AVIAN CONSERVATION & ECOLOGY

VOLUME 19, ISSUE 2, ARTICLE 7

Leipold, E. A., C. N. Gower, and L. McNew. 2024. Using an ensemble approach to predict habitat of Dusky Grouse (*Dendragapus obscurus*) in Montana, USA. Avian Conservation and Ecology 19(2):7. https://doi.org/10.5751/ACE-02697-190207

Research Paper

Using an ensemble approach to predict habitat of Dusky Grouse (*Dendragapus obscurus*) in Montana, USA

Elizabeth A. Leipold[1], *Claire N. Gower*[2] and *Lance McNew*[1]
[1]Montana State University, [2]Montana Fish, Wildlife & Parks, Bozeman MT

ABSTRACT. Dusky Grouse (*Dendragapus obscurus*) are an under-monitored game species in Montana and elsewhere across their distribution. Without population monitoring it is difficult to establish appropriate harvest regulations or understand the impact of environmental disturbances (e.g., timber harvest, climate change) on populations. As a first step toward developing methods for unbiased population monitoring, we must identify appropriate sampling sites, which requires knowledge of Dusky Grouse habitat. Our goal was to explore relationships between Dusky Grouse use and habitat characteristics, and then generate a state-wide map predicting Dusky Grouse habitat in Montana using two methods: resource selection functions and random forest classifiers. The Integrated Monitoring in Bird Conservation Regions program provided a multi-year dataset of Dusky Grouse observations, which we reduced to detected (n=132) and pseudo-absent (n=5960) locations, using geospatial datasets to obtain topographic and vegetation characteristics for each location. We evaluated the predictability of the two models using receiver operating characteristics and area under the curve (ROC/AUC) with k-fold cross validation and classification accuracy of an independent dataset of incidental Dusky Grouse locations. We found both models to be highly predictive and multiple habitat characteristics were found to help predict relative probability of use such as proportion of trees with a height of 16–20m and conifer forest vegetation types. We converted both models to binary values and used an ensemble (frequency histogram) approach to combine the models into a final predictive map. Consensus between the resource selection function and random forest models was high (93%) and the ensemble map had higher predictive accuracy when classifying the independent dataset than the other two models. Our results show that our ensembled model approach was able to accurately predict potential Dusky Grouse habitat and therefore can be used to delineate areas for future population monitoring of Dusky Grouse in Montana.

Utilisation d’une approche par ensemble pour prédire l’habitat du Tétras sombre (*Dendragapus obscurus*) dans le Montana, aux États-Unis

RÉSUMÉ. Le Tétras sombre (*Dendragapus obscurus*) est une espèce chassée peu surveillée dans le Montana ainsi que dans l’ensemble de son aire de répartition. Sans suivi des populations, il est difficile d’établir des réglementations de prélèvements appropriées ou de comprendre l’impact de perturbations environnementales (comme l’exploitation du bois d’œuvre ou le changement climatique) sur les populations. La première étape de l’élaboration de méthodes de suivis des populations non biaisés consiste à identifier des sites d’échantillonnage appropriés, ce qui nécessite une connaissance de l’habitat du Tétras sombre. Notre objectif était d’explorer l’usage de l’habitat du Tétras sombre, en fonction des caractéristiques de celui-ci, puis de générer une carte à l’échelle de l’État permettant de prédire l’emplacement de l’habitat du Tétras sombre dans le Montana à l’aide de deux méthodes : les fonctions de sélection des ressources [resource selection functions] et la classification par forêts aléatoires (arbres décisionnels). Le programme de Surveillance intégrée dans les Régions de conservation des oiseaux [Bird Conservation Region] a fourni un jeu de données d’observation du Tétras sombre sur plusieurs années, que nous avons réduit en tant que points de présence (n=132) et de pseudo-absence (n=5960). En utilisant des données géospatiales, nous avons pu obtenir les caractéristiques topographiques et de végétation pour chaque emplacement. Nous avons évalué la capacité prédictive des deux modèles en utilisant la fonction d’efficacité du récepteur aussi appelée courbe ROC [Receiver Operating Characteristics] et la surface sous la courbe [Area Under the Curve] (ROC/AUC) avec la méthode de la validation croisée k-fold et l’évaluation de l’exactitude de la classification sur un jeu de données indépendant de localisations opportunistes du Tétras sombre. Nous avons constaté que les deux modèles étaient hautement prédictifs et que plusieurs caractéristiques de l’habitat contribuaient à prédire la probabilité relative d’utilisation, telles que la proportion d’arbres d’une hauteur de 16 à 20 m et la végétation de type forêt de conifères. Nous avons converti les deux modèles en valeurs binaires et utilisé une approche par ensemble (histogramme de fréquence) pour combiner les modèles en une carte prédictive finale. La concordance entre les modèles de fonction de sélection des ressources et de classification par forêts aléatoires était élevée (93 %) et la carte ensembliste avait une précision prédictive plus élevée que les deux autres modèles lors de la classification du jeu de données indépendant. Nos résultats montrent que notre approche par ensembles multimodèles a été capable de prédire avec précision l’habitat potentiel du Tétras sombre et peut donc être utilisée pour délimiter les zones de surveillance future de la population de Tétras sombres dans le Montana.

Key Words: *habitat model; random forest; resource selection functions; species distribution model*

Corresponding author: Elizabeth A Leipold, Elizabeth.Leipold@montana.edu

INTRODUCTION

Dusky Grouse (*Dendragapus obscurus*) are an under-monitored galliform found in the interior mountain ranges of western North America (Aldrich 1963, Zwickel and Bendell 2004). Unlike other upland gamebird species, standardized survey protocols have not yet been developed for Dusky Grouse. Although presumed to occupy most of their historical distribution (Zwickel and Bendell 2004), the lack of basic information on population status, current distributions, and regional habitat associations hinders effective population management, the ability of managers to know if populations are declining, and to set appropriate harvest regulations. Further necessitating the development of targeted and robust population and habitat monitoring protocols, increasing environmental stressors (e.g., wildfire, exurban development, timber harvests, beetle kill, climate change) are occurring in Dusky Grouse habitats (Martinka 1972, Redfield 1973, Chan-McLeod and Bunnell 2003, Youtz et al. 2022).

Accurate species distribution models (SDMs) provide information on distribution and habitat associations that are useful for conservation and management. Species distribution models can be used to direct survey locations toward potential habitat thus increasing the likelihood of species detection, assist with determining conservation status, delineate areas for conservation, and help predict a species response to management actions or climate change (Guisan et al. 2006, Sofaer et al. 2019). Recently the number of techniques available for creating species distribution models has grown, and now includes more classic techniques such as resource selection functions (RSFs), as well as newer machine learning methods (Elith and Graham 2009, Grenouillet et al. 2011, Shoemaker et al. 2018, Picardi et al. 2020). Each approach has its own limitations, and no one method is universally best for all applications (Araùjo and New 2007, Elith and Graham 2009). Unfortunately, the subjective choice of modeling technique can influence the predictions of species habitat (Araùjo and New 2007, Pearson et al. 2006). Because conservation planning is increasingly reliant upon SDMs (Guisan and Thuiller 2005, Guisan et al. 2013), it is imperative that predictions of species distribution are accurate. Recent research has highlighted the benefits of ensemble modeling, where predictions of species distributions are produced with several statistical techniques, which can improve accuracy and reduce uncertainty in SDM predictions (Araújo et al. 2005, Marmion et al. 2009, Grenouillet et al. 2011). In addition, consensus within the suite of models on predicted habitat may increase certainty in the model's accuracy, while variation in predictions across models may serve as an index of uncertainty in species distribution (Latif et al. 2013).

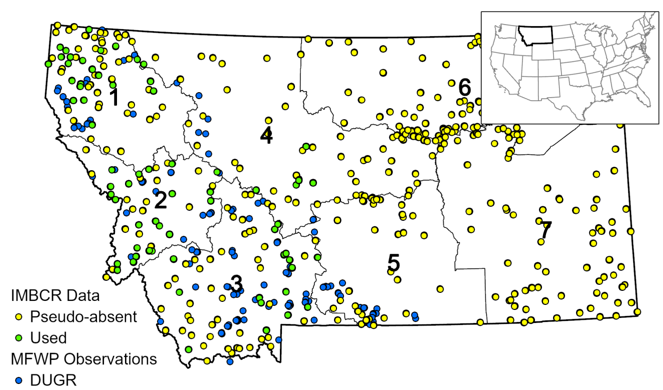
We used an ensemble modeling approach to understand habitat associations and predict the distribution of Dusky Grouse in Montana, USA. Our specific objectives were to (1) explore relationships between Dusky Grouse presence and landscape-level habitat characteristics, (2) generate and test the accuracy of predictions of Dusky Grouse habitat using resource selection functions and randomized classification trees (e.g., Random Forest), and (3) compare the performance of an ensemble model to our individual species distribution models. Our goal is to provide an accurate state-wide prediction of Dusky Grouse habitat that can be used to guide monitoring efforts by being able to predict appropriate population survey sites at scales relevant to management.

METHODS

Study area

Our study area included the entire state of Montana, USA, which consists of 7 Montana Department of Fish, Wildlife & Parks (MFWP) administrative regions (Fig. 1). Dominant vegetation types in western Montana (Regions 1–3) are mainly conifer forests, intermixed with shrublands and intermountain foothills grasslands (LANDFIRE 2016a). Eastern Montana (Regions 6–7) are dominated by grasslands, shrublands, and cultivated lands, and Dusky Grouse habitat is thought to be less abundant and more isolated (LANDFIRE 2016a). Central Montana (Regions 4–5) transitions from vegetation types found in western Montana to those found in eastern Mountain and is dominated by both mountainous and coniferous forests in the west and grasslands, shrublands, and cultivated lands toward the east (LANDFIRE 2016a). Elevation varies from 550 to 3897 m, with lower elevations in the east and higher elevations in the mountainous areas in the west (U.S. Geological Survey 2017).

