary reedgrass (*Phalaris arundinacea*). New populations for this species were found primarily on raised alder tussocks in standing water.

Comparison of Polygon-to-Points and Centroid Models

The performance of point-to-polygon and centroid datasets for common camas was compared in MaxEnt. The point-to-polygon dataset contained 314 points, while the centroid dataset contained 103 points constructed using the centroid of each polygon. MaxEnt settings included a regularization parameter of 1.0.

Output statistics and predicted area for the centroid dataset can be seen in Figure 12. The centroid dataset (103) had an AUC of 0.959 (± 0.010). Output statistics and predicted area for the polygon-to-points dataset can be seen in Figure 12. The polygon-to-points dataset (314) had an AUC of 0.966 (± 0.006). The polygon-to-points dataset had a higher AUC and lower standard deviation values than the centroid dataset and was selected for modeling efforts.

As seen in Figures 12 and 13, predicted habitat varied for these species as well; the centroid dataset predicted more habitat than the polygon-to-point dataset. In addition, the variables contributing the most to the overall models varied between datasets. Elevation had the highest percent contribution to the polygon-to-points version, where geology had the highest contribution to the centroid approach.

LiDAR

This analysis aimed to investigate whether the addition of LiDAR-derived covariates can improve model accuracy. One species (crested shieldfern) was tested to determine the efficacy of LiDAR in improving model performance. Crested shieldfern was selected because it occurs in ephemeral wetlands usually surrounded by densely forested areas, and canopy cover was hypothesized to improve delineation of these areas. Two models were run on the abridged

LiDAR project area (28% of total project area): one without the LiDAR-derived covariate and one with the LiDAR-derived canopy cover variable.

The model without LiDAR had an AUC of 0.991 (±0.004) and the model with LiDAR had an AUC of 0.990 (±0.005) (Table 7). Both models had AUC values above 0.9, meaning they performed well, and both models had similarly low standard deviation values (Figures 16 and 17). Elevation contributed about 50% to the both models. While both models performed similarly statistically, two new population of crested shieldfern were found when field assessing the model that utilized LiDAR-derived canopy cover as an additional variable. This suggests that the LiDAR data is useful for improving the predictive power of such models.

Due to the results of this analysis, LiDAR would likely be helpful in improving model performance for other species listed in this analysis. Common camas also occurs in similar habitats; LiDAR-derived canopy cover might improve delineation of habitat for this species as well. Several other LiDAR derivatives are available or can be constructed from LiDAR data. Future analyses should test the performance of other LiDAR-derived variables in SDM modeling efforts.

Regularization Parameter

All models were run with regularization coefficients of 1.0 and 0.01, and model outputs were compared. Omission and commission errors as well as AUC were compared for each prediction and the superior model was selected as the final output (Table 7).

Figure 14 shows predicted habitat for the polygon-to-points data set for common camas run with a regularization parameter of 0.01. Figure 15 shows predicted habitat for the polygon-to-points dataset for common camas run with a regularization parameter of 1.0. The AUC value for the model with a regularization parameter of 1.0 was 0.966 (± 0.006). The AUC value for the model with a regularization parameter of 0.01 was 0.964 (± 0.070). For the model with a regularization parameter of 1.0, the omission rate was closer to the predicted omission which is desirable (Philips 2006).

Differences in the regularization parameter did not seem to affect overall accuracy but dramatically affected the amount of potential habitat (Figures 12, 13, 14, 15). The clustered lady's-slipper, western pearlflower, Howell's gumweed, and crested shieldfern without LiDAR models utilized a regularization parameter of 0.01. Common camas and crested shieldfern with LiDAR models had regularization parameters of 1.0. There were no noticeable trends between overall accuracy of the models with a regularization parameter of 0.01 versus models with a regularization parameter of 1.0. Despite its small regularization parameter, the clustered lady's-slipper model had an overall accuracy of over 60%. Yet, the western pearlflower model had an overall accuracy of only 26%. The model predicted almost double the amount of potential habitat for clustered lady's-slipper than western pearlflower (Table 7).

Overall Accuracy

The created shieldfern model without LiDAR had the highest overall accuracy and the second lowest percentage of total project area predicted as habitat (Table 7). The second best model was the crested shieldfern model with LiDAR, which had an overall accuracy of 66% and the field survey identified two new populations of crested shieldfern. The major difference between these two models was the regularization parameter and the extent; the crested shieldfern model with LiDAR used a regularization parameter of 1.0, while the one without LiDAR used a regularization parameter of 0.01 (Table 7). The poorest performing model was the western pearlflower model, which had an overall accuracy of only 26%. The Howell's gumweed model predicted the largest percentage of potential habitat across the project area.

Five of the six models ranked elevation as having the highest contribution to the model. Elevation is cited as directly affecting temperature, however it can also influence precipitation (Moore et al. 1991). The clustered lady's-slipper model ranked geology as having the most importance to the model.

Species Overlay Map

Figure 11 is shows areas of congruence for one, two, three, and four species. Single species had moderate or high probability habitat over 52,682 acres (21,320 hectares) across the project area including a 2.5 mile buffer. The two species had moderate or high probability habitat over 6,375 acres (2,580 hectares) in the project area including the 2.5 mile dispersal buffer. Three species had moderate or high probability habitat over 94 acres (38 hectares) of the project area including the 2.5 mile dispersal buffer. Four species had moderate or high probability habitat over 1 acre (0.4 hectares) of the project area including the 2.5 mile dispersal buffer. There are no areas where all six models had high probability habitat. The least amount of habitat was predicted in the Lincoln RD.

Sample Size

A variety of studies have shown that MaxEnt outperforms other models when using small sample sizes, which is one of the reasons it is so widely used. Wisz et al. (2008) found that MaxEnt had the highest AUC and intermediate variance with sample size of 10, when compared to 11 other species distribution modeling algorithms. However, no model is reliable with extremely small sample sizes (Wisz et al. 2008, Franklin 2009). For clustered lady's-slipper, the sample size was less than five (Table 7). This was an unanticipated result of data cleaning techniques. When species occurrence data were initially acquired from herbarium, federal, and state databases the sample sizes for each species appeared to be much larger (15 + occurrences). When redundancies were removed between databases, this number decreased significantly. In addition, several observations were discarded because the location accuracy was deemed too imprecise (page 13). This occurred in the case of historic specimens whose locational information included an insufficient description of where they were found. Best efforts were made to specify a location for these collections, but in most cases it was not possible. This is one significant drawback of using historic herbarium data in such studies.

Transect Selection

Suitable habitat was determined based on species data collected from expert sources such as dichotomous keys, previous transects, and known occurrence sites. Using each species' suitable habitat description, transects were then selected in a GIS using associated aerial imagery characteristics. However, this was a difficult process given constraining specifications (such as distance from road) and in some cases severely limited placement of transects. Transects were only placed in high and moderate potential habitat. The field technician and the author collected data while traveling to and from transect locations. Therefore, some data were collected in no probability habitat, but not as intentionally as would have been ideal.

Weather & Timing

Overall accuracy of each model was dependent upon determination by the author and the field technician as to whether the surveyed area was suitable habitat. Attempts were made to visit known populations of each species, but this was not possible for clustered lady's-slipper and western pearlflower. These species were not in bloom when the known locations were checked, therefore, surveyors had no fresh field-based impressions of these species and their associated habitat (Table 2). Ten transects were surveyed for each species and because of phenology, there was a short window for viewing each species in flower. These circumstances may have also led to low observed accuracies for western pearlflower and clustered lady's-slipper.

Optimal windows for performing surveys for each species was decided using collected data from expert sources as well as averaging herbarium dates for those specimens in flower. Weather affects the phenology of plant species, especially in cooler mountainous environments. During the survey season in June 2016, unusually cold weather likely retarded flowering of all species, making detection difficult.

CHAPTER 5: CONCLUSION

This study successfully implemented SDMs for rare plants at the landscape scale across the 1.2-million acre (4,856,228 hectares) BSLRP study area as using the MaxEnt algorithm, and investigated several less common and unique techniques for improving SDMs in general. It also demonstrated that this methodology was effective way to assess rare plant habitat for a project within a single season.

This analysis presented a novel approach for incorporating dispersal mechanisms into SDMs by assigning an estimated dispersal distances based on species information. The project area was buffered accordingly by this distance for each species. Additionally, this study presented a method for incorporating polygon data into SDM modeling while preserving environmental data for that population. Comparisons between the centroid and polygon-to-point datasets (Table 7) showed that the predicted habitats from the later approach had higher AUC values and lower standard deviations.

A suite of non-correlated available environmental variables were selected for this analysis. These seven environmental variables included elevation, geology, precipitation, solar insolation, topographic wetness index, land cover, and LiDAR-derived canopy cover (Table 3). Six of these variables were utilized in the five main species models (Table 3). This analysis also hoped to look at how efficacious LiDAR is at improving model accuracy. The crested shieldfern model with LiDAR utilized the LiDAR-derived canopy-cover layer in addition to the six other environmental variables. LiDAR-derived canopy-cover layer was shown to improve predictive power of the model. Two new populations of crested shieldfern were found using the LiDAR-derived canopy-cover layer in the model.

This analysis also hoped to address methods for refining the amount of area predicted as suitable habitat for rare plants. Models for each species were run using regularization parameters of 0.01 and 1.0; outputs produced by each run were compared and the highest performing model was selected. Reducing the regularization parameter to 0.01 potentially resulted in overfitting for the clustered lady's-slipper, western pearlflower, Howell's gumweed, and the crested shieldfern without LiDAR models. However, these models were still able to successfully predict likely habitat as determined by field accuracy assessment (Table 6). Field assessments were performed for each model. Likely habitat, as ranked by surveyors, varied from 29 to 68%. Overall accuracy varied from 26 to 69% (Table 7). Potential species distribution maps were overlaid to display areas of congruence. These maps can be used by resource managers to identify areas for surveying prioritization (Figure 11).

Recommendations

Challenges encountered during this analysis provided insight into recommendations for improving upon SDMs for future analyses.

Environmental Variables

The low (< 50%) overall accuracies for common camas, western pearlflower, and Howell's gumweed suggest that the environmental variables used in these models were not able to fully explain their distribution. The Natural Resource Conservation Service (NRCS) produces soils maps for much of the US. The SSURGO soil survey database houses data on soils surveys for the US but does not have data on soils in wilderness areas. Because BSLRP contains wilderness areas, this layer could not be utilized in analysis of the full study area. Future analyses should attempt to include soils layers.

Data Quality

Data quality was the greatest limiting factor is this study as is common in SDMs. Future SDM studies should focus on selecting species with ten or more observations in order to improve model accuracy. Extreme care should be taken when cleaning observation data to ensure all records meet specified accuracy and taxonomy standards. Modelers need to ensure that they fully clean the data to determine sampling size before offering species to resource managers as potential species for modeling. A minimum of at least ten, if not more, observations should be used when modeling (Wisz et al. 2008).

Transects

Transects were used to accuracy assess model outputs in this analysis. This method posed difficulties during implementation. Selecting appropriate transect locations that were located in potential habitat for these species was difficult using aerial imagery. It was hard to determine if habitat was appropriate because the imagery was taken either too early (while there was still some snow) or too late (when ephemeral wetlands had dried out). Additionally, this method of accuracy assessment proved time consuming due to transects being located in difficult to reach areas. Lastly, this method of accuracy assessment preferentially sampled more high probability habitat, limiting sampling in lower probability areas.

A more appropriate sampling scheme would sample more evenly high and moderate probability habitat, as well as areas not predicted as habitat by the model. For example, sampling random plots placed across the landscape and quickly determining whether they are habitat or not might provide more data on locations not predicted as habitat by the model.

Planning Applications

Models do not represent reality, they are a mere reflection of it. The predictions created as a result of this effort do not represent the actual spatial distribution of the species of interest. Rather they are an estimate of where that species might occur based on selected environmental variables, input parameters, and random chance. As proven in this analysis, the distribution and extent of predicted habitat can vary significantly based solely on a single input parameter. Thus, all predictions should be utilized in combination with expert knowledge of species ecology and occurrence across the landscape. They should not be used as a replacement for expert knowledge, but instead as a decision-making aid for experts that makes it easier to come to informed conclusions, especially in the face of time and budgetary constraints. Concomitantly, they should not be used as a replacement for field surveys. These models should be used by resource managers to identify areas of high potential presence of rare plants, such as here where they were used to prioritize field work for the rare plant program in the BSLRP project.

Future Research

Future research should focus on improving the quality of observational and environmental data. For species with few known occurrences and high locational uncertainty, technicians should assess the accuracy of known observations in the field before committing to modeling it. In addition, botanists should utilize data collection devices capable of gathering precise coordinates on species occurrence. This data should be checked for accuracy and entered into relevant state and federal databases. Data quality should be of primary importance to any program manager who is responsible for acting as data steward for their respective program.

Program managers should ensure that data is being collected according to a standardized framework.

Future research may expand these modeling efforts to other species within this project area that were not chosen by the botanists as species of interest, but nonetheless may prove important for NEPA analyses. These species include English sundew *(Drosera anglica),* Rannoch-rush *(Scheuchzeria palustris),* and northern adderstongue *(Ophioglossum pusillum).* In addition, members of the *Botrychium* genus are typically underrepresented in field surveys and might benefit from modeling exercises to improve detection. Future research should also continue to test different polygon-to-point techniques and for additional species.

