Species Distribution Modeling for Covered Species 2025

Kenneth Nussear
Eric Simandle

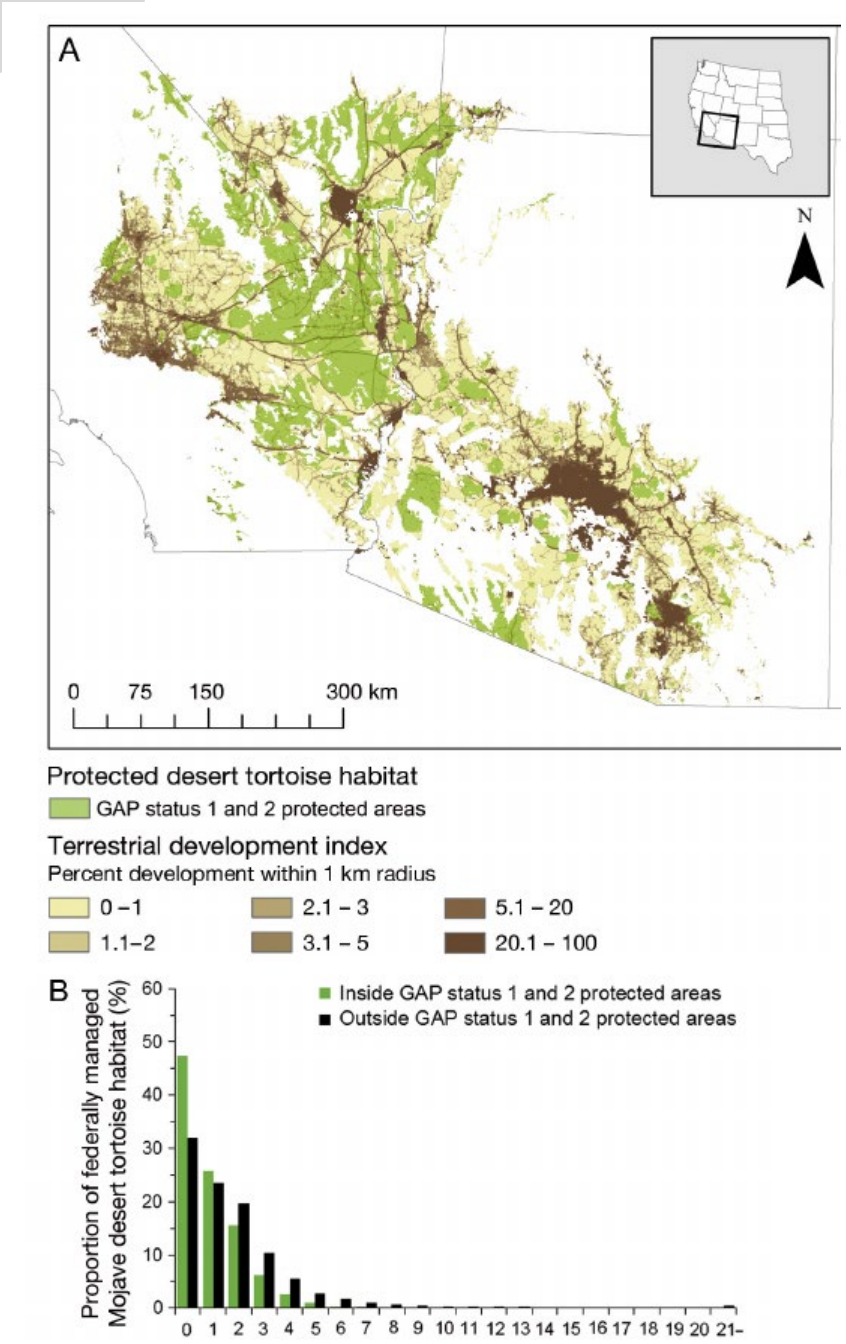
University of Nevada, Reno

Species Distribution Modeling

- Species Distribution Models (SDMs) predict the geographic distribution of species based on statistical associations between known occurrences and environmental variables.
- Early concepts of Niche
- Grinnell (1917) those factors that influence where one would find a species (what it needs).
- Elton (1927) Niche is a species functional role in the biotic community,
- Hutchinson (1957) ’n’ dimensional hyper-volume defined by the environmental dimensions within which that species can survive and reproduce
  - fundamental niche - the response of species to environment (resources) in absence of biotic interactions (competition, predation, facilitation),
  - realized niche – the environmental dimensions in which species can survive and reproduce, including the effects of biotic interactions
- Pulliam (1988) - species may be found outside of the Realized Niche - and in unsuitable habitat - e.g. in “sink” or transitional habitat

