IDAHO DEPARTMENT OF FISH AND GAME

Ed Schriever, Director

Project F19AF00802

Northern Idaho Ground Squirrel Occupancy Modeling
Cooperative Endangered Species Conservation

Interim and Final Report

Performance Period
July 1, 2019 to December 31, 2020

Compiled and edited by: Diane Evans Mack, Sonya Knetter, and Leona Svancara

September 2020
Boise, Idaho

Idaho Department of Fish and Game (IDFG) adheres to all applicable state and federal laws and regulations related to discrimination on the basis of race, color, national origin, age, gender, disability or veteran's status. If you feel you have been discriminated against in any program, activity, or facility of IDFG, or if you desire further information, please write to: Idaho Department of Fish and Game, PO Box 25, Boise, ID 83707 or US Fish and Wildlife Service, Division of Wildlife and Sport Fish Restoration Program, 5275 Leesburg Pike, MS: WSFR, Falls Church, VA 22041-3803, Telephone: (703) 358-2156.

Findings in this report are preliminary in nature and not for publication without permission of the Director of the Idaho Department of Fish and Game.

Please note that IDFG databases containing this information are dynamic. Records are added, deleted, and/or edited on a frequent basis. This information was current as of August 2020. Raw data do not have the benefit of interpretation or synthesis by IDFG.

IDFG requests that you direct any requests for this information to us rather than forwarding this information to third parties.

This publication will be made available in alternative formats upon request. Please contact IDFG for assistance.

ESA TRADITIONAL SECTION 6
INTERIM AND FINAL PERFORMANCE REPORT

1. **State:** Idaho

   Grant number: F19AF00802

   Grant name: Northern Idaho Ground Squirrel Occupancy Modeling

2. **Report Period:** July 1, 2019 to June 30, 2020

   Report due date: September 28, 2020

3. **Location of work:** Applies to northern Idaho ground squirrel habitat in Adams and Valley counties

4. **Objectives**

   This project will develop a robust occupancy model for long-term population monitoring. Objectives are to:
   a) compile spatial layers for a suite of environmental variables
   b) test a suite of environmental variables in occupancy models to identify which provide the most robust prediction of occurrence based on our annual northern Idaho ground squirrel (NIDGS) surveys
   c) identify the most parsimonious model and incorporate into Standard Operating Procedures for annual NIDGS data analysis

5. **If the work in this grant was part of a larger undertaking with other components and funding, present a brief overview of the larger activity and the role of this project.**

   This project is part of the overall recovery program for the northern Idaho ground squirrel (*Urocitellus brunneus*; NIDGS) and, more specifically, an integral part of monitoring changes in the abundance and distribution of NIDGS. Partners to the recovery effort include the U.S. Fish and Wildlife Service, Payette National Forest, Idaho Cooperative Fish and Wildlife Research Unit at the University of Idaho, species experts, and private landowners. Annual population monitoring focuses on 2 metrics: an estimate of abundance and an estimate of the probability of occupancy (i.e., spatial occurrence; Evans Mack et al. 2013). The strength of the occupancy modeling component lies in the use of covariates to allow the prediction of occupancy across the NIDGS range using relationships of those covariates to where NIDGS were detected. This project was a robust exploration of a suite of potential covariates (environmental and behavioral metrics) that potentially influence where NIDGS occur, and ultimately identified a list of variables that will be used each year in occupancy models to track distribution and assess progress toward recovery (U.S. Fish and Wildlife Service 2003).

   This project provides the foundation to a follow-up study that will model future NIDGS distribution under different climate projections (i.e., 2 emission scenarios). The follow-up work is funded by the U.S. Fish and Wildlife Service Idaho Field Office. The current model and projected future distribution models will contribute to a Species Status Assessment anticipated in 2021, and to a subsequent revision of the Recovery Plan.

6. Describe how the objectives were met.

During this reporting period, we used Grant F19AF00802 to model potential NIDGS habitat during the active season (approximately March to September) across the species' range (Figure 1). We used the modeling program Maxent (Phillips et al. 2017, Phillips et al. 2006, Phillips and Dudík 2008), which employs a set of occurrence locations (observations) and predictor variables (GIS layers) related to habitat suitability to characterize environmental conditions at occupied sites. It then compares that niche to environmental conditions that are available across the study area, resulting in a model that reflects the relative likelihood of species presence based on the degree of similarity to occupied sites.