Fig. 1. Map of study area with Integrated Monitoring in Bird Conservation Regions (IMBCR) survey sites (n = 6092) and Montana Department of Fish, Wildlife & Parks (MFWP) incidental observations of Dusky Grouse (*Dendragapus obscurus*; n = 194). At the IMBCR sites, 132 were classified as used (Dusky Grouse detected) and 5,960 as pseudo absent (Dusky Grouse not detected). MFWP divides Montana into seven administrative regions for conservation and management. FWP regions are outlined in gray and labeled by region number (regions 1–3 on the left, 4–5 in the center, and 6–7 on the right).



Grouse observations

We used two independent datasets of Dusky Grouse observations: a dataset with used and pseudo-absent locations for training and testing our models, and an independent dataset with presence-only locations for additional validation tests. For training our SDMs, we obtained Dusky Grouse observation data from the IMBCR program administrated by the Bird Conservancy of the Rockies. IMBCR survey sites occur over much of the mid-western and western parts of the United States, with the surveyed land broken up into different strata that are based on land ownership and topographic features (Pavlacky et al. 2017, Woiderski et al. 2018). A grid of 1-km² cells is generated for each stratum, with 16 survey points evenly spaced every 250 m within each cell (Woiderski et al. 2018). Each spring during May–July, a spatially balanced sampling algorithm called generalized random-tessellation stratification (GRTS), is used to randomly select a minimum of 2 sample units (cells) within each stratum, within which observers conduct 6-minute point-counts at each of the 16 sites (Woiderski et al. 2018). We obtained observation data from spring (May–June) surveys conducted during 2009–2020 for a total of 25,654 surveys conducted across 6092 sites in Montana. Surveys occurred at similar dates across all MFWP Regions reducing the potential for temporal differences in probability of detecting a Dusky Grouse to bias pseudo-absence. Dusky Grouse were detected (observed ≥ 1 time) at 132 sites and not detected at 5960 sites (Table 1; Fig. 1). Because the detection of Dusky Grouse is imperfect (Leipold 2023), the observation data fit a standard used vs available sampling design (Design 1, sampling protocol A) rather than a used vs unused sampling design (Manly et al. 2002). We defined our pseudo-absent locations using a common strategy for species distribution modeling where pseudo-absent locations are defined as surveyed sites where the target species was not detected (Lütolf et al. 2006, Hanberry et al. 2012). We included all IMBCR sites within Montana in our analysis; both used and available points are considered at the statewide population level as per Design 1 of Manly et al. (2002), which is analogous to Johnson's (1980) first order of selection.

Table 1. Summary of location data. For the Integrated Monitoring in Bird Conservation Regions (IMBCR) dataset, the number of used and pseudo-absent points in total and per Montana Department of Fish, Wildlife & Parks (MFWP) region. For the MFWP incidental data, the number of Dusky Grouse (*Dendragapus obscurus*; DUGR) observations in total and per MFWP region. NAs are the result of points being located either just outside or on the border of Montana.

Region	IMBCR		MFWP
	Used	Pseudo-absent	DUGR
Region 1	41	893	26
Region 2	50	494	22
Region 3	30	815	86
Region 4	11	832	12
Region 5	0	452	47
Region 6	0	1241	0
Region 7	0	1171	0
NA	0	62	1
Total	132	5960	194

MFWP field staff recorded the geographic locations of incidental Dusky Grouse observations observed during unrelated field activities conducted in the springs (April–June) of 2017–2021. During the 4-year period, 194 Dusky Grouse locations were recorded (Fig. 1), which we used as an independent dataset to further evaluate the predictive accuracy of our models.

Environmental predictors

Dusky Grouse make repeated softer hoots that are difficult to detect at distances greater than 100 m, but wing flutters, wing claps, and the louder single hoot can be heard up to 250 m (Harju 1974; *personal observation*). We examined the estimated distances to individuals recorded in the IMBCR data, and 95% of all Dusky Grouse observed were estimated to be within 250 m of the IMBCR point-count location. We calculated the mean statistic or proportion of a habitat characteristic within the 250-m radii circle centered on the survey point using geospatial modeling environment (GME), remotely sensed geospatial datasets, and the spatial analyst tools in ArcGIS Pro (Environmental Systems Research Institute, Redlands, CA; Appendix 1; Beyer 2015). Because Dusky Grouse tend to be found at higher elevations, may prefer east to south-facing slopes, and brood-rearing females have been found to prefer more moderate slopes, we extracted average elevation, slope, and proportions of different facing (N, NE, E, etc.) aspects from a digital elevation model (Stauffer and Peterson 1986, U.S. Geological Survey 2017, Farnsworth 2020). In addition, we calculated the mean distance to the nearest stream and to the nearest road using spatial analyst tools from ArcGIS pro (Environmental Systems Research Institute, Redlands, CA) and road and stream geospatial data layers (Montana Spatial Data Infrastructure 2017, 2018). Because the attribute data for the geospatial road layer was incomplete, we necessarily grouped unimproved and paved roads together including highways, secondary roads, and forest service roads with and without seasonal closures. We used the LANDFIRE geospatial data with a 30 × 30 m spatial resolution to describe existing vegetation type (EVT), which is the type of plant community present, projected canopy cover in 1% increments (EVC) for the vegetation type for that pixel, and average height of the dominant vegetation given in 1-m increments (EVH; LANDFIRE 2016a, b, c, 2019, 2020). Both EVC and EVH data are treated as categorical variables by LANDFIRE. We condensed the 1% increments for the canopy vegetation to 10% increments for each category of vegetation type (i.e., tree, shrub, herbaceous) and the 1-m increments for tree height to 5-m increments, and the 0.1-m increments for shrub and herbaceous heights to 0.5-m increments to reduce the number of variables evaluated. For both EVC and EVH data, there were also several categories of developed habitat or barren habitat that did not have cover percentages or heights that were grouped into two categories: developed and sparse vegetation. In addition, we created a conifer forest layer based on the descriptions for the different LANDFIRE vegetation types. We used this layer to calculate the average distance to the edge of the conifer forest from within the forest and outside of the forest (Appendices 1, 2; LANDFIRE 2016a). We removed variables from consideration if they occurred at less than 1% of the survey sites or showed no relationship with use by Dusky Grouse, resulting in 90 environmental predictor variables (Appendix 1).

Habitat associations and ensemble of predictions

Resource selection functions

We fit RSFs using general linear mixed models (GLMMs) with a logit link function, binomial error distribution, and the "bobyqa" optimizer with a maximum of 100,000 iterations for approximating beta coefficients using the "lme4" package in program R (Bates et al. 2015, R Core Team 2017). Our GLMMs included the binomial response variable of whether a Dusky Grouse was detected (1) or not detected (0) at each site, combinations of environmental predictors, and a random intercept for unique IMBCR transects to account for potential spatial autocorrelation in the observation data because of the survey points being grouped along IMBCR survey routes (Zurr 2009, Woiderski et al. 2018).

Prior to model fitting, we explored possible non-linear responses of Dusky Grouse to habitat variables using univariate generalized additive models (GAMs) and linear equations hypothesized to represent the linear and nonlinear forms such as a quadratic form [$x + x^2$] and pseudolinear threshold (ln[x + 0.001]; Franklin et al. 2000, Guisan and Zimmerman 2000, Guisan et al. 2002, Dugger et al. 2005, McNew et al. 2015). We performed the preliminary screenings of the three functional responses using univariate models built using GLMMs with a logit link function and binomial error distribution, evaluating the model support using Akaike's Information Criterion for small sample size (AIC_c; Burnham and Anderson 2002). After preliminary screenings of the functional responses, we assessed multicollinearity in the remaining habitat predictor variables for all possible pairings using Spearman-rank correlations. We considered variables to be correlated when $|r| > 0.7$. If variables were correlated, we first used knowledge of general Dusky Grouse habitat to evaluate which variable was more biologically relevant to Dusky Grouse. If no previous information about the habitat characteristic was available, using the univariate models with the best performing functional response for the correlated variables we evaluated model support using AIC_c, and whichever correlated variable was least supported was removed from our analysis (Burnham and Anderson 2002, Aldridge et al. 2012). If a variable was correlated with more than one variable, we evaluated correlations based upon most correlated to least, removing variables until only one uncorrelated variable remained in the dataset. After assessing for correlation, we had 66 remaining variables.

We first evaluated predictors in groups based on variable type: aspect, other non-vegetation variables (i.e., slope, elevation, distance to variables), conifer vegetation type (divided into two groups that were later combined because of model convergence issues), hardwood vegetation type, grassland vegetation type, shrubland vegetation type, riparian vegetation type, other vegetation type, tree canopy cover, shrub canopy cover, herbaceous vegetation cover (also divided into multiple groups that were later combined because of model convergence issues), other vegetation canopy cover, and vegetation height. Within each group, we used backwards elimination to determine the top performing variables for inclusion in the final candidate model set, with variable removal based on the p-value calculated using the "lme4" package in program R from asymptotic Wald tests, where the variable with the highest p-value > 0.05 was removed (Hosmer et al. 2013, Bates et al. 2015, Heinze et al. 2018). Backwards selection continued until all variables within the model had p-values < 0.05 (Heinze et al. 2018). The top performing variables from each group were then added to a global model and again evaluated using backwards elimination to determine a final set of variables for predicting Dusky Grouse habitat. We calculated the 95% confidence intervals for the beta coefficients using the Wald method, which estimates the fixed-effects confidence intervals, using the "lme4" package in program R (Bates et al. 2015, R Core Team 2017). We used the high and low confidence intervals to create two additional maps to examine uncertainty in the predictions (Appendix 3).