Applications

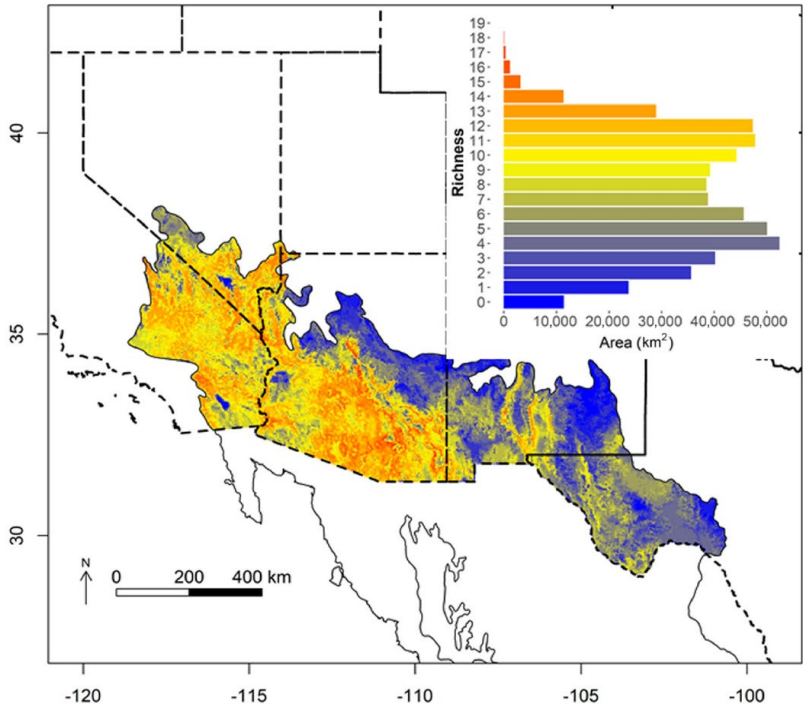
- Conservation planning, ecological forecasting, and risk assessments.
- Identify suitable habitat for rare or endangered species.
- Assess impacts of climate change on species ranges.
- Inform land use and conservation decisions.

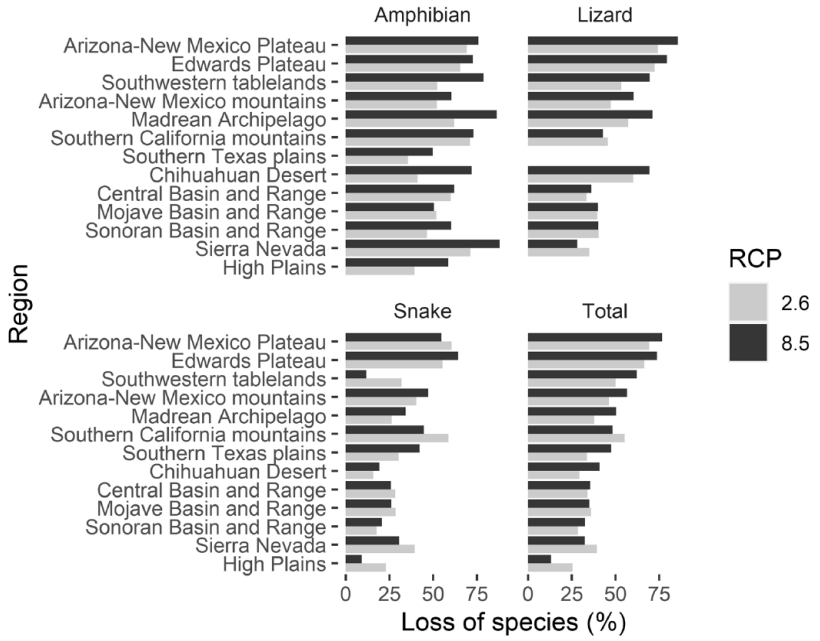


Carter et al. 2020

Applications

- Conservation planning, ecological forecasting, and risk assessments.
- Identify suitable habitat for rare or endangered species.
- Assess impacts of climate change on species ranges.
- Inform land use and conservation decisions.