Observations. We compiled 25,069 verified, trusted, and spatially accurate occurrence records of NIDGS stemming from formal surveys and incidental sightings from 2001–2019 (IDFG unpublished data). To reduce sampling bias, we randomly subsampled observations with a minimum distance of 90 m using SDMtoolbox (Brown et al. 2018, Brown et al. 2017), resulting in 1,543 presence points used in modeling (Figure 1).

Environmental Variables. To predict habitat suitability, we selected a suite of variables reflective of land cover, topography, climate, phenology, and soil properties, all of which were assumed *a priori* to influence NIDGS distribution (Table 1). We manipulated all rasters in ArcGIS 10.6.1 (ESRI 2017), ensured spatial data were in a common coordinate system, spatial resolution, and extent, and then exported variables as ASCII files for input into program Maxent 3.4.1 (Phillips et al. 2017) and R 3.5.2 (R Core Team 2018). We constructed a pairwise correlation matrix for all variables to note which were highly correlated, then explored the relationships between species locations and predictor variables, checking that each variable was a plausible predictor (i.e., frequency distribution at presence locations is not constant across environmental gradients).

Model Development. To characterize environmental conditions across our study, we randomly generated 15,000 background points (pseudoabsences) that were >270 m from presence points, >270 m apart, and outside of waterbodies, maintaining a 1:10 ratio of presence to background points (Barbet-Massin et al. 2012). We extracted the values of all variables to both presence and background points, ran Maxent in samples-with-data (SWD) format, and selected the Cloglog output, which places values on a probability scale between 0 and 1.

We optimized 2 Maxent parameters – the regularization multiplier and feature types – using the R package enmSdm v0.3.4.6 (Smith 2019). The regularization parameter imposes a penalty for designing overly-complex or over-fit models. Instead of using the default program value of 1, we tested a range of regularization multipliers from 0.5 to 20. Feature types are mathematical transformations of the covariates used in a model (i.e., linear, quadratic, product, hinge, and threshold) that allow complex relationships to be modeled (Merow et al. 2013, Elith et al. 2011). We selected the best performing combination of the regularization multiplier of a combination of feature types based on AICc (Warren and Seifort 2011) and constructed a full model inclusive of all 51 variables. We tested the model using 5-fold cross-validation, which means we ran the model 5 times, each time Maxent used 4-folds (80% of the presence points) to train the model, and 1 held out fold (20%) to test the model using the area under the receiver operating characteristics curve (AUC). AUC is a direct measure of a model's ability to discriminate the relative probability of a point being a presence (1) (i.e., sensitivity) or background location (0) (i.e., specificity). We used jackknifing to measure the importance of each variable to the resulting model and ranked variables based on their permutation importance across the 5 model runs. We then removed variables that contributed very little to the overall prediction (<1% permutation importance) and highly correlated variables (Pearson's correlation >|0.8|), tossing out those with smaller permutation importance values. With a reduced set of variables, we then repeated the process of model optimization, construction, variable ranking, and reduction until there were no highly correlated variables and all variables achieved at least 1% permutation importance. The final (reduced) model represented the average of 5 model runs using the optimized model parameters and the most important variables. Lastly, we categorized the final model into 3 habitat suitability classes (non-habitat, suitable, highly suitable) by applying thresholds to the continuous model probability values.

Results. Our model accurately predicted potential NIDGS habitat (Figure 2a, 2b) with an AUC of 0.953, which was based on 11 variables, a regularization multiplier of 1.0, and linear, product, and hinge feature types (Table 2). The most influential variable was precipitation of the driest month (bio14), followed by temperature seasonality (bio4), minimum temperature of the coldest month (bio4), tree canopy cover (tree), and percentage of silt from 30-100 cm. Partial effects plots are included in Figure 3, which show how the predicted probability of presence changes as each environmental variable is varied.

We applied 2 Maxent-generated thresholds to categorize habitat into 3 suitability classes, namely non-habitat, suitable, and highly suitable habitat (Figure 2b), which provides a more practical visualization/version for conservation management applications. The lowest threshold (0.04) set the limit between non-habitat and suitable habitat, which balanced training omission with predicted area. While it may over-predict habitat in some areas, we did not want to risk missing any population areas. We selected the maximum test sensitivity plus specificity threshold (0.16) for the limit between suitable and highly suitable habitat. This upper suitability class encompassed 97.5% of all NIDGS presences. Selecting appropriate threshold is somewhat of a subjective process. The appropriate threshold will depend on the model application and necessarily involves balancing the costs of omission (i.e., missing habitat) and commission (i.e., over-prediction).