Random forest

We developed random forest (RF) models by creating a large number of classification trees where majority voting was used to decide on the classification result (Breiman 2001) using the *train* and *trainControl* functions and the "rf" model from the "caret" package in R (Kuhn 2008, R Core Team 2017). The machine learning algorithm trained each randomized classification tree on a bootstrap sample of the training data where for each split a random selection of the total predictors were chosen and the best predictor of those chosen was used to partition the data (Evans et al. 2011, Kuhn and Johnson 2013). Random forest models are sensitive to unbalanced datasets, such as our IMBCR dataset where the number of pseudo-absent locations greatly outnumbered the used locations (Evans and Cushman 2009, Kuhn and Johnson 2013, Kuhn 2019). To account for our dataset being unbalanced, we used the down-sampling function within the caret package to rarify the random sampling data to a 1:1 ratio with the used locations (Evans et al. 2011, Kuhn and Johnson 2013, Kuhn 2019). We tuned our model by varying the number of trees and the number of variables to possibly split at each node (Kuhn 2008, Kuhn and Johnson 2013, R Core Team 2017). The number of variables to possibly split at each node, "mtry," was tested with the square root of the number of predictors, the square root of the number of predictors divided by 2, and the square root of the number of predictors times 2 (Breiman 2001, Kuhn and Johnson 2013). The number of trees tested were 300, 500, 800, 1000, and 2000. After we tuned the model, we trained it with repeated cross validation, with 5 folds and 500 repeated k-fold cross validation iterations. We generated variable importance using the "randomForest" package, which calculates the impact of removing a variable on the model or mean decrease in accuracy (Liaw and Wiener 2002, R Core Team 2017).

Predictions of Dusky Grouse habitat

We developed separate statewide predictions of relative use from each model. We used a 250-m moving window to create spatial layers for each variable upon which to predict our models. We first used slope coefficients from our top GLMM to fit an RSF (Manly et al. 2002). Second, to evaluate Dusky Grouse occurrence across Montana using the random forest model, we used the predict function in R with our 250-m circular moving window layers to construct a predictive map of potential Dusky Grouse habitat (Kuhn 2008, R Core Team 2017). We then rescaled RSF and random forest predictive maps of potential Dusky Grouse habitat in Montana to be between 0 and 1.

Model evaluation

We evaluated the performance of our models in two ways: (1) repeated k-fold cross validation with 5 folds where 80% of the IMBCR data used to train the model and 20% used to test the model, and (2) with an independent dataset. We conducted a simulation with 500 iterations (externally for the RSF and within the model fitting process for the RF) where we used k-fold cross validation and threshold-independent ROC/Area Under Curve (AUC) to evaluate mean model performance with the IMBCR training dataset (Zweig and Campbell 1993, Fielding and Bell 1997). We calculated the average AUC value with a 95% confidence interval. An AUC greater than 0.7 indicates that the model has acceptable predictive power and that the performance is better than that of pure chance (Fielding and Bell 1997, Boyce et al. 2002, Hosmer et al. 2013, Bohnett et al. 2020).

In addition, we tested model predictions using an independent dataset of incidental grouse observations collected by MFWP and comparing it with the presence locations from the IMBCR dataset. For each MFWP and IMBCR Dusky Grouse observation, we calculated the RSF value and the RF value. For each model type (RSF, RF), we used the IMBCR dataset to categorize the values into 5 quantile bins of equal size (20% of the data in each bin) that represented increasing relative probability of a point being classified as a site used by Dusky Grouse (Boyce et al. 2002, Johnson et al. 2006, McNew et al. 2013). The bins represented low (lowest 20% of raw RSF values), medium-low (20–40%), medium (40–60%), medium-high (60–80%), and high (highest 20%) probability of relative use. We then regressed the observed proportion of grouse locations from the MFWP or test dataset in each quantile bin with the observed proportions of grouse locations in each quantile bin from the IMBCR or training dataset. We used linear regression to compare the training and testing datasets, and we considered a good model fit to have a high R^2 value, a slope close to 1, and an intercept close to 0 (Johnson et al. 2006, McNew et al. 2013).

Calculating potential Dusky Grouse habitat in Montana

We used the quantile bin that correctly predicted 75% of the used points in the training and test datasets as our threshold for delineating habitat and created a binary map of habitat and non-habitat for each model. To evaluate the accuracy of the threshold, we conducted a simulation with 500 iterations, where we calculated the average percent of correctly predicted locations with a 95% confidence interval for a subset (80%) of the MFWP dataset. We conducted this simulation for the state-wide MFWP data and for each MFWP region. Because predictive accuracies of both models were similar (see Results) we added the binary maps of our two individual models to obtain a final map created from the ensembled prediction of potential habitat of Dusky Grouse in Montana (a frequency histogram approach; Araùjo and New 2007, Le Lay et al. 2010, Stohlgren et al. 2010). The ensemble map displayed values ranging from 0 to 2, where the pixel value was related to the number of models that predicted it to be habitat. A value of 2 indicates that both models predicted that pixel to be habitat, while a value of 1 indicates that only one model predicted that area to be habitat and a value of 0 indicates that both models predicted the area to not be habitat. We considered areas with a value of 2 to be high probability of being habitat (consistently predicted to be habitat) and areas with a value of 1 to be medium-high probability of being habitat (inconsistently predicted to be habitat).

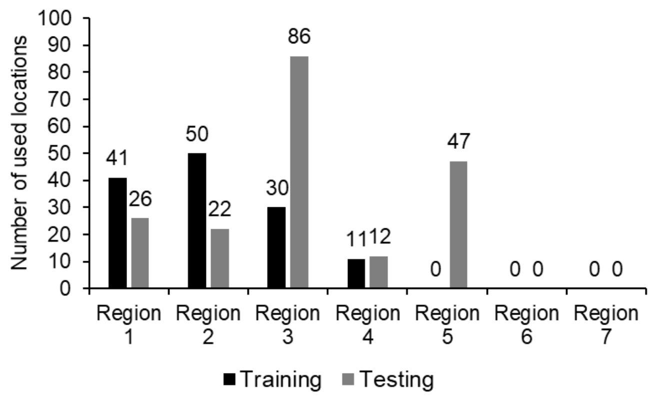
We calculated the amount of area within each MFWP administrative region and Montana that was predicted to be medium-high and high probability of being habitat. We calculated the amount of Dusky Grouse habitat in each category by summing the number of pixels predicted to be within the categories and multiplying by pixel size (0.0009 km^2). The amount of Dusky Grouse habitat in Montana and within each MFWP administrative region was calculated as ranging between the high probability of being habitat category and total potential habitat (the sum of both the medium-high and high probability of being habitat categories).

RESULTS

Location data

The majority (69%) of the used locations for the IMBCR dataset were in MFWP Regions 1 and 2, while the majority of the presence locations for the MFWP dataset were in Regions 3 (44%) and 5 (24%; Fig. 2). Of the regions with presence locations, the fewest presence locations for both datasets were in MFWP Region 4 (Fig. 2). There were no used points in the IMBCR dataset in Regions 5, 6, 7 (Fig. 2).

Fig. 2. Total number of used (Dusky Grouse, *Dendragapus obscurus*, detected) locations for the training (Integrated Monitoring in Bird Conservation Regions) and the testing (Montana Department of Fish, Wildlife & Parks) datasets for each MFWP region. No used locations were observed in Regions 6 and 7 for either dataset.



Dusky Grouse habitat associations

From the RSF model we found 7 habitat predictors to affect Dusky Grouse occurrence (Table 2). Two variables had a quadratic relationship with relative probability of use: average distance to nearest stream (β = 7.40 ± 2.11SE, β = -7.49 ± -2.70) and proportion of northern rocky mountain foothill conifer wooded steppe (β = 216.70 ± 32.83, β = -5557.00 ± 137.60; Table 3, Fig. 3). Average slope had a positive linear relationship with relative probability of use (β = 1.03 ± SE 0.26). Proportion of inter-mountain basins montane sagebrush steppe (β = 0.16 ± SE 0.06), and the proportion of trees with a height of 16–20 m (β = 0.32 ± SE 0.08) had positive pseudolinear threshold relationships with relative probability of use by Dusky Grouse and the proportion of trees with a height of 1–5 m (β = -0.63 ± SE 0.24) and distance to road (β = -0.31 ± 0.14) had negative pseudolinear threshold relationships with relative probability of use (Table 3, Fig. 3). Conditional and marginal R^2 values for this model were 0.69 and 0.66, respectively, indicating that most of the variation in the response data from our model is described by the fixed effects, with only an additional 3% associated with our points being clustered along IMBCR survey routes.

We also examined the variables of importance from the Random Forest model, and the top 10 in decreasing order from most important to least important were: proportion of trees with a height of 16–20 m, average slope, average elevation, proportion of Douglas fir (*Pseudotsuga menziesii*) forest and woodland, proportion of trees with a height of 21–25 m, proportion of montane-foothill deciduous shrubland, proportion of montane mixed conifer forest, proportion of area with 30–39% shrub canopy cover, proportion of trees with a height of 1–5 m, and proportion of area with big sagebrush steppe (Fig. 4). We evaluated the marginal effect of a variable on the random forest's

Table 2. Definitions for variables in the final resource selection function habitat model for predicting Dusky Grouse (*Dendragapus obscurus*) occurrence. We calculated the mean statistic or proportion of a habitat characteristic within a 250-m radii circle centered on the survey point. Relationship form represents the marginal relationship between a variable and probability of occurrence and is evaluated using a univariate model examining potential linear, quadratic, and pseudo-linear threshold relationships using linear equations.

Variable	EVT code	Definition	Vegetation Physiognomy	Relationship Form	Direction
Distance to Road	N/A	Average distance to nearest road (km)	N/A	linear	negative
Slope	N/A	Average slope	N/A	nonlinear: pseudo-linear threshold	positive
Distance to stream	N/A	Average distance to nearest stream (km)	N/A	nonlinear: quadratic	positive, then negative
Foothill Conifer Wooded Steppe	EVT 7165	Proportion of northern rocky mountain foothill conifer wooded steppe	Conifer	nonlinear: quadratic	positive, then negative
Montane Sagebrush Steppe	EVT 7126	Proportion of inter-mountain basins montane sagebrush steppe	Shrubland	nonlinear: pseudo-linear threshold	positive
Tree Height 1–5m	N/A	Proportion of trees with a height of 1–5m	N/A	nonlinear: pseudo-linear threshold	negative
Tree Height 16–20m	N/A	Proportion of trees with a height of 16–20m	N/A	nonlinear: pseudo-linear threshold	positive

Table 3. Slope estimates for all terms in the final resource selection function habitat model for predicting Dusky Grouse (*Dendragapus obscurus*) occurrence with 95% confidence intervals.