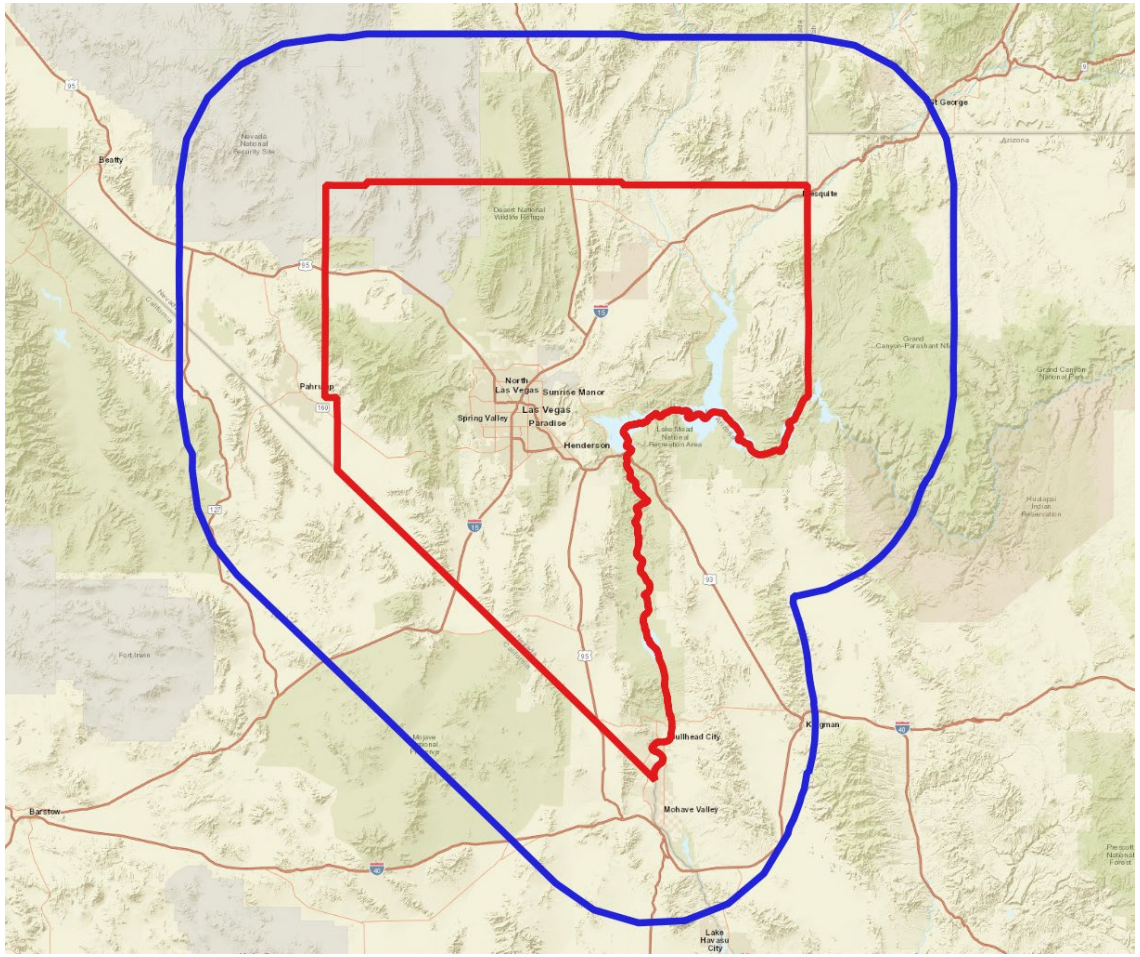




Inman et al. 2022

Study Area

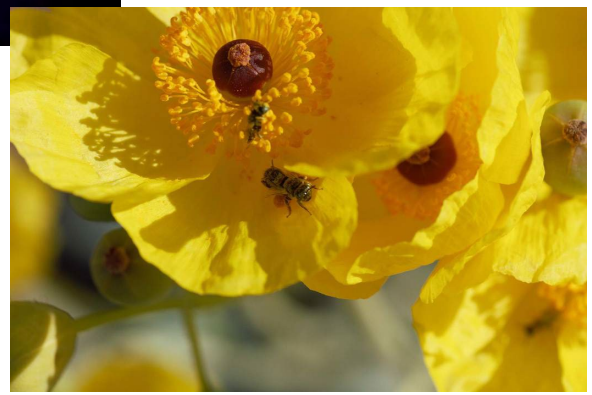
- Clark County, Nevada with 50 km buffer.
- EPSG:26911 (NAD83 UTM Zone 11N).
- 250 m resolution for all rasters.