CHAPTER 6: REFERENCES

Agee, J. K. 1997. Fire management for the 21st century. Pages 191-202 *in* KA. Kohm and J.F. Franklin, editors. Creating a forestry for the 21st century. Island Press, Washington D.C., USA.

Aguilar-Soto V., A. Melgoza-Castillo, F. Villarreal-Guerrero, C. Wehenkel, and C. Pinedo-Alvarez. 2015. Modeling the potential distribution of *Picea chihuahuana Martinez,* an endangered species on the Sierra Madre Occidental, Mexico. Forests 6:692–707.

Andersen, H. E., R. J. McGaughey, and S. R. Reutebuch. 2005. Estimating forest canopy field parameters using LIDAR data. Remote Sensing of Environment 94:441-449.

Arditti J. 1967. Factors affecting the germination of orchid seeds. Botanical Review 33:1-97.

Babyak, M.A. 2004. What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. Psychosomatic Medicine 66:441-421.

Barber, J., S. R. Brown Jr., and R. Ahl. 2012. Mid-level and base-level databased of the R1 existing vegetation (VMap) products. Geospatial Group – Engineering Staff. USDA Forest Service, Region 1. Missoula, Montana, USA.

Barry, S. and J. Elith. 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43:413-423.

Bassman, J. H. 2015. Terrestrial ecological restoration in the BSLRP Project Area: A proposed framework and rationale. USDA Forest Service Northern Region, Regional Office, Missoula, Montana, USA.

Beauregard, F., and S. de Blois. 2014. Beyond a climate-centric view of plant distribution: Edaphic variables add value to distribution models. PLoS ONE 9:e92642.

Beckwith B. 2004. The queen root of this clime: ethnoecological investigations of the blue camas (C. leichtlinii (Baker) Wats. C. quamash (Pursh) Greene; Lilicaeae) and its landscapes on Southern Vancouver Island, BC. Dissertation, University of Victoria, British Columbia, Canada.

Beven, K. J. and M. J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences 24:43-69.

Bradley, A. P. 1997. The use of the Area Under the ROC curve in evaluation of machine learning algorithms. Pattern Recognition 7:1145-1159.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124:255−279.

Brown, J. H., D. W. Mehlman, and G. C. Stevens. 1995. Spatial variation in abundance. Ecology 76:2028−2043.

Burbach, T. 2011. The influence of environmental variables on predicting rare-plant habitat in the Nez Perce National Forest. Thesis, University of Montana, USA.

Burgman, M. A. 1989. The habitat volumes of scarce and ubiquitous plants – a test of the model of environmental control. American Naturalist 133:228–239.

Collis, R. T. H. 1966. Lidar: a new atmospheric probe. Quarterly Journal of the Royal Meteorological Society 92:220-230.

Correll, D. S. 1978. Native orchids of North America. Stanford University Press, Stanford, California, USA.

Doherty, J. W. 1997. The genus cypripedium: Part 1. North American Native Orchid Journal 3:5-116.

eFloras 2008. Missouri Botanical Garden, St. Louis, MO & Harvard University Herbaria, Cambridge, MA. Dryopteris cristata. < http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=200004410>. Accessed 25 January 2017.

Egan, D. and E. A. Nielsen. 2014. The History of the Four Forest Restoration Initiative: 1980s–2010. ERI White Papers: Issues in Forest Restoration. Ecological Restoration Institute, Northern Arizona University, Arizona, USA.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43-57.

Elith, J. and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677-697.

Elzinga, C. L., D. W. Salzer, and J. W. Willoughby. 1998. Measuring & Monitoring Plant Populations. U.S. Bureau of Land Management Paper 17. Denver, Colorado, USA.

Emery, D. E. 1988. Seed propagation of native California plants. Santa Barbara Botanic Garden, Santa Barbara, California, USA.

Engler, R., A. Guisan, and L. Rechsteiner. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudoabsence data. Journal of Applied Ecology 41:263–274.

Environmental Systems Research Institute, Inc. 2013. ArcMap 10.2.2. Redlands, California, USA.

Epstein, R. 1984. The principle of parsimony and some applications in psychology. The Journal of Mind and Behavior 5:119-130.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38–49.

Finney, M.A., and J. D. Cohen. 2003. Expectation and evaluation of fuel management objectives. Pages 353-66 *in* Fire, Fuel Treatments, and Ecological Restoration: Conference Proceedings. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, USA.

Franklin, J. 2009. Mapping species distributions: Spatial inference and prediction. Cambridge University Press, Cambridge, UK.

Fourcade, Y., J. O. Engler, D. Rödder, and J. Secondi. 2014. Mapping species distributions with MaxEnt using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. PLoS ONE 5: e97122.

Fretwell, S.D., and H.L. Lucas. 1969. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheoretica 19:45-52.

Gaston, K.J. 1994. Rarity. Chapman & Hall, London, UK.

Gatziolis, D. and H.E. Andersen. 2008. A guide to LiDAR data acquisition and processing for the Forests of the Pacific Northwest. USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.

Giesler, R., M. Hogberg, and P. Hogberg. 1998. Soil chemistry and plants in Fennoscandian boreal forest as exemplified by a local gradient, Ecology, 79:119–137.

Gogol-Prokurat, M. 2011. Predicting habitat for rare plants at local spatial scales using a species distribution model. Ecological Applications 21:33–47.

Goodchild, M. F. 1996. The spatial data infrastructure of environmental modeling. Pages 11-15 *in* M.F. Goodchild., L. T. Steyaert, B. O. Parks et al., editors. GIS and environmental modeling: progress and research issues. GIS World Books, Fort Collins, Colorado, USA.

Guillera-Arroita, G., J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentinin, M. A. McCarthy, R. Tingley, and B. A. Wintle. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography 24:276-292.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147-186.

Gurevitch, J., S. M. Schneider, and G. A. Fix. 2002. The ecology of plants. First Edition. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, USA.

Hall, L.S., P. R. Krausman, and M. L. Morrison. 1997. The habitat concept and a plea for standard terminology. Wildlife Society Bulletin 25:173-182.

Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the Area Under a Receiver Operating Characteristic (ROC) Curve. Radiology 1:29-36.

Harrod R. J. 1993. Studies relating to Cypripedium fasciculatum. USDA Forest Service, Region 6 Wenatchee National Forest, Leavenworth Ranger District, Washington, USA.

Hutchinson, G.E. 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22:415-427.

International Union for Conservation of Nature [IUCN]. 2017. IUCN Species – About. < https://www.iucn.org/theme/species/about>. Accessed November 12, 2016.

Jennings, S.B., N. D. Brown, and D. Sheil. 1999. Assessing forest canopies and understory illumination: canopy closure, canopy cover and other measures. Forestry 1:59-73.

Jenny, H. 1941. Factors of Soil Formation. McGraw Hill, New York, New York, USA.

Kochler, A. W. 1973. Problems in classifying and mapping vegetation for ecological regionalization. Ecology 54:512-523.

Kramer-Schadt S., J. Niedballa, J. D. Pilgrim, B. Shroder, J. Lindenborn, V. Reinfelder, M. Stillfried, I. Heckmann, A. K. Scharf, D. M. Augeri, S. M. Cheyne, A. J. Hearn, J. Ross D. W. Macdonald, J. Mathai, J. Eaton, A. J. Marshall, G. Semiadi, R. Rustam, H. Bernard, R. Alfred, H. Samejima, J. W. Duckworth, C. Breitenmoser-Wuersten, J. L. Belant, H. Hofer, and A. Wilting. 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. Diversity and Distributions 19:1366–1379.

Kunin, W.E. and K. J. Gaston. 1997. The biology of rarity: causes and consequences of rare–common differences. Chapman & Hall, London, UK.

Lackschewitz, K. 1991. Vascular plants of west-central Montana: identification guidebook. USDA Forest Service, Intermountain Research Station, Missoula, Montana, USA.

Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding. 2002. LiDAR remote sensing for ecosystem studies. BioScience 1:19-300.

Lesica, P., M. T. Lavin, and P. F. Stickney. 2012. Manual of Montana vascular plants. BRIT Press, Fort Worth, Texas, USA.

Leyequien, E. J. Verrelst, M. Slot, G. Schaepman-Strub, I. M. A. Heitkonig, and A. Skidmore. 2007. Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity. International Journal of Applied Earth Observation and Geoinformation 9:224-224.

Mackey, B. G. 1993. Predicting the potential distribution of rain-forest structural characteristics. Journal of Vegetation Science 4:43-54.

Montana Natural Heritage Program [MTNHP]. 2017a. Montana Field Guide. Common Camas — Camassia quamash. <http://FieldGuide.mt.gov/speciesDetail.aspx?elcode=PMLIL0E040>. Accessed 24 January 2017.

Montana Natural Heritage Program [MTNHP]. 2017b. Montana Field Guide. Clustered Lady's-slipper — Cypripedium fasciculatum. <http://FieldGuide.mt.gov/speciesDetail.aspx?elcode=PMORC0Q060>. Accessed 24 January 2017.

Montana Natural Heritage Program [MTNHP]. 2017c. Montana Field Guide. Western Pearl-flower — Heterocodon rariflorum. <http://FieldGuide.mt.gov/speciesDetail.aspx?elcode=PDCAM08010>. Accessed 24 January 2017.

Montana Natural Heritage Program [MTNHP]. 2017d. Montana Field Guide. Howell's Gumweed — Grindelia howellii. <http://FieldGuide.mt.gov/speciesDetail.aspx?elcode=PDAST470E0>. Accessed 24 January 2017.

Montana Natural Heritage Program [MTNHP]. 2017e. Montana Field Guide. Crested Shieldfern — Dryopteris cristata. <http://FieldGuide.mt.gov/speciesDetail.aspx?elcode=PPDRY0A090>. Accessed 24 January 2017.

Moore, I. D., R. Grayson, and A. Ladson. 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes 5:3-30.

Morin, N.R. 2017. Heterocodon rariflorum, in Jepson Flora Project (eds.) Jepson eFlora, <http://ucjeps.berkeley.edu/cgi-bin/get_IJM.pl?tid=28069>. Accessed 25 January 2017.

National Environmental Policy Act (NEPA) of 1969 § 102, 42 U.S.C. § 4332 (1994).

NatureServe. 2015. NatureServe Explorer: An online encyclopedia of life. Version 7.1. Grindelia howellii. < http://explorer.natureserve.org/servlet/NatureServe?searchName=Grindelia%20howellii> . Accessed 25 January 2017.

National Interagency Fire Center [NIFC]. 2015. Federal Firefighting Costs (Suppression only). <https://www.nifc.gov/fireInfo/fireInfo_documents/SuppCosts.pdf>. Accessed 27 January 2017.

Nock, E. E. 2008. A simple GIS approach to predicting rare plant habitat: North central Rocky Mountains, United States Forest Service, Region One. Thesis, University of Montana, USA.

Peck, J. H., C. J. Peck, and D. R. Farrar. 1990. Influences of life history attributes on formation of local and distance fern population. American Fern Society 80:126-142.

North, M. P., S. L. Stephens, B. M. Collins, J. K. Agee, G. Aplet, J. F. Franklin, and P. Z. Fule. 2015. Reform forest fire management**.** Science 6254: 1280-1281.

Phillips, S. 2005. A brief tutorial on Maxent. <http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc>. Accessed 2 November 2015.

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

Phillis, S. J. and M. Dudik. Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. Ecography 31:161-175.

Phillips, S. J., M. Dudík, J. Elith,, C. H.Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19:181–197.

PRISM Climate Group, Oregon State University. Created 2015. <www/prism.oregonstate.edu/products/>. Accessed 2 March 2016.

Rabinowitz, D. 1981. Seven forms of rarity. Pages 205-217 *in* H. Synge, editor. The Biological Aspects of Rare Plant Conservation. John Wiley & Sons, Chichester, UK.

Raven, P.H., R. F. Evert, and S. E. Eichorn. 1999. Biology of Plants. Sixth edition. W.H. Freeman & Company Worth Publishers, New York, New York, USA.

Satterlund, D. R., and J. E. Means. 1978. Estimating solar radiation under variable cloud conditions. Forest Science 24:363- 373.

Seevers, J. and F. Lang. 1998. Bureau of Land Management. Management recommendations for clustered lady slipper orchid (Cypripedium fasciculatum Kellogg ex. S. Watson) v. 2.0. <https://www.blm.gov/or/plans/surveyandmanage/MR/VascularPlants/section9.htm>. Accessed 24 January 2017.

Sorensen, R., U. Zinko, and J. Siebert. 2006. On the calculation of the topographic wetness index: Evaluation of different methods based on field observations. Hydrology and Earth System Sciences Discussions, European Geosciences Union 1:101-112.

Southwest Crown Collaborative. 2010. Southwest Crown of the Continent Landscape Restoration Strategy. Submitted to the Collaborative Forest Landscape Restoration Program (CFLRP).

USDA Plant Guide. 2006. Great Camas *Camassia quamash* (Pursh) Greene *ssp. quamash*. USDA NRCS National Plant Data Center & Corvallis (OR) Plant Center, Corvallis, Oregon, USA. < https://plants.usda.gov/core/profile?symbol=CAQU2>. Accessed 25 January 2017.

