NatureServe Species Habitat Model Overview

Anaxyrus exsul Black Toad

NatureServe Unique Identifier: ELEMENT_GLOBAL.2.101917

Model Creation Date: 2019-04-25

General Information

This habitat model was produced for NatureServe to represent the current habitat distribution for *Anaxyrus exsul* (Black Toad).

Habitat models provide estimates of species habitat, created using predictive methods to understand species/environment relationships and map areas of likely occurrence. They cannot guarantee either species presence or absence. This model was funded by Esri as an input to the Map of Biodiversity Importance (MoBI) in the continental U.S. Additional details are provided in the attached model metadata.

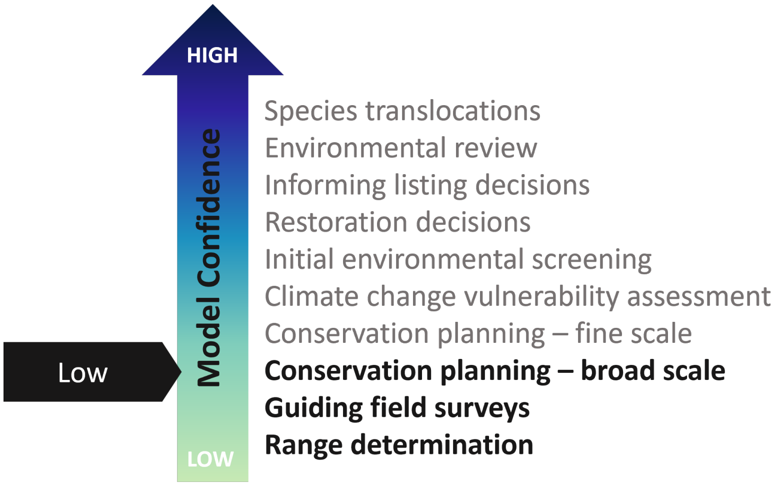
For more information about this model, please contact `data_science@NatureServe.org`. For more information on the assignment of confidence scores, see the **NatureServe Network Habitat Model Standard**.

Model Confidence & Recommended End Use

The habitat model has been assessed to have an overall confidence level of `LOW`.

Recommended Uses

Based on the overall confidence level for this model, we conclude that the model is appropriate for the recommended uses in bold in the figure on the right. However, we cannot recommed that this model be used for the applications in light grey.



Model Confidence Details

Overall confidence for this model is determined based on three factors: the integrity of inputs and methods, model validation statistics, and expert review. Details on each component are provided below.

Inputs and Methods This model has been assigned a confidence score for Inputs & Methods of `Medium` based on the factors below.

Expert Review NatureServe Review

No network review

Validation Statistics This model has been assigned a validation statistic confidence score of `Acceptable`. The model was validated using the `TSS` statistic for the high probability habitat category with a value of `0.089`.

Model Interpretation & Thresholds

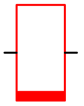
Most inductive habitat models result in a raw prediction from 0 to 1. We apply statistically defined thresholds, adjusted based on expert review when necessary, to convert the raw prediction into categorical data better suited to guiding management. Habitat categories and the thresholds used to define them for this model are reported below.

Habitat Category	Description	Value
High Probability	Habitat is optimal for species occurrence. Environmental conditions at the location are nearly identical to conditions where the species is frequently observed or where the location is known to be high-quality species habitat. The probability that an individual of the species would be found in this location is high.	`0.73–1`
Medium Probability	Environmental conditions at the location are similar to conditions where the species is regularly observed. The probability that an individual of the species would be found in this location is moderate.	`0.71–0.73`
Low Probability	Environmental conditions at the location are similar to conditions where the species is rarely observed or where observations are known to be in poor quality habitat. The probability that an individual of the species would be found in this location is low.	`0.71–0.71`
Non-habitat	Environmental conditions at the location are dissimilar to conditions wherespecies observations have occurred. The habitat is considered unsuitable, and the species is not expected to be found in this location unless incidentally passing through.	`0–0.71`

The detailed metadata pages that follow were produced when the model was originally published for the Map of Biodiversity Importance (2019). Any discrepancies in model review and model confidence information reflect reviews received since and/or updated Standards for assigning confidence scores.