Study Species

20 Species total

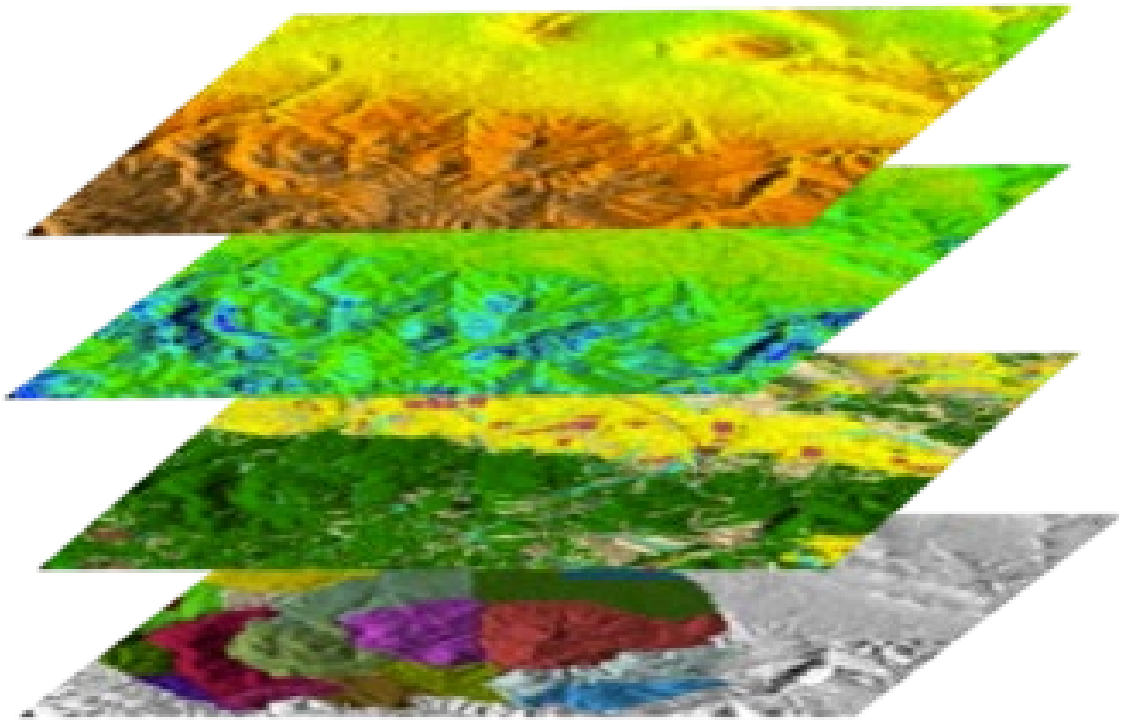
- Bats - 2
- Birds – 3
- Lizards – 2
- Amphibians – 1
- Insects - 2
- Plants - 10



CommonName	ScientificName
Townsend's big-eared bat	*Corynorhinus townsendii*
Spotted bat	*Euderma maculatum*
Yellow-billed cuckoo	*Coccyzus americanus*
Southwestern willow flycatcher	*Empidonax traillii extimus*
Ridgway's rail	*Rallus obsoletus yumanensis*
Banded Gila monster	*Heloderma suspectum cinctum*
Monarch butterfly	*Danaus plexippus*
Mojave poppy bee	*Perdita meconis*
Akali mariposa lily	*Calochortus striatus*
Blue Diamond cholla	*Cylindropuntia multigeniculata*
Parish phacelia	*Phacelia parishii*
St. George blue-eyed grass	*Sisyrinchium radicatum*
Eastern Joshua tree	*Yucca jaegeriana*
Chuckwalla	*Sauromalus ater*
Arizona Toad	*Anaxyrus microscaphus*
Straw Milkvetch	*Astragalus lentiginosus var. stramineus*
Rosy Two-toned beardtongue	*Penstemon bicolor ssp. roseus*
Yellow Two-toned Beardtongue	*Penstemon bicolor ssp. bicolor*
Clarke Phacelia	*Phacelia filiae*
Polished Blazing Star	*Mentzelia polita*