USDA Forest Service. 2005. FSM 2600 – Wildlife, fish, and sensitive plant habitat management, Chapter 2670 – threatened, endangered, and sensitive plants and animals. Amendment 2600-2005-1. Forest Service Manual National Headquarters (WO), Washington, D.C., USA.

USDA Forest Service. 2015. Fiscal year 2016 budget overview. USDA Forest Service, Washington, D.C., USA.

USDA Forest Service. 2017a. Collaborative Forest Landscape Restoration Program Overview. < https://www.fs.fed.us/restoration/CFLRP/overview.shtml>. Accessed 25 March 2016.

USDA Forest Service. 2017b. Blackfoot Swan Landscape Restoration (BSLR) Project. < https://www.fs.usda.gov/detail/r1/landmanagement/?cid=FSEPRD480785>. Accessed 2 February 2016.

Vandekerckhove, J., D. Matzke, and E. Wagenmakers. 2014. Model comparison and the principle of parsimony. Pages 1-29 *in* J.R. Busemeyer, Z. Wang, J.T. Townsend, A. Eidels, editors. The Oxford Handbook of Computational and Mathematical Psychology. DOI: 10.1093/oxfordhb/9780199957996.013.14. < http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199957996.001.0001/oxfordhb-9780199957996-e-14>. Accessed 4 March 2016.

Vaughan, I. P., and S.J. Ormerod. 2003. Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. Conservation Biology 17:1601-1611.

Vendrusculo, L. G., and A. L. Kaleita. 2013. Terrain analysis and data mining techniques applied to location of classic gully in a watershed. 2013 Agricultural and Biosystems Engineering Conference Proceedings and Presentations, 325.

Watson, R. D. 1988. Interview data on Camas growing in Moscow, Idaho. Interviewed (4-18-88) by Alson V. Thoms. Notes on file, Center for Northwest Anthropology, Washington State University, Pullman, Washington, USA.

Warren, D. L., and S. N. Seifert. 2011. Ecological niche modeling in MaxEnt: The importance of model complexity and the performance of model selection criteria. Ecological Applications 21:335-342.

Wells, T. C. E., and J.H. Williams. 1991. Population ecology of terrestrial orchids. SPb Academic Publishing, The University of Michigan, Ann Arbor, Michigan, USA.

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

Woodward, F. I., and B.G. Williams. 1987. Climate and plant distribution at global and local scales. Vegetation 69:189-197.

Zhang, Z., G. Hu, J. Ni. 2013. Effects of topographical and edaphic factors on the distribution of plant communities in two subtropical karst forests, southwestern China. Journal of Mountain Science 1:95-104.

Zientek, M. L., P. D. Derkey, R. J. Miller, J. D. Causey, A. A. Bookstrom, M. H. Carlson, G. N. Green, T. P. Frost, D. E. Boleneus, K. V. Evans, B. S. Van Gosen, A. B. Wilson, J. C. LarsenF. Kayser, W. N. Kelley, and K. C. Assmus. 2005. Spatial databases for the geology of the Northern Rocky mountains – Idaho, Montana, and Washington. US Geological Survey, USA.

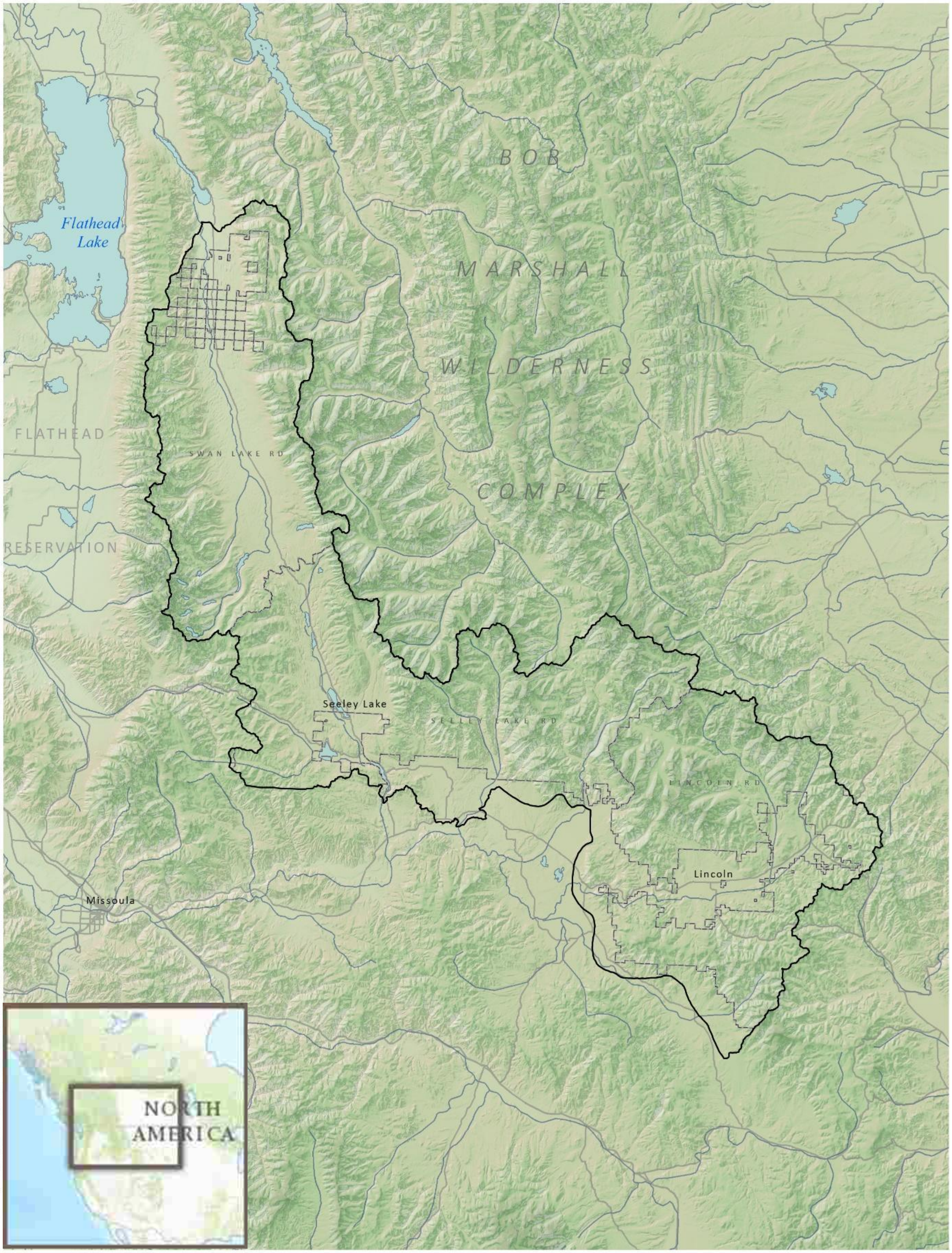

FIGURE 1. Map of Blackfoot Swan Landscape Restoration Project (BSLRP) study area.

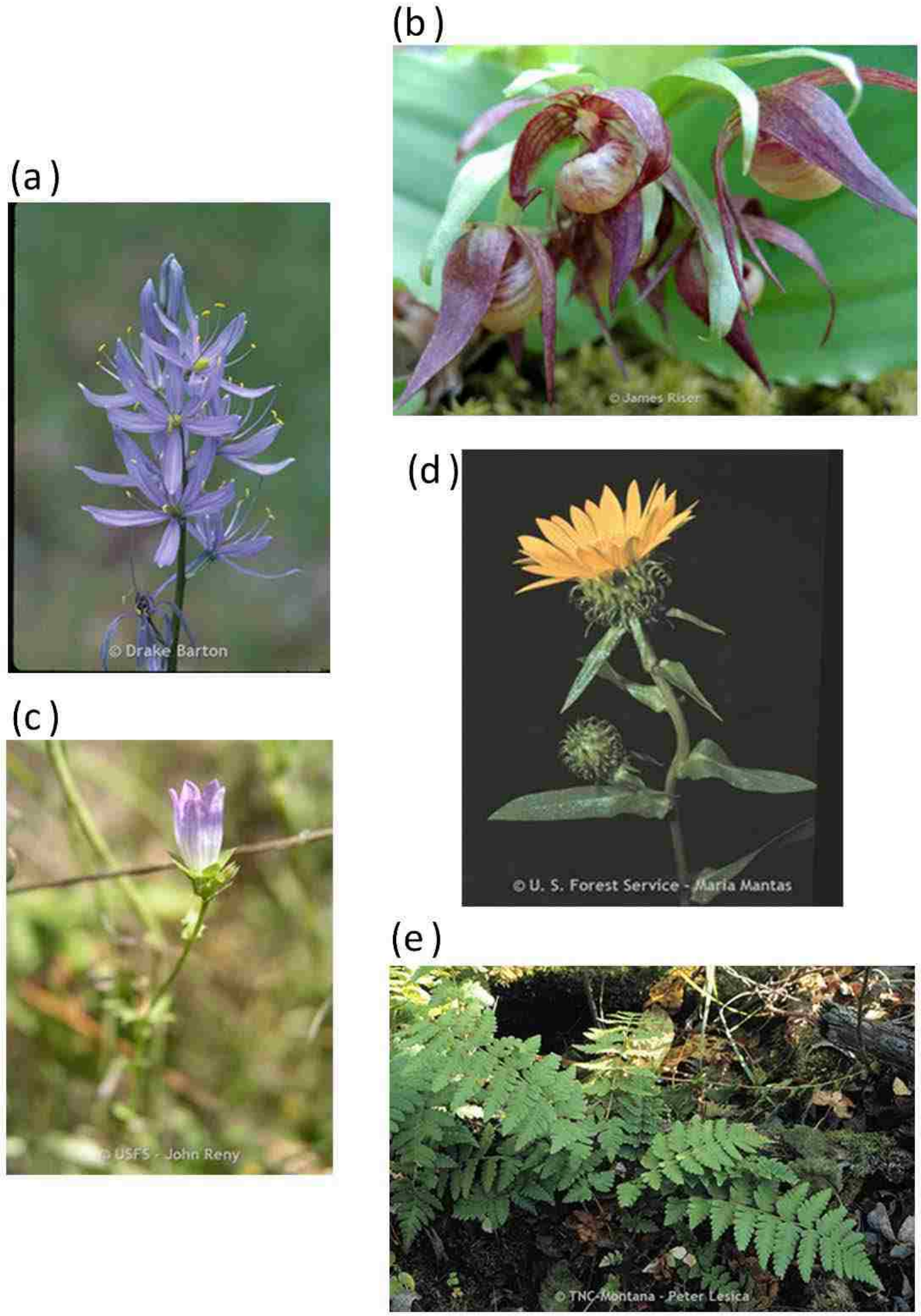

FIGURE 2. (a) Picture of common camas (*Camassia quamash)* in flower; photo by Drake Barton, Montana Natural Heritage Program, (b) picture of clustered lady's-slipper *(Cypripedium fasciculatum)* in flower; photo by James Riser, Montana Natural Heritage Program, (c) picture of western pearlflower (*Heterocodon rariflorum)* in flower; photo by John Reny, US Forest Service, (d) picture of Howell's gumweed *(Grindelia howellii)* in flower; photo by Maria Mantas, US Forest Service, (e) picture of crested shieldfern *(Dryopteris cristata)* in flower; photo by Peter Lesica, The Nature Conservancy.

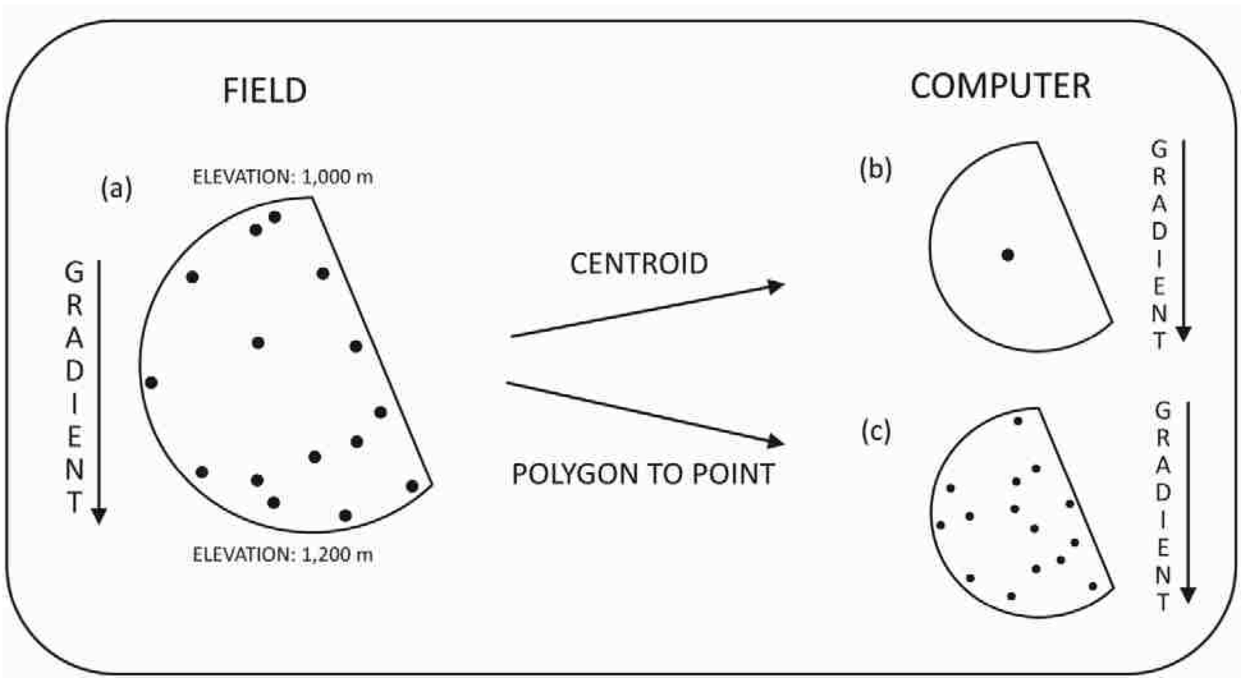

FIGURE 3. Diagram comparing the point-to-polygon and centroid methods in the field and on the computer. (a) A polygon is drawn around a plant population containing multiple individuals. This population potentially spans a gradient of changing environmental conditions. (b) Using the centroid method on a polygon population feature produces a single point at the center of that polygon. The population is now represented by a single set of environmental variables. (c) Using the polygon to point method, points are randomly assigned inside of the polygon area, with each point assigned the values of the environmental variables found at that location. The population is now represented by a range of values for each environmental variable.