Variable	Estimated slope (β_i)	Lower 95% CI	Upper 95% CI
Distance to Road	-0.31	-0.58	-0.03
Distance to Stream	7.40	3.26	11.53
Distance to Stream^2	-7.49	-12.79	-2.19
Foothill Conifer Wooded Steppe	216.70	152.32	281.03
Foothill Conifer Wooded Steppe^2	-5557.00	-5826.86	-5287.36
ln(Slope + 0.001)	1.03	0.52	1.54
ln(Montane Sagebrush Steppe + 0.001)	0.16	0.05	0.27
ln(Tree Height 1–5m + 0.001)	-0.68	-1.14	-0.22
ln(Tree Height 16–20m + 0.001)	0.32	0.15	0.48

predictions using partial dependency plots. Proportion of trees with a height of 16–20 m, slope, elevation, proportion of Douglas fir forest and woodland, proportion of trees with a height of 21–25m, proportion of montane-foothill deciduous shrubland, proportion of montane mixed conifer forest, proportion of 30–39% canopy shrub cover, and proportion of big sagebrush steppe all have positive nonlinear relationships, while proportion of trees with a height of 1–5 m and proportion of 30–39% canopy herb cover have negative nonlinear relationships (Fig. 5).

Model evaluation

The average AUC values for the RSF and RF models were 0.89 (95% CI: 0.85–0.93) and 0.87 (95% CI: 0.83–0.92), respectively (Fig. 6). The RSF model correctly classified 150/193 (78%) of the independent grouse locations into the medium-high and high categories of relative probability of use. Linear regression produced an intercept close to zero (95% CI: -0.40–0.18), a slope of 1.55 (95% CI: 0.45, 2.65), and a high R^2 value (0.87), indicating high predictive accuracy (Fig. 5). The RF model also had high predictive accuracy, with 94% of the independently detected Dusky Grouse locations correctly classified into the medium-high and high category bins of relative probability of use. Linear regression produced an intercept close to zero (95% CI: -0.30, 0.23), a slope of 1.17 (95% CI: 0.30, 2.06), and a R^2 value of 0.86 (Fig. 7).

Because both models had similarly high predictive accuracy, we used the 60–80% quantile bin (medium-high relative probability of use) as a threshold to create two binary maps to obtain ensembled estimates of spatially-explicit habitat suitability for Dusky Grouse in Montana. The percent of locations correctly classified as habitat from the MFWP data was highest in Region 1 for the RSF model, in Regions 1, 2, and 4 for the RF model and ensemble predictions, and was lowest in Region 5 for all three predictions (Table 4).

Calculating potential Dusky Grouse habitat in Montana

Slight differences existed in the predicted habitat between the RSF and RF models, with the RSF model having more conservative estimates and the RF model predicting higher amounts of habitat across the majority of the regions (exception is MFWP Region 3; Table 5, Fig. 8). Despite the slight differences in the amounts of predicted habitat (a 7% difference in total habitat and non-habitat predicted), there was a high amount of agreement (93%) between the RSF and RF models on whether an area was predicted habitat or non-habitat. Across both models, MFWP Regions 1, 2, and 3 had the highest amounts of potential Dusky Grouse habitat (Table 5, Fig. 8). Using our ensembled map we predicted 83,160–109,125 km² in Montana to be potential Dusky Grouse habitat (Table 5) with the majority (99%) of the habitat occurring in MFWP regions 1–5 (Fig. 8). Of the predicted habitat for the ensembled map, 76% was predicted to be high probability of being habitat and 24% was predicted to be medium-high probability of being habitat (Table 5, Fig. 8).

Fig. 3. Predicted relative probability of use for covariates in the resource selection function (RSF) model with 95% confidence intervals (dashed lines) while all other covariates are held at their average value.

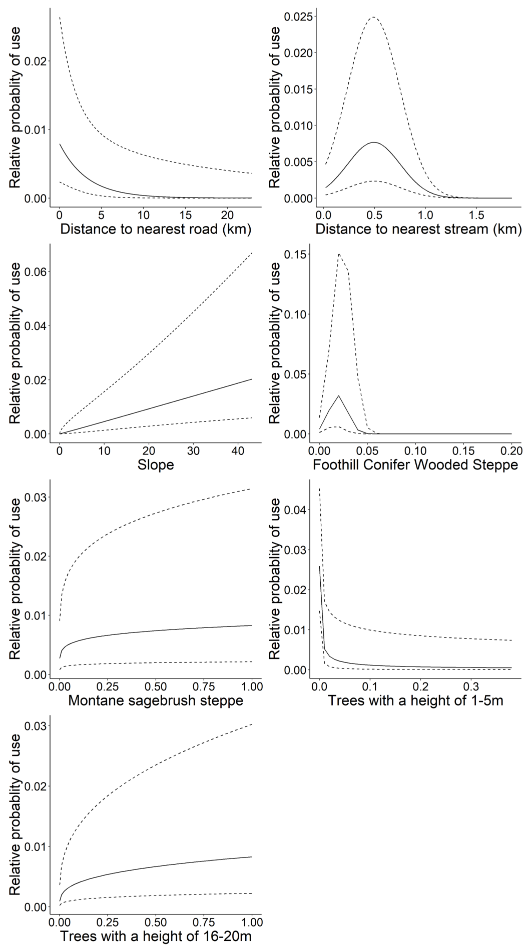


Fig. 4. Variable importance plot for the top 10 important variables from the random forest model. Variable importance was calculated as the impact of removing a variable on the model or mean decrease in accuracy.

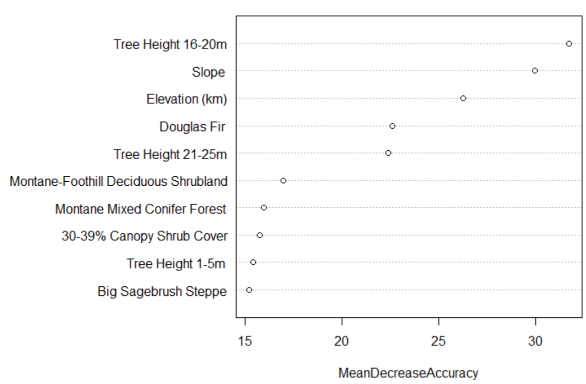


Fig. 5. Partial dependency plots for the variables of greatest importance for fitting the random forest model to evaluate the marginal effect of a variable on the random forest's predictions.

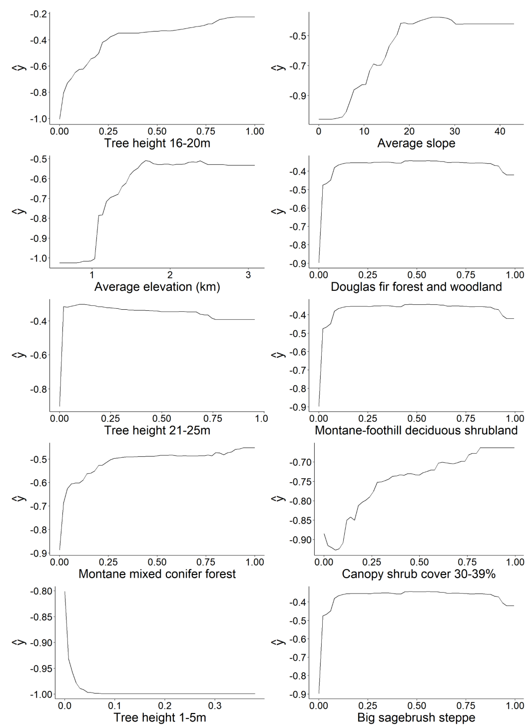


Fig. 6. Histogram of the Area Under Curve (AUC) values from the repeated k-fold cross validation for the resource selection model (top) and random forest model (bottom). Average AUC for the resource selection function model was 0.89 (95% CI: 0.85–0.93) and for the random forest model was 0.87 (95% CI: 0.82, 0.92).

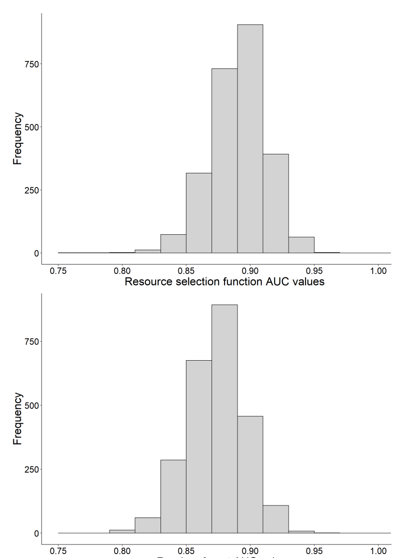
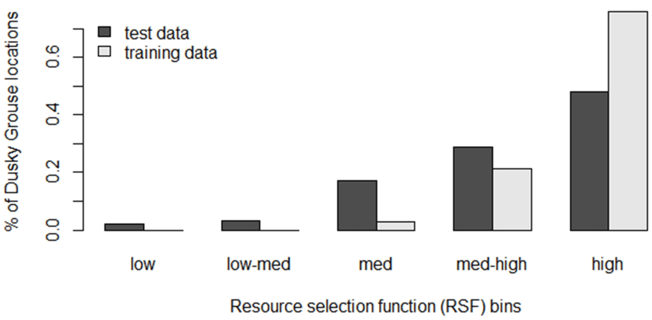
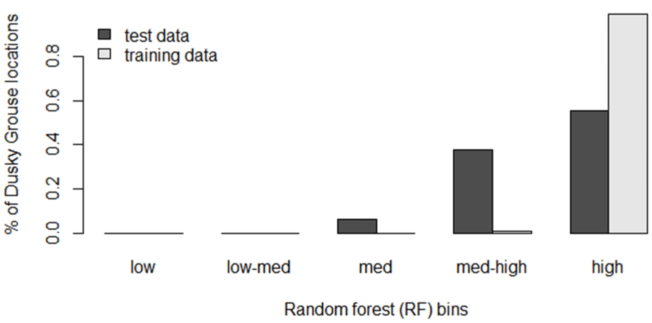


Table 4. Percent of simulated data correctly classified for all of Montana and each Montana Department of Fish, Wildlife & Parks region for the independent dataset. Percent correctly classified is calculated with 95% confidence intervals for the three models: resource selection function model (RSF), random forest model (RF), and the ensemble model.