Anaxyrus exsul

Species Distribution Model (SDM) assessment metrics and metadata
Common name: Black Toad
Grank: G1 - Critically Imperiled
Date: 17 Mar 2020
Code: anaxexsu (EGT_ID: 101917)



The following metadata describes the SDM for one species of 2,700 included in a Map of Biodiversity Importance (MoBI) in the continental U.S. developed by NatureServe and the Network of Natural Heritage Programs and funded by ESRI. This SDM incorporates the number of known and background locations indicated in Table 1, modeled with the random forests routine[1,2] in the R statistical environment[3,4]. We validated the model by jackknifing (also called leave-one-out[5,6,7]) by spatial grouping for a total of 4 groups. The statistics in Table 2 report the mean and variance of validation statistics for these jackknifing runs.

Table 1. Input statistics. Presence points are points placed in polygon-based location information or point-based observations. Groups describe groupings of points based on polygon data or spatial grouping of observations. Background points are placed throughout model area excluding known species locations.

Name	Number
Presence points	497
Groups	4
Background points	34170

Table 2. Validation statistics for jackknife trials. Overall Accuracy = Correct Classification Rate, TSS = True Skill Statistic, AUC = area under the ROC curve[8,9,6].

Name	Mean	SD	SEM
Overall Accuracy	0.54	0.10	0.05
Specificity	0.99	0.01	0.01
Sensitivity	0.10	0.19	0.10
TSS	0.09	0.20	0.10
Kappa	0.09	0.20	0.10
AUC	0.91	0.12	0.06

Validation runs used 79 environmental variables, the most important of 137 variables (top 75 percent). Each tree was built with 8 variables tried at each split (mtry) and 1000 trees built. The final model was built using 2000 trees, all presence and background points, with an mtry of 8, and 79 environmental variables.

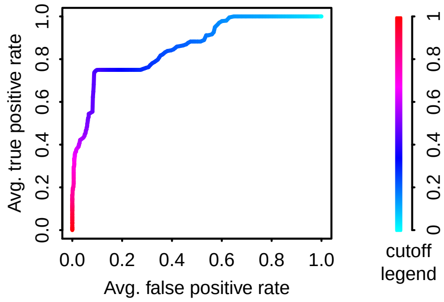


Figure 1. ROC plot for all 4 validation runs, averaged along cutoffs.

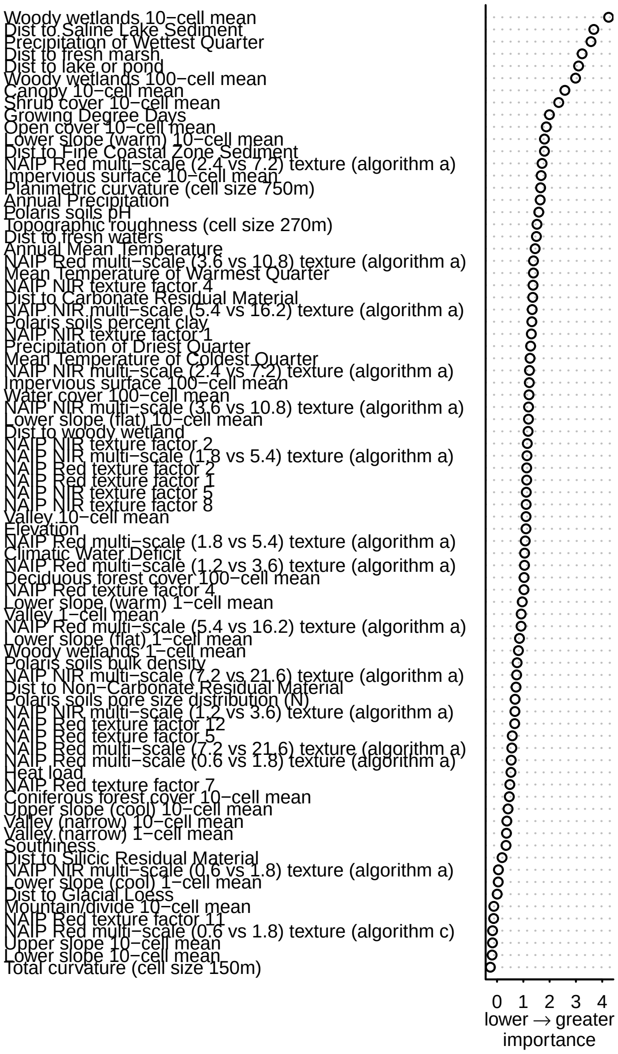


Figure 2. Relative importance of each environmental variable based on the full model using all sites as input. Importance values (mean decrease in accuracy) are extracted from the randomForest[3] function. See Appendix 1 for variable descriptions.