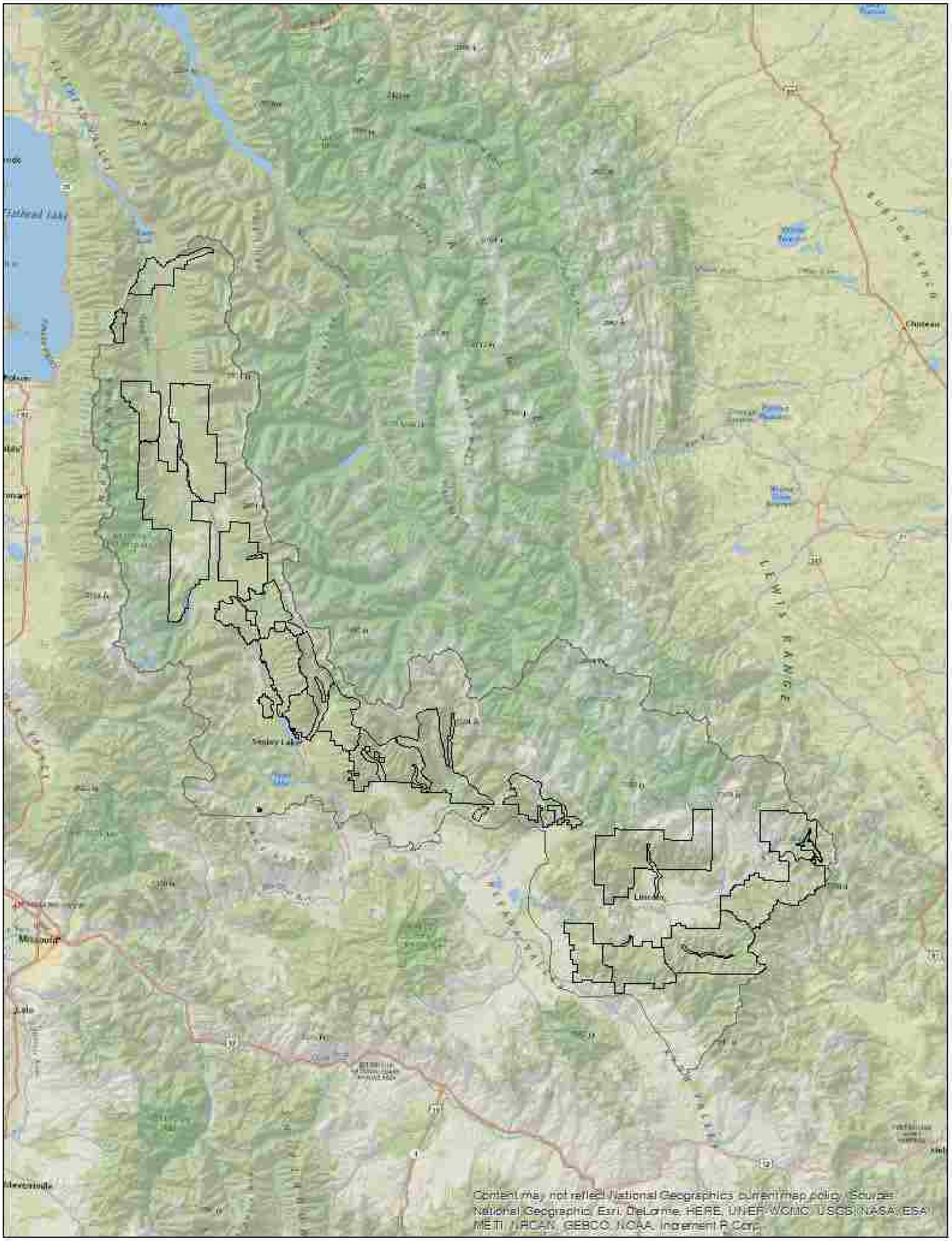

FIGURE 4. Map of LiDAR project area, which is 28% of the total BSLRP study area.

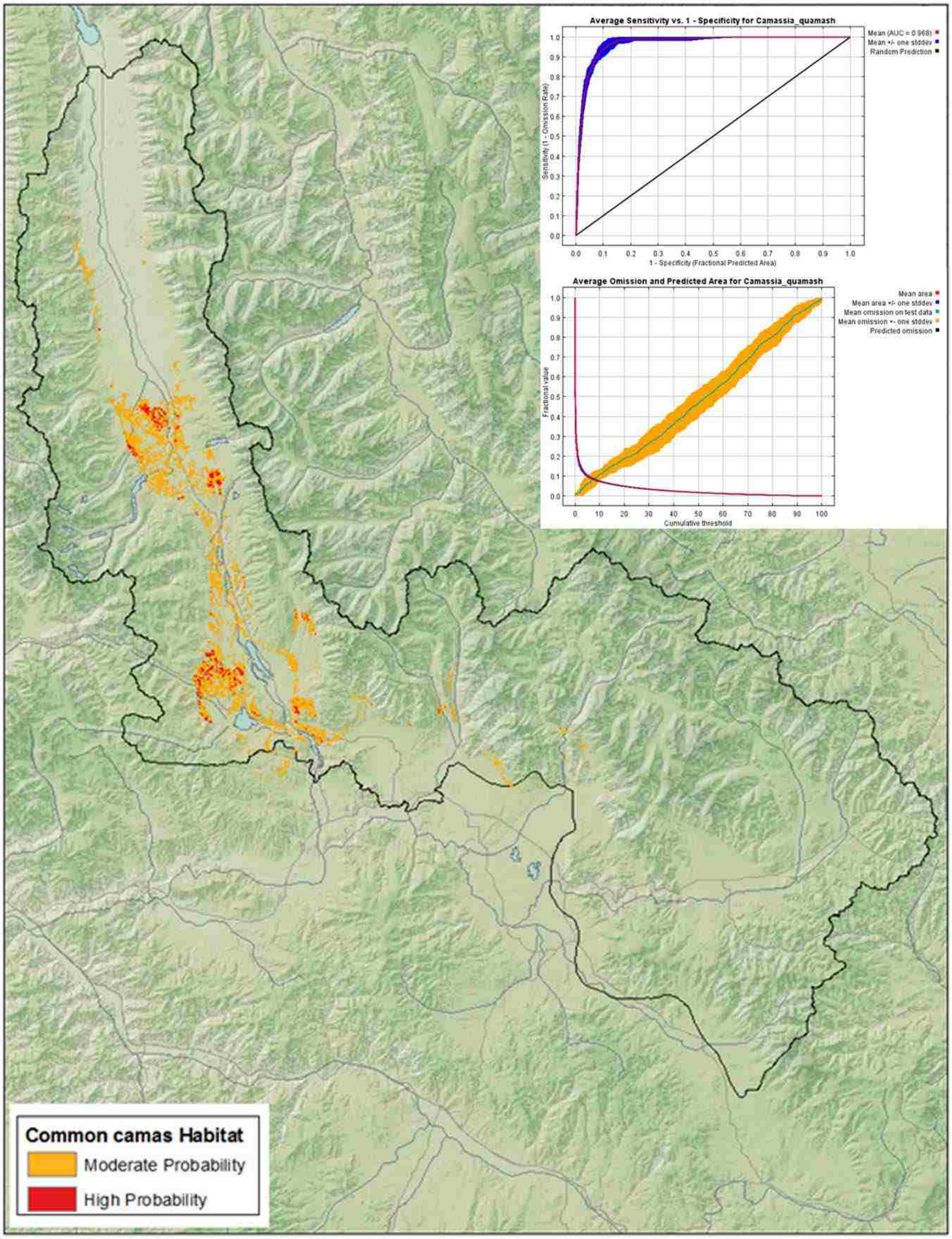

FIGURE 5. Map of predicted moderate and high probability habitat, graph of average sensitivity versus (1 – specificity), and graph of average omission for common camas.

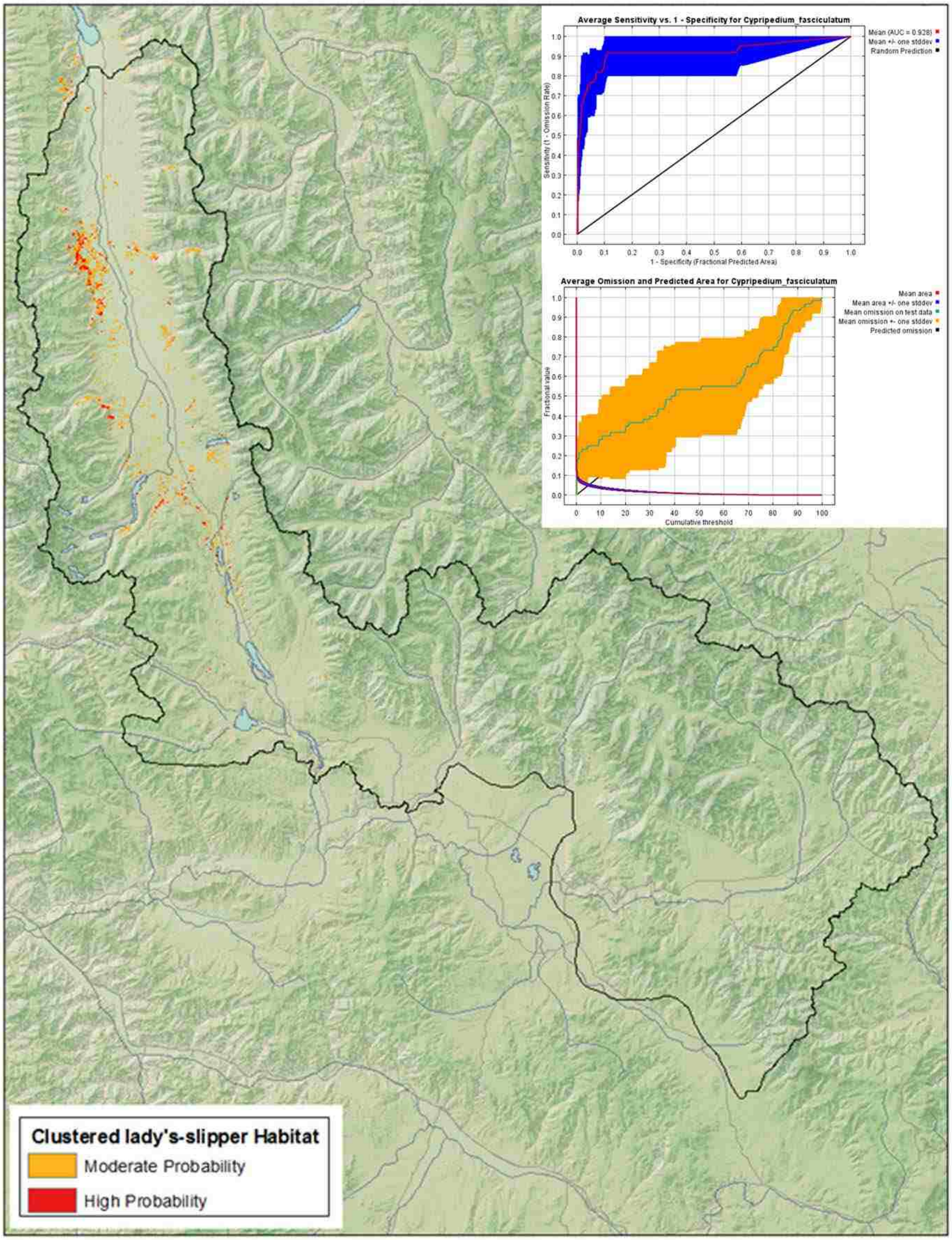

FIGURE 6. Map of predicted habitat, average sensitivity versus 1 – specificity graph, and average omission graph for clustered lady's-slipper.