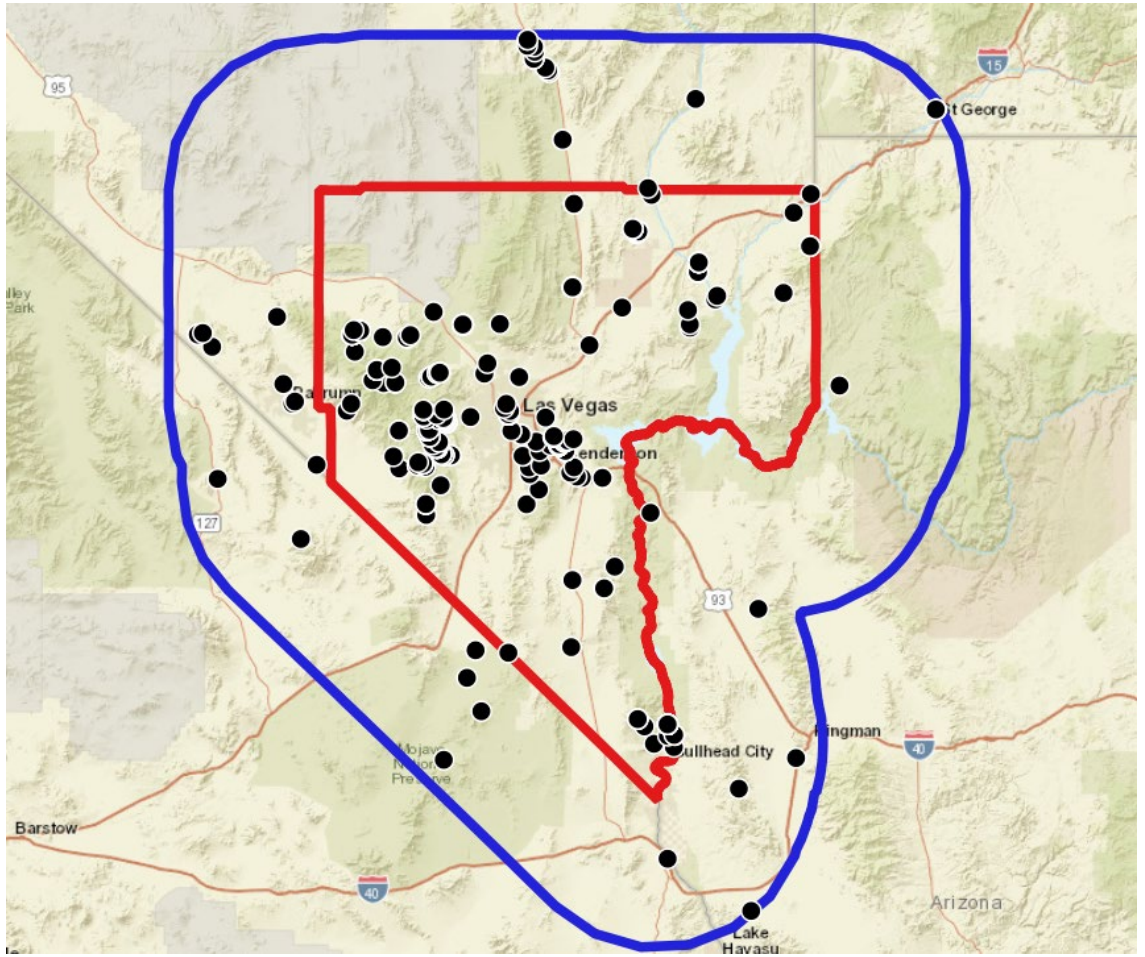
Environmental Variables

- Climate: Avg/Extreme Temp, Precipitation (PRISM, 30-year), variability of these
- Topography: Slope, Cliffs, Flow Accumulation.
- Vegetation: NDVI metrics, Shrub cover, Vegetation Type
- Soils: Sand, Silt, Clay, Gypsom, Silica, Bare ground fraction, Depth to Bedrock (SoilGrids), Ph
- Vapor pressure deficit (indicator of water stress), distance to water



Species Occurrence Data

- Data Locality Sources:
- Global Biodiversity Information Facility (GBIF) - sources with precision matching our resoulstion or better
- iNaturalist (Research Grade)
- Clark County databases
- ...