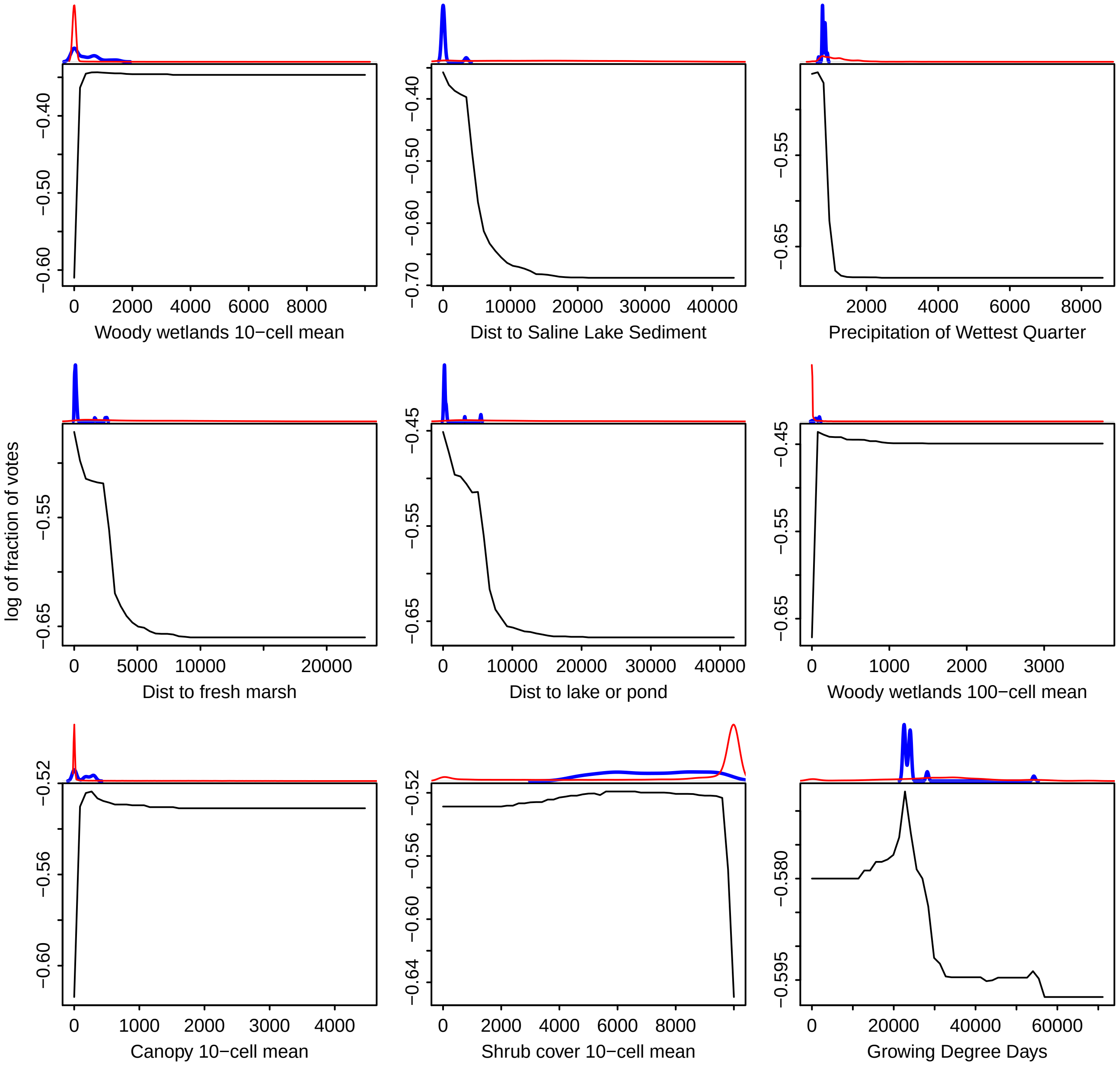


Figure 3. Partial dependence plots for the 9 environmental variables with the most influence on the model. Each plot shows the effect of the variable on the probability of appropriate habitat with the effects of the other variables removed[3]. The x-axis covers the range of values for the variable assessed; the y-axis represents the effect between the variable and model response. Peaks in the black line indicate where this variable had the strongest influence on predicting appropriate habitat. Decreasing lines from left to right show a negative relationship overall; increasing lines, positive. The distribution of each category (thin red = Background points, thick blue = Presence points) is depicted at the top margin. See Appendix 1 for variable descriptions.

Species distribution model outputs display the probability (0-1) of a location (i.e. stream reach or raster cell) having similar environmental conditions in comparison to known presence locations. No model will ever depict sites where a targeted element will occur with certainty, it can *only* depict locations it interprets as appropriate habitat for the targeted element. The delineation of suitable habitats is made by the selection of a threshold value, where locations with values above the threshold are designated as likely suitable habitat, and those with values below the threshold may be unsuitable. Threshold values are often statistically calculated. SDMs can be used in many ways and the depiction of appropriate habitat should be varied depending on intended use. For targeting field surveys, an SDM may be used to refine the search area; users should always employ additional GIS tools to further direct search efforts. A lower threshold depicting more area may be appropriate to use in this case. For a more conservative depiction of suitable habitat that shows less area, a higher threshold may be more appropriate. Different thresholds for this model (full model) are described in Table 3.

Table 3. Thresholds[11,12] calculated from the final model. The Value column reports the threshold; Groups indicates the percentage (number in brackets) of groups within which at least one point was predicted as suitable habitat; Pts indicates the percentage of PR points predicted having suitable habitat. Total numbers of groups and presence points used in the final model are reported in Table 1.

Threshold	Value	Groups	Pts	Description