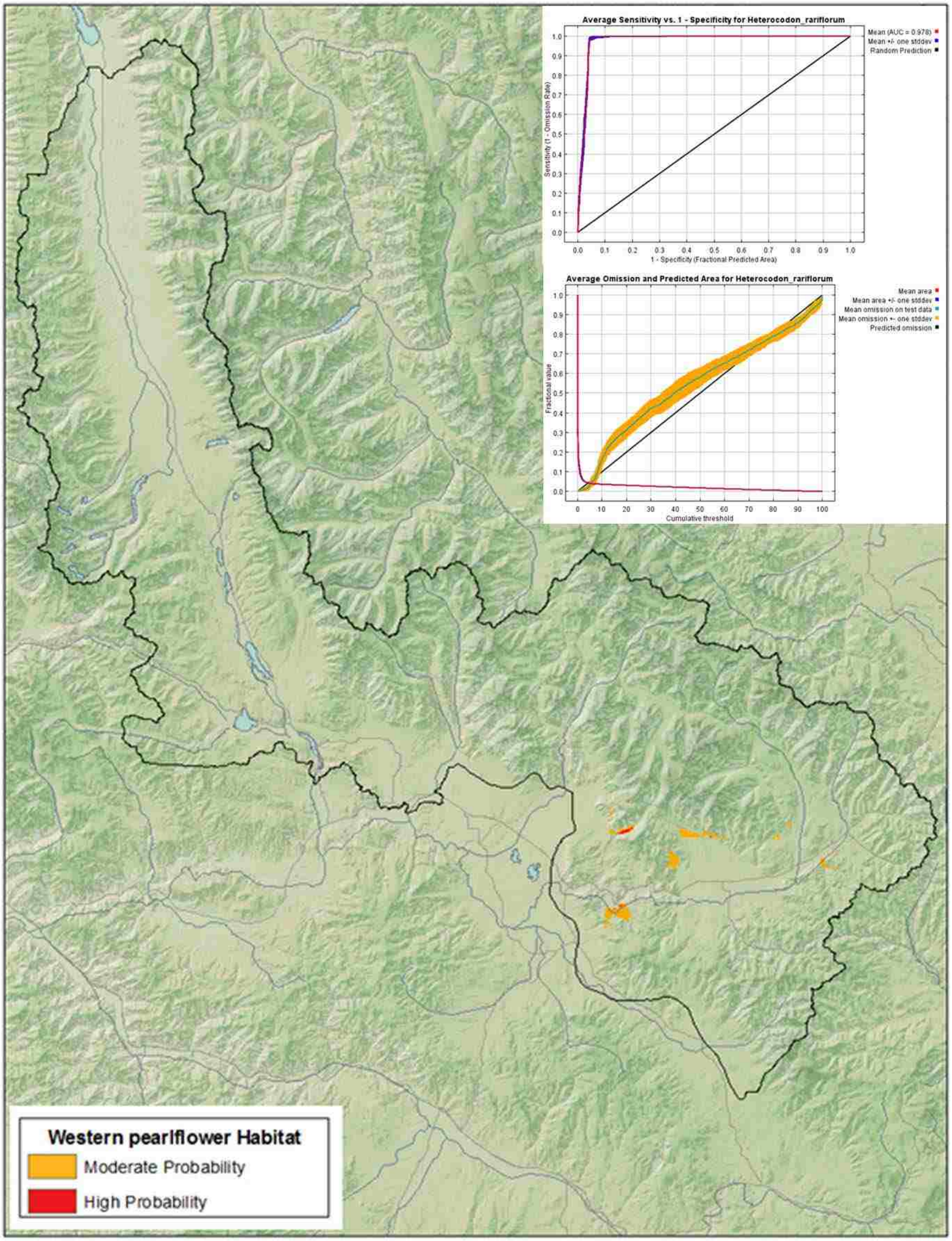

FIGURE 7. Map of predicted habitat, average sensitivity versus 1 – specificity graph, and average omission graph for western pearlflower.

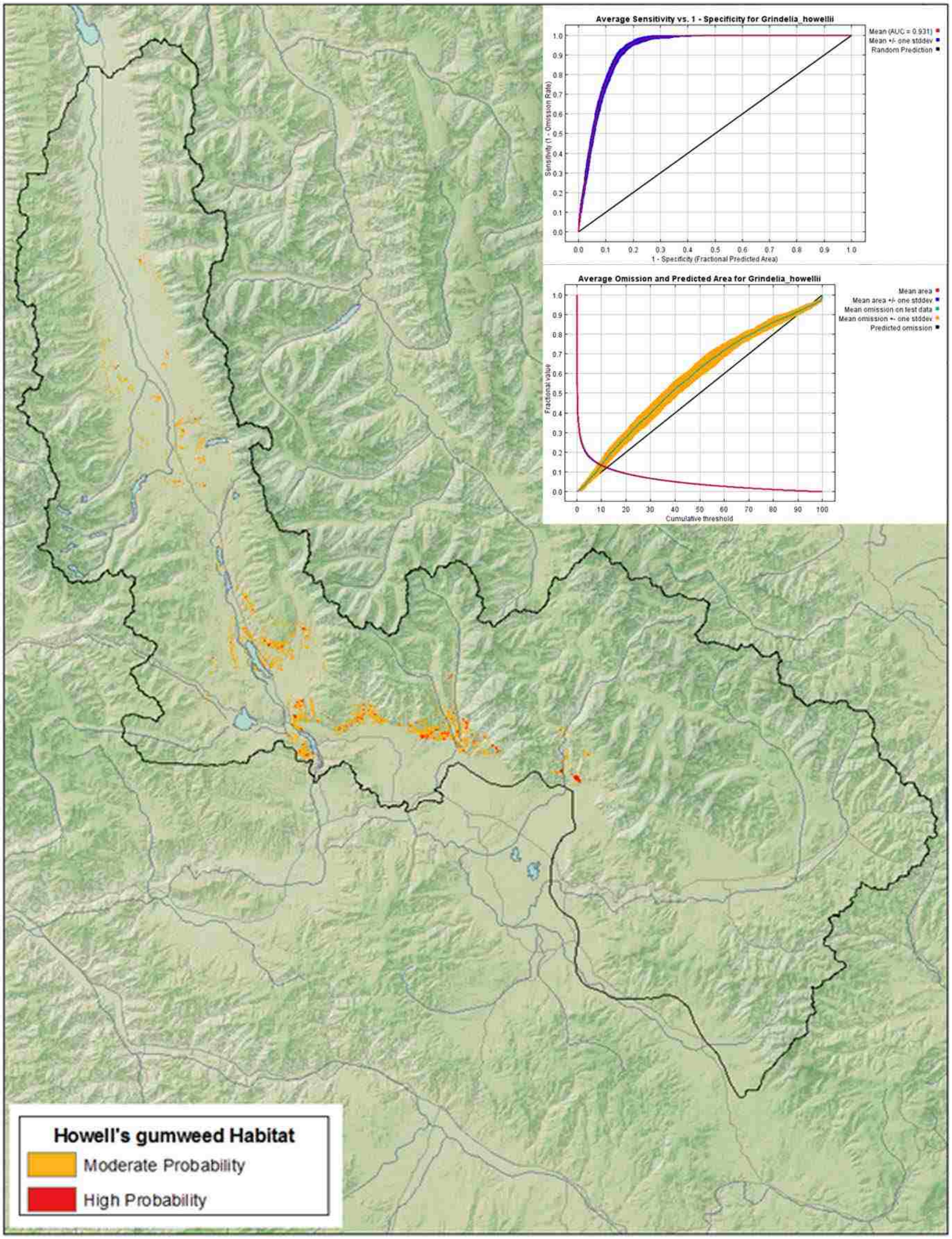

FIGURE 8. Map of predicted habitat, average sensitivity versus 1 – specificity graph, and average omission graph for Howell's gumweed.

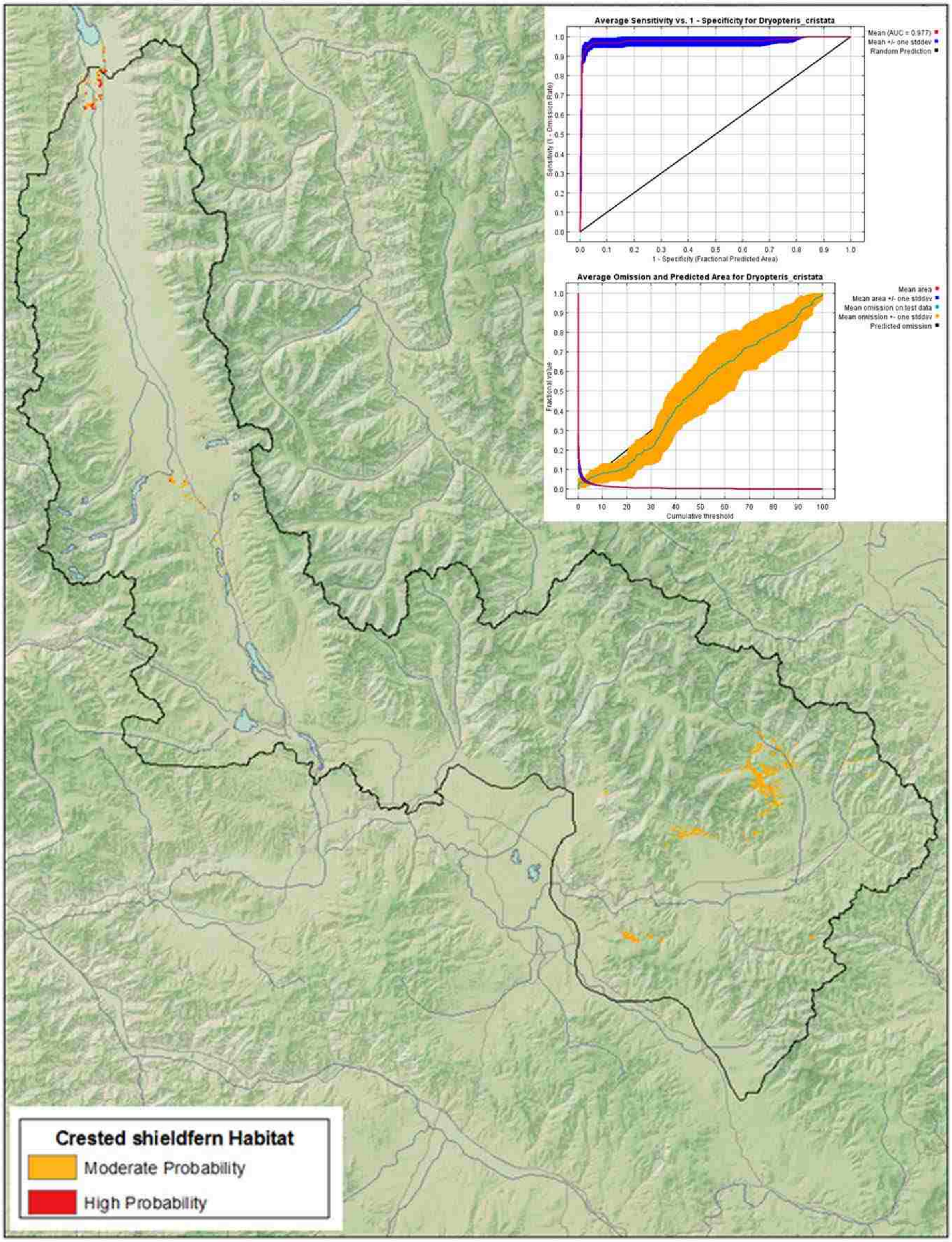

FIGURE 9. Map of predicted habitat, average sensitivity versus 1 – specificity graph, and average omission graph for crested shieldfern.

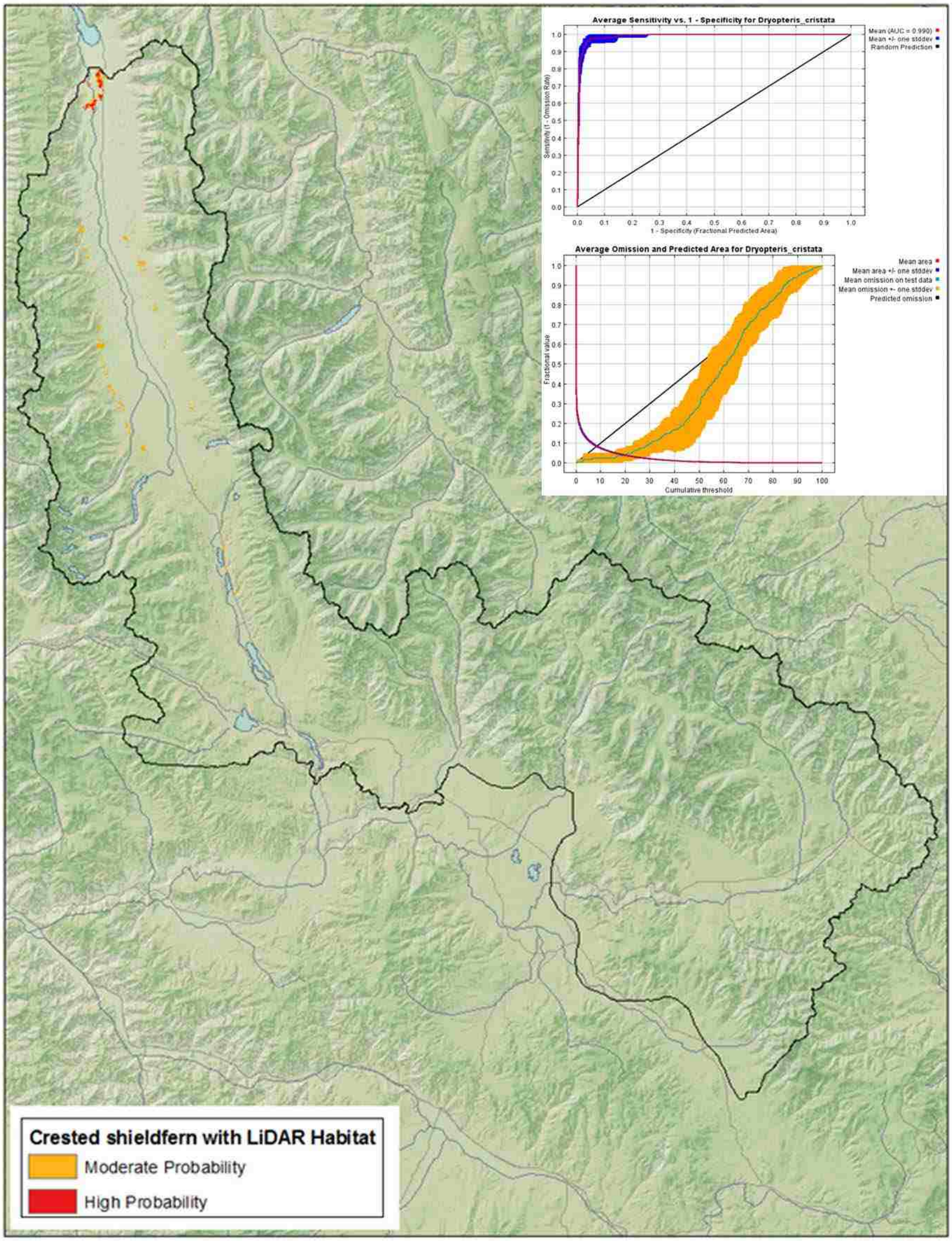

FIGURE 10. Map of predicted habitat, average sensitivity versus 1 – specificity graph, and average omission graph for crested shieldfern with LiDAR-derived canopy cover.