Model caveats and applications. This model estimates the likelihood of suitable habitat, but does not measure directly the probability of occurrence. Other factors, such as the proximity to occupied habitat, intraspecific competition (e.g., with the Columbian ground squirrel [*Urocitellus columbianus*]), disease, and stochastic events may influence occupancy, as could other fine-scale habitat quality factors that either were not considered or not well represented by remotely-sensed data, particularly if it cannot be resolved in 30 m data.

We built this model with the intent to inform future occupancy surveys, including the potential to survey novel areas to help refine our knowledge of occupied habitat and validate the model results.

Table 1. Environmental variables used to model potential northern Idaho ground squirrel habitat.

Type	Variable	Description	Data Source (resolution); Citation
Land cover	shrub	Shrub cover (%)	National Land Cover Database (NLCD) 2016 Shrubland Fractional Components (30m); USGS (2019*a*)
	sage	Sagebrush cover (%)	
	herb	Herbaceous cover (%)	
	tree	Tree canopy cover (%)	NLCD 2016 USFS Tree Canopy Cover (CONUS) (30m); USGS (2019*b*)
	NShrubHerb	Proportion of 3 x 3 neighborhood mapped as shrub/scrub or herbaceous; proxy for "meadow habitat"	NLCD 2016 Land Cover (CONUS) (30m); USGS (2019*c*)
	landcov	Land cover type (categorical)	
Topography	elev	Mean elevation	1 Arc-second Digital Elevation Models (DEMs) (10m). HLI integrates solar radiation, slope, and aspect and was calculated using the Geomorphology and Gradient Metrics Toolbox (Evans et al. 2016) and a DEM; USGS (2017)
	heat	Head Load Index (HLI)	
Climate	bio1	Annual mean temperature (°C)	Monthly temperature data (30-year normals) from 1981–2010 (250m); Holden et al. (2015)
	bio2	Mean diurnal range (mean of monthly (max temp - min temp))	
	bio3	Isothermality (bio2/bio7) (* 100)	
	bio4	Temperature seasonality (standard deviation *100)	
	bio5	Max temperature of warmest month	
	bio6	Min temperature of coldest month	
	bio7	Temperature annual range (bio5-bio6)	
	bio8	Mean temperature of the wettest quarter	
	bio9	Mean temperature of driest quarter	
	bio10	Mean temperature of warmest quarter	
	bio11	Mean temperature of coldest quarter	
	bio12	Total (annual) precipitation (mm)	Monthly precipitation data (30-year normals) from 1981–2010 (800m); PRISM Climate Group (2012)
	bio13	Precipitation of wettest month	
	bio14	Precipitation of driest month	
	bio15	Precipitation seasonality (coefficient of variation)	
	bio16	Precipitation of wettest quarter	
	bio17	Precipitation of driest quarter	
	bio18	Precipitation of warmest quarter	
	bio19	Precipitation of coldest quarter	
	bio1_0309	Mean temperature, March - Sept	
	bio2_0309	Mean diurnal range, March - Sept	
	bio3_0309	Isothermality, March - Sept (bio2/bio7) (* 100)	
	bio4_0309	Temperature seasonality, March - Sept	
	bio7_0309	Temperature range, March - Sept	
	bio12_0309	Precipitation, March - Sept	
	bio15_0309	Precipitation seasonality (coefficient of variation)	

Table 1. cont.