Area	RSF Model			RF Model			Ensemble Model		
	Percent	Lower 95% CI	Upper 95% CI	Percent	Lower 95% CI	Upper 95% CI	Percent	Lower 95% CI	Upper 95% CI
Montana	77.7	74.7	81.2	93.8	92.2	95.5	96.9	96.1	98.1
Region 1	96.2	95.0	100	100	100	100	100	100	100
Region 2	85.6	81.3	93.8	100	100	100	100	100	100
Region 3	87.2	83.8	91.2	94.3	92.6	97.1	96.5	95.6	98.5
Region 4	83.6	77.8	100	100	100	100	100	100	100
Region 5	45.9	37.8	51.4	87.3	83.8	91.9	93.7	91.9	97.3

Fig. 7. Proportion of Dusky Grouse (*Dendragapus obscurus*) locations in five bins of increasing relative probability of use values for resource selection function values (top) and random forest model values (bottom) that we used to train (n = 132) and test (n = 193; 1 location was outside Montana) the different models of predicted Dusky Grouse habitat. A good predictive model will assign most of the training and test Dusky Grouse locations to medium-high or high categories of predicted use.





DISCUSSION

We used IMBCR's dataset of Dusky Grouse observations and an ensemble modeling approach to evaluate habitat associations and create statewide predictions of Dusky Grouse habitat in Montana. Our RSF and RF models found somewhat different landscape metrics to be important for predicting habitat suitability, which resulted in slight differences in predicted habitat. However, both models had high predictive accuracy and predictions of potential Dusky Grouse habitat were generally similar among methods. By using an ensemble of models, created by summing 2 binary maps, we identified some uncertainty, that is areas of disagreement between models, in our predictions and created a robust prediction of habitat suitability for Dusky Grouse that can be used to inform habitat and population management programs in Montana.

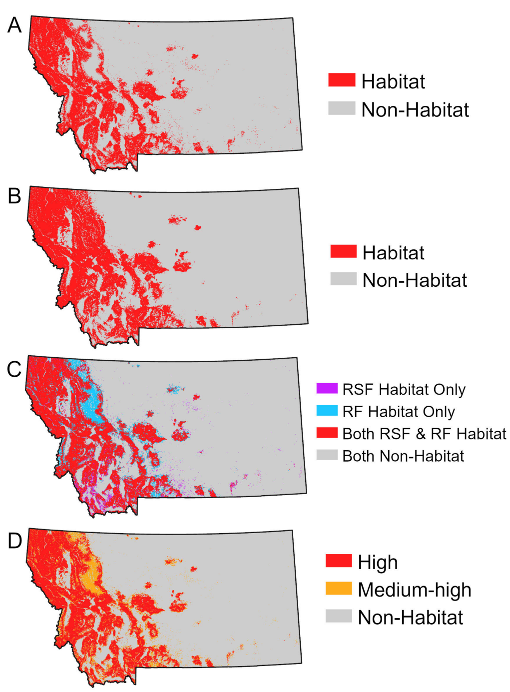
Agreement between our landscape-level evaluation of habitat associations and previously reported field-based habitat associations of Dusky Grouse supports the validity of our predictive habitat suitability models (Johnsgard 2016). Dusky Grouse were strongly associated with coniferous and mountainous forests, and the coverages of forest height classes were important predictors of relative habitat suitability during the breeding season in both models. In addition, we found relative habitat suitability increased with the coverage of mid-old growth coniferous forest (tree heights of 16–25 m; Cade and Hoffman 1993). Also consistent with previous field work, our RF model indicated strong positive associations of breeding Dusky Grouse with forests dominated by Douglas fir, lodgepole pine (*Pinus contorta*), and ponderosa pine (*Pinus ponderosa*; Marshall 1946, Martinka 1972, Cade and Hoffman 1990). During the reproductive season, Dusky Grouse have also been found to prefer more open forests compared to more closed forests during the winter (Stauffer and Peterson 1986). Support for partial coverages of three vegetation types (foothill conifer wooded steppe, montane foothill deciduous shrubland, and montane sagebrush steppe) suggest the importance of transitional vegetation types as Dusky Grouse migrate from high elevation forested habitat in the winter to open canopy and herbaceous nesting habitats (Mussehl 1960, 1963, Zwickel 1973, Johnsgard 2016).

A used:available RSF and a categorical RF are conceptually different models, where the RF is a discrete classifier and the used:available RSF is not (Breiman 2001, Manly et al. 2002). Therefore, as expected, the RSF and RF models had similar, but subtly different, spatial predictions of potential Dusky Grouse habitat. Generally, the RF model predicted more hectares of habitat than the RSF model. Discrepancies between model predictions (7% difference in predicted total habitat and non-habitat) were largely due to the difference in the importance of the effect of road proximity, with the largest differences in roadless areas (e.g., the Bob Marshall Wilderness). The top RSF model included an estimated negative linear effect of distance to road on the relative probability of use, whereas this effect was ranked second to last in importance by the RF model resulting in an extremely low impact on the RF's prediction. The supported positive effect of road proximity on Dusky Grouse use is not intuitive but may be

Table 5. Estimated area (km²) of potential Dusky Grouse habitat for Montana FWP administrative regions for the 3 predictive maps. The RSF and RF models are divided into a binary map of habitat and non-habitat based on the 60% quantile, while the ensemble map is based on an ensemble frequency histogram where consensus between the models on predicted habitat resulted in high probability of being habitat and areas of unagreed upon predicted habitat between the RSF and RF models resulted in medium-high probability of being habitat, and consensus between the models on predicted non-habitat resulted in non-habitat. Total habitat is the sum of the medium high and high probability of being habitat categories.

Region	RSF: Non-Habitat:	RSF: Habitat	RF: Non-Habitat	RF: Habitat	E: Non-Habitat	E: Med. High	E: High	E: Total Habitat
Region 1	9489	25045	5402	29133	4714	5463	24357	29821
Region 2	7112	20195	4498	22809	4073	3464	19770	23234
Region 3	22589	27509	22669	27429	19077	7104	23917	31021
Region 4	60602	10725	55435	15892	54751	6535	10041	16576
Region 5	40171	5456	39471	6157	38538	2566	4524	7089
Region 6	71873	581	71590	865	71459	544	451	995
Region 7	78732	351	78944	139	78694	289	100	390
Total	290570	89862	278008	102423	271306	25966	83160	109125

Fig. 8. Predicted Dusky Grouse (*Dendragapus obscurus*) habitat (red) for the resource selection function map (A) and random forest map (B). Areas of consensus and differences (C) in predicted Dusky Grouse habitat between the resource selection function model (RSF) and random forest (RF) models, where areas both models predict habitat are red, where only RSF predicted habitat are purple, areas where only RF predicted habitat are blue, and areas where both models predict non-habitat are gray. Predicted Dusky Grouse habitat for the ensemble model (D) where red represents high probability of being habitat (consistently predicted to be habitat), orange represents medium-high probability of being habitat (inconsistently predicted to be habitat), and gray represents non-habitat. MFWP administrative regions are delineated in gray (left top to bottom: Regions 1, 2, 3; center top to bottom: Regions 4, 5; and right top to bottom: 6, 7).



the result of how we treated the road layer used to estimate Euclidean distance to road. Because of incomplete attribute data in the available geospatial layer, we could not differentiate between unimproved roads such as low-traffic forest service roads or roads with seasonal closures, and paved roads, which included highways and other high-traffic roads. Most used points were closer in proximity to unimproved forest service roads with relatively light traffic and that often experience seasonal closures. Dusky Grouse males are known to select breeding territories on old logging roads (Martinka 1972), and this behavior with the lack of differentiation between forest service roads and other road types may have resulted in a positive effect of road proximity that may not hold true for higher-trafficked roads with non-native surfaces. Despite the areas of disagreement among models, the two models agreed 93% of the time, and 76% of the predicted habitat in the ensemble map was predicted by both the RSF and RF models, and statewide predictive accuracy of both holdout training and independent Dusky Grouse observations was high.

Transferability of model predictions to unsampled areas outside of Montana may be limited by regional variation in habitat relationships. For example, the coverage of trees with a height of 16–20 m was an important predictor of relative habitat suitability in Montana where this tree height corresponded to preferred conifer forests at elevations of 560–3401 m but may not be a good indicator of selection where this relationship does not occur (Araùjo and Guisan 2006, Randin et al. 2006, Heikkinen et al. 2012). At finer spatial extents (e.g., MFWP administrative regions), heterogeneity in landscape metric relationships may reduce the predictive accuracy of our statewide models in areas where training data sample sizes were small (i.e., where Dusky Grouse observations were few). Though we found the predictive accuracy of our individual models to be high (≥ 77%) for the state of Montana, predictive performance of our statewide models varied across MFWP administrative regions. For the RSF predictive map, we correctly classified 88% of the independent locations within MFWP regions 1–4, but only 46% were correctly classified for Region 5, an area with no presence locations within the training dataset. We found that despite the lack of observed grouse locations, our RF model had almost twice the predictive accuracy of the RSF model for Region 5 (87% vs 43%). Differences in the predictive performance of RSF and RF models within and outside of (e.g., extrapolated) study areas further justify an ensemble approach to species distribution modeling (Marmion et al. 2009, Duque-Lazo et al. 2016). By combining multiple models, ensemble models improve accuracy and predictive performance over individual models, depending upon the accuracy of the individual models used (Marmion et al. 2009, Stohlgren et al. 2010, Grenouillet et al. 2011). Indeed, predictive accuracy improved to 94% in Region 5 when ensembled predictions of habitat suitability were used, and overall accuracy for the entire state of Montana improved to 97%. Other studies have also found increased robustness and predictive power when using an ensemble of models (Marmion et al. 2009, Stohlgren et al. 2010, Latif et al. 2013, Fuller et al. 2018) and in our case, it did so despite relatively few presence locations in some administrative regions.