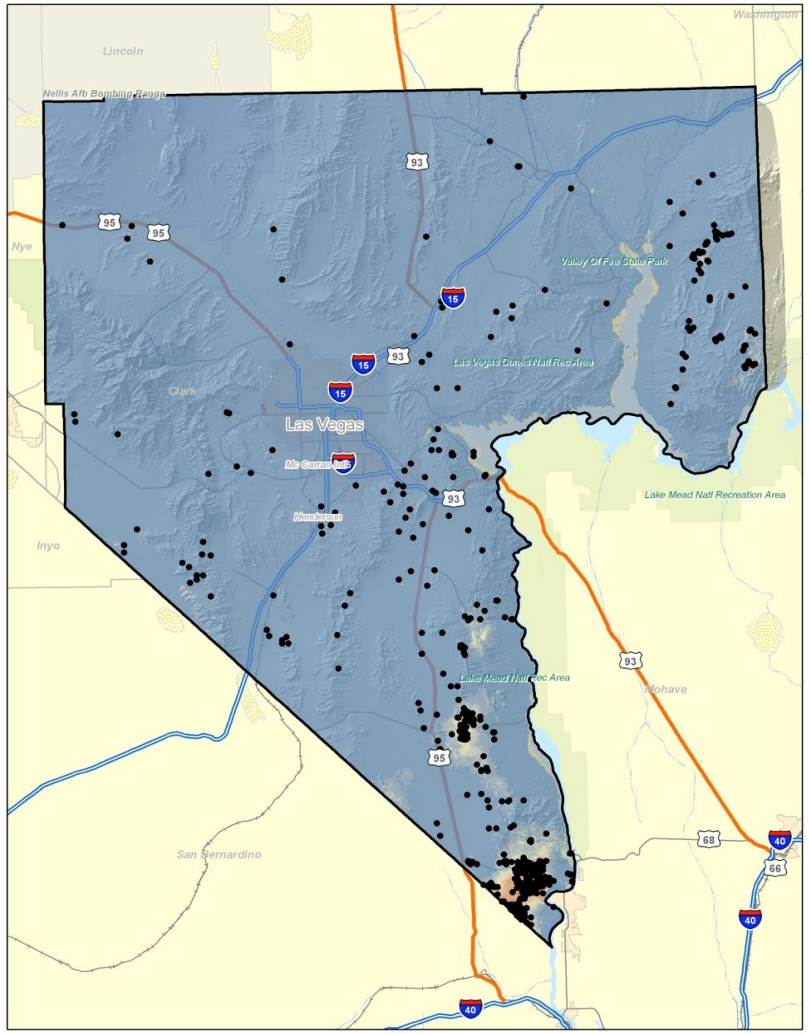


Monarch Butterfly

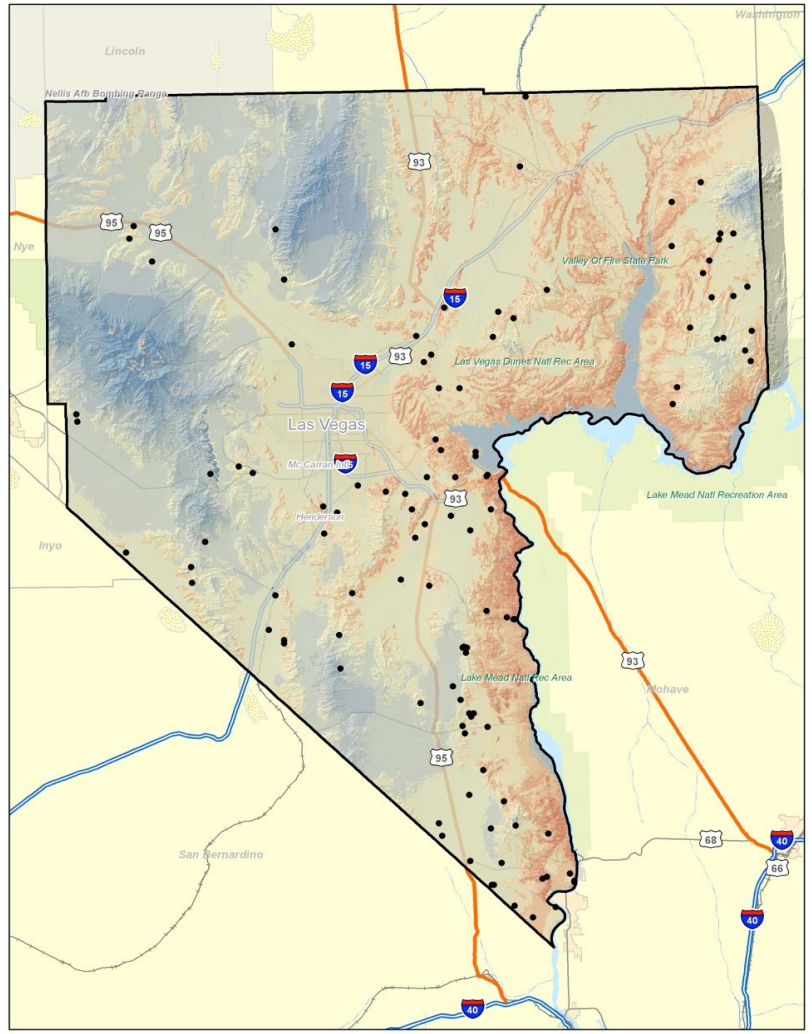
Species Occurrence Data

- Spatial filtering applied to reduce sampling bias
  - One observation per pixel
  - Thinned to 4 within 16 adjacent pixels