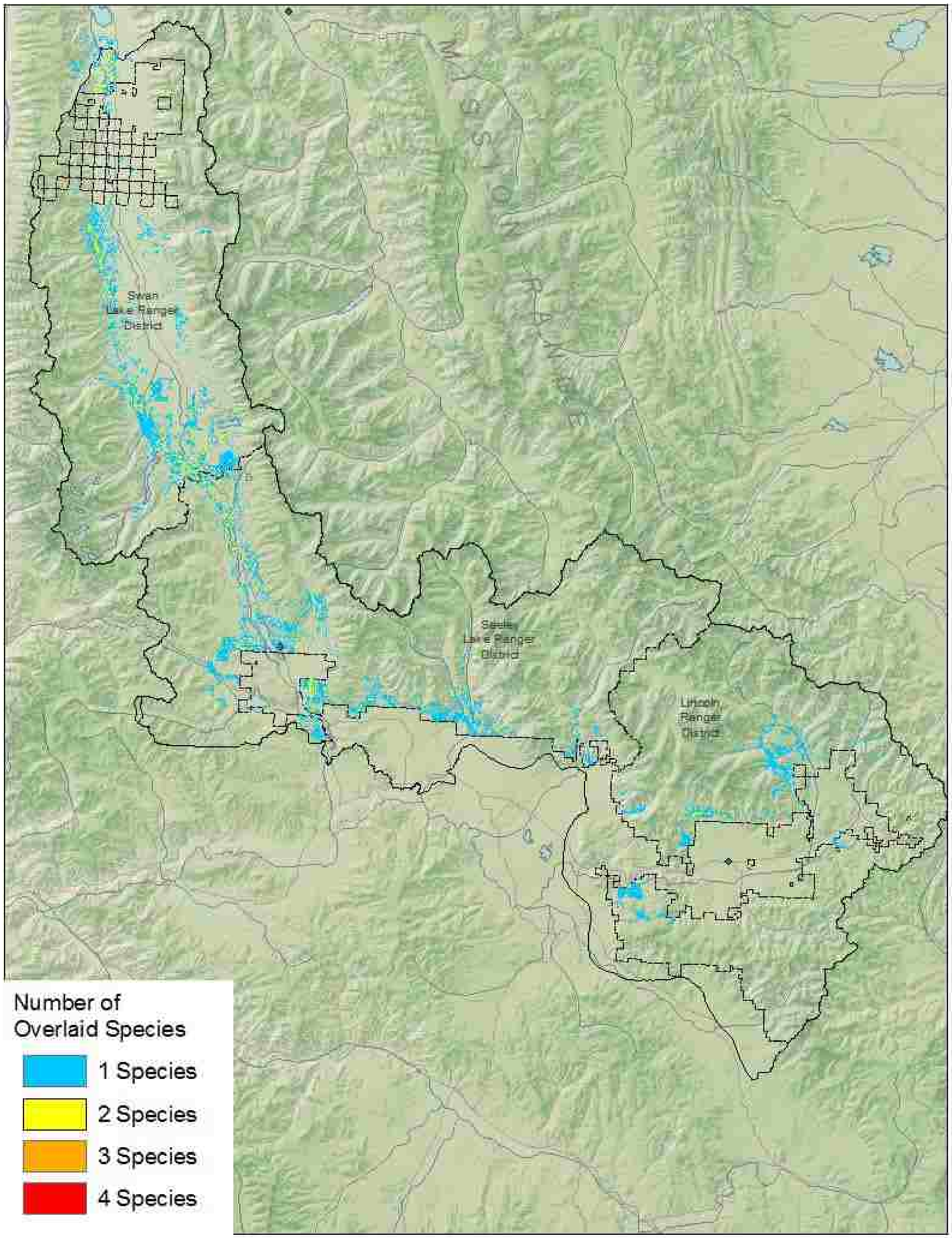

FIGURE 11. Map of predicted moderate and high probability habitat for all models overlaid to show areas of congruence for one, two, three, and four species.

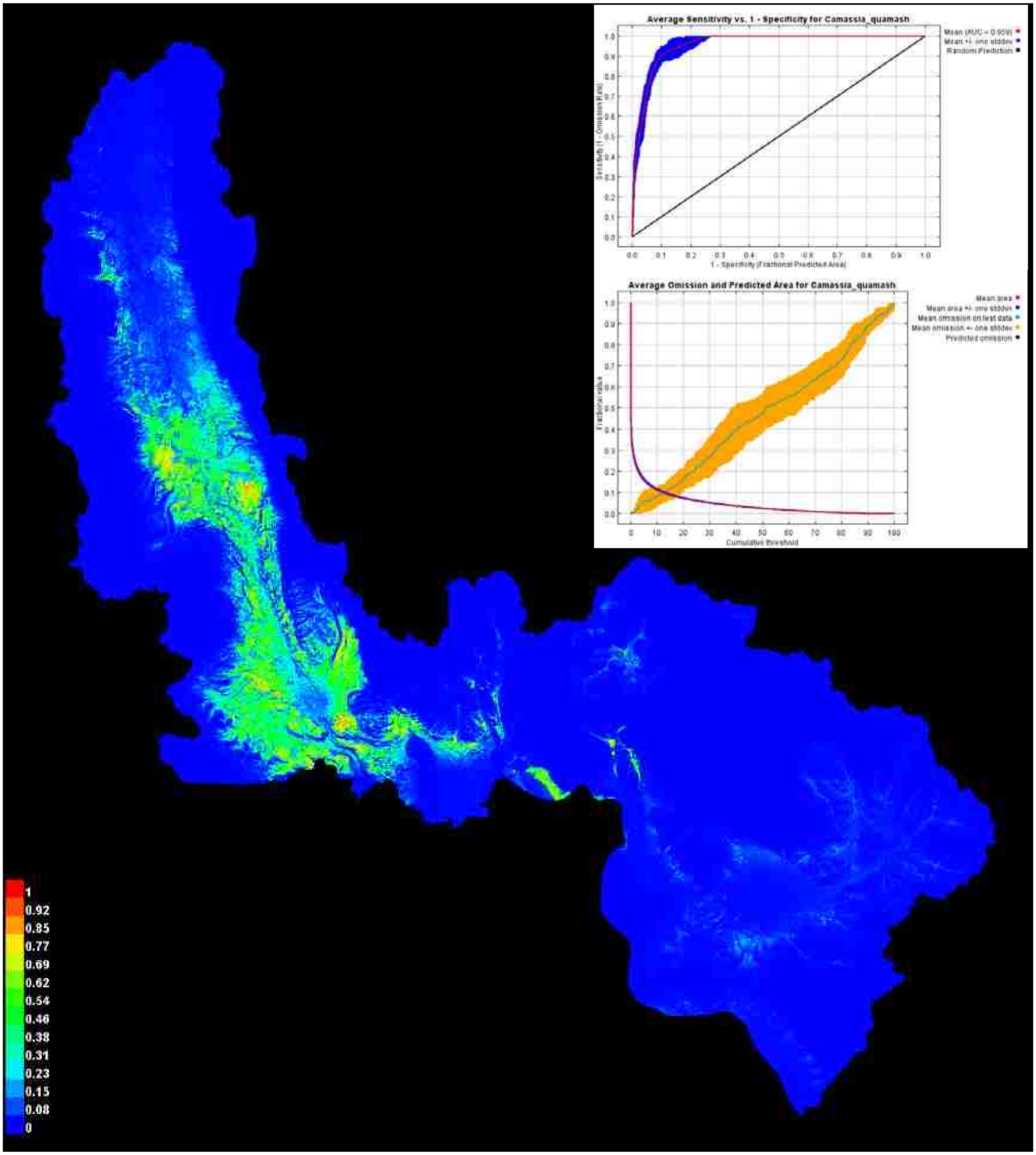

FIGURE 12. Map of predicted habitat, average sensitivity vs. specificity – 1 graph, and average omission graph for centroid dataset for common camas.

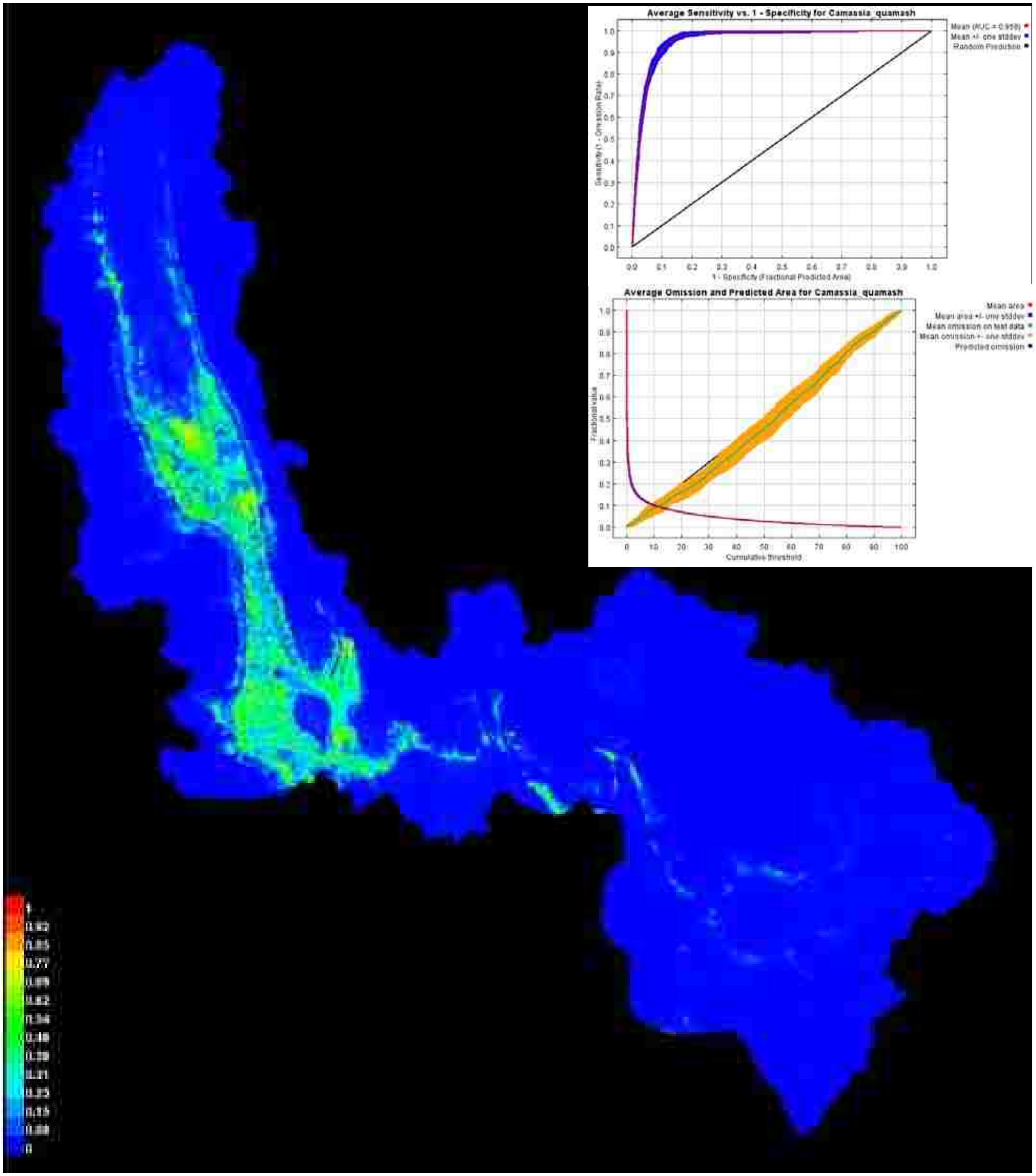

FIGURE 13. Map of predicted habitat, average sensitivity vs. specificity – 1 graph, and average omission graph polygon-to-points dataset for common camas.

