Final Project Report

Covered Species Model Updates 2023

Project Number: 2017-UNR-1782C

July 31, 2024

Prepared by: Ken Nussear,.

Anjana Parandhaman, and Dahlia Villalobos

knussear@unr.edu

University of Nevada, Reno

1664 N. Virginia Street

Reno, NV. 89557

Study Area

The study area used for modeling was incorporated from earlier modeling efforts (Nussear and Simandle 2019), and consisted of a 50km buffer area around the boundary for Clark County NV. The projection used was a NAD83, UTM Zone 11N projection (corresponding to the EPSG 26911). Raster resolution for all input and output layers was set at a 250 m grid resolution. All environmental data, and point data were cropped, masked and re-sampled to this resolution and extent. Envrionmental Layers used include those from Nussear and Simandle 2019, SWECO 2018. These are given in Table 1 below.

Table 1. Environmental covariate names and their source.

Name	Source
Ave Max Temp	Average of the maximum monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Ave Min Temp	Average of the maximum monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Clay	Downloaded from the Soil Grids 250m project. Hengl et al. 2017
Coarse fragments	Downloaded from the Soil Grids 250m project. Hengl et al. 2017
CV Max Temp	Coefficient of Variation of the maximum monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
CV Min Temp	Coefficient of Variation of the maximum monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Dist to cliffs	Distance of Cliffs - from Inman et al. 2014
Extreme Max Temp	Extreme Maximum of monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Extreme Min Temp	Extreme Minimum of monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Flow Accum	Inman et al. 2014
NDVI Amplitude	USGS Phenology network - https://www.usgs.gov/land-resources/eros/phenology/science/deriving-phenological-metrics-ndvi?qt-science_center_objects=0#qt-science_center_objects
NDVI Length of Season	USGS Phenology network - https://www.usgs.gov/land-resources/eros/phenology/science/deriving-phenological-metrics-ndvi?qt-science_center_objects=0#qt-science_center_objects
NDVI Max	USGS Phenology network - https://www.usgs.gov/land-resources/eros/phenology/science/deriving-phenological-metrics-ndvi?qt-science_center_objects=0#qt-science_center_objects
Sand	Downloaded from the Soil Grids 250m project. Hengl et al. 2017

Name	Source
Silt	Downloaded from the Soil Grids 250m project. Hengl et al. 2017
Slope	Calculated from USGS National Map. https://www.usgs.gov/core-science-systems/national-geospatial-program/national-map
Start of Season (day)	USGS Phenology network - https://www.usgs.gov/land-resources/eros/phenology/science/deriving-phenological-metrics-ndvi?qt-science_center_objects=0#qt-science_center_objects
Winter Precip	Average of the cumulative annual winter precipitation (October - March) for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
CV Winter Precip	Coefficient of Variation for the cumulative annual winter precipitation (October - March) for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Surface roughness	Inman et al. 2014
Average Spring Max Temp	Average of the maximum monthly temperatures for March - May for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
CV Average Spring Max Temp	Coefficient of Variation for the maximum monthly temperatures for a 30-year normal period between 1988 and 2018 calculated from monthly PRISM data at 800m resolution and downscaled to a 250 m resolution with bicubic spline interpolation using gdal-warp in python.
Percent washes	Calculated from USGS National Map. https://www.usgs.gov/core-science-systems/national-geospatial-program/national-map
Absolute depth to bedrock	Downloaded from the Soil Grids 250m project. Hengl et al. 2017

Species locality data

Species locality data were obtained from our earlier modeling efforts (Nussear and Simandle 2019, SWECO 2018), and were updated from current searches at INaturalist using research grade observations without obscured locality data, as well as with data updates from Clark County NV which were supplied to us by John Ellis in the form of 1 excel file (Clark County Species Observations.xlsx), and 1 geodatabase (CC_Data_Deliverable.gdb) that we received on August 23.

Modeling Methods

Modeling updates were conducted using an ensemble modeling approach that incorporated four different algorithms commonly used in species distribution modeling. These were: generalized additive models (GAM; using the “mgcv” method Wood 2006), random forests (RF; implemented in the R package “randomForest,” Liaw and Wiener 2002), MaxEnt (version 3..4.1, Phillips et al. 2006) implemented in the Maxnet algorithm in R (maxnet v 0.1.4, Phillips 2021) and Generalized boosted regression models (GBM) implemented in the ‘gbm’ package (version 2.1.8.1, Greenwell et al 2022). All models were executed using custom species distribution modeling code developed by Nussear for an upcoming package for R, (Nussear et al in Prep 2023). The use of multi-algorithm ensembles renders predictions less susceptible to the biases, assumptions, or limitations of any individual algorithm, while broadening the types of environmental response functions that can be identified (Araujo and New 2006). Moreover, empirical evaluations have found GAM, RF, MaxEnt, and GBM to be consistently strong performers among habitat distribution modeling algorithms (Franklin 2010). All modeling was conducted in R version 4.3 (R Core Team 2023).

True absence points were not available for any of the study species at this time. For this reason, all models were fit using randomly generated background points (pseudo-absences). Random selections of background points are considered a reliable method for regression techniques, and are a widely used method (Wisz and Guisan 2009; Barbet-Massin et al. 2012). Background points were randomly drawn from a bioclimatic envelope model executed in the bioclim algorithm from the dismo package (v 1.3, Hijmans et al. 2023) selecting points with the same frequency of occurrences (Barbet-Massin et al. 2012) from areas with a model value below 0.3.

To keep models interpretable and to improve their generalization across the study area, we also did not include interaction terms. Because presence points tended to be spatially aggregated, which can lead to substantial bias in model predictions, we first rasterized the presence points to the modeling resolution (i.e., such that only one presence point could occur within each grid cell) and subsequently applied a geographically-weighted resampling procedure in which a maximum of three observations could be sampled from cells on a uniform grid at a spatial resolution 4 times larger than the modeling extent (e.g., 1 km^2 for a 250 m^2). This systematic grid sampling approach for spatial thinning of presence points can be effective at reducing spatial bias under a variety of conditions (Fourcade et al. 2014). To further reduce bias in our predictions, we used cross-validations to fit and evaluate all habitat models. In this process, each algorithm was fit across 20 samples of randomly selected, spatially thinned presence points, with a 20% random sample (without replacement) withheld for model evaluation at each iteration (i.e., 80 % of presence points were used in model fitting, and 20% in model evaluation). Background points were also randomly drawn for each cross-validation.

This modeling effort included a vegetation layer that was provided in a shapefile format (Vegetation_USNVC_Divisions_20240423_LAME_CC.shp),consisting of plant association/alliance group polygons, and was to be evaluated for inclusion in the modeling efforts. This created changes in our modeling approach, and the implementation of additional modeling techniques. First the vegetation layer was a smaller extent than the buffered study area for the initial models, restricted to the Clark County boundary, with sections missing in the northwestern extent on the Nevada National Test Site, and a few smaller patches, where the Lake Mead area was included in the final V3 model versions. In addition the data are categorical with 16 levels of vegetation associations (Table 3). Given the relatively narrow distribution of some species with respect to vegetation communities this can create difficulties modeling, as all of the categories need to be represented in the presence and absence data. To achieve this requirement we augmented absence data with random samples stratified within each of the strata, adding up to 5 points per strata – and treated them as pseudo absences. For these modeling efforts the GAM algorithm was not used as there were problems with model convergence, and the MaxEnt models all converged on either all present, or all absent outputs, and were not used for ensemble modeling. Thus the final models including the vegetation layer relied on the RF and GBM algorithms .

Metrics of model prediction accuracy were calculated based on the evaluation data for each of the cross-validation runs, and subsequently averaged across runs. Performance metrics included several threshold-independent measures: AUC (the area under the receiver operating characteristic; Fielding and Bell 1997), the Boyce Index (BI; Boyce et al. 2002; Hirzel et al. 2006), and the True Skill Statistic (TSS; Allouche et al. 2006). TSS takes into account both omission and commission errors and is insensitive to data prevalence (Allouche et al. 2006).

Habitat distribution models vary in their ability to effectively discriminate different classes of habitat along the full range of habitat suitability values (0 – 1; Hirzel et al. 2006). To evaluate this property, we calculated the continuous Predicted / Expected (P/E) ratio curves based on the BI (Hirzel et al. 2006) using the *ecospat* package (v 3.0) in R. These curves reflect how well each model deviates from random expectation, and inform the interpretation of biologically meaningful suitability categories by indicating the effective resolution of suitability scores for each model (i.e., the model's ability to distinguish different classes of suitability; Hirzel et al. 2006).

To generate predictive layers of habitat suitability for each species (Table 2), we selected the top candidate models from each algorithm, based upon model performance metrics across cross-validation runs where the AUC was greater than the mean of all models. Ensemble predictions for individual algorithms were generated by taking the weighted average among candidate models for all algorithm types (i.e., one ensemble prediction each for GAM, RF, GBM, and MaxEnt models), with the weights determined by TSS scores for each of the included models. Layers representing the standard error of the overall ensemble habitat suitability layer were calculated as the standard deviation in model predictions across all candidate models, divided by the square root of the number of candidate models considered).

Quantitative model interpretation

To facilitate biological interpretations of the ensemble models, we410 calculated the relative importance of environmental predictors across candidate models for each algorithm. To illustrate the shape of the relationships between predicted habitat suitability and important environmental covariates, we derived partial response curves for the top 4 environmental parameters for each of the algorithms. Partial response curves show the predicted habitat suitability across a single covariate's range of values, while holding all other covariates at their mean value (e.g., Elith et al. 2005). To indicate the overall distribution of covariate values across the study region, we overlaid the response curve plots with histograms representing each environmental covariate. These histograms were calculated from the combined presence and pseudo absence locations.

Habitat Models

Models with and without the new vegetation layer were completed for 17 species – which are given in the table below (Table 2).

Table 2. Species "codes", common names, and scientific names for species covered in this modeling effort.

SPECIES CODE	COMMON NAME	SCIENTIFIC NAME
ANLE	Sticky Ringstem	Anulocaulis leiosolenus
AQCH	Golden eagle	Aquila chrysaetos
ARCA	Las Vegas Bearpoppy	Arctomecon californica
ASGETR	Three Corner Milkvetch	Astragalus geyeri var. triquetrus
ATCU	Burrowing Owl	Athene cunicularia

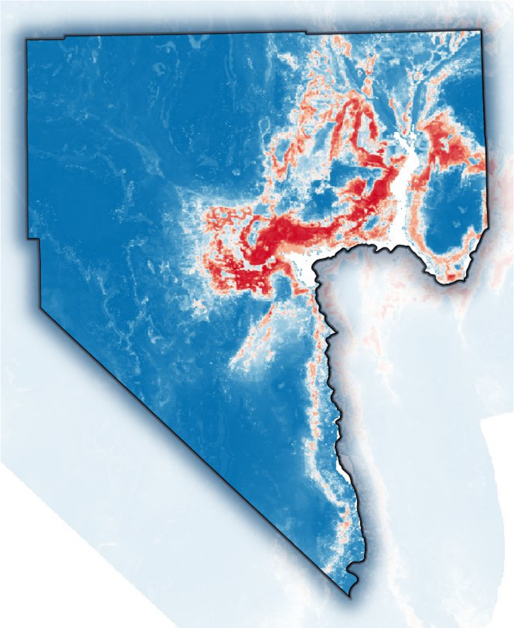
SPECIES CODE	COMMON NAME	SCIENTIFIC NAME
CHPE	Desert Pocket Mouse	Chaetodipus penicillatus
COCH	Gilded Flicker	Colaptes chrysoides
ENAR	Silverleaf Sunray	Enceliopsis argophylla
ERBI	Pahrump Valley Buckwheat	Eriogonum bifurcatum
ERCO	Las Vegas Buckwheat	Eriogonum corymbosum var. nilesii
ERVI	Sticky Buckwheat	Eriogonum viscidulum
GOAG	Mojave Desert tortoise	Gopherus agassizii
LALU	Loggerhead Shrike	Lanius ludovicianu
PEAL	White-margined Beardtongue	Penstemon albomarginatus
TOBE	Bendire's thrasher	Toxostoma bendirei
TOLE	Le Conte's thrasher	Toxostoma lecontei
VIBE	Loggerhead shrike	Lanius ludovicianus

Table 3. Vegetation layer Groups, "Tmp Names" and Rasterized values used in modeling for the vegetation included models.

Raster Value	GROUP	"Tmp Name"
0	Californian Forest & Woodland	Californian Broadleaf Forest and Woodland
1	Developed	Land Use and Development
2	North American Warm Desert Scrub & Grassland	North American Warm Desert Ruderal Grassland
3	Rocky Mountain Forest & Woodland	Inter-Mountain Basins Subalpine Limber-Bristlecone Pine Woodland
4	Southwestern North American Warm Desert Freshwater Marsh & Bosque	North American Warm Desert Riparian Low Bosque and Shrubland
5	Urban Interface Mojave Desert Scrub	Urban Interface Mojave Desert Scrub
6	Vacant	Transisitional Lands
7	Vacant or Cleared	Land Use and Development
8	Water	Canals and Other Man-made Watercourses
9	Western North American Alpine Tundra	North American Desert Alkaline-Saline Marsh and Playa
10	Western North American Cool Semi-Desert Scrub & Grassland	Great Basin-Intermountain Tall Sagebrush Steppe and Shrubland
11	Western North American Grassland & Shrubland	Southern Rocky Mountain Mountain-mahogany - Mixed Foothill Shrubland
12	Western North American Interior Chaparral	Western Madrean Chaparral
13	Western North American Interior Flooded Forest	Western Interior Riparian Forest and Woodland and Interior West Ruderal Riparian Forest and Scrub
14	Western North American Pinyon - Juniper Woodland & Scrub	Colorado Plateau - Great Basin Juniper Open Woodland
15	Western North American Temperate Freshwater Marsh, Wet Meadow & Shrubland	Rocky Mountain Alpine-Montane Wet Meadow

Model outputs and performance tables for each are given below..

ANLE – Sticky Ringstem



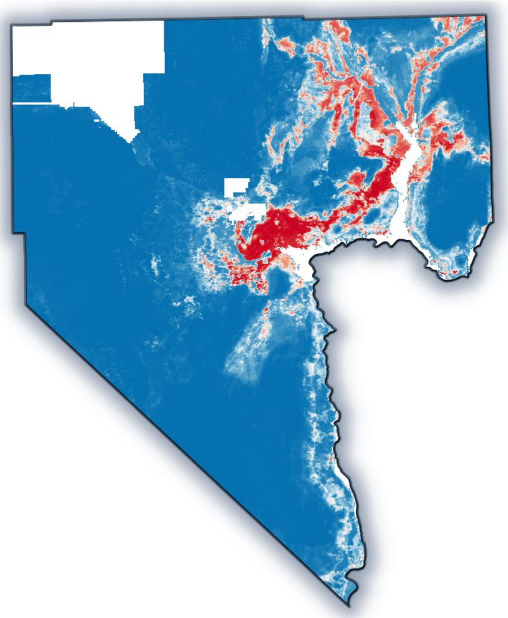


Figure 1 – Models for ANLE – with
environmental raster layers only (left) and with the inclusion of the vegetation layer (right).

The Sticky Ringstem model had high overall performance – where AUC, BI, and TSS all had high values with minimal losses between training and testing data. (Table 4). With the inclusion of vegetation the model performance appeared high (perhaps too high) (Table 5). The models both indicated higher habitat values in the Muddy River/Moapa area, and near Lake Mead (where visible in the vegetation model), but the vegetation influenced model had reduced predicted habitat in the area west of Las Vegas (Figure 1). This reduction appears to be more closely aligned with the foot print of the localities used in modeling in this case (Figure 2).

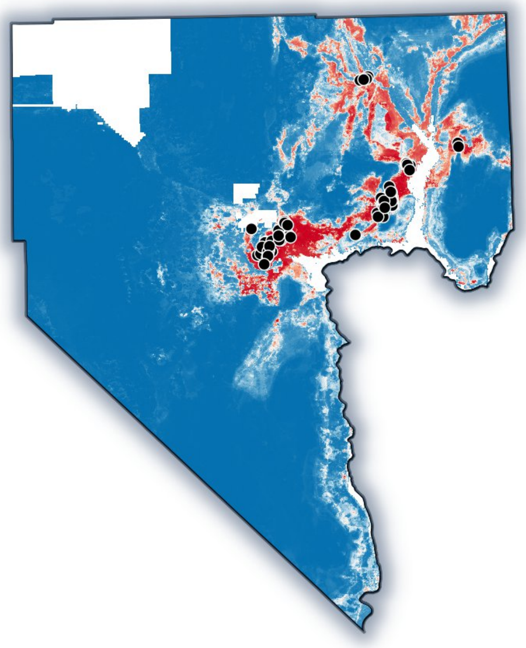


Figure 2 – Model for ANLE with the inclusion of the vegetation layer showing localities for the species used for modeling and testing. Note the localities in the Lake Mead area with no underlying values could not be used in modeling.

Table 4 performance metrics for ANLE. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.98	0.8	0.97	1	0.83	0.97
GAM	1	0.95	0.99	0.49	0.48	0.58	1	0.83	0.95
RF	1	0.98	1	0.97	0.9	0.96	1	0.83	0.97
MX	0.99	0.98	0.99	0.98	0.96	0.99	0.91	0.88	0.89
GBM	1	0.93	0.99	0.85	0.57	0.84	1	0.83	0.97

Table 5 performance metrics for ANLE with vegetation model. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.87	0.86	0.96	1	0.95	0.99
RF	1	1	1	0.93	0.96	0.97	1	0.95	0.99
GBM	1	1	1	0.86	0.85	0.91	1	0.95	0.99

Importance of the individual environmental layers indicated different variables among the different algorithms, highlighting the importance of using the ensemble of multiple algorithms. The final ensemble model (Figure 1) was composed from weighted models of 12 RF, 12 Maxent, 12 GBM, and 6 GAM models. Slope was the lowest performing variable, and was below 10% importance in all algorithms (Table 6). Minimum and maximum temperatures, as well as the soil gypsum content were the variables that had the greatest influence (Table 5).

The models including vegetation (20 RF and 20 GBM) indicated shifts in variable importance relative to the base models. The GBM, which showed very little importance of the vegetation layer (Table 7) had the highest importance for Ave Spring Max Temp, and an increase in Soil Gypsum, which dominated the models. The RF importance showed 2.4% importance for the vegetation layer, with increased importance of Average Minimum Temperature, Spring Temperatures, Soil Gypsum, and NDVI Amplitude (Table 7).

Table 6. Relative importance of the input variables used in modeling for ANLE

Variable	GAM	GBM	RF	MX
Ave Min Temp	37.6	6.7	16.8	42.7
Ave Spring Max Temp	35.9	27	28.4	5.1
Soil gypsum	5.3	65.5	31.9	25.9
NDVI Amplitude	2	0.7	14.5	9.8
Silt	10.7	0	4	8.2
Slope	8.4	0	4.3	8.5

Table 7. Relative importance of the input variables used in modeling for ANLE with the vegetation layer.

Variable	GBM	RF
Ave Min Temp	0	10.6
Ave Spring Max Temp	32.4	27.4
Soil gypsum	61.8	29.2
NDVI Amplitude	5.8	19.2
Silt	0	5.5
Slope	0	5.8
Veg	0	2.4

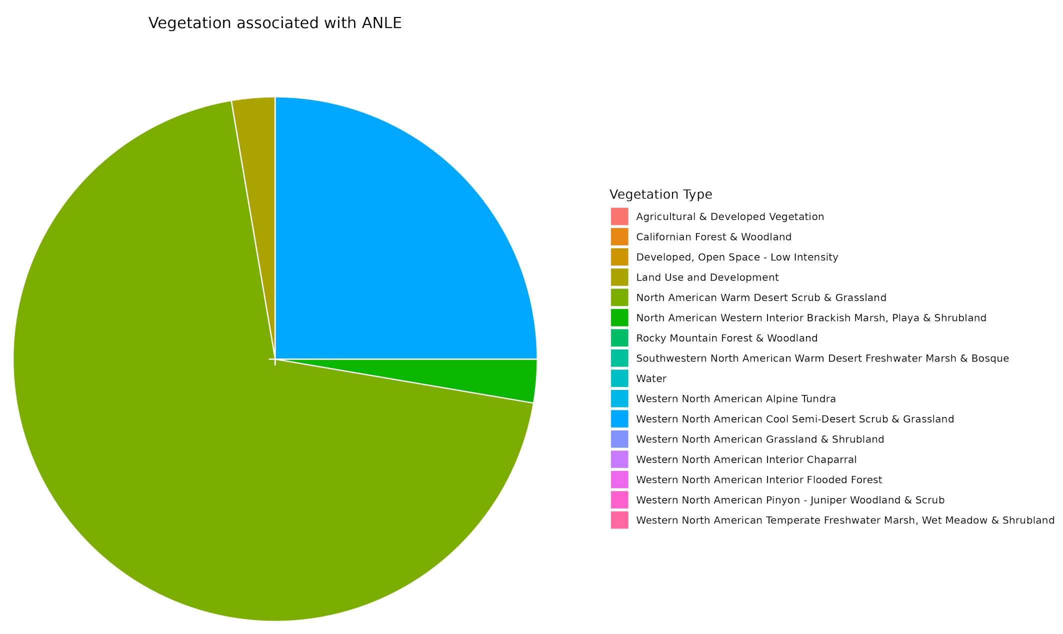


Figure 3. Vegetation types associated with ANLE Point locations.

ANLE points were largely located with in The North American Warm Desert Scrub & Grassland followed by Western North American Cool Semi-Desert Scrub & Grassland (Figure 3).