All Data



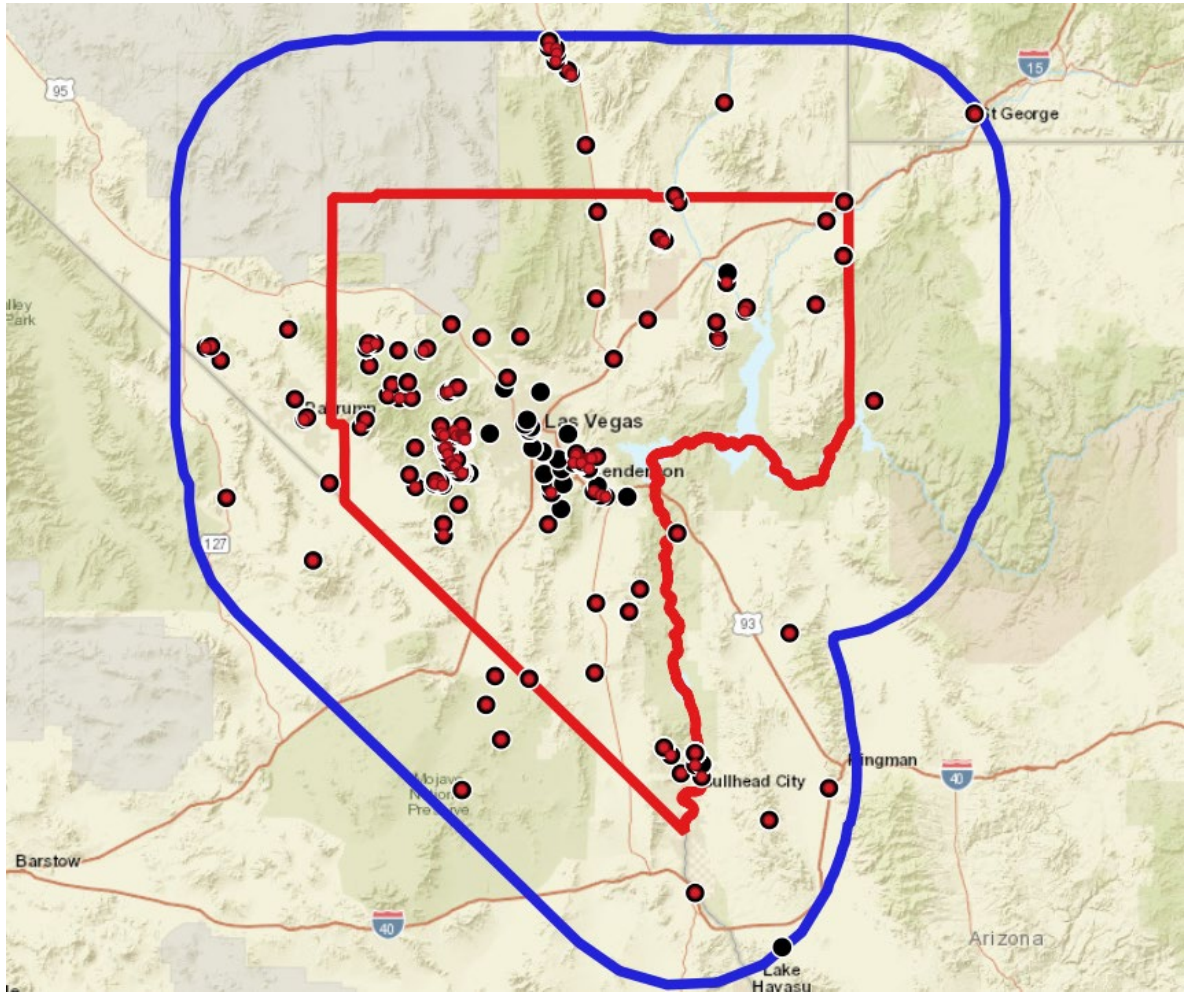
Thinned Data



Species Occurrence Data

- Spatial filtering applied to reduce sampling bias
  - One observation per pixel
  - Thinned to 4 within 16 adjacent pixels