Type	Variable	Description	Data Source (resolution); Citation
Phenology	SOST	Start of Season - Time	C6 Aqua Western U.S. 250 m eMODIS Remote Sensing Phenology Data, averaged from 2003–2018; USGS (2018)
	MAXT	Time of maximum photosynthetic in the canopy	
	AMP	Amplitude - maximum increase in canopy photosynthetic activity above the baseline	
Soil Properties	sand_030	Sand percentage in first 30 cm	POLARIS soils data (30m); Chaney et al. (2016)
	sand_30100	Sand percentage from 30-100 cm	
	silt_030	Silt percentage in first 30 cm	
	silt_30100	Silt percentage from 30-100 cm	
	clay_030	Clay percentage in first 30 cm	
	clay_30100	Clay percentage from 30-100 cm	
	bd_030	Bulk density in first 30 cm, g/cm^3	
	bd_30100	Bulk density from 30-100 cm, g/cm^3	
	om_030	Organic matter in first 30 cm, log10(%)	
	om_30100	Organic matter from 30-100 cm, log10(%)	
	ths_030	Saturated soil water content in first 30 cm, m^3/m^3	
	ths_30100	Saturated soil water content from 30-100 cm,	
	ph_030	Soil pH in first 30 cm	
	ph_30100	Soil pH from 30-100 cm	

Table 2. Environmental variables used in the final NIDGS habitat model, along with the relative contributions and model fit, using optimized parameters. Percent contribution measures the relative contribution of each variable to the Maxent prediction. Permutation importance (PI) measures the resulting drop in AUC when covariates at training and background points are randomly permuted. A large PI value means the model is more sensitive to that variable.

Variable	Percent contribution (%)	Permutation importance (%)
bio14	15.9	25.0
bio4	14.6	13.5
bio6	9.0	12.3
tree	3.4	11.9
silt_30100	15.0	11.5
bio18	2.7	9.7
bio3	14.6	6.2
bio15	0.9	3.7
SOST	1.4	2.8
heat	1.7	1.9
Nshrubherb	20.7	1.4
AUC	0.953	
Regularization multiplier	1.0	
Feature types	linear, product, hinge	

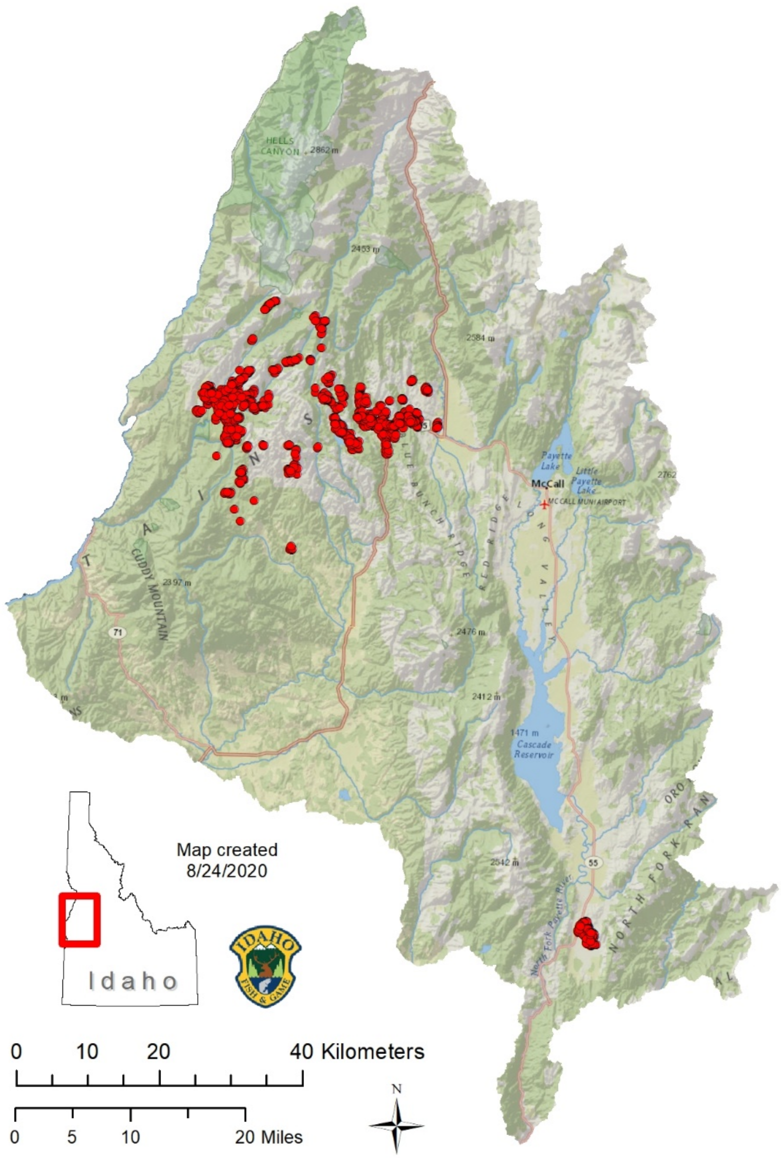

Figure 1. Known presence locations of northern Idaho ground squirrels (red dots) used to model suitable habitat (IDFG unpublished data). The model extent represents a collection of 10-digit hydrological units (HUC 5) that encompassed all species locations and the species' current and historic range.