AQCH – Golden eagle

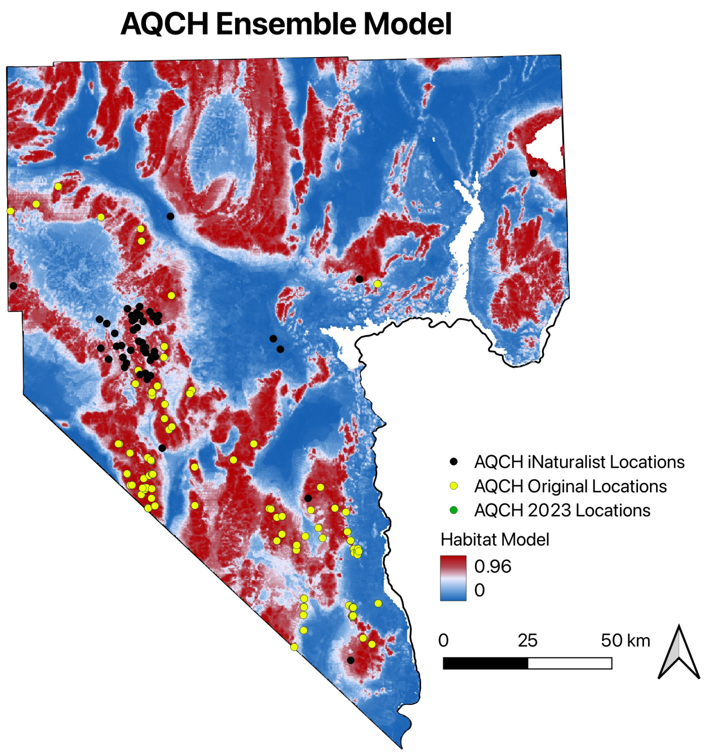


Figure 4. Ensemble Model for AQCH. Without the Vegetation Layer.

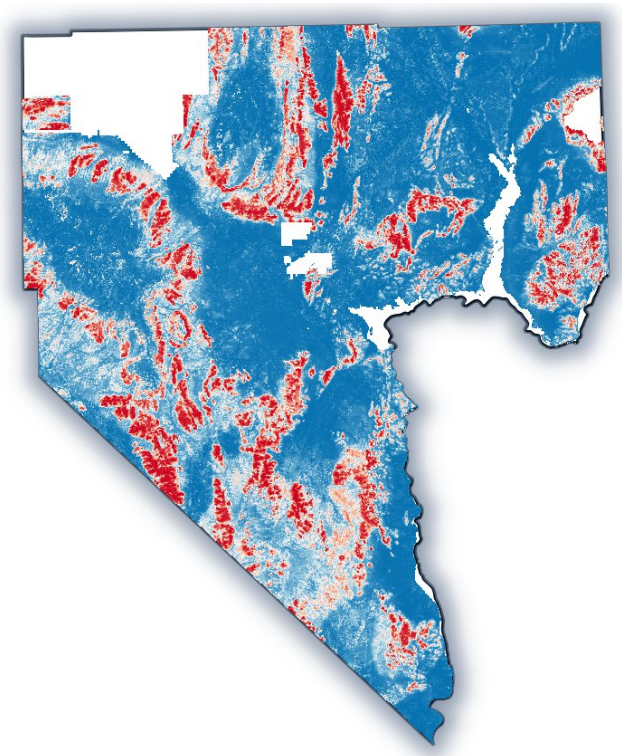


Figure 5. Ensemble Models for AQCH. Including the Vegetation Layer

The models for Golden Eagle nesting sites generally performed well, although the GBM and GAM models had a lower score for the Boyce index for both the training and especially testing datasets (Table 8). The models with vegetation include had very similar performance statistics (Table 9), but showed a far more restrained habitat prediction relative to the non vegetation models (Figures 4 and 5). The habit values in the "No Veg" models at the localities trended lower, (P - 0.01), where the models including vegetation had more points with max model values attributed (Figure 6).

Table 8. Performance metrics for AQCH. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC Training	AUC Testing	AUC All	BI Training	BI Testing	BI All	TSS Training	TSS Testing	TSS All
EM	1	0.98	1	0.98	0.91	0.98	0.97	0.89	0.94
GBM	1	0.98	0.99	0.79	0.53	0.9	0.96	0.89	0.92
RF	1	0.97	1	0.98	0.93	0.99	0.99	0.87	0.97
GAM	1	0.98	0.99	0.83	0.41	0.91	0.96	0.87	0.94
MX	0.98	0.96	0.97	0.94	0.95	0.96	0.84	0.78	0.82

Table 9. Performance metrics for AQCH models including vegetation. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC Training	AUC Testing	AUC All	BI Training	BI Testing	BI All	TSS Training	TSS Testing	TSS All
EM	1	0.93	0.99	0.98	0.84	0.94	1	0.79	0.96
RF	1	0.94	0.99	0.99	0.91	0.99	1	0.85	0.97
GBM	1	0.92	0.99	0.97	0.55	0.89	1	0.74	0.95

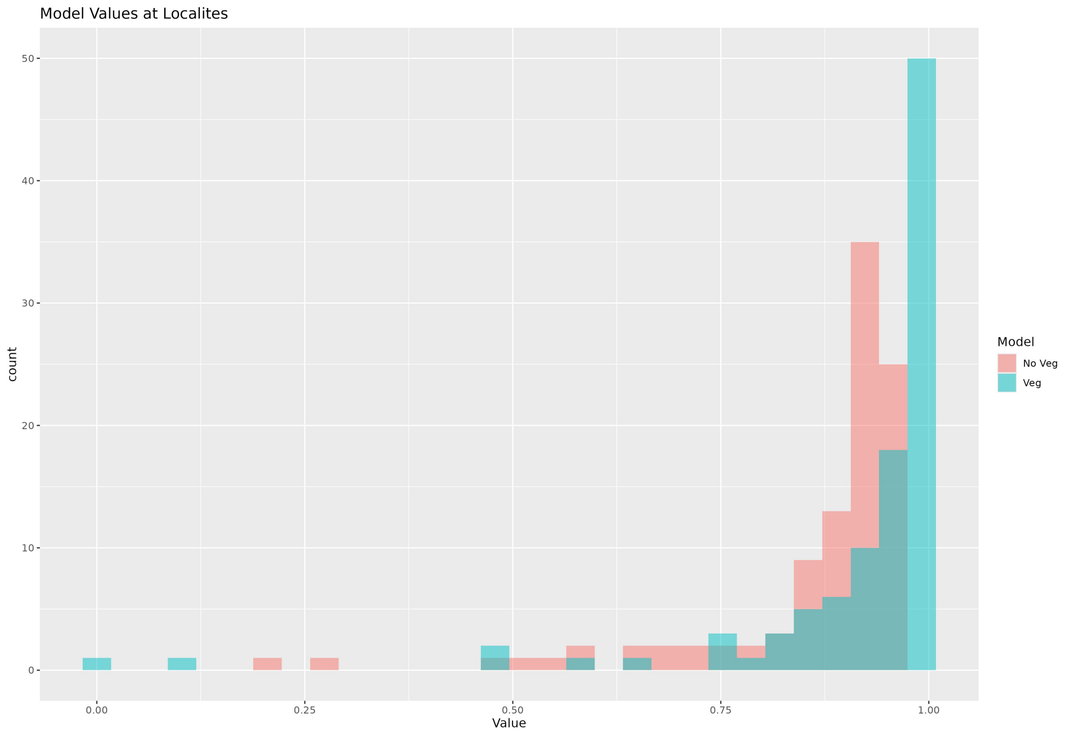


Figure 6. Modeled habitat values for the AQCH Ensemble Models with and without the inclusion of the Vegetation layer.

All of the algorithms generally had a good spread of variable inclusion (Table 10). Variable importance indicated that the Distance to Cliffs variable had the lowest contributions, which was opposite of our expectation. Each of the remaining input variables had contributions over 15% for at least one of the algorithms, and Average Spring Max Temp, and slope and the two temperature measures had contributions of 39% or higher. The models including vegetation had moderate dependence on the vegetation layer, with 8 and 7% importance, and the GBM models (N=20) showed increased importance of Topographic index, but with inclusion of all variables with the exception of Minimum Temperature, Distance to Cliffs, and Depth to Bedrock (Table 10). The Random forest models (N=20) had a more balanced inclusion of the environmental variables (Table 11).

Table 10. Relative importance of the input variables used in modeling for AQCH for the models without the vegetation layer.

Variable	GBM	RF	GAM	MX
Ave Min Temp	0	12.4	20.6	32.5
Average Spring Max Temp	30	22.5	18.8	28.8
Silt	0	6.2	16.1	12.3
Slope	40.2	21.7	16.1	4
Topographic Index	27.4	18.3	8.9	21.4
Depth to Bedrock	2.4	15.6	17.2	0
Distance to Cliffs	0	3.3	2.5	1

Table 11. Relative importance of the input variables used in modeling for AQCH for the models including the vegetation layer.

Variable	GBM	RF
Ave Min Temp	0	9.4
Average Spring Max Temp	6.6	16.3
Silt	10.7	13.2
Slope	11.5	15.1
Topographic Index	62.9	20.1
Depth to Bedrock	0.3	10.4
Distance to Cliffs	1	7.4
Vegetation	7	8.1

The ensemble model without vegetation was comprised of 18 Random Forest models, 5 GAM models, with 10 MaxEnt, and 11 GBM models contributing. The ensemble model for the models with vegetation included contained 13 RF and 2 GBM models.

Vegetation associated with AQCH localities was largely composed of North American Warm Desert Scrub & Grassland, and Western North American Cool Semi-Desert Scrub & Grassland (Figure 7).

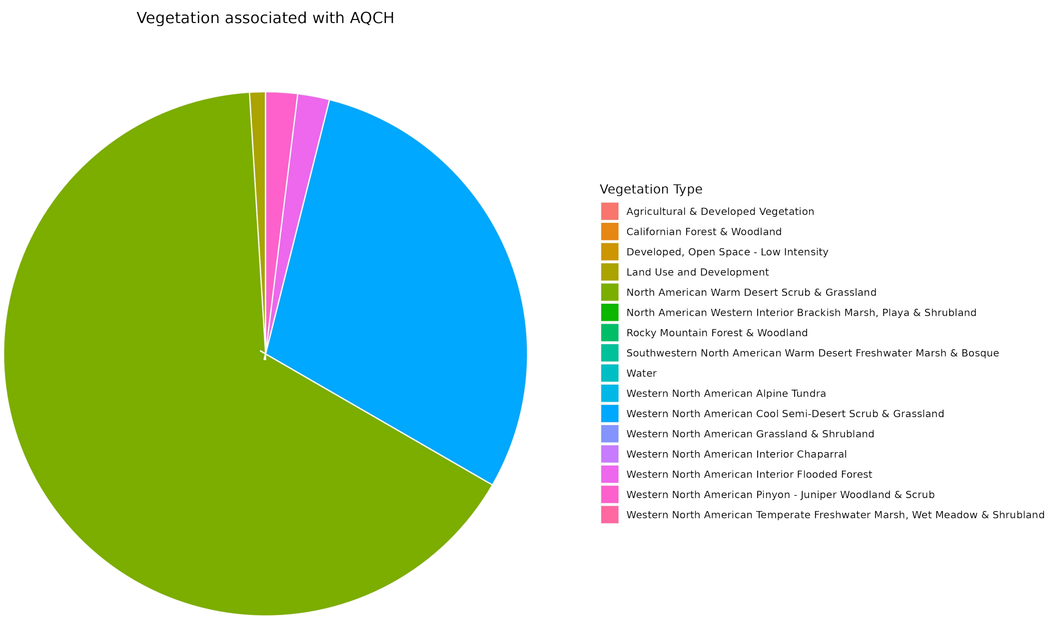


Figure 7. Relative frequency of vegetation associations at the locality point locations.

ARCA - Las Vegas Bearpoppy

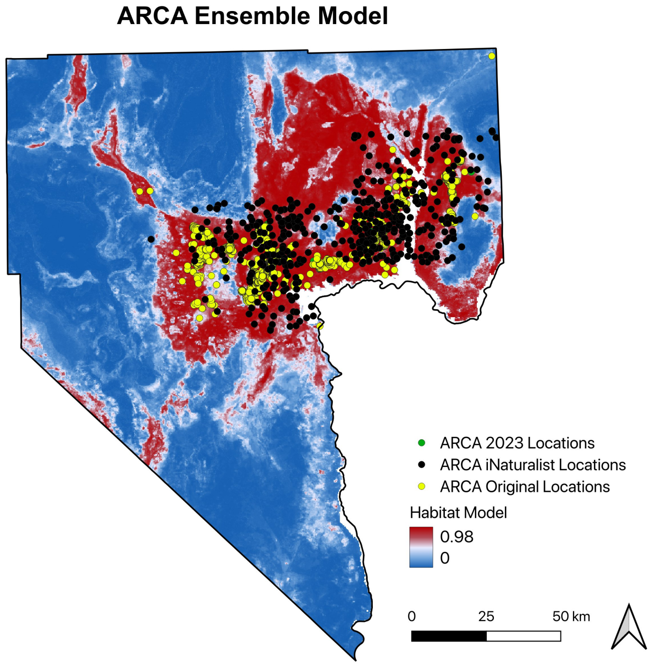


Figure 8 – Ensemble Model for ARCA

The Las Vegas Bearpoppy model had extremely high performance measures for both training and testing evaluations, and across all three performance metrics (Table 12). The models with the vegetation layer included had similarly high performance (Table 13), and the predicted habitat appeared to be very similar between models when comparing the areas that had vegetation information where prediction was possible (Figure 9).

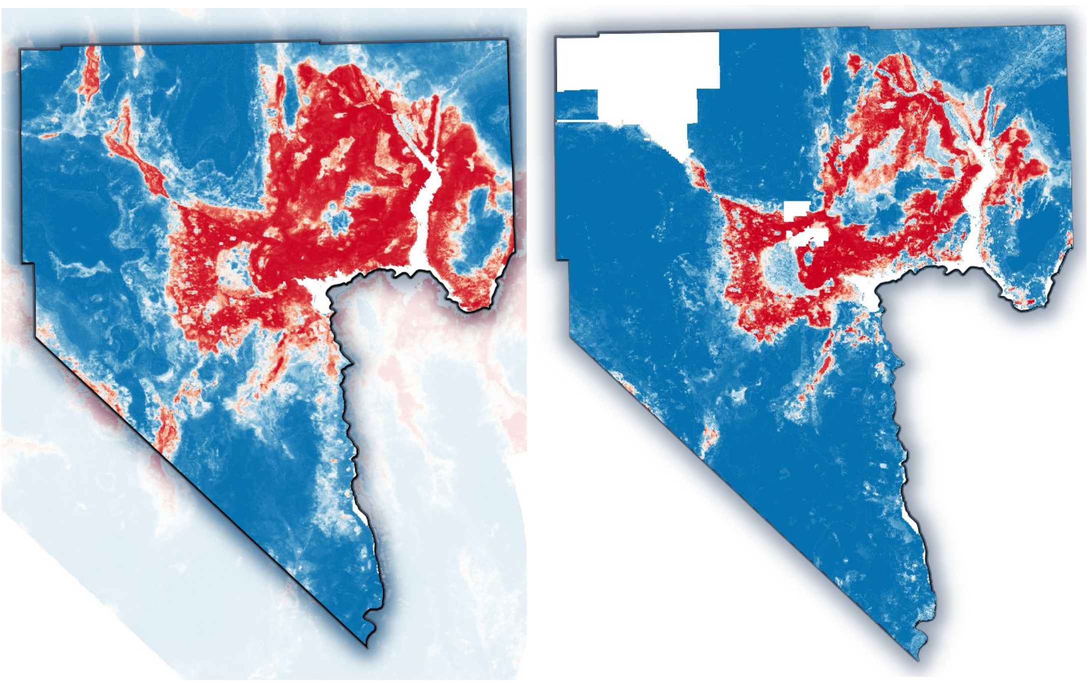


Figure 9. Ensemble model (Left), and the model with the vegetation layer included (Right) for ARCA.

Table 12. Performance metrics for ARCA. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.99	1	1	0.96	1	0.96	0.94	0.95
GAM	0.98	0.99	0.98	0.98	0.97	0.99	0.89	0.9	0.89
RF	1	0.99	1	0.99	0.93	0.97	1	0.95	0.98
MX	0.98	0.99	0.98	0.99	0.99	0.99	0.88	0.9	0.88
GBM	0.99	0.99	0.99	0.99	0.98	0.99	0.91	0.93	0.92

Table 13. Performance metrics for ARCA with the vegetation layer included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.99	0.95	0.97	1	0.93	0.99
RF	1	1	1	0.99	0.95	0.97	1	0.93	0.99

Model importance showed that each of the selected variables had high contribution in at least one of the models (e.g. silt in the GAM model), The Average Spring Max Temperature, and the Soil Gypsum content were among the highest contributors, although the MaxEnt models depended heavily on the variability in winter precipitation (Table 14).

Table 14. Relative importance of the input variables used in modeling for ARCA.

Variable	GAM	GBM	RF	MX
Average Spring Max Temp	21.5	39	26.2	8.5
Soil gypsum	8.7	56.3	28.9	9.8
NDVI Amplitude	18.7	1.8	14	1.3
Silt	20.9	0	7.4	2.3
CV Winter Precip	30.3	2.9	23.5	78.1

The high performing models with vegetation included only those using Random Forest. The RF model showed a relatively even inclusion of the Environmental variables, however, the vegetation layer showed little importance to the models (Table 15) (Table 15).

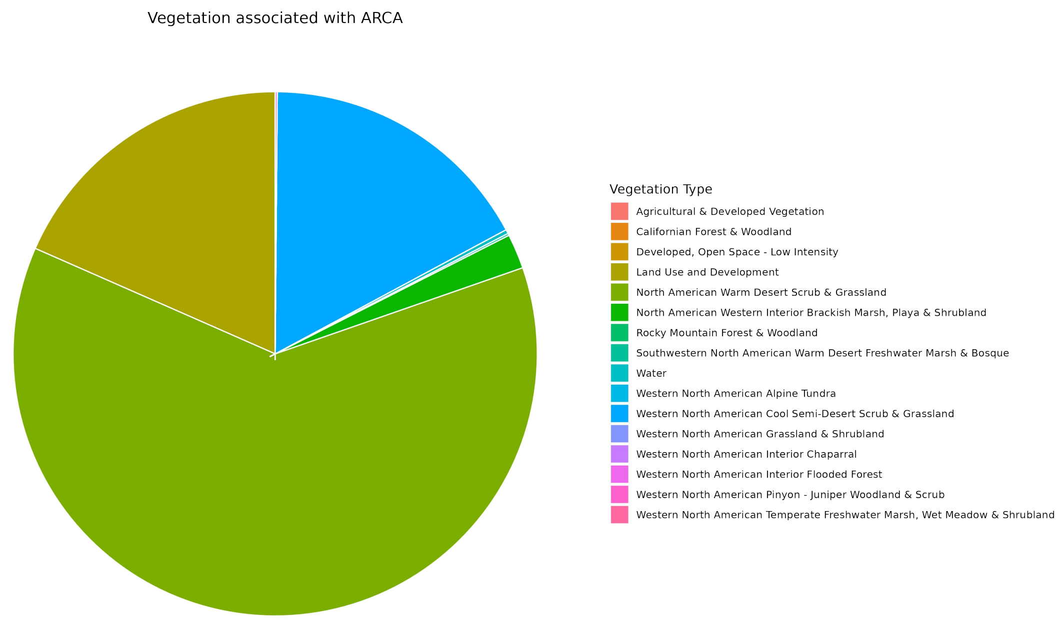


Figure 10. Relative frequency of vegetation associations at the locality point locations.

Table 15, Relative importance of the input variables used in modeling for ARCA with the vegetation layer included.

Variable	RF
Average Spring Max Temp	23.70
Soil gypsum	33.06
NDVI Amplitude	19.52
Silt	8.58
CV Winter Precip	13.54
Veg14	1.60

ARCA locations were largely located within North American Warm Desert Scrub & Grassland , with lower presence in Western North American Cool Semi-Desert Scrub & Grassland and lands classified as Development (Figure 10).

There were 19 random forest, 6 GAM, 10 gbm, and 5 maxent models that contributed to the ensemble model. The vegetation based models included20 RF models.

ASGETR – Three Corner Milkvetch

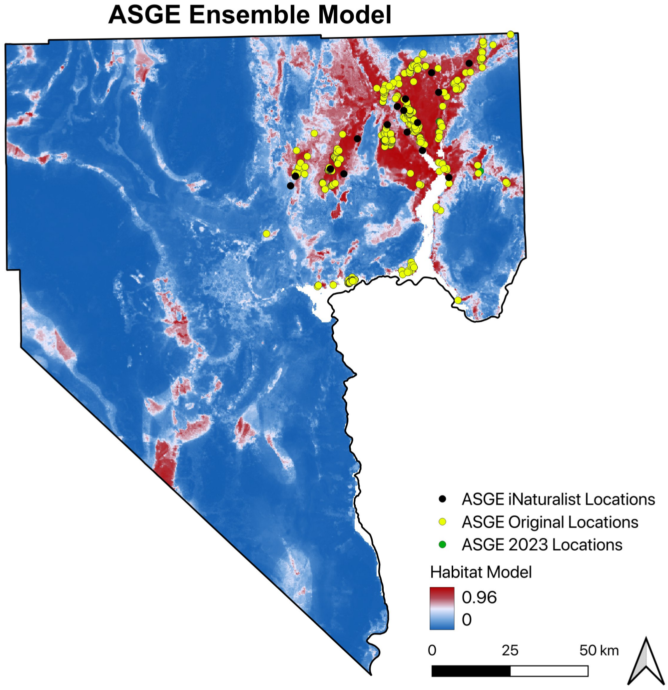


Figure 11 – Ensemble Model for ASGETR – without the vegetation layer included.