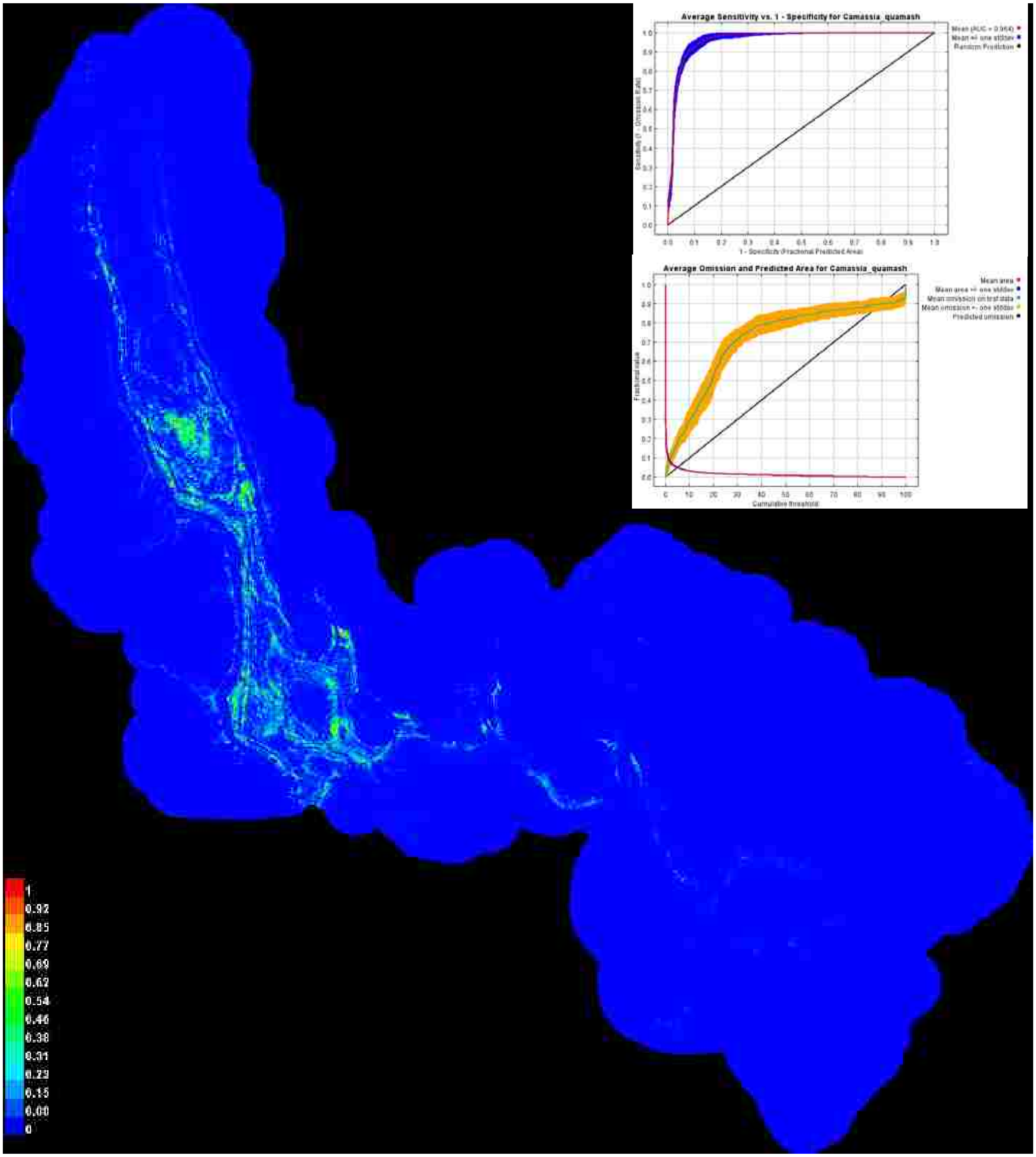

FIGURE 14. Average omission and predicted area, average sensitivity vs. specificity **–** 1, and output predicted distribution for common camas with a regularization parameter of 0.01.

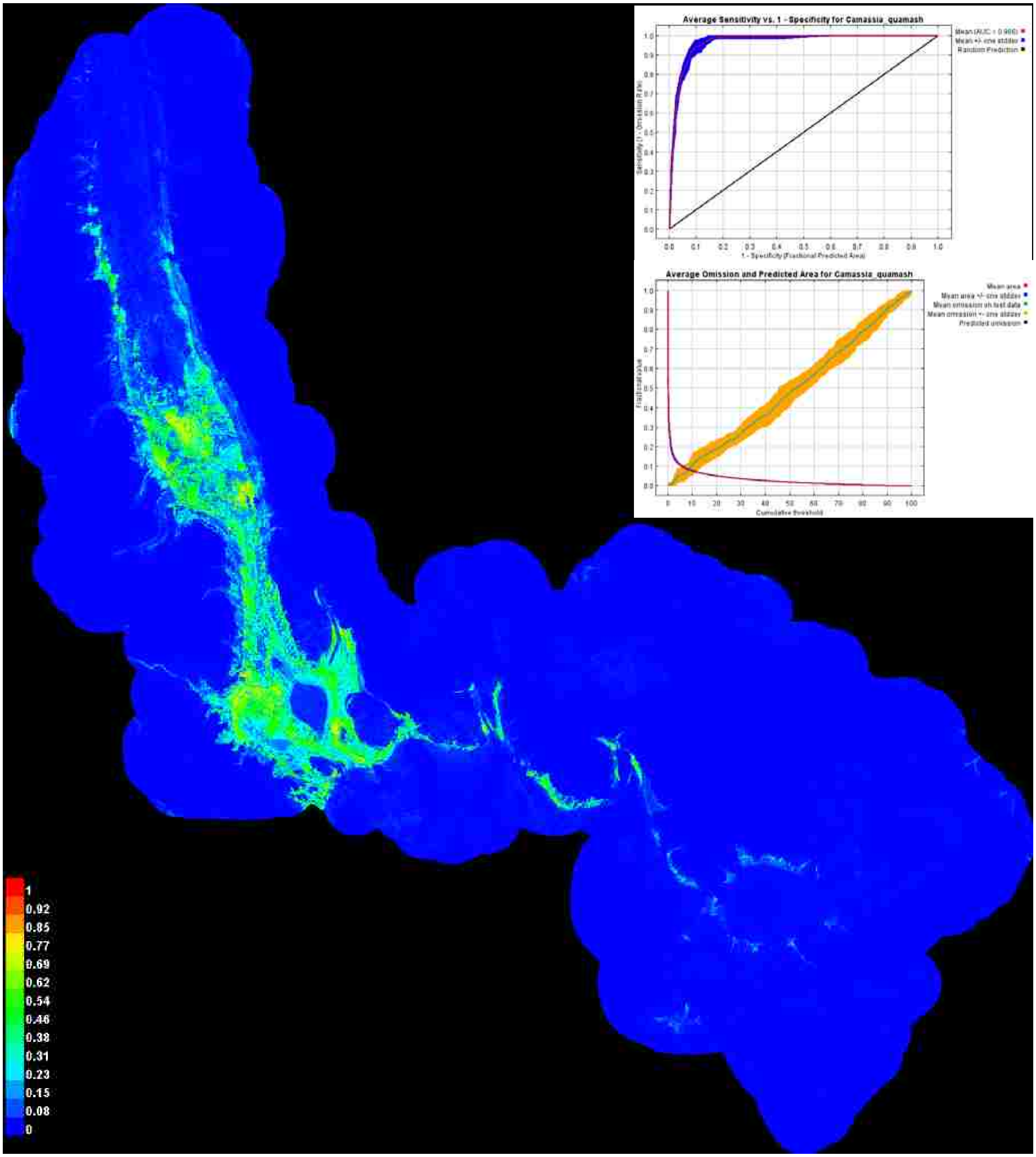

FIGURE 15. Average omission and predicted area, average sensitivity vs. specificity – 1, and output predicted distribution for common camas with a regularization parameter of 1.0.

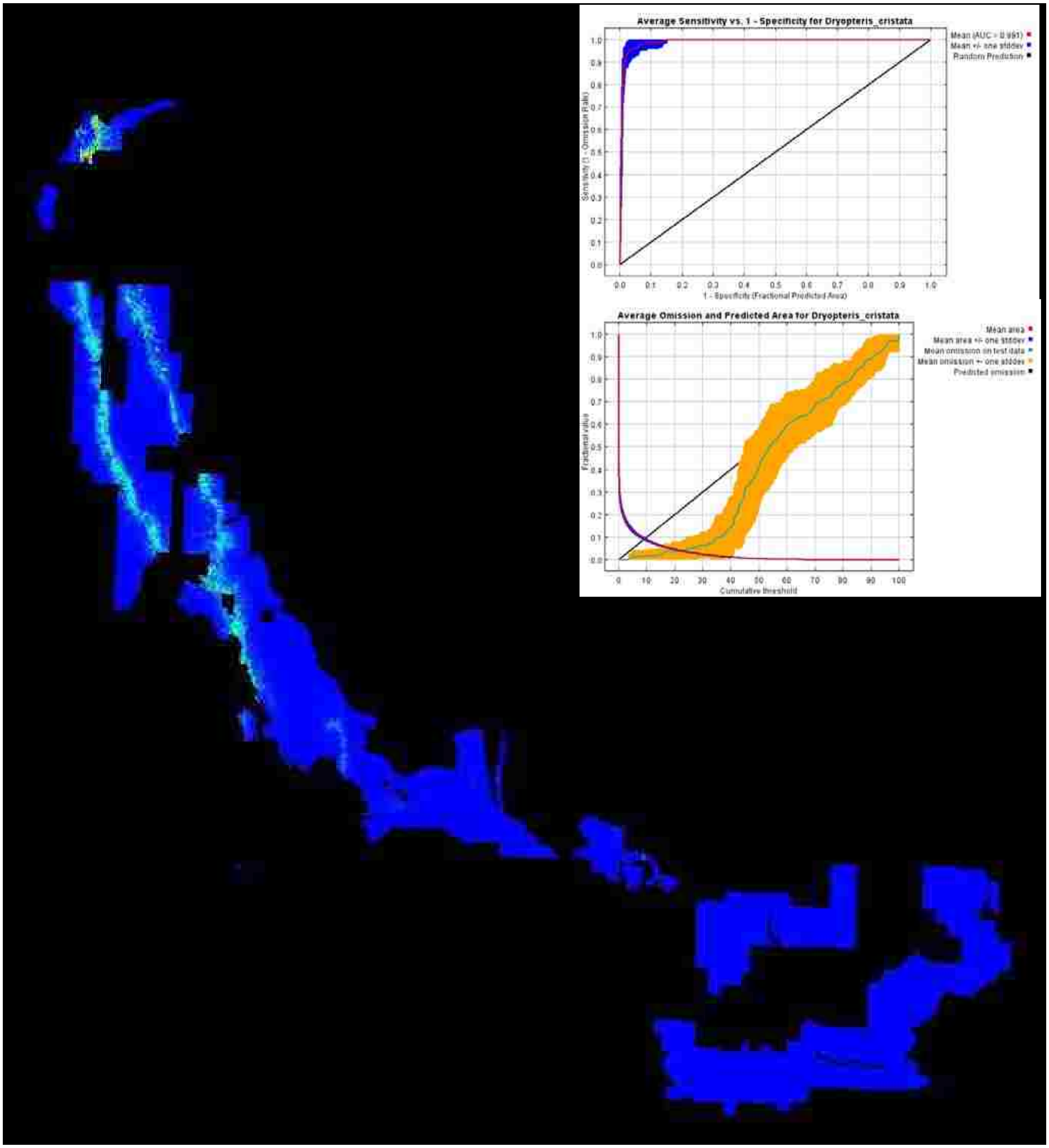

FIGURE 16. Average omission and predicted area, average sensitivity vs. specificity – 1, and output predicted distribution for crested shieldfern without LiDAR-derived canopy cover.