Species distribution models are often used to address many conservation and management objectives, including surveillance or monitoring, finding new populations of a species, and designating areas for conservation (Guisan and Thuiller 2005, Williams et al. 2009, Le Lay et al. 2010, Stohlgren et al. 2010, Crall et al. 2013, Guisan et al. 2013, Fuller et al. 2018, Sofaer et al. 2019). For all purposes, it is important that model predictions are accurate and identify areas of predictive uncertainty. By using an ensemble of models, we accurately and confidently predicted areas of habitat versus non-habitat for Dusky Grouse in Montana, even in administrative areas that had few to no presence locations in the training dataset.

CONCLUSION

Our spatially resolute and statewide predictions of relative habitat suitability derived from observations produced by the rigorous and randomized IMBCR survey program are (1) an advancement on previous ad hoc delineations of Dusky Grouse distribution based on incidental publicly reported observations, and (2) provide justifiable strata for prioritizing and planning Dusky Grouse population monitoring. Using multiple modeling techniques and an ensemble approach to prediction we were able to understand relative habitat use and predict potential Dusky Grouse habitat. Our results provide baseline information on Dusky Grouse habitats in Montana that can be used to inform conservation planning and future research. In addition, our predictive map can be useful for developing optimal sampling for long-term population monitoring. By identifying suitable survey areas, our study represents a first step in developing robust statewide population-level assessments of Dusky Grouse.

Acknowledgments:

Funding was provided by Montana Fish, Wildlife & Parks. We would like to thank Christian Meny, Cara Staab, and the Bird Conservancy of the Rockies for collecting and contributing observation data for dusky grouse. We would also like to thank Montana, Fish, Wildlife & Parks for contributing an independent dataset that we could use to verify our models. In addition, we would like to thank the Wildlife Habitat Ecology lab at Montana State University for their support.

LITERATURE CITED

Aldrich, J. W. 1963. Geographic orientation of American Tetraonidae. Journal of Wildlife Management 27:528-545. https://doi.org/10.2307/3798463

Aldridge, C. L., D. J. Saher, T. M. Childers, K. E. Stahlnecker, and Z. H. Bowen. 2012. Crucial nesting habitat for Gunnison Sage-Grouse: a spatially explicit hierarchical approach. Journal of Wildlife Management 76:391-406. https://doi.org/10.1002/jwmg.268

Araùjo, M. B., and A. Guisan. 2006. Five (or so) challenges for species distribution modeling. Journal of Biogeography 33:1677-1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Araùjo, M. B., and M. New. 2007. Ensemble forecasting of species distributions. Trends in Ecology and Evolution 22:42-47. https://doi.org/10.1016/j.tree.2006.09.010

Araùjo, M. B., R. J. Whittaker, R. J. Ladle, and M. Erhard. 2005. Reducing uncertainty in projections of extinction risk from climate change. Global Ecology and Biogeography 14:529-538. https://doi.org/10.1111/j.1466-822X.2005.00182.x

Bates, D., M. Machler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1):1-48. https://doi.org/10.18637/jss.v067.i01

Beyer, H. L. 2015. Geospatial modeling environment. Version 0.7.4.0.

Bohnett, E., D. Hulse, B. Ahmad, and T. Hoctor. 2020. Multi-level, multi-scale modeling and predictive mapping for jaguars in the Brazilian Pantanal. Open Journal of Ecology 10:243-263. https://doi.org/10.4236/oje.2020.105016

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling 157:281-300. https://doi.org/10.1016/S0304-3800(02)00200-4

Breiman, L. 2001. Random forests. Machine Learning 45:5-32. https://doi.org/10.1023/A:1010933404324

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.

Cade, B. S., and R. W. Hoffman. 1990. Winter use of Douglas-fir forests by Blue Grouse in Colorado. Journal of Wildlife Management 54:471-479. https://doi.org/10.2307/3809661

Cade, B. S., and R. W. Hoffman. 1993. Differential migration of Blue Grouse in Colorado. Auk 110:70-77.

Chan-McLeod, A., and F. L. Bunnell. 2003. Potential approaches to integrating silvicultural control of mountain pine beetle with wildlife and sustainable management objectives. Pages 267-277 in T. L. Shore, J. E. Brooks, J. E. Stone, editors. Mountain Pine Beetle Symposium: Challenges and Solutions, Keowna, British Columbia, Canada. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, Canada.

Crall, A. W., C. S. Jarnevich, B. Panke, N. Young, M. Renz, and J. Morisette. 2013. Using habitat suitability models to target invasive plant species surveys. Ecological Applications 23:60-72. https://doi.org/10.1890/12-0465.1

Dugger, K. M., F. Wager, R. G. Anthony, and G. S. Olson. 2005. The relationship between habitat characteristics and demographic performance of northern Spotted Owls in southern Oregon. Condor 107:863-878. https://doi.org/10.1093/condor/107.4.863

Duque-Lazo, J., H. van Gils, T. A. Groen, and R. M. Navarro-Cerrillo. 2016. Transferability of species distribution models: the case of Phytophthora cinnamomi in southwest Spain and southwest Australia. Ecological Modelling 320:62-70. https://doi.org/10.1016/j.ecolmodel.2015.09.019

Elith, J., and C. H. Graham. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. Ecography 32:66-77. https://doi.org/10.1111/j.1600-0587.2008.05505.x

Evans, J. S., and S. A. Cushman. 2009. Gradient modeling of conifer species using random forests. Landscape Ecology 24:673-683. https://doi.org/10.1007/s10980-009-9341-0

Evans, J. S., M. A. Murphy, Z. A. Holden, and S. A. Cushman. 2011. Modeling species distribution and change using random forest. Pages 139-159 in C. A. Drew, Y. F. Wiersma, and F. Huettmann, editors. Predictive species and habitat modeling in landscape ecology. Springer, New York, New York, USA. https://doi.org/10.1007/978-1-4419-7390-0_8

Farnsworth, S. Y. 2020. Forest grouse ecology and management in the Bear River Range northern Utah. Thesis. Utah State University, Logan, Utah, USA.

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. https://doi.org/10.1017/S0376892997000088

Franklin, A. B., D. R. Anderson, R. J. Gutiérrez, and K. P. Burnham. 2000. Climate, habitat quality, and fitness in northern Spotted Owl populations in northwestern California. Ecological Monographs 70:539-590. https://doi.org/10.1890/0012-9615(2000)070[0539:CHQAFI]2.0.CO;2

Fuller, L., M. Shewring, and F. M. Caryl. 2018. A novel method for targeting survey effort to identify new bat roosts using habitat suitability modeling. European Journal of Wildlife Research 64:31. https://doi.org/10.1007/s10344-018-1191-0

Grenouillet, G., L. Buisson, N. Casajus, and S. Lek. 2011. Ensemble modelling of species distribution: the effects of geographical and environmental ranges. Ecography 34:9-17. https://doi.org/10.1111/j.1600-0587.2010.06152.x

Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. Zimmerman. 2006. Using niche-based models to improve the sampling of rare species. Conservation Biology 20:501-511. https://doi.org/10.1111/j.1523-1739.2006.00354.x

Guisan, A., T. C. J. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157:89-100. https://doi.org/10.1016/S0304-3800(02)00204-1

Guisan, A., R. Tingley, J. B. Baumgartner, I. Naujokaitis-Lewis, P. R. Sutcliffe, A. I. Tulloch, T. J. Regan, L. Brotons, E. McDonald-Madden, C. Mantyka-Pringle, T. G. Martin, J. R. Rhodes, R. Maggini, S. A. Setterfield, J. Elith, M. W. Schwartz, B. A. Wintle, O. Broennimann, M. Austin, S. Ferrier, M. R. Kearney, H. P. Possingham, and Y. M. Buckley. 2013. Predicting species distributions for conservation decisions. Ecology Letters 16:1424-1435. https://doi.org/10.1111/ele.12189

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8:993-1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guisan, A., and N. E. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147-186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hanberry, B., H. S. He, and B. J. Palik. 2012. Pseudoabsence generation strategies for species distribution models. PLoS ONE 7(8):e44486. https://doi.org/10.1371/journal.pone.0044486

Harju, H. J. 1974. An analysis of some aspects of the ecology of Dusky Grouse. Thesis. University of Wyoming, Laramie, Wyoming, USA.