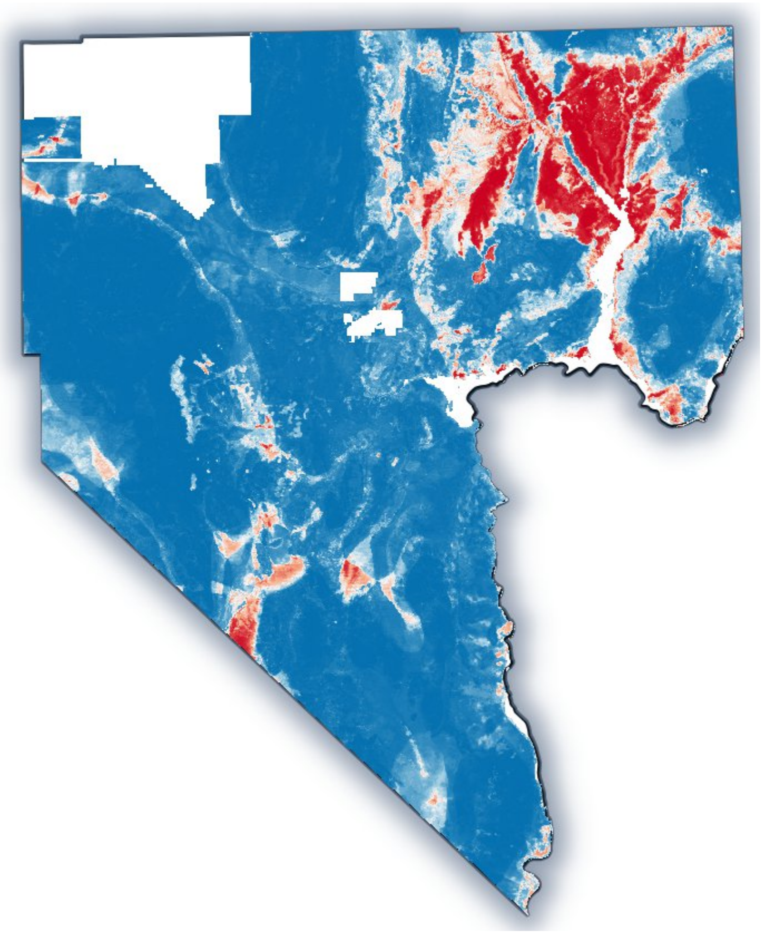


Figure 12 – Ensemble Model for ASGETR with the vegetation layer included.

The Ensemble model for ASGETR had the majority of habitat predicted for the area in the northeastern extent of Clark County – especially in the Virgin/Muddy River area, Mormon Mesa, and Moapa (Figure 11). The models with the vegetation layer included had very similar habitat projections (Figure 12).

Performance was mixed for ASGETR – with very high AUC/BI, and TSS scores for some algorithms, and with poor performance in others (e.g. GAMs, Table 16). Performance was similarly high for AUC and TSS in the models where the vegetation layers were included (Table 17). However, the Boyce index for the testing data had much lower scores (Table 17).

Table 16. Performance metrics for ASGETR. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.99	0.95	0.99	1	0.93	0.98
GAM	1	0.99	1	NA	-1	-1	0.99	0.96	0.98
RF	1	0.99	1	0.98	0.84	0.97	1	0.94	0.99
MX	0.99	0.99	0.99	0.99	0.95	0.99	0.93	0.94	0.93
GBM	1	0.99	1	0.84	0.74	0.93	1	0.93	0.98

Table 17. Performance metrics for ASGETR with the vegetation layer included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.96	0.57	0.89	1	0.98	0.99
RF	1	1	1	0.99	0.8	0.98	1	0.97	0.99
GBM	1	1	1	0.82	0.31	0.9	0.99	0.97	0.98

Variable importance indicated that all of the 6 variables considered had > 10 percent importance for at least one algorithm. The MaxEnt appeared to have a poor fit, despite high performance metrics, as it essentially relied only on the Silica Index (Table 18). This variable also performed high in the GBM model, which also considered winter precipitation, while the GAM and RF algorithms had more even consideration of the variables used for modeling. In the models including vegetation the GBM model included only the Silica Index, with ~ 1% inclusion of the winter precipitation layer. The RF model had 4% importance attributed to the vegetation layer, with larger importance attributed to several other variables (Table 19).

Table 18. Relative importance of the input variables used in modeling for ASGETR

Variable	GAM	GBM	RF	MX
Winter precipitation	13.5	8.7	22.2	0
Winter minimum temperature	26.7	0	10.9	0
NDVI amplitude	11	1.9	13.7	0
Slope	16.1	0	4.6	0.2
Silica index	17.1	89.4	37.6	99.7
Sandy soils	15.8	0	11.1	0.1

Table 19. Relative importance of the input variables used in modeling for ASGETR with vegetation models included.

Variable	GBM	RF
Winter precipitation	1.1	20.8
Winter minimum temperature	0	6.5
NDVI amplitude	0	9.3
Slope	0	7.2
Silica index	98.9	41.1
Sandy soils	0	11.1
Vegetation	0	4

The ASETR ensemble model was comprised of 14 Random Forest, 11 Maxent, 14 GBM, and 1 GAM model (which explains the poor performance metrics for GAM). The models including the vegetation layer were composed of 14 RF and 10 GBM models.

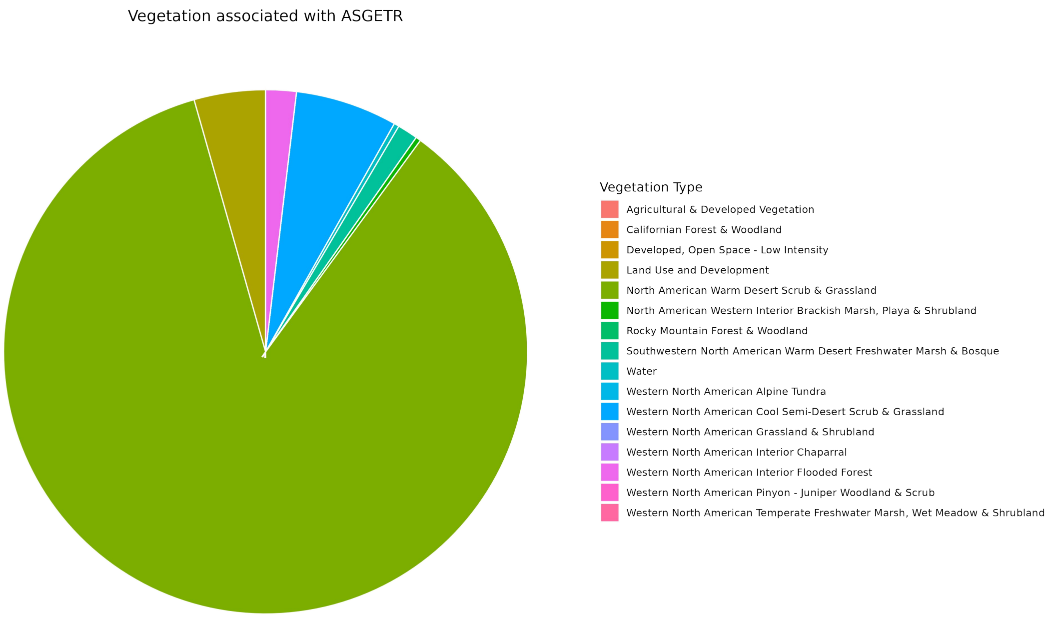


Figure 13. Relative frequency of vegetation associations at the locality point locations for ASGETR.

Vegetation associated with ASGETR localities was largely within North American Warm Desert Scrub & Grassland (Figure 13).

ATCU – Burrowing Owl

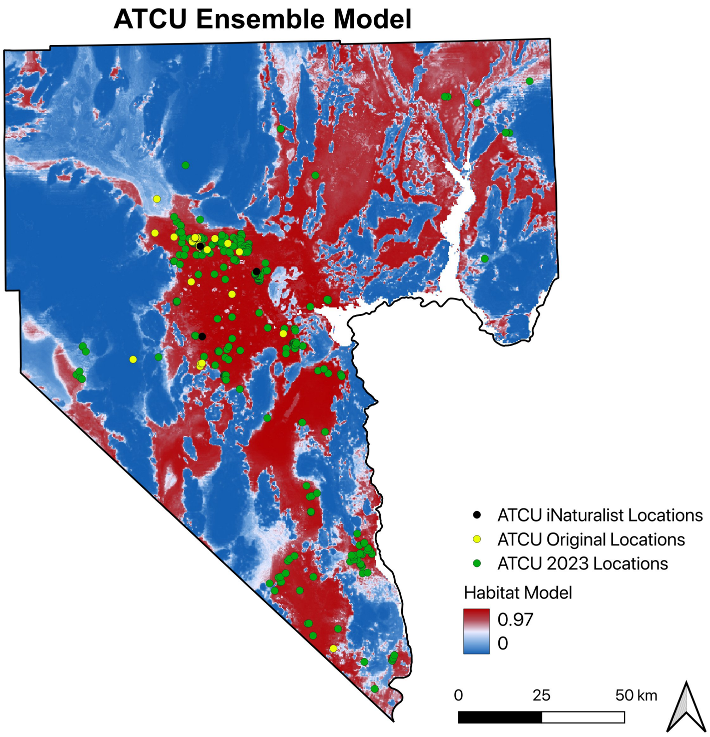


Figure 14 – Model for ATCU without vegetation included.

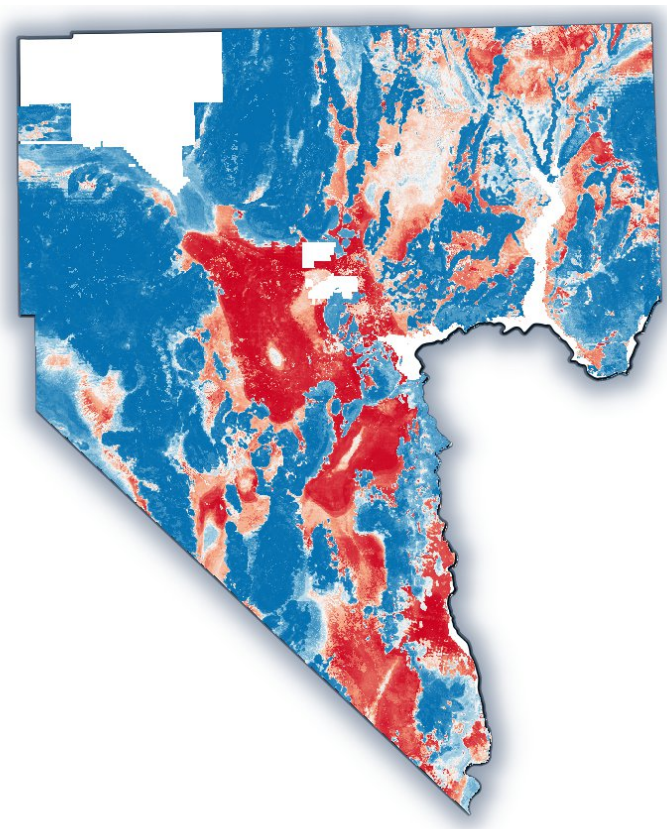


Figure 15 – Model for ATCU with the vegetation layer included.

The Burrowing Owl model showed a similar habitat prediction for both the models with and without vegetation (Figures 15 and 14 respectviely). Model performance for the burrowing owl models was mixed, with high performance for training data in all algorithms, but with lower testing performance for the Boyce index for the GAM model, and with lower TSS for testing sets in the GAM and Maxent models (Table 20). The GBM model had lower performance for the models including vegetation (Table 21), with lower AUC, and BI for the Testing dataset. The RF model performed well for most metrics (Table 21).

Table 20. performance metrics for ATCU. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.95	0.99	0.88	0.68	0.94	0.98	0.8	0.93
GAM	0.99	0.92	0.98	0.76	0.32	0.75	0.9	0.73	0.85
RF	1	0.96	0.99	0.81	0.72	0.93	1	0.86	0.95
MX	0.98	0.92	0.97	0.9	0.76	0.9	0.9	0.73	0.85
GBM	1	0.95	0.99	0.89	0.65	0.93	0.98	0.84	0.94

Table 21. performance metrics for ATCU models including the vegetation layer. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.95	0.99	0.97	0.9	0.98	1	0.86	0.97
RF	1	0.95	0.99	0.99	0.9	0.96	1	0.86	0.97
GBM	1	0.94	0.99	0.96	0.67	0.96	0.94	0.79	0.91

Variable importance indicated high inclusion of most variables, although with different importance among algorithms (Table 22). The GBM model essentially modeled using three of the six variables, while the MaxEnt model largely relied on only three. The models including vegetation had moderate importance attributed to vegetation (Table 23). The GBM model was largely slow based, while the RF had a more balanced importance among the environmental layers, and 10% attributed to vegetation (Table 23).

Table 22. Relative importance of the input variables used in modeling for ATCU

Variable	GAM	GBM	RF	MX
Winter precipitation	7.3	32.6	26.8	1.3
Winter minimum temperature	23.6	20.7	26.4	39
NDVI amplitude	17.3	0	6.9	13.8
Slope	17.9	46.7	32.8	43.5
Coarse Fragments	33.9	0	7.1	2.5

Table 23. Relative importance of the input variables used in modeling for ATCU with the vegetation layer included

Variable	GBM	RF
Winter precipitation	4.5	20.8
Winter minimum temperature	8.1	17.1
NDVI amplitude	0	6.9
Slope	85.4	37
Coarse Fragments	0	8
Vegetation	1.9	10.1

Localities for Burrowing Owls had the highest association with North American Warm Desert Scrub & Grassland, and interestingly Land Use and Development was the second highest, followed by Western North American Cool Semi-Desert Scrub & Grassland (Figure 16). Several other vegetation types had limited association with the localities.

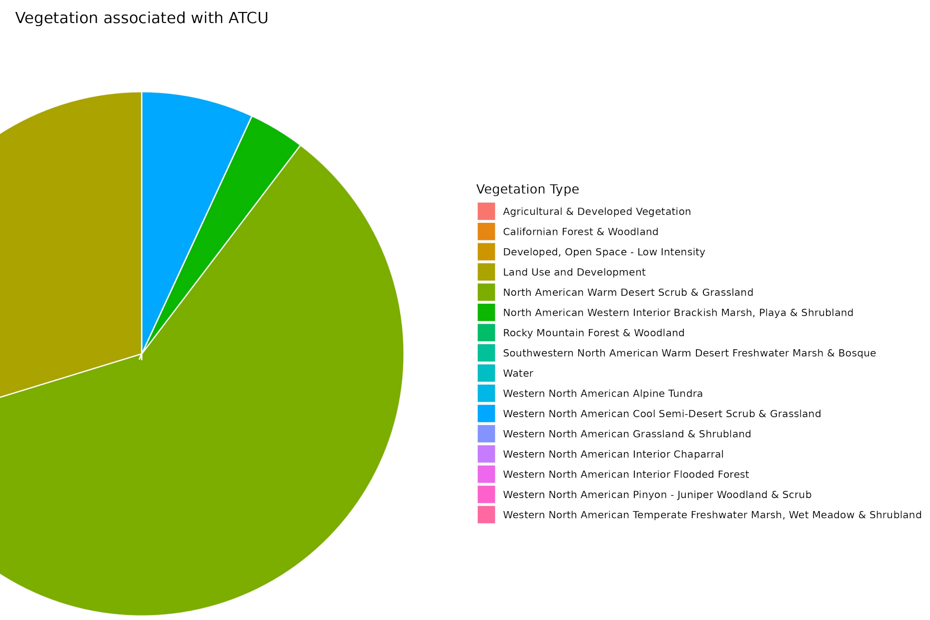


Figure 16. Relative frequency of vegetation associations at the locality point locations for ATCU.

The model contribution toward the ensemble model was comprised of 17 random forest models,7 MaxEnt models, 5 GAM models, and 13 GBM models. The models using vegetation included 20 RF, and 3 GBM models.

CHPE – Desert Pocket Mouse

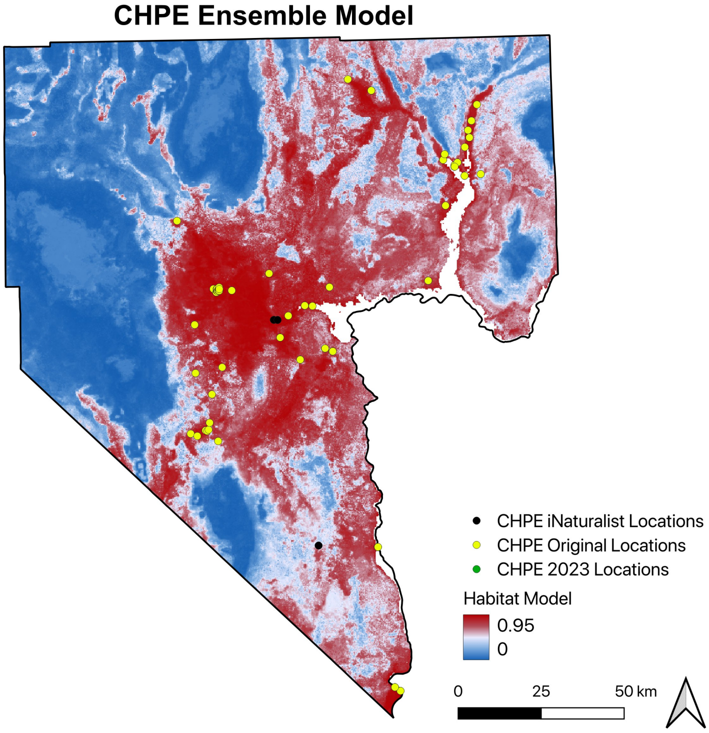


Figure 17 – Ensemble Model for CHPE – for models not including the vegetation layer.

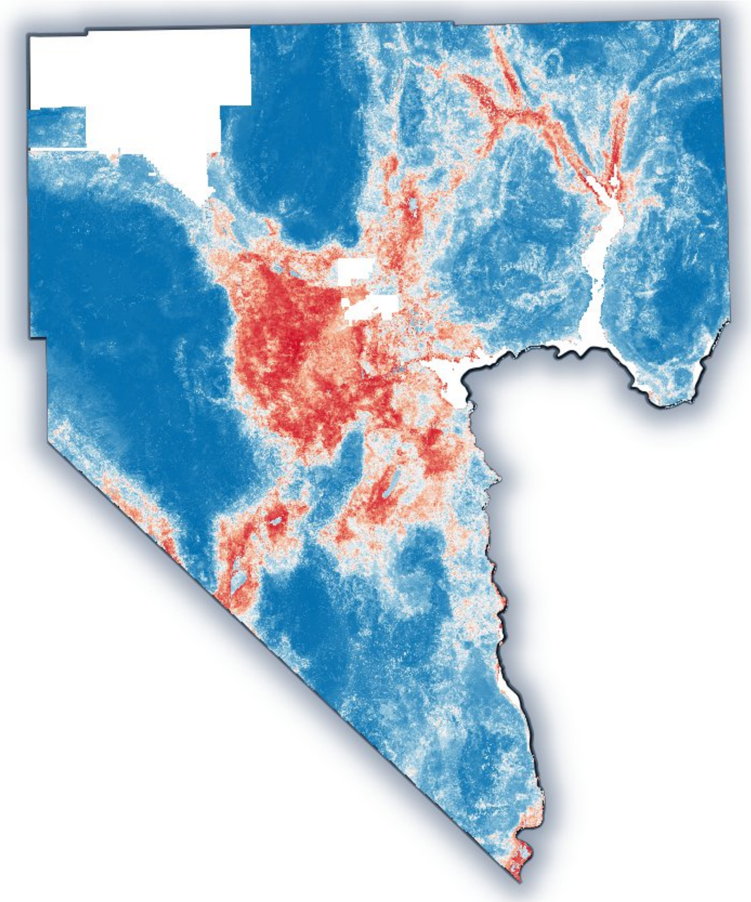


Figure 18 – Ensemble Model for CHPE – for models including the vegetation layer.

The models including the vegetation layer for CHPE had similar predicted habitat relative to the base models (Figures 17 and 18). There were notable reductions west of the confluence of the Muddy and Virgin rivers, and a significant retraction in the southern part of the county (Figures 17 and 18). There was also a reduced level of prediction throughout the county, although the overall suitability levels were not different between models (p= 0.85).