Equal sensitivity and specificity	0.709	100(4)	99.8	The probability at which the absolute value of the difference between sensitivity and specificity is minimized.
Maximum of sensitivity plus specificity	0.709	100(4)	100	The probability at which the sum of sensitivity and specificity is maximized.
Minimum Training Presence	0.709	100(4)	100	The highest probability value at which 100% of input presence points remain classified as suitable habitat.
Minimum Training Presence by Group	0.950	100(4)	8.9	The highest probability value at which 100% of input groups have at least one presence point classified as suitable habitat.
Tenth percentile of training presence	0.790	100(4)	89.9	The probability at which 90% of the input presence points are classified as suitable habitat.
F-measure with alpha set to 0.01	0.709	100(4)	100	The probability value at which the harmonic mean of precision and recall, with strong weighting towards recall, is maximized.

Model Evaluation and Intended Use

All SDMs are sensitive to data inputs and methodological choices. Table 4 summarizes ratings by reviewers of the model output. Table 5 presents scoring of modeling factors based on the model evaluation rubric presented in Sofaer et al. 2019[13]. Models developed for the MoBI project are intended to inform creation of a national map of biodiversity value, and we recommend additional refinement and review before these data are used for more targeted, species-specific decision making. In particular, many MoBI models would benefit from greater consideration of species data and environmental predictor inputs, a more thorough review by species experts, and iteration to address comments received.

Table 4. Expert review results. The following is a summary of the reviews by NatureServe Network biologists received to date. Biologists assigned an overall score to the model (1 = poorest performance; 5 = highest performance) and also had an opportunity to make specific suggestions related to modeling extent. Summaries of overall scores are reported here.

Mean Star Rating	Number of reviews	Max rating	Min rating	Median rating
★★★★★ (-)	0	-	-	-

While not reviewed and ranked by network experts, this species was reviewed by NatureServe staff on the MoBI Team. Output was deemed acceptable for this project.

Table 5. Model evaluation results based on Sofaer et al. 2019. Scores can be attributed as ideal, acceptable, or interpret with caution.