Heikkinen, R. K., M. Marmion, and M. Luoto. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography 35:276-288. https://doi.org/10.1111/j.1600-0587.2011.06999.x

Heinze, G., C. Wallisch, and D. Dunkler. 2018. Variable selection - a review and recommendations for the practicing statistician. Biometrical Journal 60:431-449. https://doi.org/10.1002/bimj.201700067

Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. Applied logistic regression. Wiley, New York, New York, USA. https://doi.org/10.1002/9781118548387

Johnsgard, P. A. 2016. The North American grouse: their biology and behavior. Zea Books, Lincoln, Nebraska, USA. https://doi.org/10.13014/K22Z13FS

Johnson, C. J., S. E. Nielsen, E. H. Merril, T. L. McDonald, and M. S. Boyce. 2006. Resource selection functions based on use-availability data: theoretical motivation and evaluating methods. Journal of Wildlife Management 70:347-357. https://doi.org/10.2193/0022-541X(2006)70[347:RSFBOU]2.0.CO;2

Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 61(1):65-71. https://doi.org/10.2307/1937156

Kuhn, M. 2008. Building predictive models in R using the caret package. Journal of Statistical Software 28(5). https://doi.org/10.18637/jss.v028.i05

Kuhn, M. 2019. The caret package. https://topepo.github.io/caret/index.html

Kuhn, M., and K. Johnson. 2013. Applied predictive modeling. Springer, New York, New York, USA. https://doi.org/10.1007/978-1-4614-6849-3

LANDFIRE. 2016a. LANDFIRE Existing Vegetation Canopy Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. https://www.landfire.gov/viewer/

LANDFIRE. 2016b. LANDFIRE Existing Vegetation Height Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. https://www.landfire.gov/viewer/

LANDFIRE. 2016c. LANDFIRE Existing Vegetation Type Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. https://www.landfire.gov/viewer/

LANDFIRE. 2019. LANDFIRE product descriptions with references. Geological Survey. U.S. Department of the Interior, Washington, D.C., USA.

LANDFIRE. 2020. LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA.

Latif, Q. S., V. A. Saab, J. G. Dudley, and J. P. Hollenbeck. 2013. Ensemble modeling to predict habitat suitability for a large-scale disturbance specialist. Ecology and Evolution 3:4348-4364. https://doi.org/10.1002/ece3.790

Le Lay, G., R. Engler, E. Franc, and A. Guisan. 2010. Prospective sampling based on model ensembles improves the detection of rare species. Ecography 33:1015-1027. https://doi.org/10.1111/j.1600-0587.2010.06338.x

Leipold, E. 2023. Developing a monitoring program for Dusky Grouse in Montana. Thesis. Montana State University, Bozeman, Montana, USA.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18-22.

Lütolf, M., F. Kienast, and A. Guisan. 2006. The ghost of past species occurrence: improving species distribution models for presence-only data. Journal of Applied Ecology 43:802-825. https://doi.org/10.1111/j.1365-2664.2006.01191.x

Manly, B. F. J., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies. Second edition. Kluwer Academic, New York, New York, USA.

Marmion, M., M. Parviainen, M. Luoto, R. K. Heikkinen, and W. Thuiller. 2009. Evaluation of consensus methods in predictive species distribution modeling. Diversity and Distributions 15:59-69. https://doi.org/10.1111/j.1472-4642.2008.00491.x

Marshall, W. 1946. Cover preferences, seasonal movements, and food habits of Richardson's Grouse and Ruffed Grouse in southern Idaho. Wilson Bulletin 1:42-52.

Martinka, R. R. 1972. Structural characteristics of Blue Grouse territories in southwestern Montana. Journal of Wildlife Management 36:498-510. https://doi.org/10.2307/3799081

McNew, L. B., A. J. Gregory, and B. K. Sandercock. 2013. Spatial heterogeneity in habitat selection: nest site selection by Greater Prairie-Chickens. Journal of Wildlife Management 77:791-801. https://doi.org/10.1002/jwmg.493

McNew, L. B., V. L. Winder, J. C. Pitman, and B. K. Sandercock. 2015. Alternative rangeland management strategies and the nesting ecology of Greater Prairie-Chickens. Rangeland Ecology & Management 68:298-304. https://doi.org/10.1016/j.rama.2015.03.009

Montana Spatial Data Infrastructure (MSDI). 2017. Transportation. Montana State Library, Helena, Montana, USA. http://geoinfo.msl.mt.gov/msdi/transportation/

Montana Spatial Data Infrastructure (MSDI). 2018. Hydrography. Montana State Library, Helena, Montana, USA. http://geoinfo.msl.mt.gov/msdi/hydrography

Mussehl, T. W. 1960. Blue Grouse production, movements, and populations in the Bridger Mountains, Montana. Journal of Wildlife Management 24:60-68. https://doi.org/10.2307/3797357

Mussehl, T. W. 1963. Blue Grouse brood cover selection and land-use implications. Journal of Wildlife Management 27:546-555. https://doi.org/10.2307/3798464

Pavlacky, D. C. J., J. A. Lukacs, R. C. Blakesley, D. S. Skokowsky, D. S. Klute, B. A. Hahn, V. J. Dreitz, T. L. George, and D. J. Hanni. 2017. A statistically rigorous sampling design to integrate avian monitoring and management within Bird Conservation Regions. PLoS ONE 12(10):e0185924. https://doi.org/10.1371/journal.pone.0185924

Pearson, R. G., W. Thuiller, M. B. Araùjo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson, and D. C. Lees. 2006. Model-based uncertainty in species range prediction. Journal of Biogeography 33:1704-1711. https://doi.org/10.1111/j.1365-2699.2006.01460.x

Picardi, S., T. Messmer, B. Crabb, M. Kohl, D. Dahlgren, N. Frey, R. Larsen, and R. Baxter. 2020. Predicting Greater Sage-Grouse habitat selection at the southern periphery of their range. Ecology and Evolution 10:13451-13463. https://doi.org/10.1002/ece3.6950

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Randin, C. F., T. Dirnböck, S. Dullinger, N. E. Zimmerman, M. Zappa, and A. Guisan. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography 33:1689-1703. https://doi.org/10.1111/j.1365-2699.2006.01466.x

Redfield, J. A. 1973. Demography and genetics in colonizing populations of Blue Grouse (*Dendragapus obscurus*). Evolution 27:576-592. https://doi.org/10.1111/j.1558-5646.1973.tb00707.x

Shoemaker, K. T., L. J. Heffelfinger, N. J. Jackson, M. E. Blum, T. Wasley, K. M. Stewart. 2018. A machine-learning approach for extending classical wildlife resource selection analyses. Ecology and Evolution 8:3556-3569. https://doi.org/10.1002/ece3.3936

Sofaer, H. R., C. S. Jarnevich, I. S. Pearse, R. L. Smyth, S. Auer, G. L. Cook, T. C. Edwards Jr., G. F. Guala, T. G. Howard, J. T. Morisette, and H. Hamilton. 2019. Development and delivery of species distribution models to inform decision-making. BioScience 69:544-557. https://doi.org/10.1093/biosci/biz045

Stohlgren, T. J., P. Ma, S. Kumar, M. Rocca, J. T. Morisette, C. S. Jarnevich, and N. Benson. 2010. Ensemble habitat mapping of invasive plant species. Risk Analysis 30:224-235. https://doi.org/10.1111/j.1539-6924.2009.01343.x

Stauffer, D. F., and S. R. Peterson. 1986. Seasonal microhabitat relationships of Blue Grouse in southeastern Idaho. Great Basin Naturalist 46:117-122.

U.S. Geological Survey. 2017. The national map. U.S. Geological Survey, Reston, Virginia, USA. https://nationalmap.gov/3DEP/3dep_prodserv.html

Williams, J. N., C. Seo, J. Thorne, J. K. Nelson, S. Erwin, J. M. O'Brien, and M. W. Schwartz. 2009. Using species distribution models to predict new occurrences for rare plants. Diversity and Distributions 15:565-576. https://doi.org/10.1111/j.1472-4642.2009.00567.x

Woiderski, B. J., N. E. Drilling, J. M. Timmer, M. F. McLaren, C. M. White, N. J. Van Lanen, D.C. Pavlacky Jr., and R. A. Sparks. 2018. Integrated monitoring in bird conservation regions (IMBCR): 2017 field season report. Bird Conservancy of the Rockies, Brighton, Colorado, USA.

Youtz, J., R. Goljani Amirkhiz, and J. K. Frey. 2022. Modeling the impact of climate change and wildfire on the Dusky Grouse (*Dendragapus obscurus*) in the American Southwest: implications for conservation. Avian Conservation and Ecology 17(1):35. https://doi.org/10.5751/ACE-02222-170135

Zurr, A. F. 2009. Mixed effects models and extensions in ecology with R. Springer, New York, New York, USA. https://doi.org/10.1007/978-0-387-87458-6

Zweig, M. H., and G. Campbell. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry 39:561-577. https://doi.org/10.1093/clinchem/39.4.561

Zwickel, F. C. 1973. Dispersion of female Blue Grouse during the brood season. Condor 75:114-119. https://doi.org/10.2307/1366542

Zwickel, F. C., and J. F. Bendell. 2004. Blue Grouse: their biology and natural history. NRC Research, Ottawa, Ontario, Canada.

Editor-in-Chief: Alexander L. Bond
Subject Editor: John R. Sauer

Sponsored by the Society of Canadian Ornithologists and Birds Canada

Parrainée par la Société des ornithologistes du Canada et Oiseaux Canada

Appendix 1. Description of habitat variables (LANDFIRE 2016a, b, c).