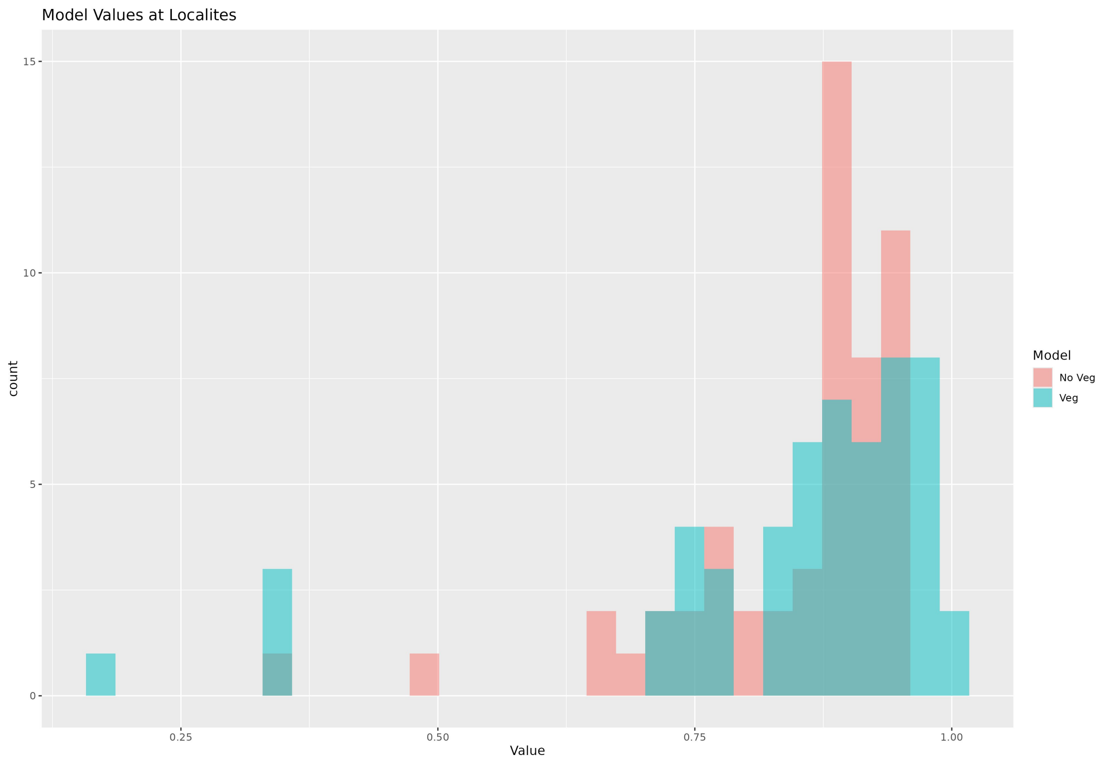


Figure 19. Model values for the Ensemble models with and without vegetation included.

The base models for the Desert Pocket Mouse generally performed well, although the GAM model had lower scored for the Boyce index (Training and Testing), and the GBM model had poor performance in the testing set for the Boyce Index (Table 24). The models including vegetation had similarly high performance,the BI for the testing data had poor performance (Table 25).

Table 24. performance metrics for CHPE models without vegetation. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.97	1	0.99	0.78	0.85	1	0.88	0.97
gam	1	0.96	0.99	0.29	0.47	0.47	0.98	0.81	0.95
rf	1	0.96	1	0.97	0.89	0.92	1	0.81	0.96
mx	0.99	0.95	0.98	0.93	0.94	0.97	0.87	0.81	0.86
gbm	1	0.96	1	0.84	0.26	0.78	1	0.81	0.96

Table 25. performance metrics for CHPE models including vegetation. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.94	1	0.99	0.3	0.98	1	0.78	0.94
RF	1	0.93	1	0.98	0.45	0.97	1	0.78	0.96
GBM	1	0.94	0.99	0.98	0.45	0.93	1	0.78	0.94

Model variable importance indicated that Silt, average minimum temperature, and the variability in winter precipitation were all important in at least one model. Despite being used in previous models the NDVI measures of start of season (an NDVI measure of plant growth initiation), and the peak value did not contribute importantly in any algorithm, and percent of clay also contributed minimally (Table 26) .Of the remaining variables Average Minimum Temperature and the CV in winter precipitation had the highest contributions, although it should be noted that the MaxEnt model again relied heavily on only one variable, which typically indicates poor model fit. The models including the vegetation layer had high importance attributed to vegetation (Table 27) – where the GBM model importance shifted away from Average Minimum Temperature, and Silt, and the importance for variables within the RF model remained more balanced (Table 27).

Table 26. Relative importance of the input variables used in modeling for CHPE

Variable	GAM	GBM	RF	MX
Winter Precip	2.7	3.8	13.2	0.3
Start of Season (day)	0	0	4.7	0.3
PPT Clay	4.1	1.2	7.3	1
CV Winter Precip	14.5	1.9	11.8	90.6
PCT Coarse frags	18	5	10.4	0.3
Ave Min Temp	30	68.4	28.2	5.9
NDVI Max	4.6	0.6	7.6	0.1
PPT Silt	26.1	19.1	16.7	1.6

Table 27. Relative importance of the input variables used in modeling for CHPE for models including the vegetation layer.

Variable	GBM	RF
Winter Precip	53.5	21.9
Start of Season (day)	0	6.2
PPT Clay	0	6.4
CV Winter Precip	1.4	9.4
PCT Coarse frags	3	9.8
Ave Min Temp	9.3	11.7
NDVI Max	1	8.6
PPT Silt	0.4	9
Vegetation	31.3	17

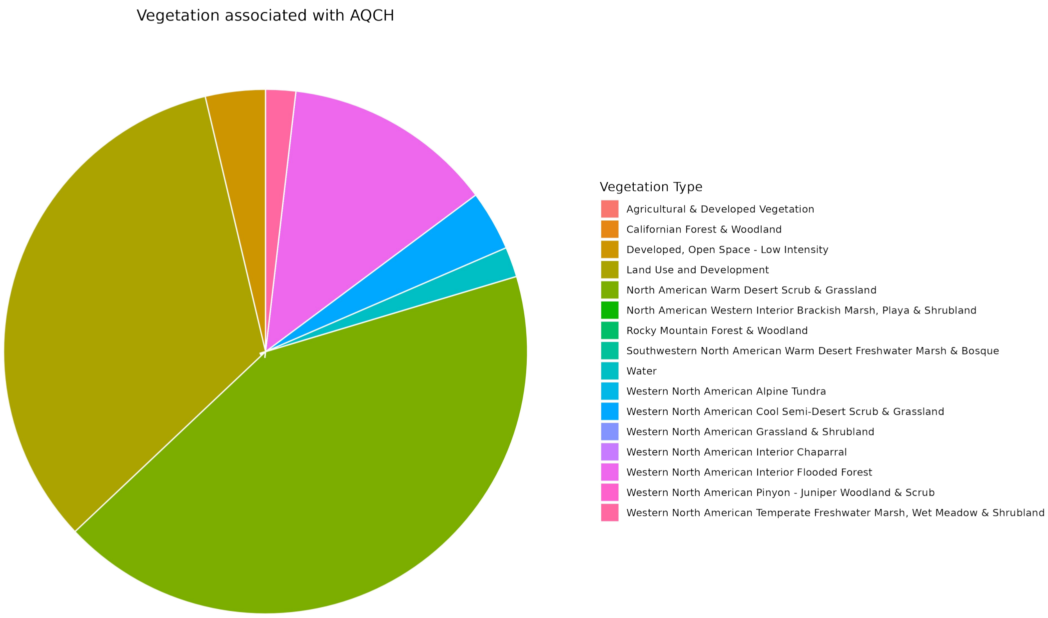


Figure 20. Relative frequency of vegetation associations at the locality point locations for CHPE.

Vegetation associated with CHPE localities was generally comprised of North American Warm Desert Scrub & Grassland, Land Use and Development, and Western North American Interior Flooded Forest to a lesser extent (Figure 20).

Contributions of algorithms toward the ensemble model consisted of 19 random forest models, 11 MaxEnt models, 2 GAM models, and 10 GBM models. Ensemble Models including vegetation consisted of 13 RF and 12 GBM models.

COCH – Gilded Flicker

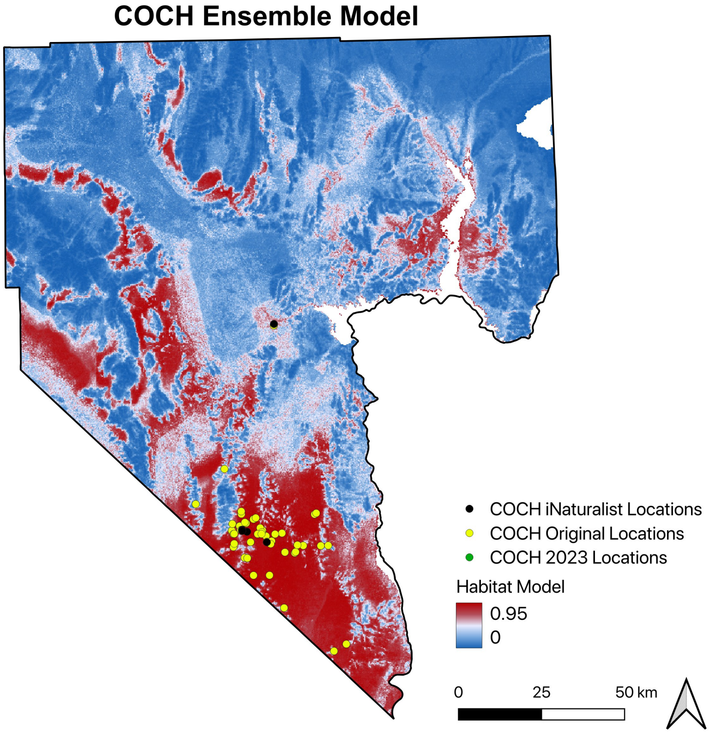


Figure 21 – Ensemble Model for Gilded Flicker

Predicted habitat differed somewhat between the base Ensemble Models (Figure 21), and those with vegetation included (Figure 22). The models withe vegetation included had a reduction in predicted habitat in the northern portions of the county, with predictions in the south – where most localities were – remaining similar between the models.

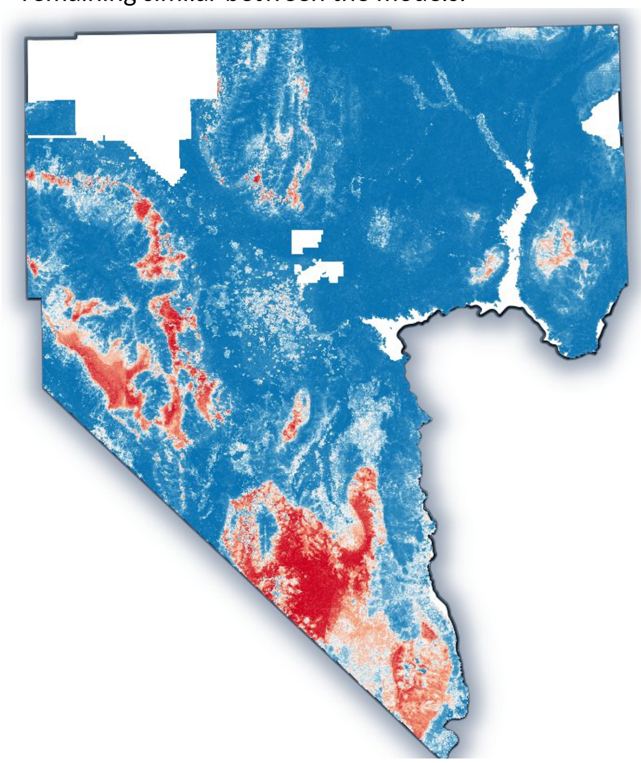


Figure 22. Ensemble Model with vegetation included for Gilded Flicker

Model performance was high for all algorithms with an extremely low score for the Boyce index for the testing dataset (Table 28). All other algorithms were very high in all areas. The models including vegetation had similar performance metrics (Table 29).

Table 28. performance metrics for COCH. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.96	0.99	0.95	0.79	0.98	1	0.81	0.95
gam	0.99	0.94	0.98	0.9	0.13	0.93	0.93	0.81	0.9
rf	1	0.98	1	0.98	0.8	0.99	1	0.87	0.97
mx	0.97	0.91	0.96	0.89	0.85	0.87	0.87	0.81	0.86
gbm	1	0.94	0.99	0.77	0.84	0.92	0.99	0.78	0.95

Table 29. performance metrics for COCH for models including vegetation. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data, and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.99	1	0.98	0.8	0.97	1	0.91	0.98
RF	1	1	1	0.92	0.88	0.97	1	1	0.98
GBM	1	0.99	1	1	0.8	0.83	1	0.91	0.98

Variable importance indicated again that NDVI measures had the lowest contributions. Maxent again focused on a single variable model, and GBM largely on two, while the GAM and Random forest had more even consideration of input variables (Table 30). Winter precipitation, and topography measures had high contributions. For the models including vegetation there was high importance associated with that layer as well (Table 31). The GBM models were essentially driven by 3/4 input variables, while the RF models were more balanced, similarly to its performance among other species modeled herein.

Table 30. Relative importance of the input variables used in modeling for COCH

Variable	GAM	GBM	RF	MX
Dist to cliffs	18.4	0	4.6	0
NDVI Length of Season	4.5	0	5.5	0
NDVI Max	3.4	0.6	8.6	0
Winter Precip	27.6	11.1	17.7	0.1
CV Winter Precip	33.3	80	35	98.7
Slope	7.5	4	15.4	1.2
Flow Accum	5.4	4.4	13.1	0

Table 31. Relative importance of the input variables used in modeling for COCH for models including vegetation.

Variable	GBM	RF
Dist to cliffs	0	3.1
NDVI Amplitude	7.6	14.9
NDVI Length of Season	0	3.2
NDVI Max	0	6.4
Winter Precip	45.7	26.8
CV Winter Precip	25.8	18.5
Slope	5	9.2
Flow Accum	1.1	5.9
Vegetation	14.8	12

The ensemble model was comprised of 15 Random Forest models, 14 GBM models, with 9 MaxEnt, and 4 GAM models contributing. The vegetation ensemble model consisted of 20 each of RF and GBM models.

The vegetation associated with the localities for COCH was largely within Western North

American Cool Semi-Desert Scrub & Grassland, and North American Warm Desert Scrub & Grassland (Figure 23).

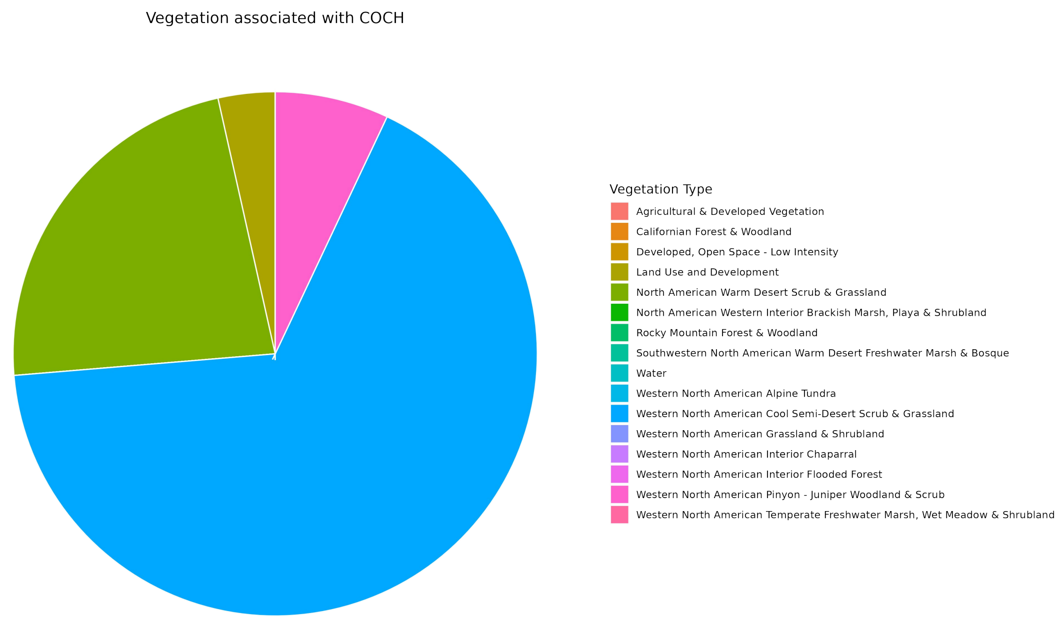


Figure 23. Relative frequency of vegetation associations at the locality point locations for COCH.

ENAR – Silverleaf sunray

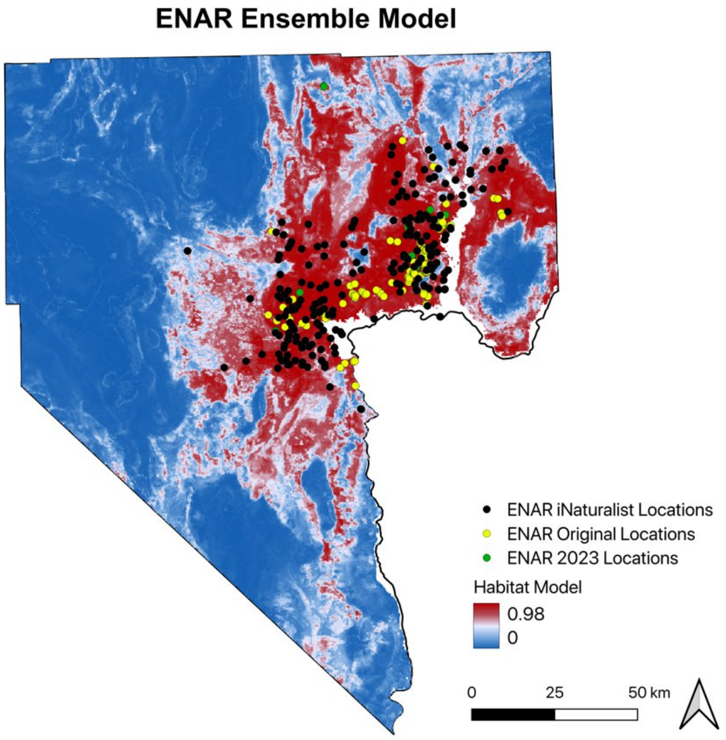


Figure 24. Ensemble Model for ENAR

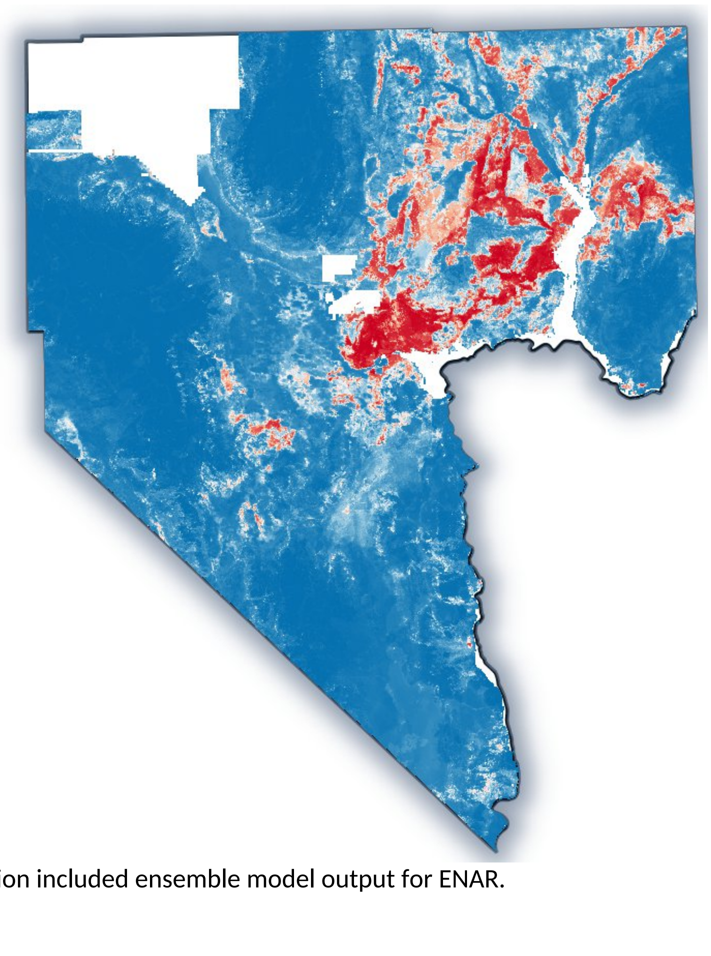


Figure 25. Vegetation included ensemble model output for ENAR.

Models for ENAR have very similar appearance for based model, and the vegetation inclusion model with predicted range reduction in southern portion of the county (Figures 24 and 25). The models generally performed well, although for the base models the GBM model had a lower score for the Boyce index (testing dataset), and the GAM model had slightly poorer performance in the testing set for the Boyce Index compared to other algorithms (Table 32). The models including vegetation also had high performance (Table 33).

Table 32. Performance metrics for the base models for ENAR. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.99	1	0.99	0.95	0.99	0.97	0.88	0.95
GAM	0.99	0.97	0.98	0.98	0.81	0.96	0.89	0.81	0.87
RF	1	0.99	1	1	0.87	0.93	1	0.89	0.98
MX	0.98	0.96	0.97	1	0.97	1	0.84	0.81	0.82
GBM	1	0.98	0.99	0.95	0.63	0.93	0.95	0.87	0.94

Table 33. Performance metrics for the vegetation inclusion models for ENAR. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.91	0.77	0.79	1	0.89	0.98
RF	1	0.98	1	0.88	0.84	0.82	1	0.89	0.98
GBM	1	0.93	0.99	0.99	0.95	0.99	1	0.89	0.98

Variable importance indicated that NDVI measures had the lowest contributions. The Maxent algorithm was based on a single variable (topographic roughness) indicating potentially poor model fit, and GBM largely on two variables (winter minimum temperature and gypsum potential), while the GAM and Random forest had more even consideration of input variables (Table 4). Winter precipitation, winter minimum temperature, and gypsum potential measures had high contributions towards GAM and Random forest models. Models including vegetation showed vegetation as an important component for the RF, but not the GBM models (Table 35). The GBM model relied heavily on Vegetation, Winter Minimum Temperature and Gypsum, while the RF had a more balanced importance across the variables, with Vegetation and Winter Minimum Temperature as the highest importance (Table 35). The GBM models were largety drive by Gypsum potential, with limited influence of NDVI and Minimum temperatures.