Category	Metric	Score	Notes
Species Data	Presence data quality	Acceptable	Heritage Network data augmented with outside data which may or may not be vetted for accuracy or weighted for spatial representation.
	Absence/Background Data	Acceptable	Background points randomly placed throughout study area excluding species locations.
	Evaluation Data	Acceptable	Models are validated by jackknifing (i.e. leave-one-out).
Environmental Predictors	Ecological and predictive relevance	Acceptable	Selection of predictor variables were based on previous modeling experience by the Natural Heritage Network. Time constraints of this project prevented making species specific selections.
	Spatial and temporal alignment	Acceptable	Reasonable attempts to align predictor and presence data were made.
Modeling Process	Algorithm choice	Acceptable	Random Forest is highly rated classification model that is well documented as suitable for modeling rare species.
	Sensitivity	Acceptable	Settings for Random Forest were adjusted to best model the species; however, different models/parameters were not tested within one model run.
	Statistical rigor	Acceptable	Collinearity of predictors recognized and addressed; presence points grouped to minimize sample bias and minimize spatial autocorrelation boost during validation; other assumptions recognized and considered.
	Performance	Acceptable	Model TSS $\geq$ 0.6. Mapped model output is evaluated for ecological plausibility by expert review.
	Model review	Acceptable	Model was reviewed by NatureServe staff.
Model Products	Mapped products	Acceptable	Single calculated threshold selected for all final models to be integrated into MoBI.
	Interpretation support products	Ideal	All standards met.
	Reproducibility	Ideal	All standards met.
	Iterative	Interpret with Caution	Model not revised.

Model Comments

Figure 4. A generalized view of the model predictions throughout the modeled area. State boundaries are depicted as a thin gray line. The modeled area is outlined in red. Basemap: CartoDB.Positron (©OpenStreetMap, contributors: ©CARTO).

This distribution model would not have been possible without data sharing among organizations. Other data sets and sources may have been evaluated, but this final model includes data from these sources:

- NatureServe, Multi-jurisdictional Database from Heritage Network Programs

This model was built using a methodology developed through collaboration among the Florida Natural Areas Inventory, the New York Natural Heritage Program, the Pennsylvania Natural Heritage Program, and the Virginia Natural Heritage Program, all member programs of the NatureServe Network. It is one of a suite of species distribution models developed using the same methods, scripts, and environmental data sets. Our goal was to be consistent and transparent in our methodology, validation, and output.

Please cite this document and its associated SDM as:
NatureServe and Heritage Network Partners. 2020. Species distribution model for Black Toad (*Anaxyrus exsul*). Created on 17 Mar 2020. Arlington, VA with Network partners from VA, PA, and NY.

References

[1] Breiman, L. 2001. Random forests. Machine Learning 45:5-32.

[2] Iverson, L. R., A. M. Prasad, and A. Liaw. 2004. New machine learning tools for predictive vegetation mapping after climate change: Bagging and Random Forest perform better than Regression Tree Analysis. Landscape ecology of trees and forests.Proceedings of the twelfth annual IALE (UK) conference, Cirencester, UK, 21-24 June 2004 317-320.

[3] Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18-22. Version 4.6-14.

[4] R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. R version 3.5.1 (2018-07-02).

[5] 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.

[6] Fielding, A. H. 2002. What are the appropriate characteristics of an accuracy measure? Pages 271-280 in Predicting Species Occurrences, issues of accuracy and scale. J. M. Scott, P. J. Helglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, F. B. Samson, eds. Island Press, Washington.

[7] Pearson, R.G. 2007. Species Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.

[8] Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43:1223-1232.

[9] Vaughan, I. P. and S. J. Ormerod. 2005. The continuing challenges of testing species distribution models. Journal of Applied Ecology 42:720-730.

[10] Sing, T., O. Sander, N. Beerenwinkel, T. Lengauer. 2005. ROCR: visualizing classifier performance in R. Bioinformatics 21(20):3940-3941.

[11] Liu, C., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385–393.

[12] Liu, C., G. Newell, and M. White. 2015. On the selection of thresholds for predicting species occurrence with presence-only data. Ecology and Evolution 6:337–348.

[13] Sofaer, H. R., C. S. Jarnevich1, 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. The development and delivery of species distribution models to inform decision-making. BioScience. BioScience 69:544-557.

Appendix 1. Descriptions for environmental variables included in model.