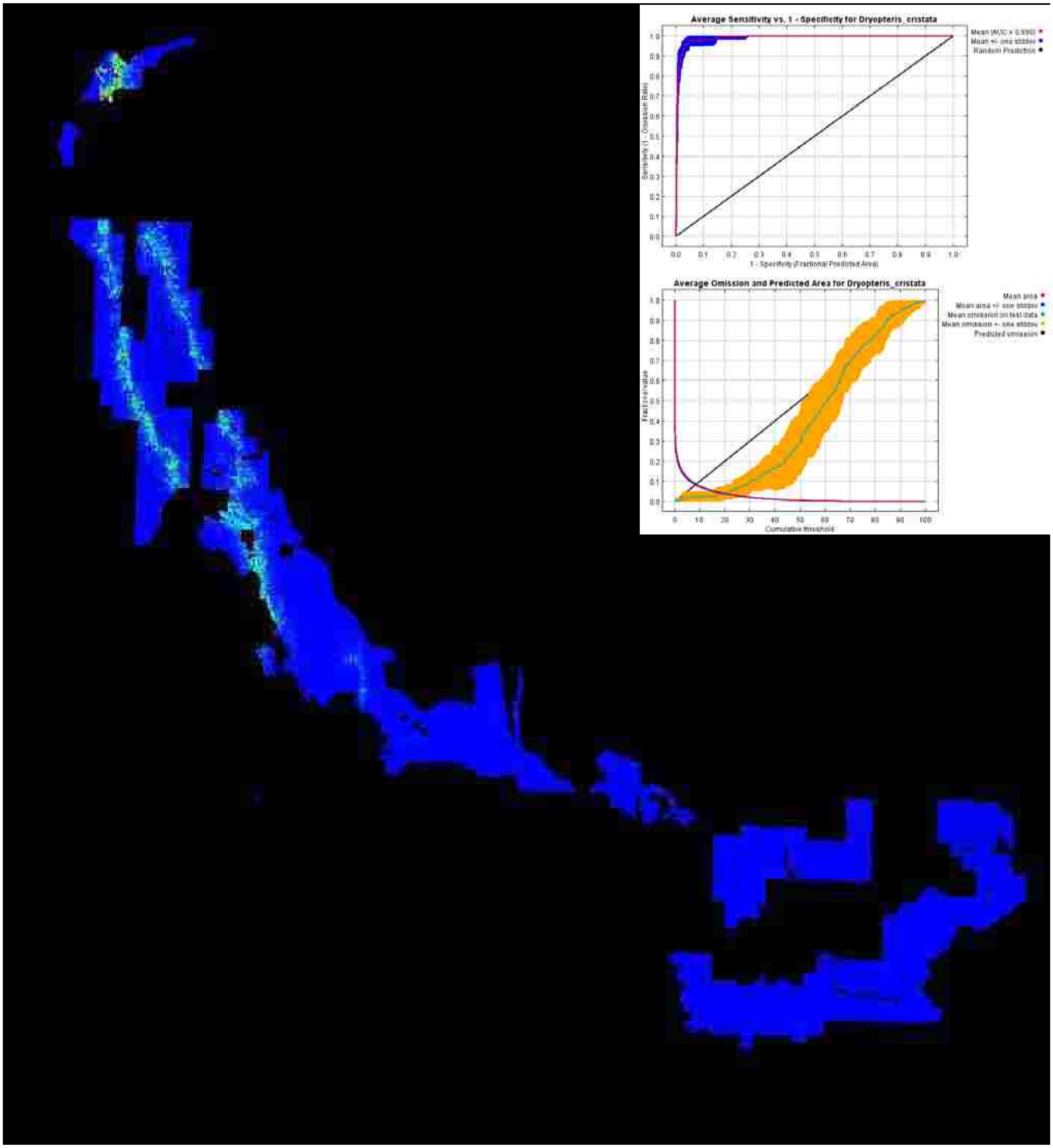

FIGURE 17. Average omission and predicted area, average sensitivity vs. specificity – 1, and output predicted distribution for crested shieldfern with LiDAR-derived canopy cover.

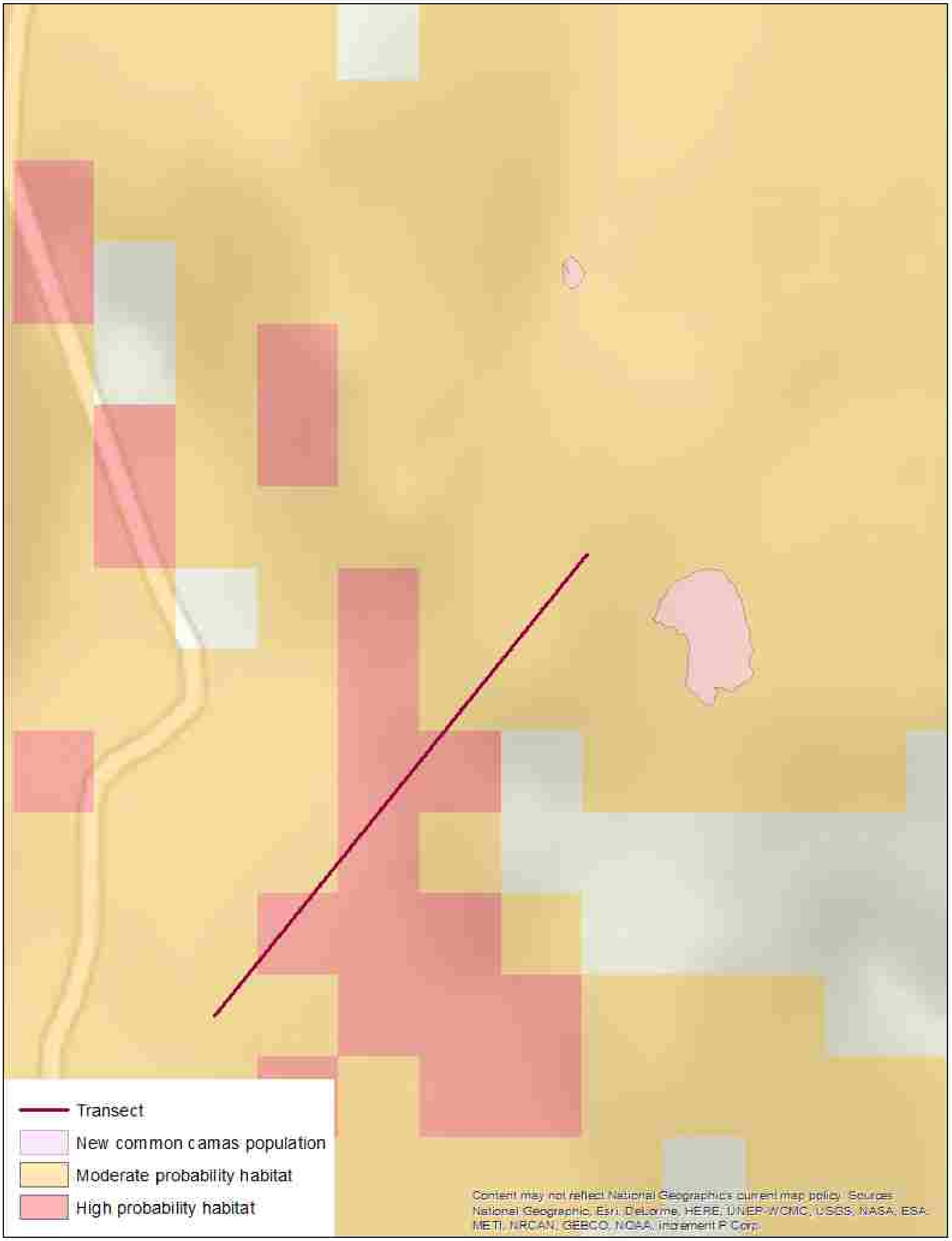

FIGURE 18. Map of a transect for common camas demonstrating how new populations were found outside of the transect.

Species	Dispersal Mechanism	Reasoning	Dispersal Buffer (miles)
Common camas	seed, bulbs	long, two mechanisms	2.5
Clustered lady's-slipper	seed	long, wind dispersal	2.5
Western pearlflower	seed	short, gravity dispersal	1
Howell's gumweed	seed	long, sticky flowering head	2.5
Crested shieldfern	spore	long, spores are robust	2.5
Crested shieldfern w/ LiDAR	spore	long, spores are robust	-

TABLE 1. Dispersal mechanism, reasoning for assigning dispersal buffer, and dispersal buffer distance for each modeled species.

Species	Number of Populations	Random Points	Regularization Parameter	Survey Dates (2016)
Common camas	103	314	1	6/8, 6/9, 6/10, 6/11, 6/15
Clustered lady's-slipper	4	18	0.01	6/13, 6/15, 6/16, 6/17
Western pearlflower	40	1191	0.01	6/23, 6/24, 6/25
Howell's gumweed	151	849	0.01	7/19, 7/20, 7/21, 7/22
Crested shieldfern	12	150	0.01	8/16, 8/17, 8/18, 8/19
Crested shieldfern with LiDAR	12	150	1	8/16, 8/17, 8/18, 8/19

TABLE 2. Displays the number of populations (after data cleaning), number of random points (using polygon-to-point method), dispersal buffer, regularization parameter, and dates transects were surveyed.

Covariate	Resolution	Source
Elevation (DEM)	30 m	NED
VMap	10 m	USFS Region 1
Solar Insolation	30 m	Derived from DEM
Geology	Vector	USGS
TWI	30 m	Derived from DEM
Precipitation	800 m	PRISM
Canopy Cover	20 m	Derived from LiDAR

TABLE 3. Environmental variables utilized in the model, native data resolution (all were resampled to 30 m), and data source.

Covariate	Geology	VMap	PRISM	Solar Insolation	TWI	DEM
Geology	1					
VMap	-0.04598	1				
PRISM	0.38872	0.02688	1			
Solar Insolation	-0.09368	-0.08135	-0.11629	1		
TWI	0.01041	-0.04282	-0.09666	0.01968	1	
DEM	-0.15172	0.01655	0.61711	0.05255	-0.19043	1

TABLE 4. Variable correlations for six environmental variables used in the models for all species. Variables were considered correlated when the value was $\geq \pm 0.7$.

Covariate	DEM	TWI	Solar Insolation	PRISM	VMap	Geology	Canopy Cover
DEM	1						
TWI	-0.15816	1					
Solar Insolation	0.14505	0.00537	1				
PRISM	0.24836	-0.01669	-0.03607	1			
VMap	0.03701	-0.0321	-0.06775	0.02375	1		
Geology	-0.5988	0.06186	-0.09265	0.28812	0.03459	1	
Canopy Cover	0.01782	-0.01179	-0.09315	0.01131	0.33557	-0.02844	1

TABLE 5. Variable correlations for seven environmental variables for the crested shieldfern model including canopy cover derived from LiDAR data. Variables were considered correlated when the value was ≥ ± 0.7.

Species	Moderate Probability Habitat Cutoff	High Probability Habitat Cutoff
Common camas	0.4	0.7
Clustered lady's-slipper	0.4	0.7
Western pearlflower	0.4	0.7
Howell's gumweed	0.01	0.03
Crested shieldfern	0.05	0.13
Crested shieldfern with LiDAR	0.02	0.16

TABLE 6. Cutoff values used to determine moderate and high probability habitat. These values were determined using Jenks natural breaks.

	MaxEnt Estimates					Field Accuracy Assessment			
Species	AUC	AUC Standard Deviation	% of PA (FS) Hi	% of PA (FS) Mod	Mod & High Probability Habitat (Acres)	Overall Accuracy (%)	% Likely Habitat	Number of New Populations	Number of Transects
Common camas	0.968	0.007	0.16	1.81	22,169	38	29	11	11
Clustered lady's-slipper	0.928	0.068	0.95	0.16	12,452	60	53	0	10
Western pearlflower	0.978	0.002	0.02	0.46	3,517	26	34	0	10
Howell's gumweed	0.931	0.005	0.08	1.31	16,344	44	31	0	10
Crested shieldfern	0.977	0.021	0.03	0.59	7,340	69	68	0	9
Crested shieldfern w/ LiDAR	0.99	0.005	0.15	0.69	2,756	66	61	2	10

TABLE 7. MaxEnt SDM statistics for all six predictions including AUC, AUC standard deviation, percent of project area (USFS only lands) that is high probability habitat, percent of project area (USFS only lands) that is moderate probability habitat, and acres of moderate and high probability habitat. Field accuracy assessment statistics for all six models including overall accuracy, percent of transects ranked as likely habitat, number of new populations found, and number of transects surveyed.

APPENDIX A:

INFORMATION TO DOCUMENT FOR A TRANSECT

GENERAL
- Date
- Name of surveyor
- Access (how you got there, what trails / roads were taken)
- Latitude / longitude / elevation
- TRS/ Quad Name for area surveyed (can add this at the office)

HABITAT
- Describe the landscape:
    - Slope & aspect
    - Landscape description
    - Soil and geology type
    - Moisture level (dry, wet, etc.)
    - General cover (open or shaded)
    - Density (Closed canopy, open canopy, etc.)
- Dominant Habitat Type for unit
- Target TES Species for survey (if known)
- Plant list of all species encountered during survey
- Dominant overstory (trees in the area)
- Dominant understory, % cover of top ~10 dominant species
- Weed populations encountered
- % Cover of weeds
- Significant features in survey area:
    - Specialized habitats (Rock outcropping, moist seeps, open sandy slopes, etc.)
    - User created trails (OHV, horses, etc)
    - Cattle damage (What kind, where it is located, etc.)
    - Fire history (Burned stumps, burn piles, down burned debris, etc.; approximate age)
    - Timber history (Old skid trails, slash piles, old stumps, etc; approximate age)
    - Historical features (Cabins, archaeological sites, etc.)
- Amount of downed debris (fuel load)
- Is this transect likely habitat for this species?
- Take a picture of transect start (looking down transect) and transect end (looking up transect)

ELEMENT OCCURRENCES (within the population, document the following)
- If sensitive species are found, list top 10 related species and detailed habitat notes and size of population
- Possible threats to population?
- Dominant overstory (trees in the area)
- Dominant understory, % cover of top ~10 dominant species
- Weed populations encountered
- % Cover of weeds