(a) (b)

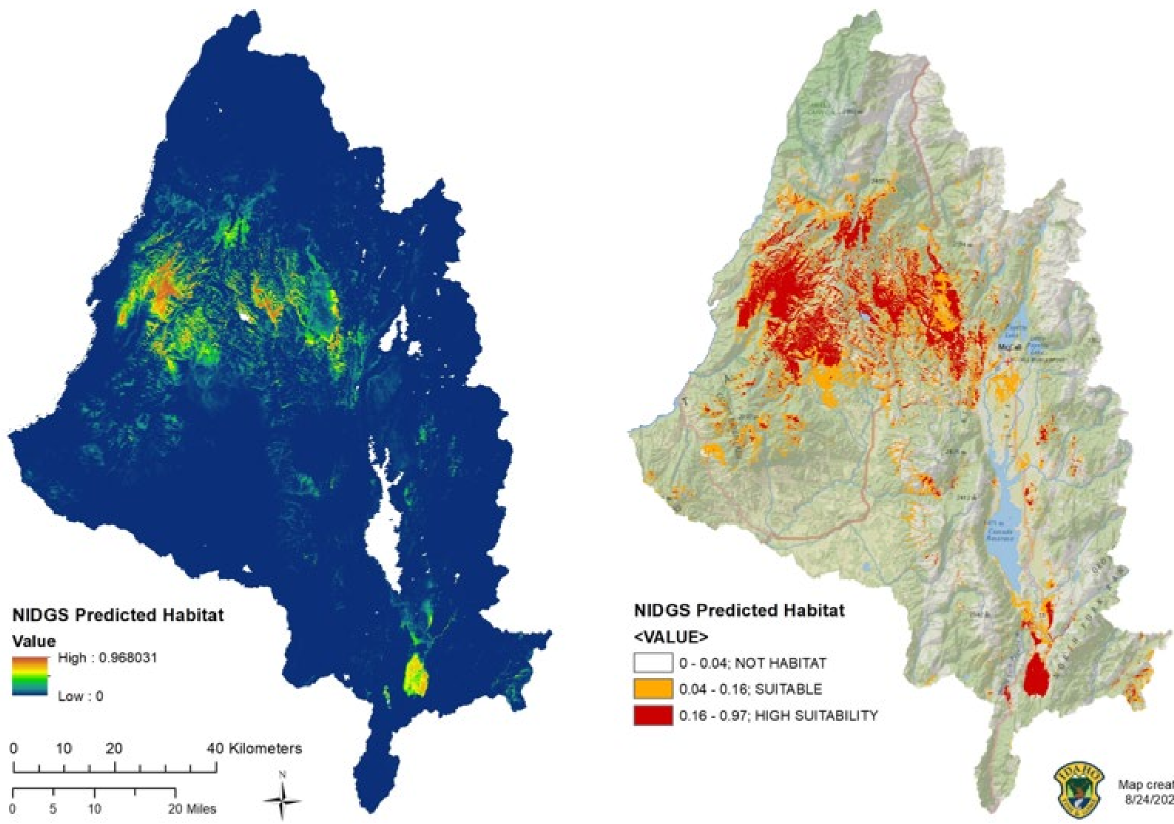

Figure 2. Predicted NIDGS habitat suitability. (a) Model values reflect the relative likelihood of species presence based on how similar environmental conditions are to conditions at presence locations (i.e., observations). Areas in red indicate the highest suitability, whereas blue areas are low. (b) This model is classified into a suitability map by applying thresholds. To separate non-habitat from suitable habitat, we selected a threshold that balanced training omission and predicted area (0.04). We selected the maximum test sensitivity plus specificity threshold to differentiate suitable from high suitability habitat (0.16). These thresholds were calculated by Maxent during model testing.