Table 34. Relative importance of the input variables used in base modeling for ENAR.

Variable	GAM	GBM	RF	MX
Winter min temperature	19.1	50	33.9	2.8
Gypsum potential	16.1	48.4	26.2	1.9
NDVI maximum	10.4	0.3	10.4	0.1
Surface texture (ATI)	13.2	0	6.6	0.1
Winter precipitation	27.4	1.4	14.7	0.1
Roughness (TRI)	13.8	0	8.2	95.1

Table 35. Relative importance of the input variables used in modeling for ENAR with the vegetation layer included..

Variable	GBM	RF
Winter min temperature	4.9	15.3
Gypsum potential	87.8	35.3
NDVI maximum	6.7	17.5
Surface texture (ATI)	0	10.7
Winter precipitation	0	6.6
Roughness (TRI)	0	7.6
Vegetation	0.6	7.1

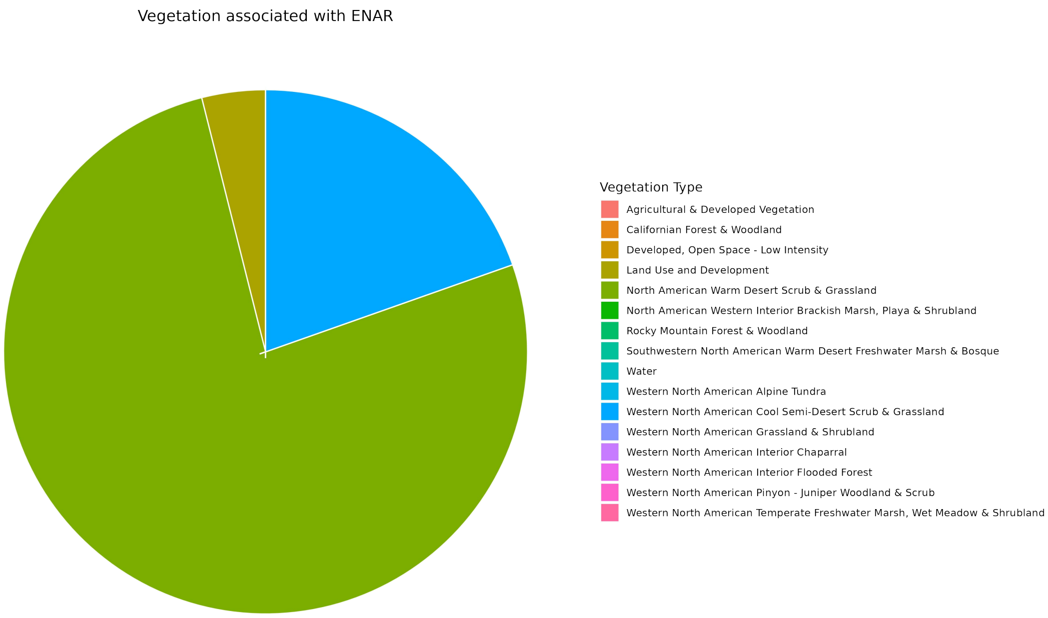


Figure 26. Relative frequency of vegetation associations at the locality point locations for ENAR.

The vegetation associated with ENAR localities was largely North American Warm Desert Scrub & Grassland, with smaller elements Western North American Cool Semi-Desert Scrub & Grassland and other elements (Figure 26).

The ensemble model was comprised of 17 Random Forest models, 7 GAM models, with 4 MaxEnt, and 12 GBM models contributing. For the vegetation inclusion models there were 20 RF and 20 GBM models used in the ensemble.

ERBI – Pahrump Valley buckwheat

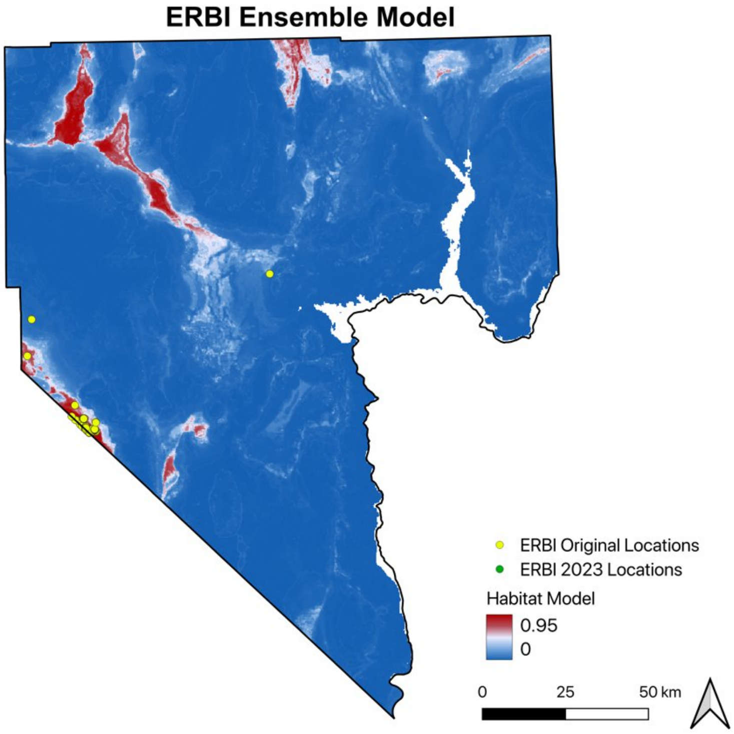


Figure 27. Ensemble base Model for ERBI.

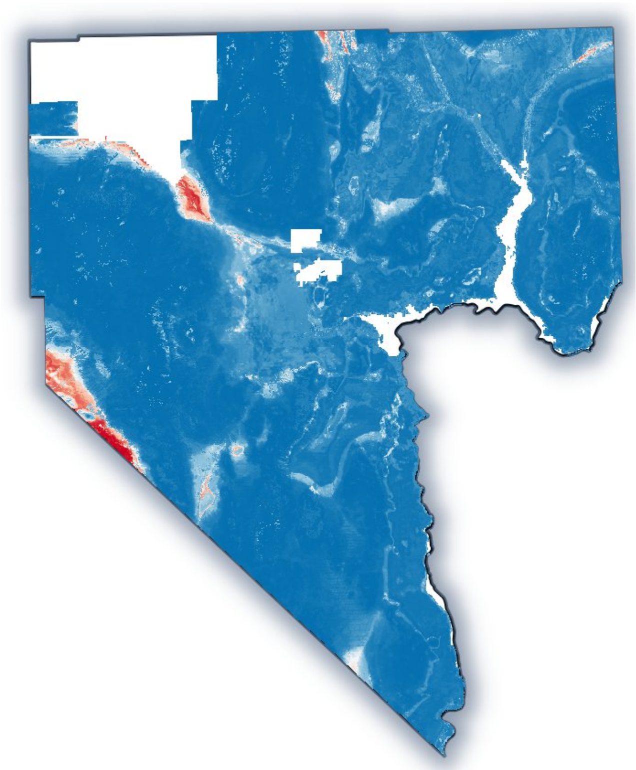


Figure 28. Ensemble base Model for ERBI.

The models for Pahrump Valley buckwheat predict similar habitat footprints (Figures 26 and 27). The ensemble model including vegetation also predicts habitat near the Corn Creek area (Figure 27), suggesting suitable habitat along the US 95 corridor as is seen in the models without vegetation (Figure 26), but the vegetation model only includes a portion of this area in the plant classification conducted to date. The models generally performed well, although the GBM in the base model had a lower score for the Boyce index (testing and training dataset), and the GAM model had much poorer performance in the testing set for the Boyce Index compared to other algorithms in the base models (Table 36). The models including vegetation showed high performance with the exception of the Boyce index for the GBM models, which had low values for both the training and testing data (Table 37). This is reflected in the Boyce Index Curves shown for this model (Figure 28), although it should be noted that the ensemble model BI remains high (Table 37), and has an excellent BI curve indicating good model discrimination (Figure 29).

Table 36. Performance metrics for ERBI. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.99	1	0.95	0.95	0.98	1	0.94	0.98
GAM	1	0.99	1	0.55	0.82	0.89	0.98	0.88	0.96
RF	1	0.99	1	0.74	0.9	0.9	1	0.94	0.99
MX	1	0.99	0.99	0.96	0.92	0.95	0.93	0.94	0.93
GBM	1	0.99	1	0.06	0.88	0.77	1	0.94	0.99

Table 37. Performance metrics for the vegetation inclusion models for ERBI. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.97	0.83	0.87	1	1	1
RF	1	1	1	0.78	0.77	0.91	1	1	0.98
GBM	1	1	1	0.91	0.29	0.82	1	1	1

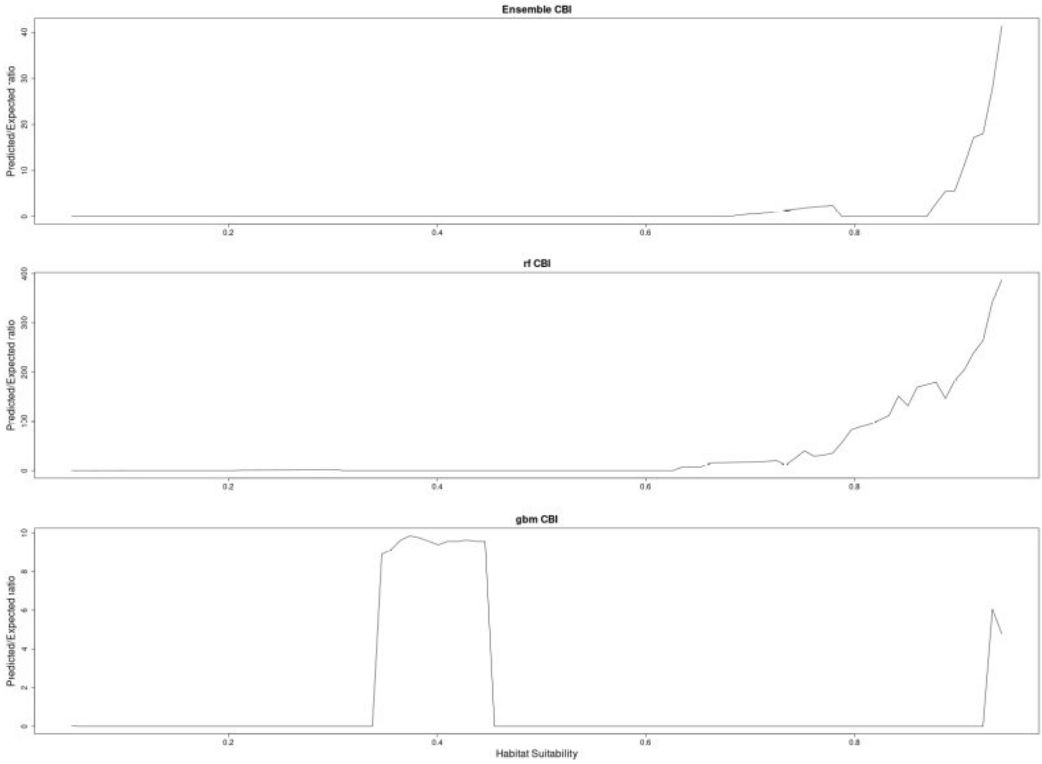


Figure 29. Continuous Boyce Index plots for the overall Ensemble vegetation model (Top), and the Random Forest (RF) models and GBM models used in the ensemble.

Variable importance for the base models indicated that the percent silt covariate had the lowest contributions across most models. Average maximum temperature, extreme minimum temperature, and slope had high contributions and were important across several algorithms (Table 38). For the models including vegetation the GBM shifted moderate importance to the vegetation layer, Slope, and the Extreme Minimum Temperature layer – with 0 importance for other layers (Table 39). The Random Forest model for the vegetation approach showed a 11% importance for the vegetation layer, with importance attributed to all other variables of at least 5% (Table 39).

Table 38. Relative importance of the input variables used in modeling for ERBI.

Variable	GAM	GBM	RF	MX
Ave Max Temp	23.6	12.8	28.1	44.2
Clay	20.5	0	5.6	5.8
Extreme Min Temp	32.6	14.4	25.1	40
Silt	14.7	0	4.1	2
Slope	8.5	72.8	37.1	8

Table 39. Relative importance of the input variables used in modeling for ERBI.

Variable	GBM	RF
Ave Max Temp	6.8	24.3
Clay	0	5
Extreme Min Temp	18.2	23
Silt	0	5
Slope	67.4	30.8
Vegetation	7.6	11.9

Vegetation associated with the ERBI localities was largely North American Warm Desert Scrub & Grassland (Figure 30).

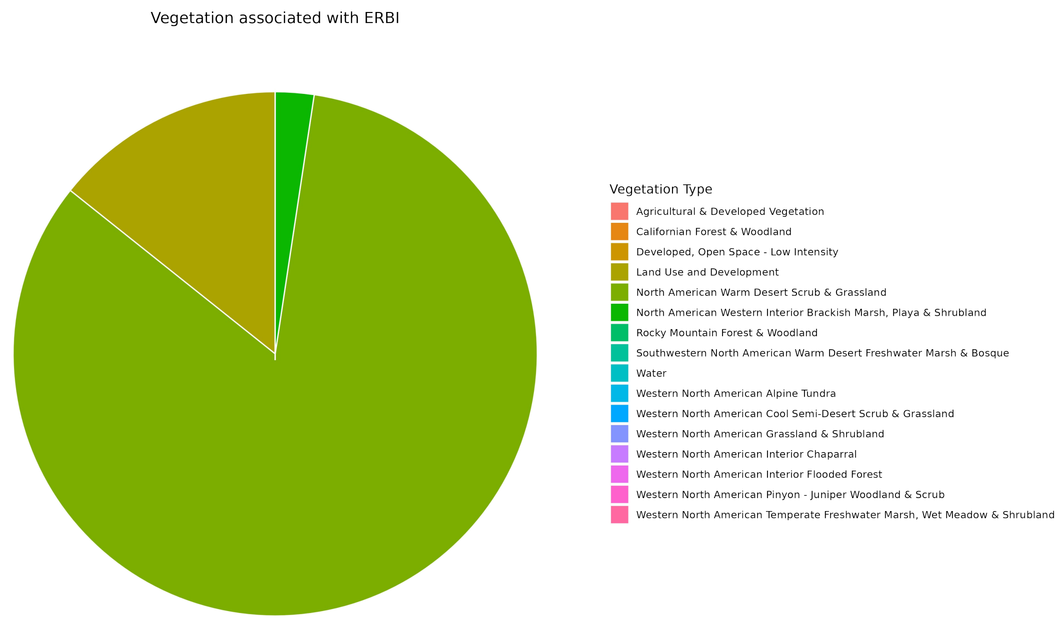


Figure 30. Relative frequency of vegetation associations at the locality point locations for ERBI.

The ensemble model was comprised of 14 Random Forest models, 6 GAM models, with 8 MaxEnt, and 12 GBM models contributing. The ensemble models including vegetation were 20 each of RF and GBM algorithms.

ERCO – Las Vegas buckwheat

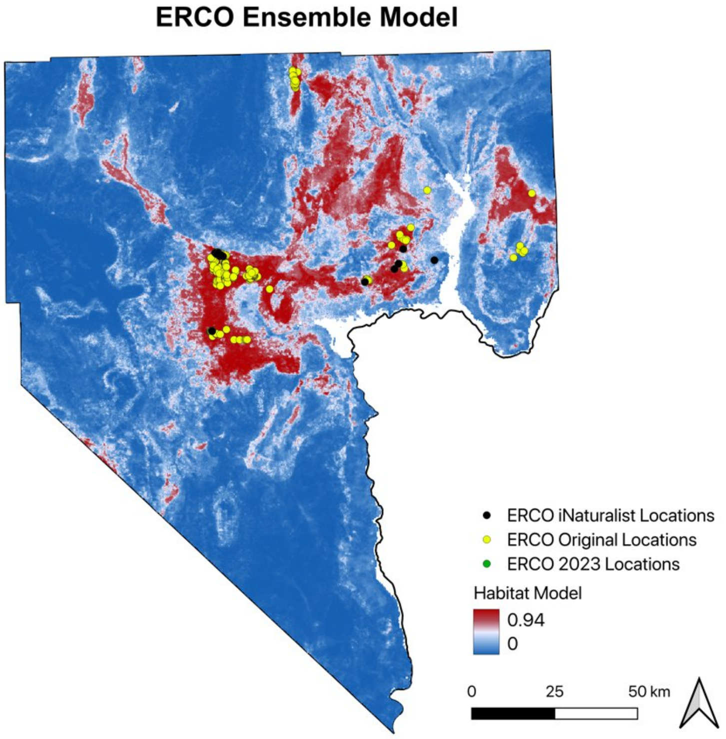


Figure 31. Ensemble Model for the base model approach for ERCO.

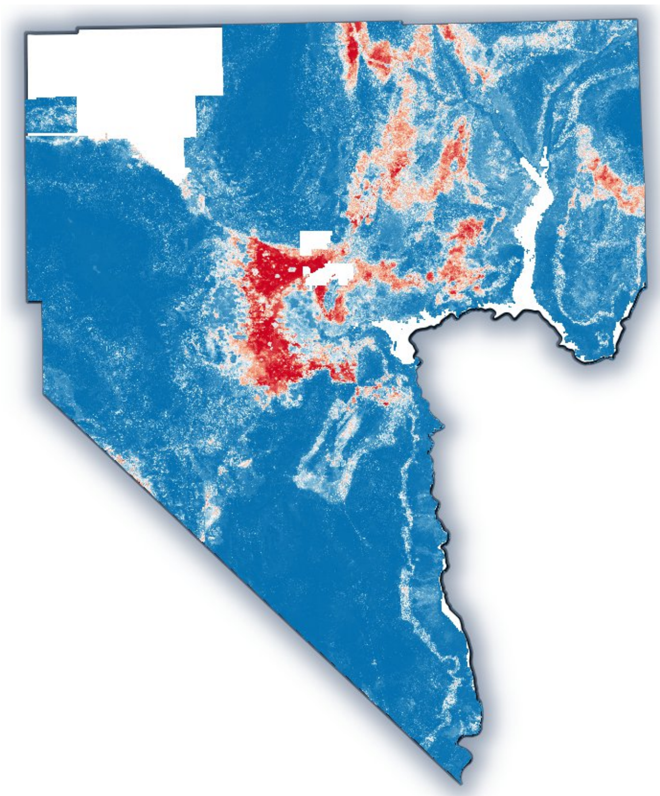


Figure 32. Ensemble Model for the vegetation inclusion models for ERCO.

The Las Vegas buckwheat model had a similar footprint of predicted habitat for both the base model (Figure 31) and the models including vegetation (Figure 32). The base model had high overall performance – where AUC, BI, and TSS all had high values with minimal losses between training and testing data (Table 40). The GAM and GBM models evaluated with the testing dataset had lower BI scores than other algorithms, with the GAM model having especially low scores (Table 40). The models including vegetation had high performance (Table 41).

Table 40. Performance metrics for ERCO. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.92	0.9	0.97	1	0.9	0.98
GAM	1	0.97	0.99	0.87	0.34	0.85	0.99	0.9	0.97
RF	1	0.99	1	0.98	0.89	0.98	1	0.88	0.98
MX	0.99	0.97	0.99	0.98	0.92	0.99	0.92	0.85	0.9
GBM	1	0.98	1	0.88	0.68	0.83	1	0.9	0.98

Table 41. Performance metrics for ERCO with the vegetation layer included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.97	0.82	0.95	1	0.97	0.99
RF	1	1	1	0.94	0.8	0.98	1	1	1
GBM	1	0.99	1	0.93	0.97	0.98	0.99	0.94	0.98

Variable importance indicated that maximum NDVI had low contributions towards all models (Table 42). The MaxEnt model was a single variable model, and GBM was largely dependent on three variables. The other models had contributions from all variables, with CV of winter precipitation and average maximum temperature contributing the most (Table 42). The models including vegetation showed mixed importance of the vegetation layer, with the GBM model showing 25%, and the RF models with 14% (Table 43). Both the GBM and RF had influence of several variables with Average Maximum Temperature and Gypsum as two of the highest performing variable for both (Table 43).

Table 42. Relative importance of the input variables used in modeling for ERCO.

Variable	GAM	GBM	RF	MX
Ave Max Temp	20.9	34.3	25.3	1.6
Soil gypsum	3.6	47	24.3	2.2
NDVI Amplitude	12.6	15.4	16.9	7.8
NDVI Max	2.1	0	5.1	0.2
Silt	7.2	0.7	9.6	4.9
Start of Season (day)	20.4	2.6	10.4	0.4
CV Winter Precip	33.2	0	8.4	82.9

Table 43. Relative importance of the input variables used in modeling for ERCO with the vegetation layer included.