Variable Name	Variable Description
Annual Mean Temperature	Annual Mean Temperature
Annual Precipitation	Annual Precipitation
Canopy 10-cell mean	mean percent canopy cover in 10-cell radius (30 meter cells)
Climatic Water Deficit	Climatic Water Deficit
Coniferous forest cover 10-cell mean	mean coniferous forest cover within 10-cell radius
Deciduous forest cover 100-cell mean	mean deciduous forest cover within 100 cell radius
Dist to Carbonate Residual Material	Euclidean distance to surficial geology type: Carbonate Residual Material
Dist to Fine Coastal Zone Sediment	Euclidean distance to surficial geology type: Alluvium and Fine-Textured Coastal Zone Sediment
Dist to Glacial Loess	Euclidean distance to surficial geology type: Eolian Sediment, Fine-Textured (Glacial Loess)
Dist to Non-Carbonate Residual Material	Euclidean distance to surficial geology type: Non-Carbonate Residual Material
Dist to Saline Lake Sediment	Euclidean distance to surficial geology type: Saline Lake Sediment
Dist to Silicic Residual Material	Euclidean distance to surficial geology type: Silicic Residual Material
Dist to fresh marsh	Distance to freshwater emergent wetland
Dist to fresh waters	Euclidean distance to nearest stream, river, or other inland waterbody (excluding estuaries)
Dist to lake or pond	Euclidean distance to nearest river or other inland waterbody (features represented by polygons)
Dist to woody wetland	Distance to forested palustrine wetland
Elevation	Elevation in decimeters (originally in meters)
Growing Degree Days	Growing Degree Days
Impervious surface 10-cell mean	mean percent impervious cover in 10-cell radius (30 meter cells)
Impervious surface 100-cell mean	mean percent impervious cover in 100-cell radius (30 meter cells)
Lower slope (cool) 1-cell mean	mean percent Lower slope (cool) in 1-cell mean (30 meter cells)
Lower slope (flat) 1-cell mean	mean percent Lower slope (flat) in 1-cell mean (30 meter cells)
Lower slope (flat) 10-cell mean	mean percent Lower slope (flat) in 10-cell mean (30 meter cells)
Lower slope (neutral) 10-cell mean	mean percent Lower slope (neutral) in 10-cell mean (30 meter cells)
Lower slope (warm) 1-cell mean	mean percent Lower slope (warm) in 1-cell mean (30 meter cells)
Lower slope (warm) 10-cell mean	mean percent Lower slope (warm) in 10-cell mean (30 meter cells)
Mean Temperature of Coldest Quarter	Mean Temperature of Coldest Quarter
Mean Temperature of Warmest Quarter	Mean Temperature of Warmest Quarter
Mountain/divide 10-cell mean	mean percent Peak/divide in 10-cell mean (30 meter cells)
NAIP NIR multi-scale (0.6 vs 1.8) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 0.6m cell size vs. 3x3 cell variance at a 1.8m cell size.
NAIP NIR multi-scale (1.2 vs 3.6) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 1.2m cell size vs. 3x3 cell variance at a 3.6m cell size.
NAIP NIR multi-scale (1.8 vs 5.4) texture (algorithm a)	Multi-scale (NDTI) texture from near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 1.8m cell size vs. 3x3 cell variance at a 5.4m cell size.
NAIP NIR multi-scale (2.4 vs 7.2) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 2.4m cell size vs. 3x3 cell variance at a 7.2m cell size.
NAIP NIR multi-scale (3.6 vs 10.8) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 3.6m cell size vs. 3x3 cell variance at a 10.8m cell size.
NAIP NIR multi-scale (5.4 vs 16.2) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 5.4m cell size vs. 3x3 cell variance at a 16.2m cell size.
NAIP NIR multi-scale (7.2 vs 21.6) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery near-infrared band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 7.2m cell size vs. 3x3 cell variance at a 21.6m cell size.
NAIP NIR texture factor 1	Single-scale texture from aerial imagery near-infrared band reduced with factor analysis, factor 1.
NAIP NIR texture factor 2	Single scale texture from aerial imagery near-infrared band reduced with factor analysis, factor 2.
NAIP NIR texture factor 4	Single scale texture from aerial imagery near-infrared band reduced with factor analysis, factor 4.
NAIP NIR texture factor 5	Single scale texture from aerial imagery near-infrared band reduced with factor analysis, factor 5.