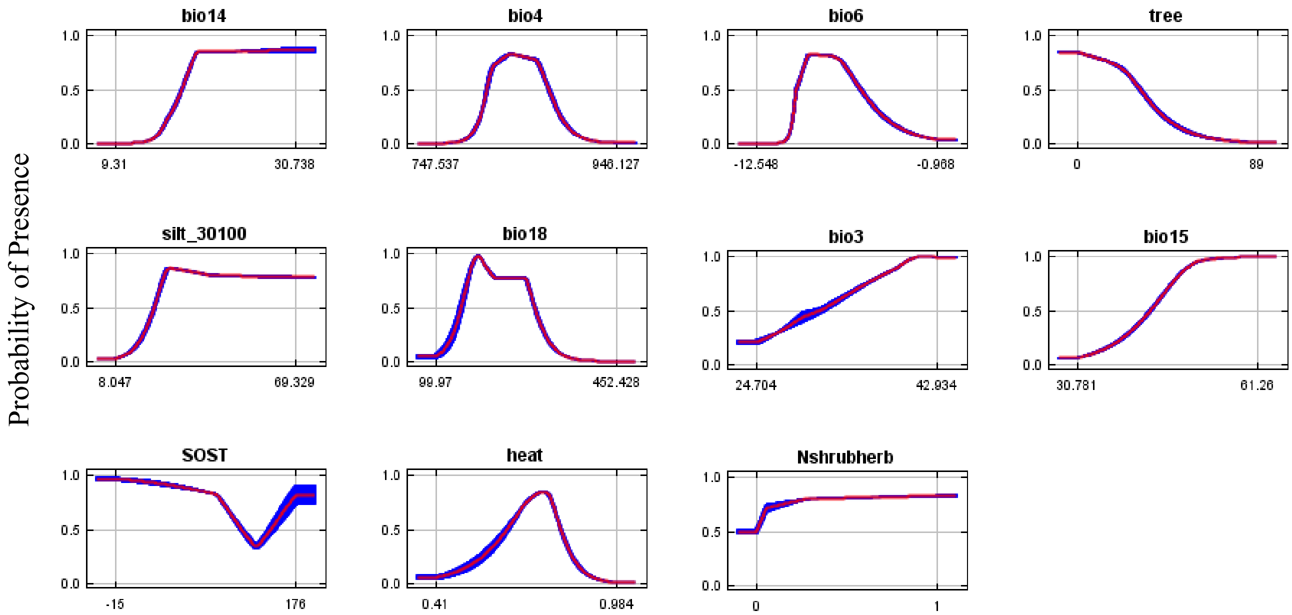

Figure 3. Partial effects plots of variables included in the final NIDGS habitat model, which show how the probability of presence (Y-axes) changes as each environmental variable is varied (X-axes), keeping all other variables at their average sample value. The red line indicates the mean response of the 5 replicate Maxent runs and the blue bands are one standard deviation error bars.

7. **Discuss differences between work anticipated in grant proposal and grant agreement, and that actually carried out with Federal Aid grant funds.**

   All proposed work was completed as anticipated within this report period. In addition, we went beyond the stated objectives and constructed individual habitat models for 3 genetic populations (i.e., West, East, and Round Valley) to assess differences in local habitat associations and how the sum of these models compared to the rangewide model described in part 6 above. As all objectives were addressed and all funds were expended by 30 June 2020, this report serves as both an interim and final report of the project.

8. **List any publications or in-house reports resulting from this work.**

   A manuscript is in preparation. In-house products include a PowerPoint presentation and a geodatabase housing all the spatial data from model development and results.

Name, title, phone number, and e-mail address of people compiling this report:

Diane Evans Mack, Regional Wildlife Biologist
(208) 634-8137
diane.evansmack@idfg.idaho.gov

Sonya Knetter, GIS Analyst II
(208)287-2737
sonya.knetter@idfg.idaho.gov

Leona Svancara, GIS Analyst III
(208)850-3774
leona.svancara@idfg.idaho.gov

Literature Cited

Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3(2):327–338.

Brown, J. L., J. R Bennett, and C. M. French. 2018. SDMtoolbox v2.3 for ArcMap 10.1–10.3. Accessed November 16, 2018. http://sdmtoolbox.org.

Brown, J. L., J. R Bennett, and C. M. French. 2017. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5:e4095. DOI 10.7717/peerj.4095.