Variable	GBM	RF
Ave Max Temp	16.3	21.3
Soil gypsum	29.5	17
NDVI Amplitude	16.9	15.2
NDVI Max	0	4.7
Silt	0.3	7.9
Start of Season (day)	12.4	12.5
CV Winter Precip	0	7.7
Vegetation	24.7	13.6

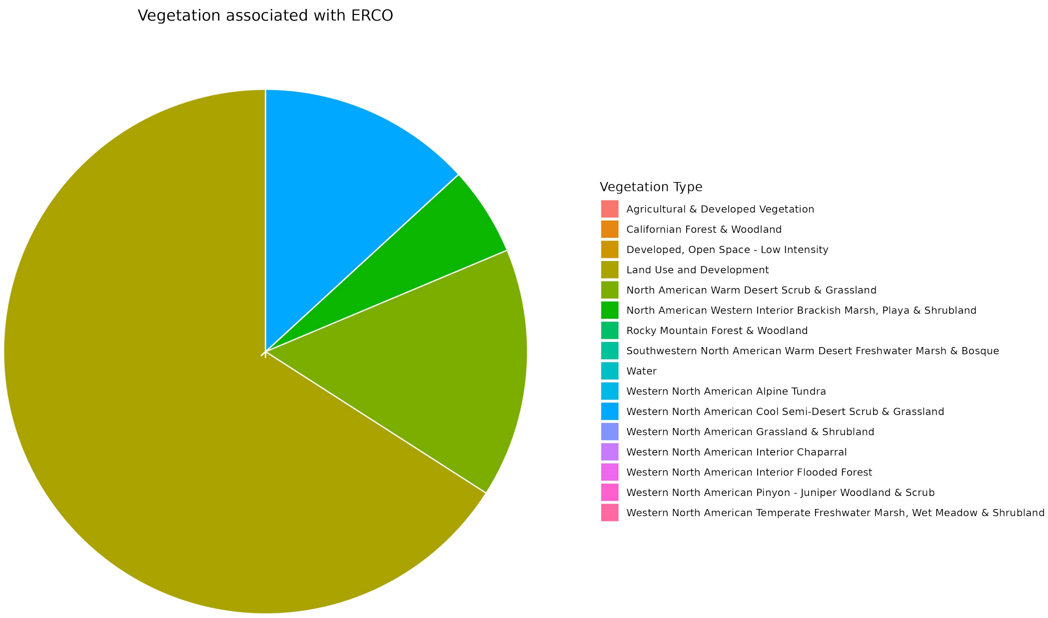


Figure 33. Relative frequency of vegetation associations at the locality point locations for ERCO.

The vegetation associated with ERCO localities was largely attributed to Land Use and Development, with secondary attribution of North American Warm Desert Scrub & Grassland, and Western North American Cool Semi-Desert Scrub & Grassland (Figure 33).

The ensemble model was comprised of 13 Random Forest models, 2 GAM, 9 MaxEnt, and 16 GBM models contributing. The vegetation inclusion ensemble model was created from a combination of 20 RF an20 GBM models.

ERVI – Sticky buckwheat

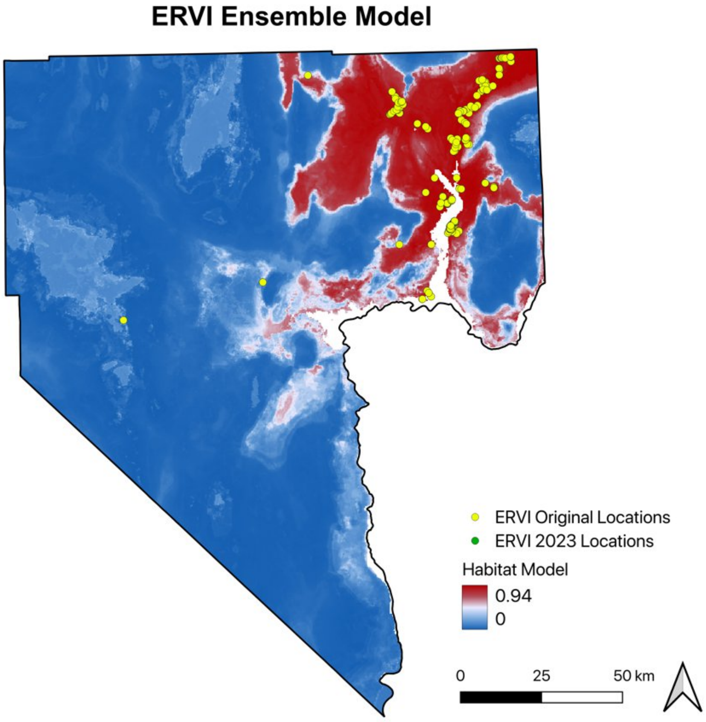


Figure 34. Ensemble Model for ERVI for the base model.

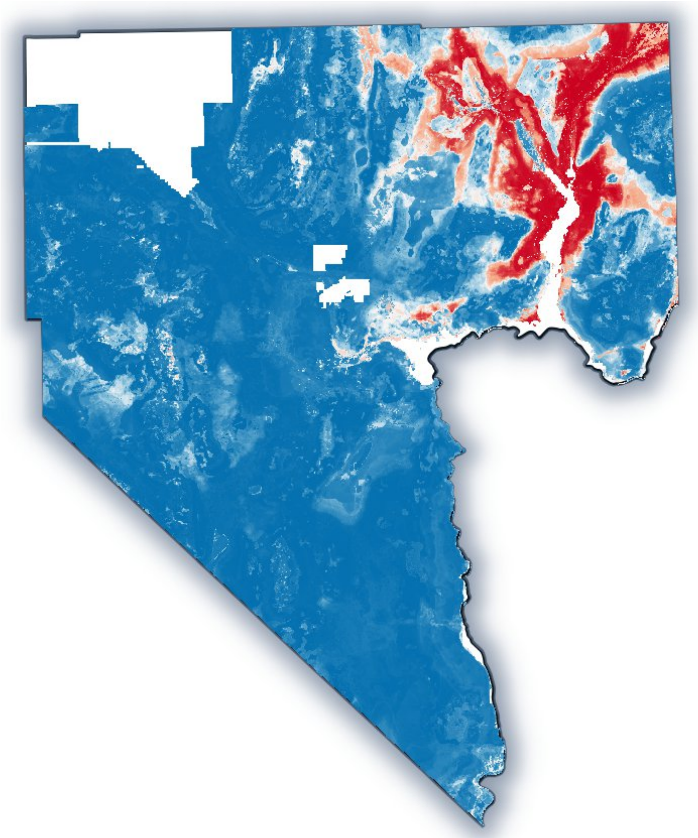


Figure 35. Ensemble Model for ERVI for the models including the vegetation layer.

The Sticky buckwheat model had somewhat similar predicted habitat for both models with and without vegetation (Figures 34 and 35), but with the vegetation model more closely associated with the Muddy and Virgin River areas (Figure 35). The models had high overall performance – where AUC, BI, and TSS all had high values with minimal losses between training and testing data (Table 44). The GAM and GBM models had lower BI scores than other algorithms, with both models having especially low scores for training, testing, and all data (Table 44). The performance for models including vegetation also had good performance for AUC and TSS, but had poor Boyce indices when evaluated with both the testing datasets (Table 45).

Table 44. Performance metrics for ERVI. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	1	1	0.92	0.53	0.94	1	1	1
GAM	1	1	1	0.14	0.59	0.59	1	0.97	0.99
RF	1	1	1	0.93	0.85	0.96	1	1	1
MX	0.98	0.99	0.98	0.97	0.97	0.97	0.93	0.97	0.94
GBM	1	1	1	0.31	0.76	0.71	1	0.97	0.99

Table 45. Performance metrics for ERVI models with the vegetation layer included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.99	1	0.9	0.74	0.93	1	0.9	0.98
RF	1	1	1	0.93	0.83	0.96	1	0.93	0.99
GBM	1	0.98	1	0.83	0.47	0.81	0.99	0.86	0.95

Average spring maximum temperature and CV of winter precipitation had generally high contributions towards all models (Table 46). The Maxent model as well as the GBM model were largely dependent on a single variable. However, GAM and Random Forest models were dependent on multiple variables (Table 46). Models including vegetation had moderate importance attributed to the vegetation layer for RF only, while GBM had no attribtution. The GBM model largely relied on three other variables (Average Spring Max Temp, coarse fragments, and variability in precipitation) – while the RF again had more balanced inclusion (Table 47).

Table 46. Relative importance of the input variables used in modeling for ERVI.

Variable	GAM	GBM	RF	MX
Average Spring Max Temp	47.8	82.2	48	3.6
Depth to bedrock	10.2	0	11.1	0
Coarse frags	7.2	2.2	13.6	0.5
Sand	2.6	0	4	0.2
CV Winter Precip	32.2	15.6	23.3	95.7

Table 47. Relative importance of the input variables used in modeling for ERVI with vegetation included.

Variable	GBM	RF
Average Spring Max Temp	57.7	31.7
Depth to bedrock	0	11.5
Coarse frags	26.9	23.6
Sand	0	6.1
CV Winter Precip	15.4	19.5
Vegetation	0	7.6

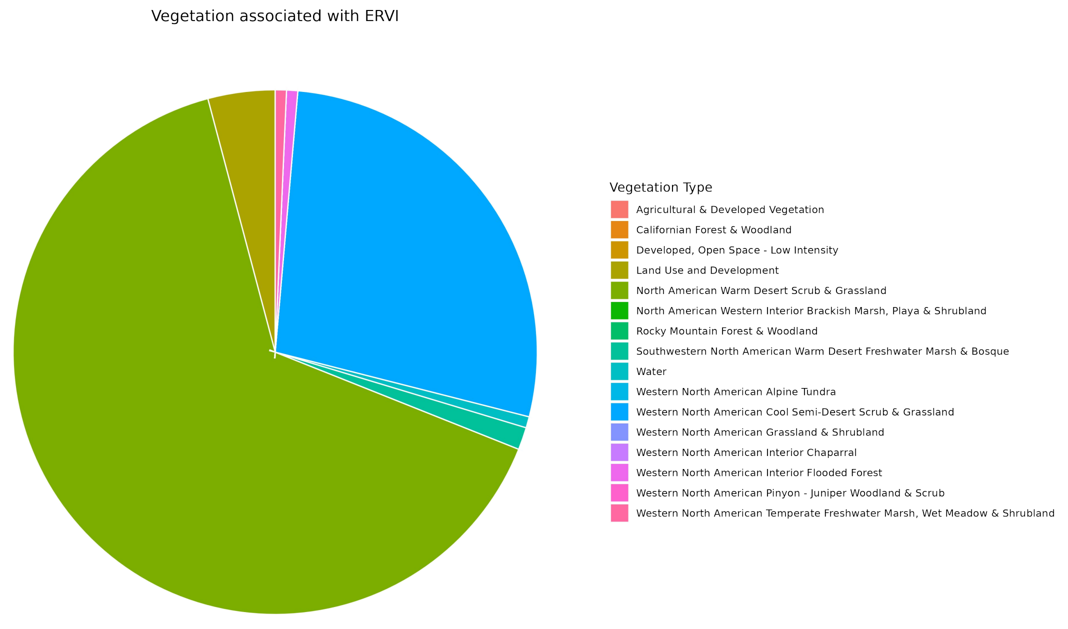


Figure 36. Relative frequency of vegetation associations at the locality point locations for ERVI.

ERVI localities were most commonly associated with North American Warm Desert Scrub & Grassland and Western North American Cool Semi-Desert Scrub & Grassland (Figure 36).

The ensemble model was comprised of 20 Random Forest models, 7 GAM, 3 MaxEnt, and 10 GBM models contributing. The ensemble model including vegetation was composed of 20 RF models and 17 GBMs.

GOAG – Mojave Desert tortoise

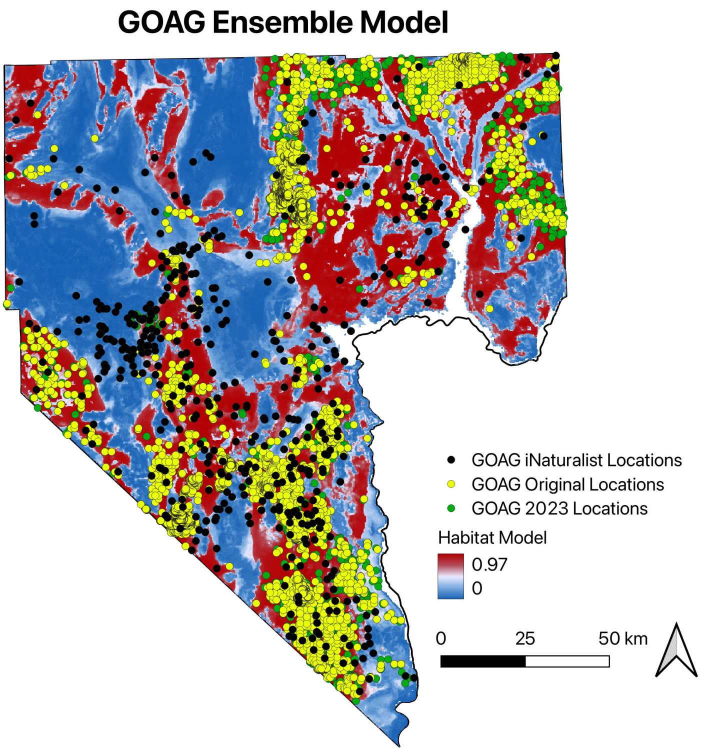


Figure 37. Ensemble Models for GOAG using base environmental measures.

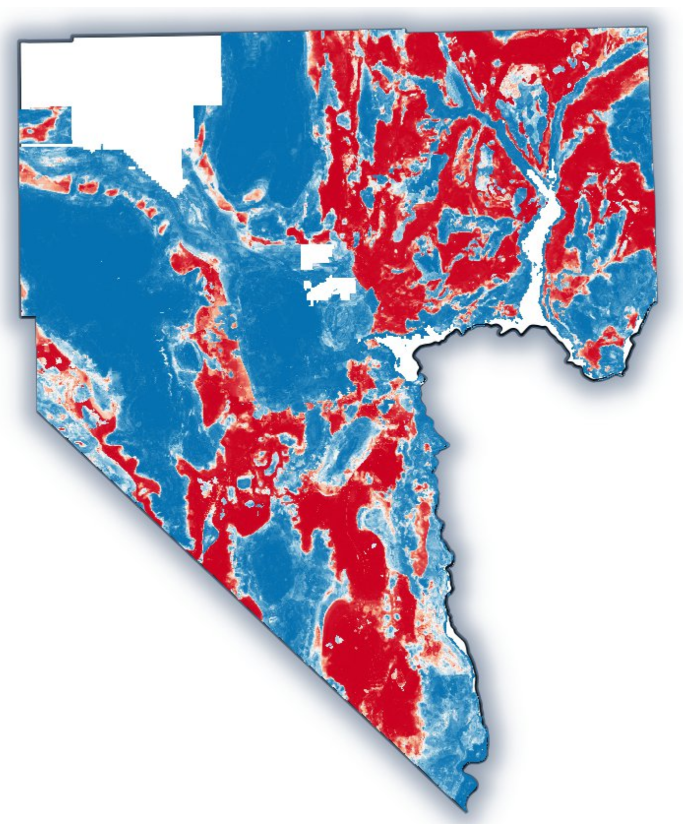


Figure 38. Ensemble Models for GOAG including the vegetation layer.

Predicted habitat for Desert Tortoises was similar between the base models, and the models including the vegetation layer (Figures 37 and 38). The base models high performance across the board. However, the RF models had a lower Boyce index for the training and testing data. (Table 48). The models including vegetation had only RF models selected as high performing. The overall performance was excellent, with all metrics high for both training and testing data (Table 49).

Table 48. Performance metrics for GOAG. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	**AUC_Training**	**AUC_Testing**	**AUC_All**	**BI_Training**	**BI_Testing**	**BI_All**	**TSS_Training**	**TSS_Testing**	**TSS_All**
EM	1	0.97	0.99	0.85	0.91	0.89	0.91	0.83	0.9
GBM	0.96	0.95	0.96	0.99	0.97	0.99	0.8	0.8	0.8
RF	1	0.97	1	0.63	0.7	0.82	0.99	0.84	0.96
GAM	0.95	0.95	0.95	1	0.97	1	0.78	0.77	0.78
MX	0.95	0.95	0.95	1	0.99	1	0.78	0.79	0.78

Table 49. Performance metrics for GOAG models including the vegetation layer. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.97	1	0.94	0.98	0.99	1	0.89	0.98
RF	1	0.97	1	0.99	0.99	1	1	0.84	0.96

Variable importance indicated that the coefficient of variation for Winter precipitation largely drove the MaxEnt models, while the other three algorithms had multiple variables contributing toward performance, with the CV of winter precip being the least important. Soil substrate variables and max temperature were important in the other three algorithms. The GBM model was largely driven by these, with no contribution of slope or winter precipitation contributing, while the GAM and RF models were supported by these variables. (Table 50). The RF models including vegetation showed only 2.7% importance of the vegetation layer, with balanced inclusion of the remaining variables (Table 51).

Table 50. Relative importance of the input variables used in modeling for GOAG.

Variable	GBM	RF	GAM	MX
Ave Max Temp	11	19.4	17.1	3.3
Depth to Bedrock	51.4	19.8	17.9	0
PPT Sand	37.6	19.6	15.3	1.3
Slope	0	12.5	16.8	1.4
Winter Precip	0	19.7	16.9	0.1
CV Winter Precip	0	9	16	93.8

Table 51. Relative importance of the input variables used in modeling for GOAG with vegetation layer included.

Variable	RF
Ave Max Temp	13.30
Depth to Bedrock	18.68
PPT Sand	23.33
Slope	12.44
Winter Precip	22.02
CV Winter Precip	7.63
Vegetation	2.61

The ensemble model was comprised of 20 Random Forest models, and 16 GBM models, with 2 GAM models, with 3 MaxEnt models contributing. Ther were 20 RF models contributing to the vegetation based ensemble model.

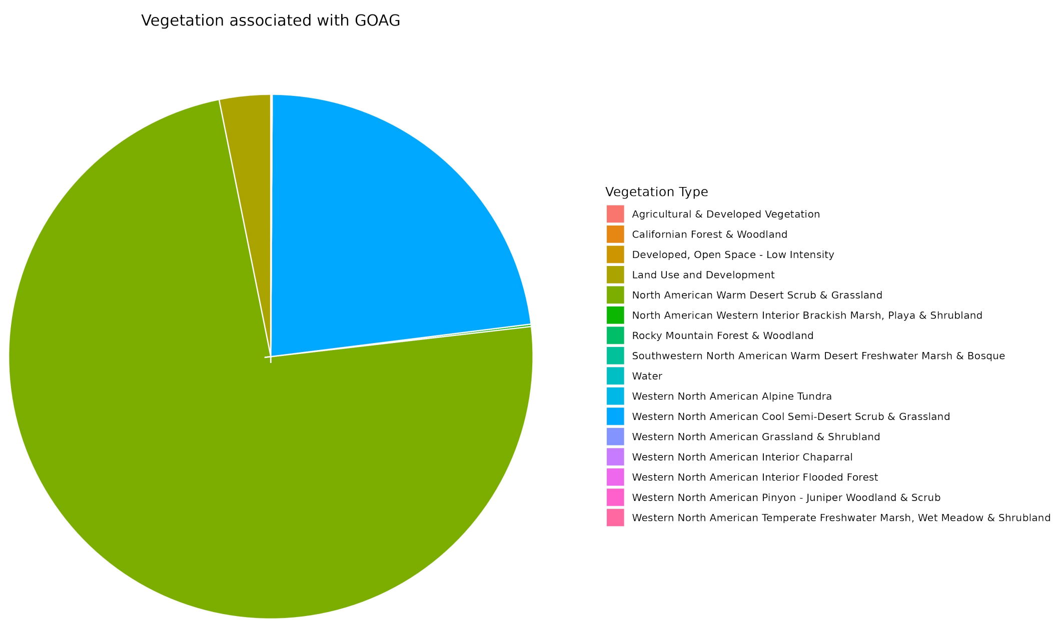


Figure 39. Relative frequency of vegetation associations at the locality point locations for GOAG.

GOAG localities were largely located within North American Warm Desert Scrub & Grassland, and Western North American Cool Semi-Desert Scrub & Grassland (Figure 39).

LALU – Loggerhead shrike

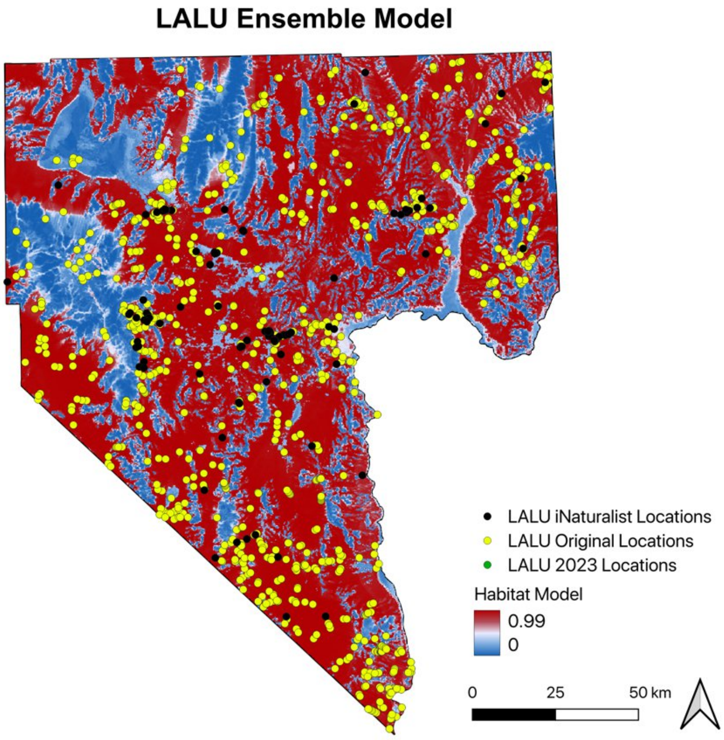


Figure 40. Ensemble Model for the LALU base models.