Variable	Source	EVT Code (Ecological Systems)	Statistic	Vegetation Physiognomy
North (N) Facing Aspect	DEM	N/A	Proportion	N/A
Northeast (NE) Facing Aspect	DEM	N/A	Proportion	N/A
East (E) Facing Aspect	DEM	N/A	Proportion	N/A
Southeast (SE) Facing Aspect	DEM	N/A	Proportion	N/A
South (S) Facing Aspect	DEM	N/A	Proportion	N/A
Southwest (SW) Facing Aspect	DEM	N/A	Proportion	N/A
West (W) Facing Aspect	DEM	N/A	Proportion	N/A
Northwest (NW) Facing Aspect	DEM	N/A	Proportion	N/A
Flat Aspect	DEM	N/A	Proportion	N/A
Slope	DEM	N/A	mean	N/A
Elevation (km)	DEM	N/A	mean	N/A
Distance (km) to Nearest Road	MSDI	N/A	mean	N/A
Distance (km) to Nearest Stream	MSDI	N/A	mean	N/A
Distance (km) to Forest Edge From Outside of the Forest	EVT	N/A	mean	N/A
Distance (km) to Forest Edge From Inside of the Forest	EVT	N/A	mean	N/A
Northern Rocky Mountain Western Larch Savanna	EVT	7010	Proportion	Conifer
Rocky Mountain Aspen Forest and Woodland	EVT	7011	Proportion	Hardwood
Northern Rocky Mountain Dry-Mesic Montane Mixed Conifer Forest	EVT	7045	Proportion	Conifer
Northern Rocky Mountain Subalpine Woodland and Parkland	EVT	7046	Proportion	Conifer
Northern Rocky Mountain Mesic Montane Mixed Conifer Forest	EVT	7047	Proportion	Conifer
Rocky Mountain Foothill Limber Pine-Juniper Woodland	EVT	7049	Proportion	Conifer
Rocky Mountain Lodgepole Pine Forest	EVT	7050	Proportion	Conifer
Northern Rocky Mountain Ponderosa Pine Woodland and Savanna	EVT	7053	Proportion	Conifer
Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	EVT	7055	Proportion	Conifer
Rocky Mountain Subalpine Mesic-Wet Spruce-Fir Forest and Woodland	EVT	7056	Proportion	Conifer
Inter-Mountain Basins Big Sagebrush Shrubland	EVT	7080	Proportion	Shrubland
Northern Rocky Mountain Montane-Foothill Deciduous Shrubland	EVT	7106	Proportion	Shrubland

Inter-Mountain Basins Big Sagebrush Steppe	EVT	7125	Proportion	Shrubland
Inter-Mountain Basins Montane Sagebrush Steppe	EVT	7126	Proportion	Shrubland
Northern Rocky Mountain Lower Montane-Foothill-Valley Grassland	EVT	7139	Proportion	Grassland
Northern Rocky Mountain Subalpine-Upper Montane Grassland	EVT	7140	Proportion	Grassland
Rocky Mountain Subalpine-Montane Mesic Meadow	EVT	7145	Proportion	Grassland
Northern Rocky Mountain Foothill Conifer Wooded Steppe	EVT	7165	Proportion	Conifer
Middle Rocky Mountain Montane Douglas-fir Forest and Woodland	EVT	7166	Proportion	Conifer
Northern Rocky Mountain Subalpine Deciduous Shrubland	EVT	7169	Proportion	Shrubland
Recently Logged-Herb and Grass Cover	EVT	7191	Proportion	Grassland
Recently Logged-Shrub Cover	EVT	7192	Proportion	Shrubland
Recently Logged-Tree Cover	EVT	7193	Proportion	Conifer
Recently Burned-Herb and Grass Cover	EVT	7195	Proportion	Grassland
Recently Burned-Tree Cover	EVT	7197	Proportion	Conifer
Open Water	EVT	7292	Proportion	Open Water
Developed-Roads	EVT	7299	Proportion	Developed-Roads
Western Cool Temperate Pasture and Hayland	EVT	7967	Proportion	Agricultural
Western Cool Temperate Wheat	EVT	7968	Proportion	Agricultural
Northern Rocky Mountain Lower Montane Riparian Woodland	EVT	9012	Proportion	Riparian
Rocky Mountain Alpine-Montane Wet Meadow	EVT	9017	Proportion	Riparian
Rocky Mountain Cliff Canyon and Massive Bedrock	EVT	9018	Proportion	Sparsely Vegetated
Rocky Mountain Subalpine-Montane Riparian Woodland	EVT	9022	Proportion	Riparian
Interior Western North American Temperate Ruderal Shrubland	EVT	9328	Proportion	Exotic Tree-Shrub
Northern Rocky Mountain Lower Montane Riparian Shrubland	EVT	9512	Proportion	Riparian
Interior Western North American Temperate Ruderal Grassland	EVT	9828	Proportion	Exotic Herbaceous
no vegetation present	EVC	N/A	Proportion	N/A
canopy cover from sparse vegetation	EVC	N/A	Proportion	N/A
canopy cover from agricultural crops	EVC	N/A	Proportion	N/A
developed areas	EVC	N/A	Proportion	N/A
Tree Canopy Cover 10-19%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 20-29%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 30-39%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 40-49%	EVC	N/A	Proportion	N/A

Tree Canopy Cover 50-59%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 60-69%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 70-79%	EVC	N/A	Proportion	N/A
Tree Canopy Cover 80-88%	EVC	N/A	Proportion	N/A
Shrub Canopy Cover 10-19%	EVC	N/A	Proportion	N/A
Shrub Canopy Cover 20-29%	EVC	N/A	Proportion	N/A
Shrub Canopy Cover 30-39%	EVC	N/A	Proportion	N/A
Shrub Canopy Cover 40-49%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 10-19%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 20-29%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 30-39%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 40-49%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 50-59%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 60-69%	EVC	N/A	Proportion	N/A
Herb Canopy Cover 70-79%	EVC	N/A	Proportion	N/A
Tree Height 1-5m	EVH	N/A	Proportion	N/A
Tree Height 6-10m	EVH	N/A	Proportion	N/A
Tree Height 11-15m	EVH	N/A	Proportion	N/A
Tree Height 16-20m	EVH	N/A	Proportion	N/A
Tree Height 21-26m	EVH	N/A	Proportion	N/A
Shrub Height 0.1-0.5m	EVH	N/A	Proportion	N/A
Shrub Height shrubs 0.6-1m	EVH	N/A	Proportion	N/A
Shrub Height 1.1-1.5m	EVH	N/A	Proportion	N/A
Shrub Height 1.6-2.0m	EVH	N/A	Proportion	N/A
Herb Height 0.1-0.5m	EVH	N/A	Proportion	N/A
Herb Height 0.6-1m tall	EVH	N/A	Proportion	N/A

Literature Cited

LANDFIREa. 2016. LANDFIRE Existing Vegetation Canopy Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. Accessed September 2020 at https://www.landfire.gov/viewer/.

LANDFIREb. 2016. LANDFIRE Existing Vegetation Height Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. Accessed September 2020 at https://www.landfire.gov/viewer/.

LANDFIRE. 2016c. LANDFIRE Existing Vegetation Type Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. Accessed September 2020 at https://www.landfire.gov/viewer/.

Appendix 2. Description of variables used to create a forest layer for Montana (LANDFIRE 2016).

EVT code (ecological systems)	Existing Vegetation Type (ecological systems name)	Vegetation Physiognomy	Collapsed Vegetation Type Name
7010	Northern Rocky Mountain Western Larch Savanna	Conifer	Western Larch Forest and Woodland
7045	Northern Rocky Mountain Dry-Mesic Montane Mixed Conifer Forest	Conifer	Douglas-fir-Ponderosa Pine-Lodgepole Pine Forest and Woodland
7046	Northern Rocky Mountain Subalpine Woodland and Parkland	Conifer	Subalpine Woodland and Parkland
7047	Northern Rocky Mountain Mesic Montane Mixed Conifer Forest	Conifer	Douglas-fir-Grand Fir-White Fir Forest and Woodland
7049	Rocky Mountain Foothill Limber Pine-Juniper Woodland	Conifer	Limber Pine Woodland
7050	Rocky Mountain Lodgepole Pine Forest	Conifer	Lodgepole Pine Forest and Woodland
7053	Northern Rocky Mountain Ponderosa Pine Woodland and Savanna	Conifer	Ponderosa Pine Forest, Woodland and Savanna
7055	Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	Conifer	Spruce-Fir Forest and Woodland
7056	Rocky Mountain Subalpine Mesic-Wet Spruce-Fir Forest and Woodland	Conifer	Spruce-Fir Forest and Woodland
7057	Rocky Mountain Subalpine-Montane Limber-Bristlecone Pine Woodland	Conifer	Limber Pine Woodland
7062	Inter-Mountain Basins Curl-leaf Mountain Mahogany Woodland	Conifer	Mountain Mahogany Woodland and Shrubland
7165	Northern Rocky Mountain Foothill Conifer Wooded Steppe	Conifer	Douglas-fir Forest and Woodland
7166	Middle Rocky Mountain Montane Douglas-fir Forest and Woodland	Conifer	Douglas-fir Forest and Woodland
7167	Rocky Mountain Poor-Site Lodgepole Pine Forest	Conifer	Lodgepole Pine Forest and Woodland
7179	Northwestern Great Plains-Black Hills Ponderosa Pine Woodland and Savanna	Conifer	Ponderosa Pine Forest, Woodland and Savanna
7193	Recently Logged-Tree Cover	Conifer	Transitional Forest Vegetation
7197	Recently Burned-Tree Cover	Conifer	Transitional Forest Vegetation
7200	Recently Disturbed Other-Tree Cover	Conifer	Transitional Forest Vegetation
7061	Inter-Mountain Basins Aspen-Mixed Conifer Forest and Woodland	Conifer-Hardwood	Aspen-Mixed Conifer Forest and Woodland
7009	Northwestern Great Plains Aspen Forest and Parkland	Hardwood	Aspen Forest, Woodland, and Parkland
7011	Rocky Mountain Aspen Forest and Woodland	Hardwood	Aspen Forest, Woodland, and Parkland
7161	Northern Rocky Mountain Conifer Swamp	Riparian	Spruce-Fir Forest and Woodland
9019	Rocky Mountain Lower Montane-Foothill Riparian Woodland	Riparian	Western Riparian Woodland and Shrubland
9022	Rocky Mountain Subalpine-Montane Riparian Woodland	Riparian	Western Riparian Woodland and Shrubland

Literature Cited

LANDFIRE. 2016. LANDFIRE Existing Vegetation Canopy Layer, LANDFIRE 2016 Remap (LF 2.0.0). Geological Survey. U.S. Department of the Interior, Washington, D.C., USA. Accessed September 2020 at https://www.landfire.gov/viewer/.

Appendix 3. Uncertainty maps for the RSF model. Map A (top) represents a map created using the high confidence intervals for the RSF's beta estimates and then converted to a binary map using the same threshold as used when creating the RSF binary map. Map B represents a map created using the low confidence intervals for the RSF's beta estimates, following the same methodology used to create map A.