Chaney, N. W., E. F. Wood, A. B. McBratney, J. W. Hempel, T. W. Nauman, C. W. Brungard, and N. P. Odgers. 2016. POLARIS: a 30-meter probabilistic soil series map of the contiguous United States. Geoderma, 274:54–67. https://doi.org/10.1016/j.geoderma.2016.03.025

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. https://doi.org/10.1111/j.1472-4642.2010.00725.x

ESRI, ArcGIS Desktop. 2017. Release 10.6.1. Redlands, CA: Environmental Systems Research Institute.

Evans, J. S., J. Oakleaf, S. Cushman, and D. Theobald. 2016. Geomorphometry & Gradient Metrics Toolbox. Version 2.0. Accessed 2 November 2016. https://www.arcgis.com/home/item.html?id=63ffcecf3b2a45bf99a84cdaedefaccf.

Evans Mack, D., C. Gillen, and L. McDonald. 2013. Long-term population monitoring of northern Idaho ground squirrel: sampling design, pilot implementation, and 2013 population estimates. Unpublished report to the U.S. Fish and Wildlife Service, threatened and endangered species projects F11AP00417 and F12AP00815 and cooperative agreement no. F11AC00175. December 31, 2013. Idaho Department of Fish and Game, Boise, Idaho, USA.

Holden, Z.A., A. Swanson, A. E. Klene, J. T. Abatzoglou, S. Z. Dobrowski, S. A. Cushman, J. Squires, G. G. Moisen, and J. W. Oyler. 2015. Development of high-resolution (250m) historical daily gridded air temperature data using reanalysis and distributed sensor networks for the US Northern Rocky Mountains. Int. J. Climatol. 36:3620–3632.

Merow, C., M. J. Smith, and J. A Silander Jr. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36:1058–1069.

Phillips, S. J., and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161–175.

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

Phillips, S. J., M. Dudík, and R. E. Schapire. 2017. Maxent software for modeling species niches and distributions (Version 3.4.1). Accessed 15 May 2018. http://biodiversityinformatics.amnh.org/open_source/maxent/.

PRISM Climate Group. 2012. 30-Year Normal Monthly Climate Data, 1981-2010 (800m), created 11 July 2012. Raster digital data. Accessed 10 June 2016. http://www.prism.oregonstate.edu/.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Smith, A. B. 2019. enmSdm: Tools for modeling species niches and distributions. R package version 0.3.4.6. URL https://github.com/adamlilith/enmSdm

USDI Fish and Wildlife Service. 2003. Recovery plan for the northern Idaho ground squirrel (Spermophilus brunneus brunneus). U.S. Fish and Wildlife Service, Portland, Oregon, USA.

U.S. Geological Survey (USGS). 2017. 1 Arc-second Digital Elevation Models (DEMs) – USGS National Map 3DEP Downloadable Data Collection. Raster digital data. Accessed 2016 May 17. https://nationalmap.gov/3DEP/.

USGS. 2018. C6 Aqua Western U.S. 250 m eMODIS Remote Sensing Phenology Data, version 2. Raster digital data. Accessed 31 January 2020. https://phenology.cr.usgs.gov/get_data_Aqua_C6_250w.php. https://doi.org//10.5066/F7PC30G1.

USGS. 2019*a*. Provisional Remote Sensing Shrub/Grass NLCD Base Products for the Western U.S. Raster digital data. Accessed November 13, 2019. https://www.mrlc.gov/data.

USGS. 2019*b*. NLCD 2016 Land Cover Conterminous United States. Raster digital data. Accessed 25 April 2019. https://www.mrlc.gov/data.

USGS. 2019*c*. NLCD 2016 USFS Tree Canopy Cover (CONUS). Raster digital data. Accessed 22 August 2019. https://www.mrlc.gov/data.

Warren, D. L., and S. N. Seifert. 2011. Ecological Niche Modeling in Maxent: the importance of model complexity and the performance of model selection. Ecological Applications 21:335–342.

Submitted by:
David Smith
Grants/Contracts Specialist
David.smith@idfg.idaho.gov

Martha Wackenhut
Federal Aid Coordinator

Approved by:

IDAHO DEPARTMENT OF FISH AND GAME

Toby Boudreau, Chief
Bureau of Wildlife