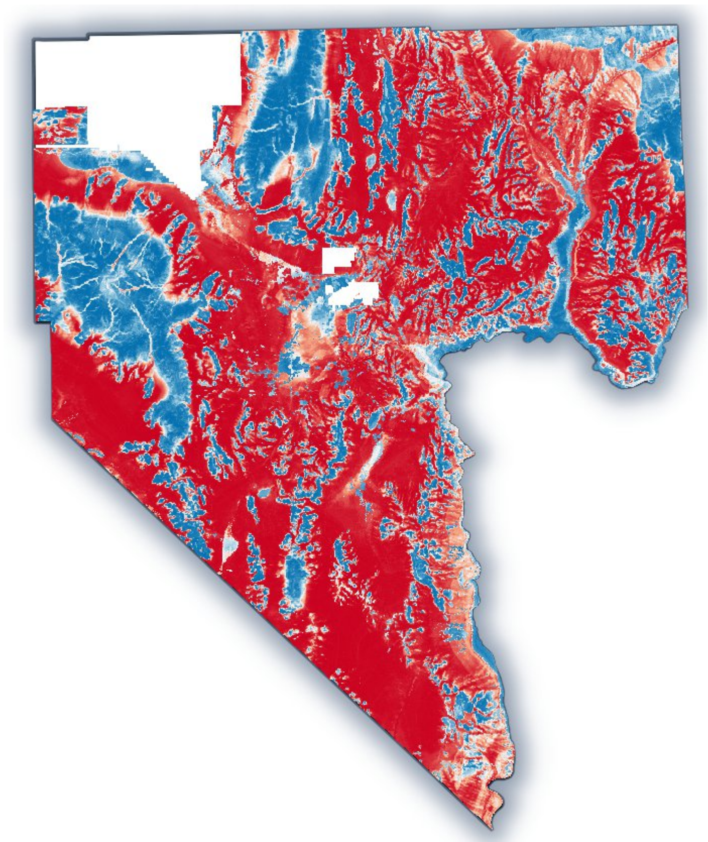


Figure 41. Ensemble Model for the LALU models including vegetation.

The habitat predictions for Loggerhead Shrike had similar footprints of predicted habitat (Figures 40 and 41), which is likely do to the low importance of the vegetation layer in the modeling effort (Table 55). For this species only GBM and RF models were in the highest performing groups. The Loggerhead shrike model had high performance for certain algorithms – where AUC and TSS all had high values with minimal losses between training and testing data (Table 52). The BI scores were generally lower, with the GBM and Random Forest models having low BI scores for training, testing, and all data. The ensemble model had low scores for training and all data. The models including vegetation had higher performance for all metric (Table 53).

Table 52. Performance metrics for LALU. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.97	0.99	0.18	0.81	0.55	0.91	0.84	0.89
GBM	0.96	0.96	0.96	0.74	0.6	0.73	0.82	0.81	0.82
RF	1	0.97	1	0.32	0.83	0.84	0.99	0.85	0.96

Table 53. Performance metrics for the vegetation inclusive models for LALU. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.96	1	0.99	0.95	0.97	1	0.88	0.98
RF	1	0.96	1	0.99	0.95	0.97	1	0.88	0.98

Flow accumulation and slope had generally high contributions towards all models (Table 54). The GBM model as well as the Random Forest model were dependent on multiple variables, with the Random Forest model having contributions from all variables included in modeling (Table 54). Inclusion of the vegetation layer influenced included only RF models, where there was 5% performance attributed to vegetation (Table 55).

Table 54. Relative importance of the input variables used in modeling for LALU for the base model set.

Variable	GBM	RF
Winter Precip	0	12.4
CV Winter Precip	2.8	12.4
Average Spring Max Temp	0.3	10.8
Slope	31.1	20.9
NDVI Start of Season	0.2	14.2
Flow Accum	65.6	29.2

Table 55. Relative importance of the input variables used in modeling for LALU for the vegetation added model set

Variable	RF
Winter Precip	12.71
CV Winter Precip	13.34
Average Spring Max Temp	11.31
Slope	22.15
NDVI Start of Season	12.39
Flow Accum	22.98
Vegetation	5.12

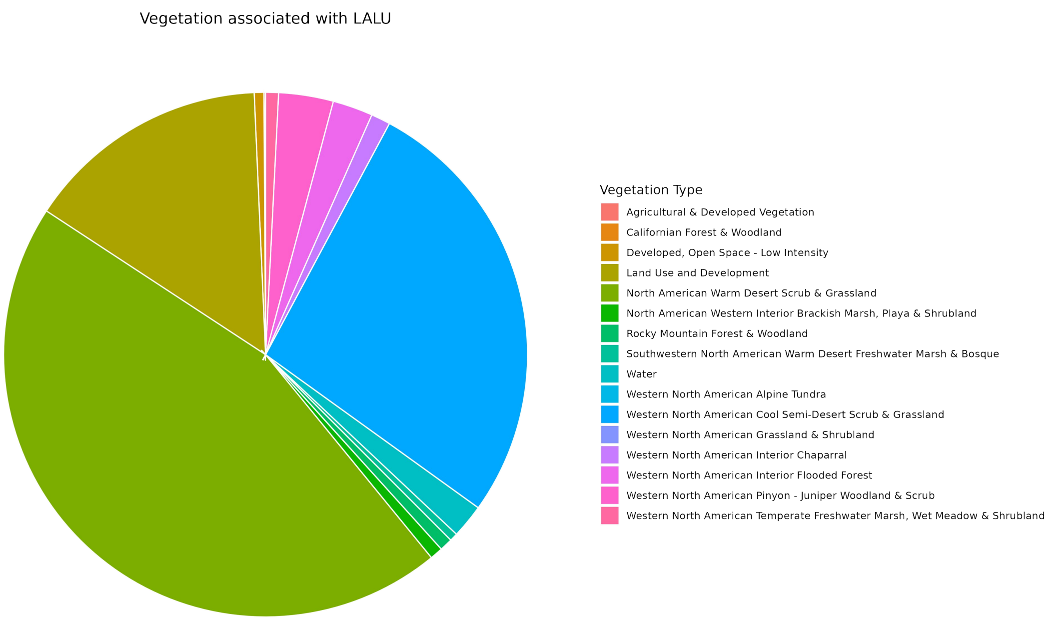


Figure 42. Relative frequency of vegetation associations at the locality point locations for LALU.

Vegetation associated with LALU localities was mostly within North American Warm Desert Scrub & Grassland, with Western North American Cool Semi-Desert Scrub & Grassland, and Land Use and Development as the second and third most common (Figure 42).

The ensemble model was comprised of 20 Random Forest models and 20 GBM models. The vegetation based model was composed of 20 RF models.

PEAL – White-margined Beardtongue

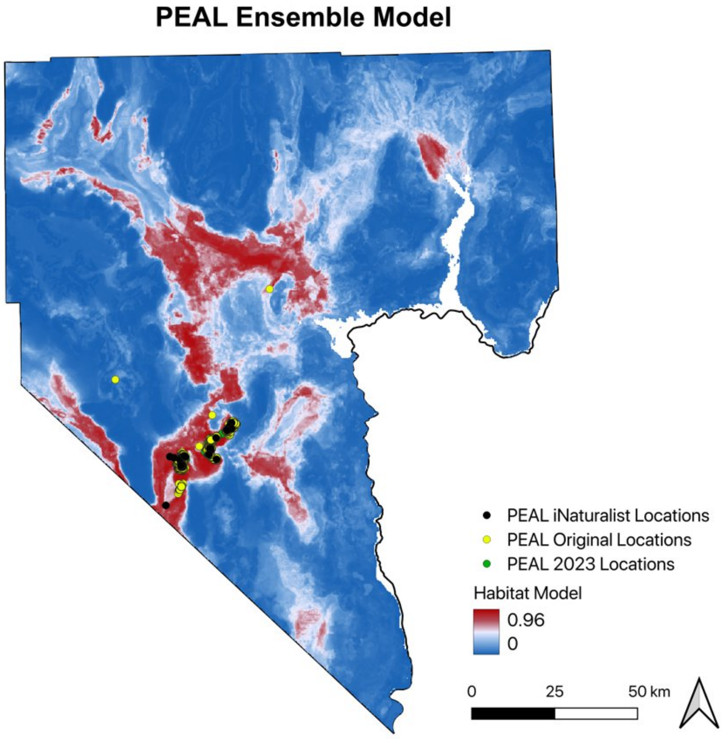


Figure 43. PEAL Ensemble Model for the base environmental layers.

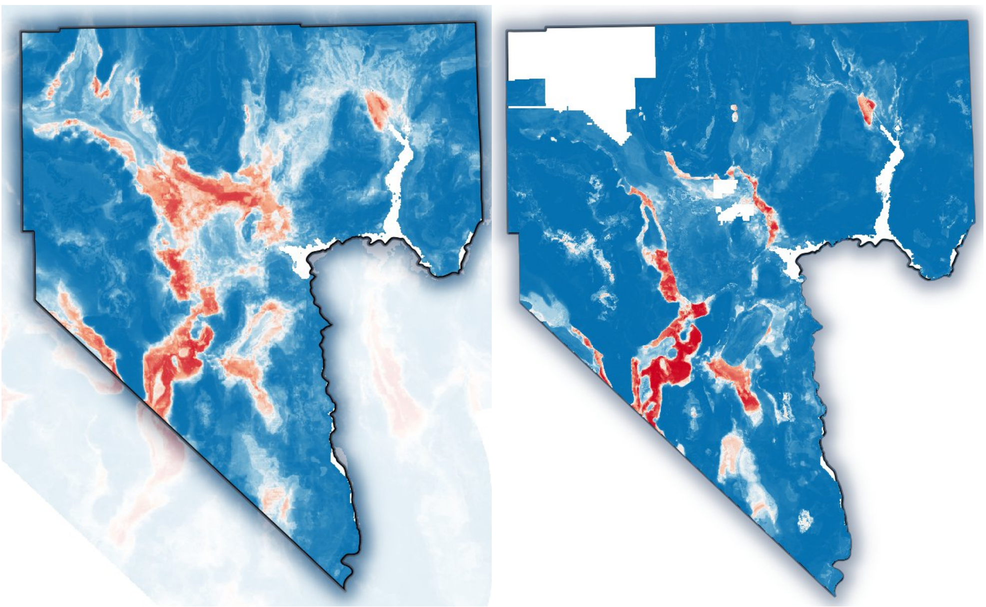


Figure 44. PEAL Ensemble Model for the base enviroonmental layers (Left) compared with the vegetation layer model (Right).

Models for White-margined Beardtongue differed considerably between the base models, and the models including the vegetation layer (Figure 44). The models including vegetation as a predictor layer showed a much more restrictive predicted habitat area, with habitat more restricted to the area surrounding the I-15 Corridor, concordant with the localities (Figure 43), and reduced area predicted around the perimeter of Las Vegas. The base model had high performance across all algorithms – where AUC, BI, and TSS all had high values with minimal losses between training and testing data (Table 56). The BI score for the Random Forest model built with the testing data was generally lower than other models (Table 56). Performance for the models including vegetation was high for all performance measures tested (Table 57).

Table 56. Performance metrics for PEAL. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.97	0.96	0.97	0.98	0.94	0.97
RF	0.99	0.98	0.99	0.94	0.81	0.91	0.92	0.91	0.92
GAM	1	0.98	1	0.96	0.79	0.97	1	0.94	0.99
MX	0.98	0.98	0.98	0.97	0.97	0.97	0.87	0.94	0.88
GBM	1	0.98	0.99	0.94	0.84	0.98	0.96	0.92	0.95

Table 57. Performance metrics for PEAL for models including vegetation. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.96	0.99	0.99	1	0.93	0.99
RF	1	0.99	1	0.98	0.98	0.99	1	0.95	0.99
GBM	1	0.96	0.99	0.85	0.98	0.94	1	0.93	0.99

Depth to bedrock, winter precipitation, and CV of winter precipitation had generally high contributions towards all models (Table 58). The MaxEnt model was again largely dependent on a single variable. The other models were more dependent on all variables included in modeling. Vegetation had good performance in the augmented models, with good representation across all variables for the RF model, but with a low reliance of Vegetation. The GBM models had no influence of vegetation, but were driven by the CV of winter precipitation , depth to bedrock and, sand content (Table 59).

Table 58. Relative importance of the input variables used in PEAL base models.

Variable	GBM	RF	GA<	MX
Depth to bedrock	5.6	13.4	25.6	0
Clay	23	14.3	9.9	7.8
Extreme Min Temp	0.3	17.4	17.4	1.6
Slope	0	4.6	14.9	1.8
Winter Precip	41.2	23.9	14	1.1
CV Winter Precip	29.8	26.4	18.2	87.7

Table 59. Relative importance of the input variables used in PEAL models with vegetation.

Variable	GBM	RF
Depth to bedrock	31.8	18.1
Sand	15.1	12.4
Extreme Min Temp	0.8	16.5
Slope	0	3.3
Winter Precip	2.3	19
CV Winter Precip	50	28.6
Vegetation	0	2.2

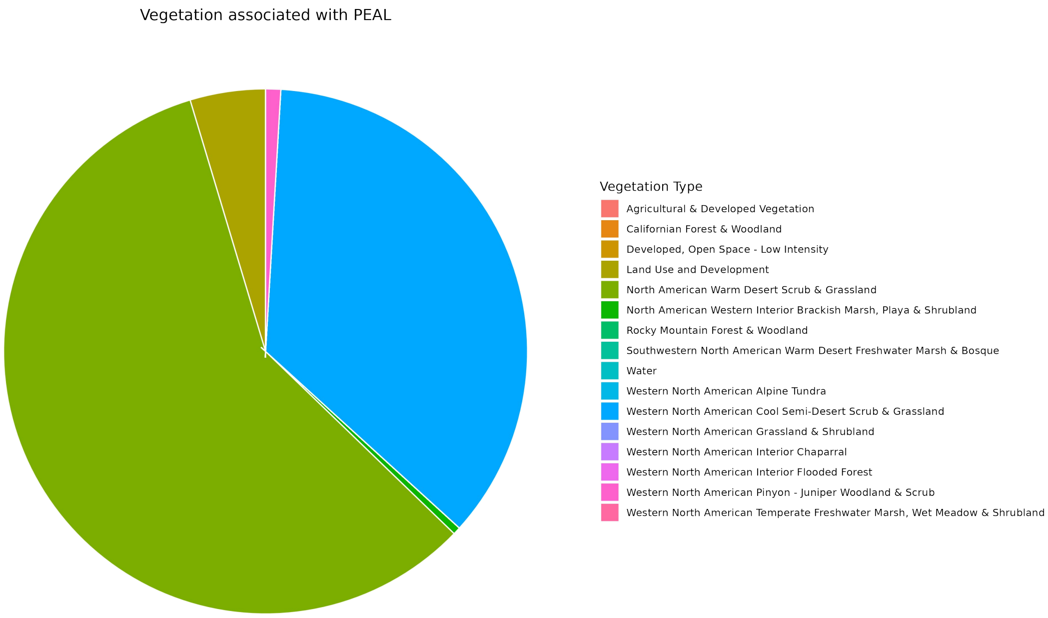


Figure 45. Relative frequency of vegetation associations at the locality point locations for PEAL.

PEAL localities were associated with North American Warm Desert Scrub & Grassland and Western North American Cool Semi-Desert Scrub & Grassland, with a fraction in Land Use and Development (Figure 45).

The ensemble model was comprised of 18 Random Forest models, 3 GAM, 11 MaxEnt, and 9 GBM models. There were 20 RF and 20 GBM included in the ensemble for the vegetation enhanced model.

TOBE – Bendire's thrasher

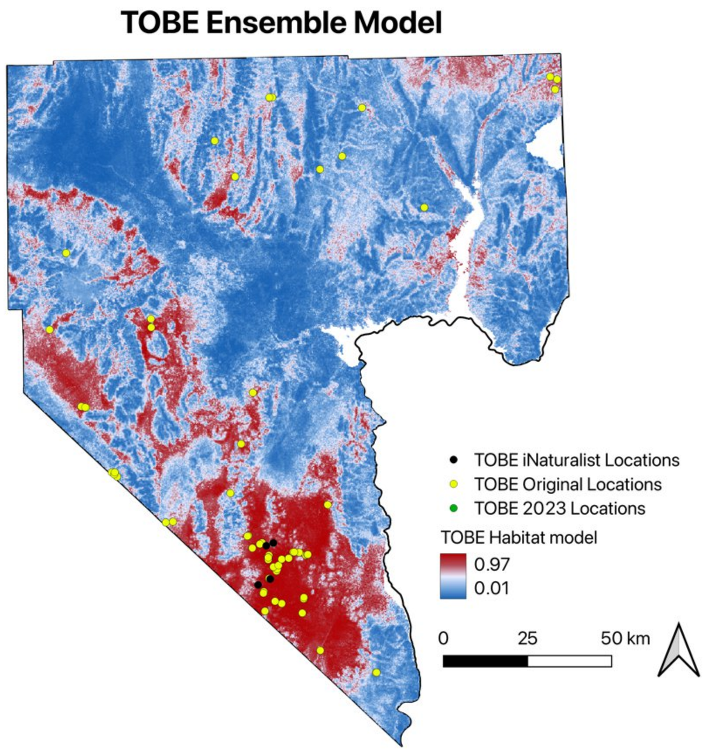


Figure 46. TOBE Ensemble Model for the base layers.

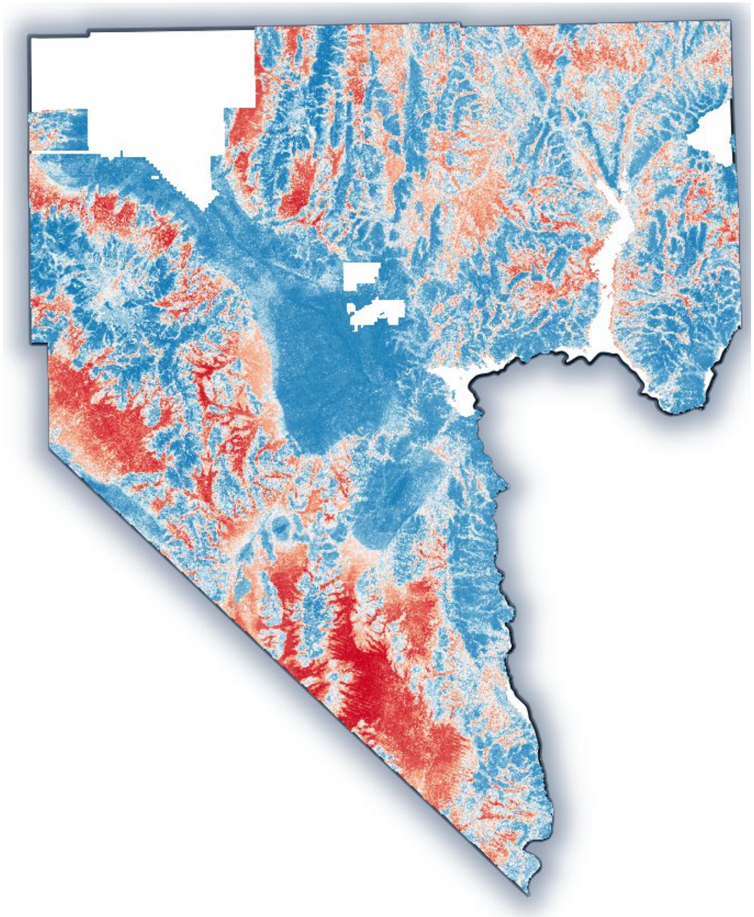


Figure 47. TOBE Ensemble Model for the vegetation inclusion model.

The Model for Bendire’s thrasher with vegetation showed a increased habitat prediction area, especially in the northeastern portion of Clark County (Figure 46 and 47). The models had excellent performance, with only the GAM model showing lower performance for the Boyce index only for the testing dataset. The other models had high AUC, BI and TSS scores, with little drop from training to testing (blind) data (Table 60). The models including vegetation also indicated high performance, with good AUC, TSS and BI scores for training data. The scores for the Boyce index did drop for the testing dataset (Table 61).

Table 60. Performance metrics for TOBE. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	0.99	0.95	0.99	0.96	0.92	0.98	0.93	0.77	0.88
GAM	0.96	0.91	0.95	0.95	0.57	0.95	0.82	0.75	0.79
RF	1	0.96	1	1	0.97	1	0.99	0.81	0.95
MX	0.94	0.93	0.93	0.98	0.75	0.99	0.74	0.77	0.74
GBM	0.99	0.95	0.98	0.92	0.86	0.95	0.91	0.75	0.87

Table 61. Performance metrics for TOBE with vegetation included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.98	1	0.91	0.78	0.93	1	0.85	0.97
RF	1	0.99	1	0.9	0.4	0.97	1	0.92	0.98
GBM	1	0.95	1	0.96	0.71	0.87	1	0.85	0.94

The flow accumulation index, which gives an index of topographic position in the landscape had the lowest performance, but still contributed 10% toward the RF model. (Table 62). The higher performing variables were winter precipitation, and its variance, while the other variables contributed relatively evenly across models, with the exception of the distance to cliffs, which performed well in the GAM models. The MaxEnt model was largely a single variable model, using only the CV of winter precipitation. For the models including vegetation the RF model had a good inclusion of all variables, including the vegetation layer at 41% and 13% for the GBM and RF models respectively (Table 63). The GBM models had influences from most variables, with three variables that had low importance, while the RF models included contributions from all variables (Table 63).

Table 62. Relative importance of the input variables used in modeling for TOBE.

Variable	GBM	RF	GAM	MX
Dist to cliffs	0	6.8	18.6	0
NDVI Amplitude	2.5	11.8	16.6	0.3
NDVI Max	8.2	13.5	4.3	0.1
Winter Precip	42.9	21.9	19.7	0.2
CV Winter Precip	24.2	20.6	22	96.7
Slope	17.3	14.7	10.8	2.6
Flow Accum	4.9	10.7	8	0

Table 63. Relative importance of the input variables used in modeling for TOBE.