NAIP NIR texture factor 8	Single scale texture from aerial imagery near-infrared band reduced with factor analysis, factor 8.
NAIP Red multi-scale (0.6 vs 1.8) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 0.6m cell size vs. 3x3 cell variance at a 1.8m cell size.
NAIP Red multi-scale (0.6 vs 1.8) texture (algorithm c)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses a summed alternating pixel convolution filter to contrast local 3x3 cell variance at a 0.6m cell size vs. 3x3 cell variance at a 1.8m cell size.
NAIP Red multi-scale (1.2 vs 3.6) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 1.2m cell size vs. 3x3 cell variance at a 3.6m cell size.
NAIP Red multi-scale (1.8 vs 5.4) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 1.8m cell size vs. 3x3 cell variance at a 5.4m cell size.
NAIP Red multi-scale (2.4 vs 7.2) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 2.4m cell size vs. 3x3 cell variance at a 7.2m cell size.
NAIP Red multi-scale (3.6 vs 10.8) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 3.6m cell size vs. 3x3 cell variance at a 10.8m cell size.
NAIP Red multi-scale (5.4 vs 16.2) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 5.4m cell size vs. 3x3 cell variance at a 16.2m cell size.
NAIP Red multi-scale (7.2 vs 21.6) texture (algorithm a)	Multi-scale (NDTI) texture from aerial imagery red band (30 meter cells). Uses local standard deviation to contrast local 3x3 cell variance at a 7.2m cell size vs. 3x3 cell variance at a 21.6m cell size.
NAIP Red texture factor 1	Single scale texture from aerial imagery red band reduced with factor analysis, factor 1.
NAIP Red texture factor 11	Single scale texture from aerial imagery red band reduced with factor analysis, factor 11.
NAIP Red texture factor 12	Single scale texture from aerial imagery red band reduced with factor analysis, factor 12.
NAIP Red texture factor 2	Single scale texture from aerial imagery red band reduced with factor analysis, factor 2.
NAIP Red texture factor 4	Single scale texture from aerial imagery red band reduced with factor analysis, factor 4.
NAIP Red texture factor 5	Single scale texture from aerial imagery red band reduced with factor analysis, factor 5.
NAIP Red texture factor 7	Single scale texture from aerial imagery red band reduced with factor analysis, factor 7.
Open cover 10-cell mean	mean open cover within 10-cell radius
Polaris soils bulk density	Polaris soils bulk density
Polaris soils pH	Polaris soils pH
Polaris soils percent clay	Polaris soils percent clay
Polaris soils pore size distribution (N)	Polaris soils pore size distribution (N)
Precipitation of Driest Quarter	Precipitation of Driest Quarter
Precipitation of Wettest Quarter	Precipitation of Wettest Quarter
Shrub cover 10-cell mean	mean shrub cover within 10-cell radius
Upper slope (cool) 10-cell mean	mean percent Upper slope (cool) in 10-cell mean (30 meter cells)
Upper slope (neutral) 10-cell mean	mean percent Upper slope (neutral) in 10-cell mean (30 meter cells)
Valley (narrow) 1-cell mean	mean percent Valley (narrow) in 1-cell mean (30 meter cells)
Valley (narrow) 10-cell mean	mean percent Valley (narrow) in 10-cell mean (30 meter cells)
Valley 1-cell mean	mean percent Valley in 1-cell mean (30 meter cells)
Valley 10-cell mean	mean percent Valley in 10-cell mean (30 meter cells)
Water cover 100-cell mean	mean open water cover within 100 cell radius
Woody wetlands 1-cell mean	mean woody wetland cover within 1 cell radius
Woody wetlands 10-cell mean	mean woody wetland cover within 10 cell radius
Woody wetlands 100-cell mean	mean woody wetland cover within 100 cell radius

Appendix 2. Model details for reproducibility

- All R Scripts are available at github
- The repository version (repo head) used for this run was: 6e7d05472163f33d1fd83c03450f9e23cea50902
- The model run name was: anaxexsu_20190425_082600
- Validation metrics requiring a threshold use MTP (minimum training presence)

- R version: R version 3.5.1 (2018-07-02)
- Random seed for full randomForest model: 425082600
- randomForest mtry: 8
- randomForest ntrees: 2000