Variable	GBM	RF
Dist to cliffs	0.6	6.4
NDVI Amplitude	1.2	8.8
NDVI Max	2	12
Winter Precip	23.5	19.6
CV Winter Precip	2.4	10.5
Slope	14.6	14.1
Flow Accum	14.2	15.3
Vegetation	41.5	13.3

The ensemble model was comprised of 13 Random Forest models,12 GBM, 8 GAM, and 7 MaxEnt models. For the models with vegetation included there were 20 RF and 20 GBM models contributing.

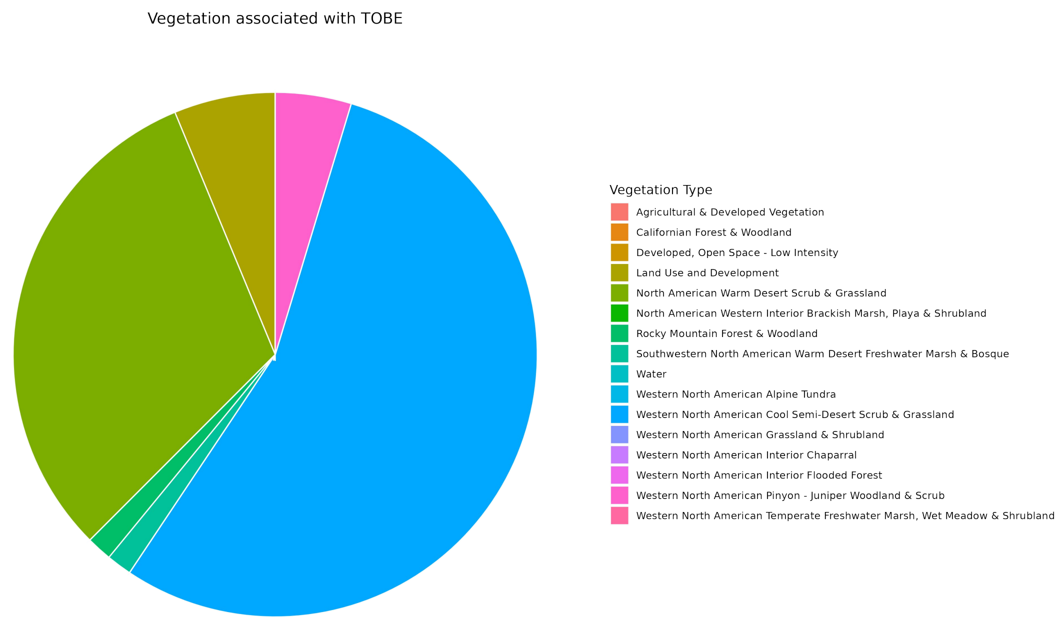


Figure 48. Relative frequency of vegetation associations at the locality point locations for TOBE.

Vegetation associated with TOBE localities was predominantly Western North American Cool Semi-Desert Scrub & Grassland, and North American Warm Desert Scrub & Grassland (Figure 48).

TOLE – Le Conte's thrasher

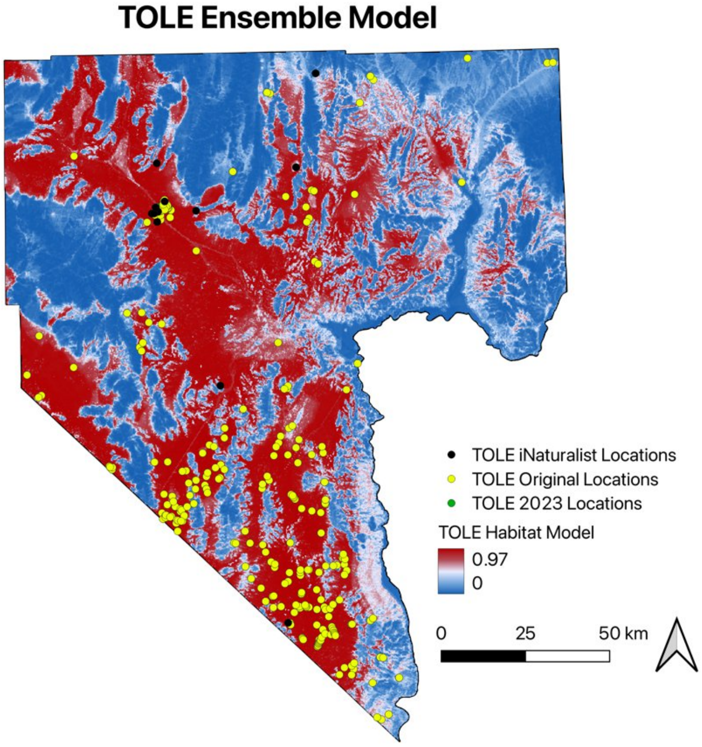


Figure 49. Ensemble Models for TOLE using base environmental layers.

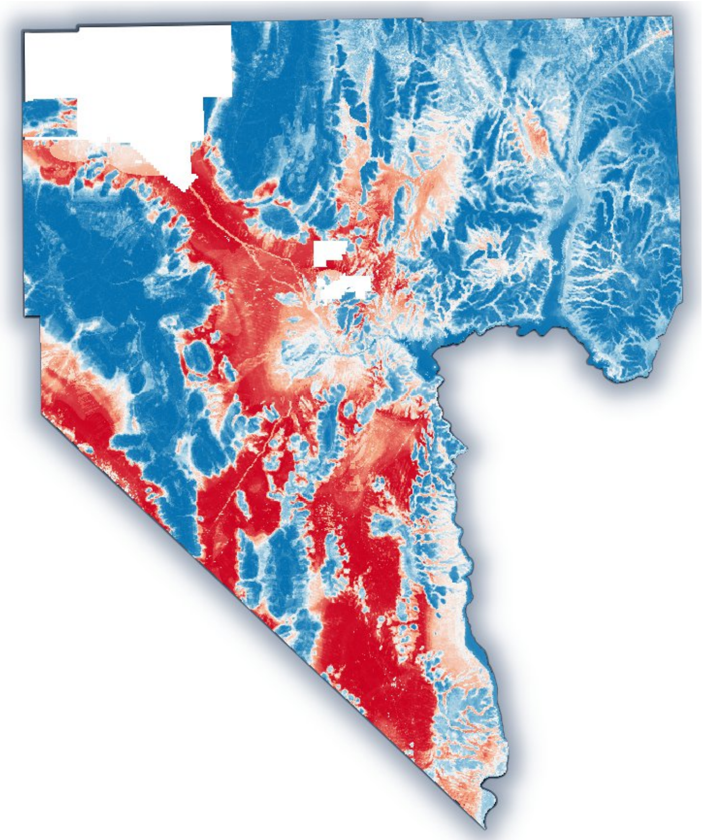


Figure 50. Ensemble Models for TOLE including the vegetation layer.

The Leconte's Thrasher model had a similar predicted habitat footprint for both the base and vegetation augmented models, but with differences in the northeastern portion of the county where the distribution was more limited in the vegetation augmented model (Figures 49 and 50). The models for the base had high overall performance – where AUC, BI, and TSS all had high values with minimal losses between training and testing data with the exception of the GBM model for the testing dataset (Table 64). The GAM and model had lower BI scores than other algorithms. For the models including vegetation only RF models performed well enough to be i8ncluded. The performance scores were higher than the base models, and metrics were high across the board (Table 65).

Table 64. Performance metrics for TOLE. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.94	0.99	0.91	0.71	0.96	0.95	0.73	0.9
GAM	0.96	0.9	0.95	0.79	0.74	0.83	0.78	0.68	0.76
RF	1	0.95	0.99	0.94	0.75	0.98	0.99	0.72	0.94
MX	0.95	0.91	0.95	0.92	0.84	0.93	0.78	0.67	0.76
GBM	0.99	0.94	0.98	0.92	0.51	0.96	0.91	0.7	0.87

Table 65. Performance metrics for TOLE with vegetation models included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.96	0.99	0.98	0.89	0.97	1	0.81	0.96
RF	1	0.96	0.99	0.98	0.89	0.97	1	0.81	0.96

The MaxEnt model converged on two variables, the GBM on three (but largely driven by the CV of winter precipitation), and the RF and GAM were more evenly balanced. (Table 66). CV Winter Precip had high contributions across all models, and all variables had higher than 10% contribution for at least one of the algorithms. For the models including vegetation – the RF model had several variables that showed high importance, and moderate importance to the vegetation layer (Table 67).

Table 66. Relative importance of the input variables used in modeling for TOLE.

Variable	GBM	RF	GAM	MX
Flow Accum	18.7	20.4	7.9	0
NDVI Length of Season	0	6.7	15.8	0
CV Winter Precip	73	39.2	45.2	23.4
CV Average Spring Max Temp	0	16.3	14.5	76.2
Slope	8.4	17.4	16.6	0.4

Table 67. Relative importance of the input variables used in modeling for TOLE with the vegetation layer included.

Variable	RF
Flow Accum	15.10
NDVI Length of Season	6.72
Winter Precip	8.06
CV Winter Precip	29.69
CV Average Spring Max Temp	13.56
Slope	19.82
Vegetation	7.05

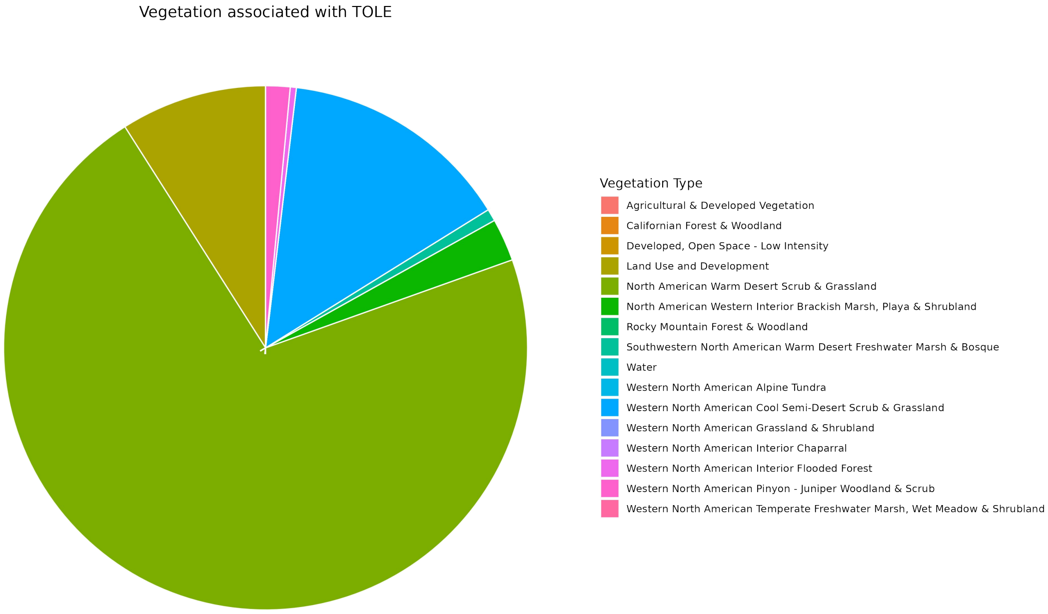


Figure 51. Relative frequency of vegetation associations at the locality point locations for TOLE.

Vegetation associated with the TOLE localities was predominantly North American Warm Desert Ruderal Grassland and North American Warm Desert Scrub & Grassland, with Western North American Cool Semi-Desert Scrub & Grassland, Land Use and Development, and several other associations at lower prevalence (Figure 51).

The ensemble model was comprised of 19 Random Forest models, 3 GAM, 4 MaxEnt, and 14 GBM models contributing. The vegetation based model had 20 RF included in the ensemble.

VIBE – Arizona Bell's vireo

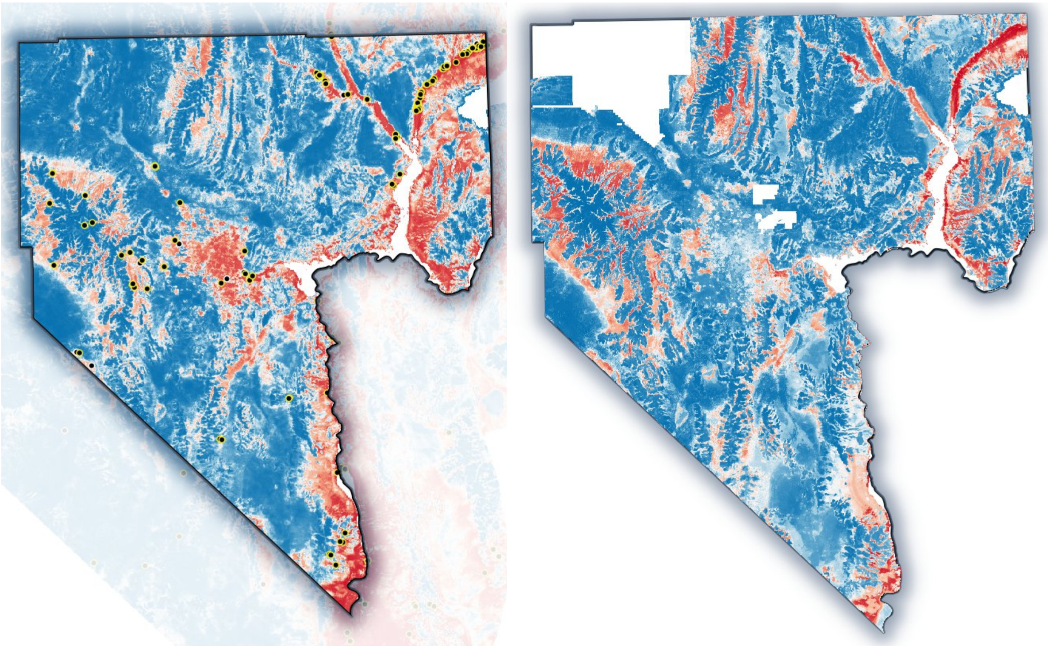


Figure 52. Ensemble Model comparing models with vegetation (right) and base environmental layers (Left) for VIBE.

The Arizona Bell's Vireo models had markedly different footprints of predicted habitat between the base models, and those including the vegetation layer, where the models including vegetation showed a reduced area of predicted habitat relative to the standard modeling using base layers alone (Figure 52), where habitat was more restricted tightly around riparian areas in Lake Mead, and some upland habitat predicted around the spring range. The models had high performance across all algorithms (Table 68). For the base model, training and testing performance remained high across all performance metrics, indicating good generalizability (Table 68). Similarly – the models including vegetation had excellent performance for both testing and training data (Table 69).

Table 68. Performance metrics for VIBE. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.97	0.99	0.97	0.89	0.97	0.94	0.87	0.93
GAM	0.98	0.96	0.98	0.97	0.82	0.95	0.88	0.84	0.87
RF	1	0.97	1	0.99	0.83	0.98	1	0.86	0.97
MX	0.97	0.97	0.97	0.99	0.94	0.99	0.86	0.87	0.86
GBM	0.99	0.97	0.99	0.96	0.85	0.97	0.95	0.9	0.94

Table 69. Performance metrics for VIBE with the vegetation layer included. AUC (Area under Curve), BI – (Boyce Index), TSS (true skill statistic) were each calculated independently for Training and Testing data and using all points.

Model	AUC_Training	AUC_Testing	AUC_All	BI_Training	BI_Testing	BI_All	TSS_Training	TSS_Testing	TSS_All
EM	1	0.96	1	0.95	0.69	0.96	1	0.89	0.98
RF	1	0.97	1	0.99	0.81	0.98	1	0.85	0.97
GBM	1	0.94	0.99	0.99	0.9	0.97	0.96	0.78	0.92

The variable contributions indicated relatively high inclusion of all variables, with Winter Precipitation contributing the least (Table 70). The GBM model consisted largely of only three variables, while the other three were more balanced across the board (Table 70). Vegetation contributed little toward variable importance for either the GBM algorithm, and only moderately for RF (Table 71). Both the GBM and RF models showed importance across many inputs, and Average Max Temperature Spring Max Temperature, Silt Content, and Topographic index were among the highest contributing variables for both modeling algorithms (Table 71).

Table 70. Relative importance of the input variables used in modeling for VIBE.

Variable	GAM	RF	MX	GBM
Ave Max Temp	0.6	16	10.9	7.7
Average Spring Max Temp	11.2	14.7	45.7	7.8
NDVI Amplitude	24.4	14.2	0.3	1.2
Winter Precip	22.2	8.1	2	0
Slope	19.1	8.4	6.1	0
TPX	12.3	10	7.8	1
Silt	10.3	28.6	27.1	82.3

Table 71. Relative importance of the input variables used in modeling for VIBE including vegetation.

Variable	GBM	RF
Ave Max Temp	12	13.4
Average Spring Max Temp	12.9	14.1
NDVI Amplitude	8.3	11.8
Winter Precip	0	9.7
Slope	0	9.9
TPX	32.6	15.8
Silt	33.6	18.3
Vegetation	0.6	7

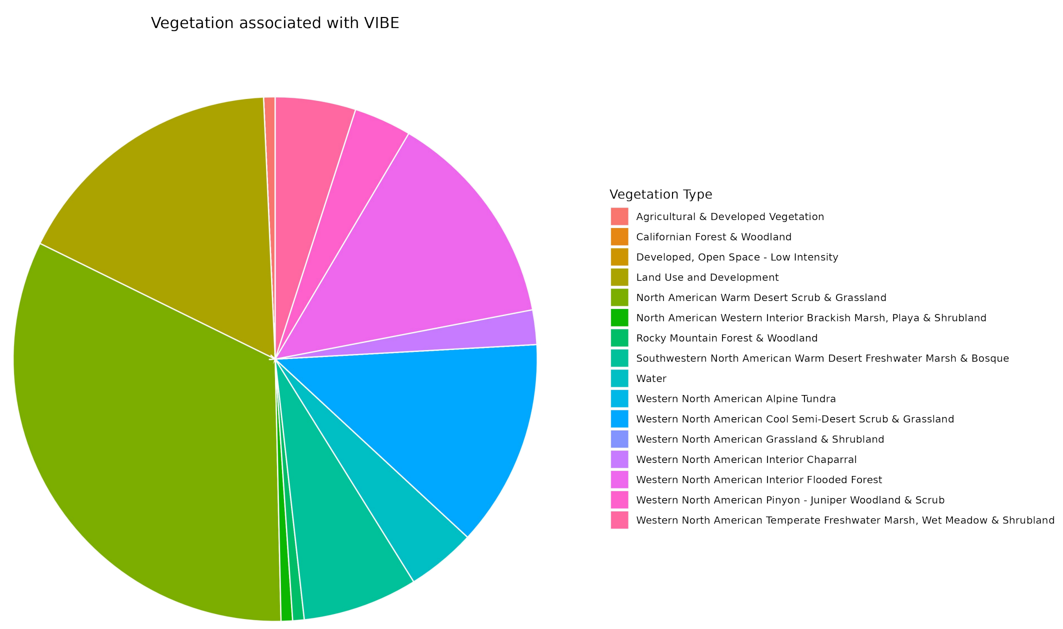


Figure 53. Relative frequency of vegetation associations at the locality point locations for VIBE.

Vegetation associated with Bells Vireo localities had among the most diverse of all habitat types (Figure 53). The largest was North American Warm Desert Scrub & Grassland .

The ensemble model was comprised of 11 Random Forest models, 10 GBM models, 9 GAM and 11 MaxEnt models, with an uncharacteristically even contribution among all algorithms. The vegetation based models were an ensemble of 20 RF and 10 GBM models.

References

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

Araujo, M.B. and M. New. 2006. Ensemble forecasting of species distributions. Trends in Ecology and Evolution 22: 42-47.

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

Boyce, M., Vernier, P., Nielsen, S., and Schmiegelow, F. 2002. Evaluating resource selection functions. Ecological modeling 157: 281–300.

Elith, J., Ferrier, S., Huettmann, F., and Leathwick, J. 2005. The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. Ecological modeling 186: 280-289

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

Fourcade, Y., Engler, J., Rodder, D., and Secondi, J. 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. PLoS One 9: e97122.

Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. Cambridge

Heaton, J.S., X. Miao, K. Von Secondorff Hoff, D. Charlet, P. Cashman, J. Trexler, A. Grimmer, and R. Patil. 2011. Ecosystem Indicators. Final Report 2005-UNR-578. Report to Clark County MSHCP 2005-UNR-578-D27.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., and Guisan, A. 2006. Evaluating the ability of habitat models to predict species presences. Ecological Modeling 199: 142-152.

Liaw, A. and Wiener, M. 2002. Classification and Regression by randomForest. R News 2(3), 18--22.

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

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Thuiller, W., B. Lafourcade, R Engler, and M Araujo. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. Ecography 32:369-373.

Watling, J.I., Brandt, L.A., Bucklin, D.N., Fujisaki, I., Mazzotti, F.J., Romanach S.S., and Speroterra, C. 2015. Performance metrics and variance partitioning reveal sources of uncertainty in species distribution models. Ecological modeling 309-310: 48-59.

Wisz, M.S., and Guisan, A. 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecology 9:8. DOI: 10.1186/1472-6785-9-8

Wood, D. A., A. G. Vandergast, K. R. Barr, R. D. Inman, T. C. Esque, K. E. Nussear, and R. N. Fisher. 2013. “Comparative phylogeography reveals deep lineages and regional evolutionary hotspots in the Mojave and Sonoran Deserts.” In: Diversity and Distributions 19.7, pp. 722–737. doi: 10.1111/ddi. 12022.