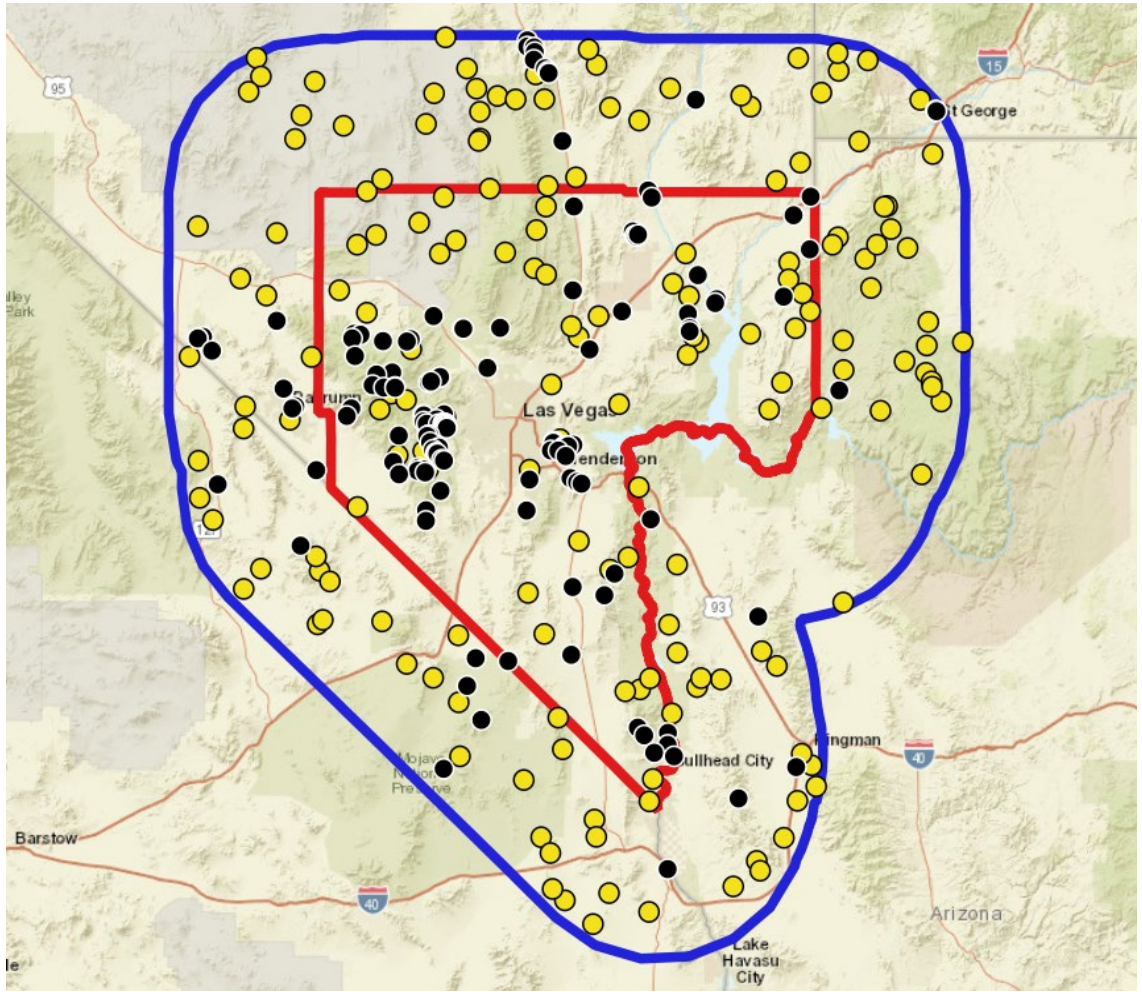
Total	Thinned
547	142



Monarch Butterfly

Species Occurrence Data

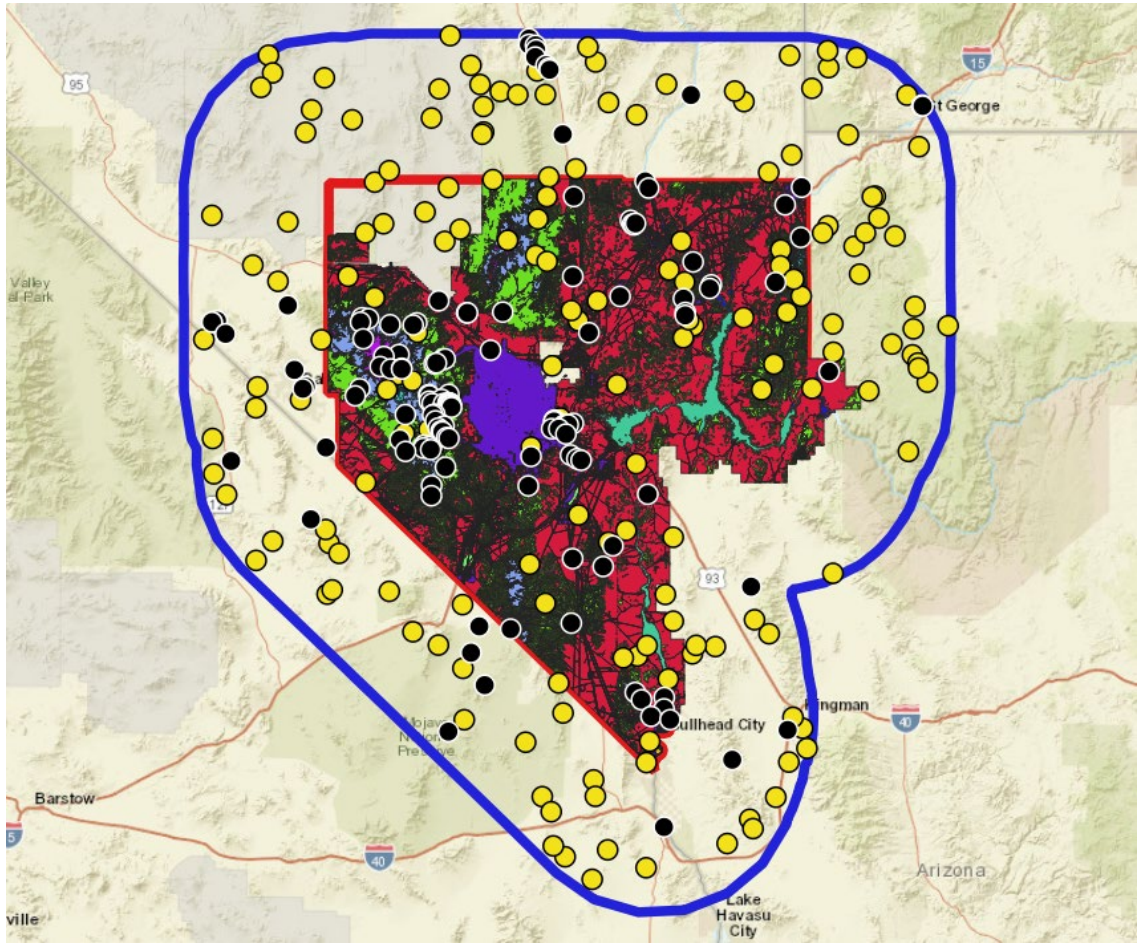
- Pseudo Absences used to sample unoccupied background areas
- Used bioclimatic envelope model to select background points.



Monarch Butterfly

CC Vegetation Layer

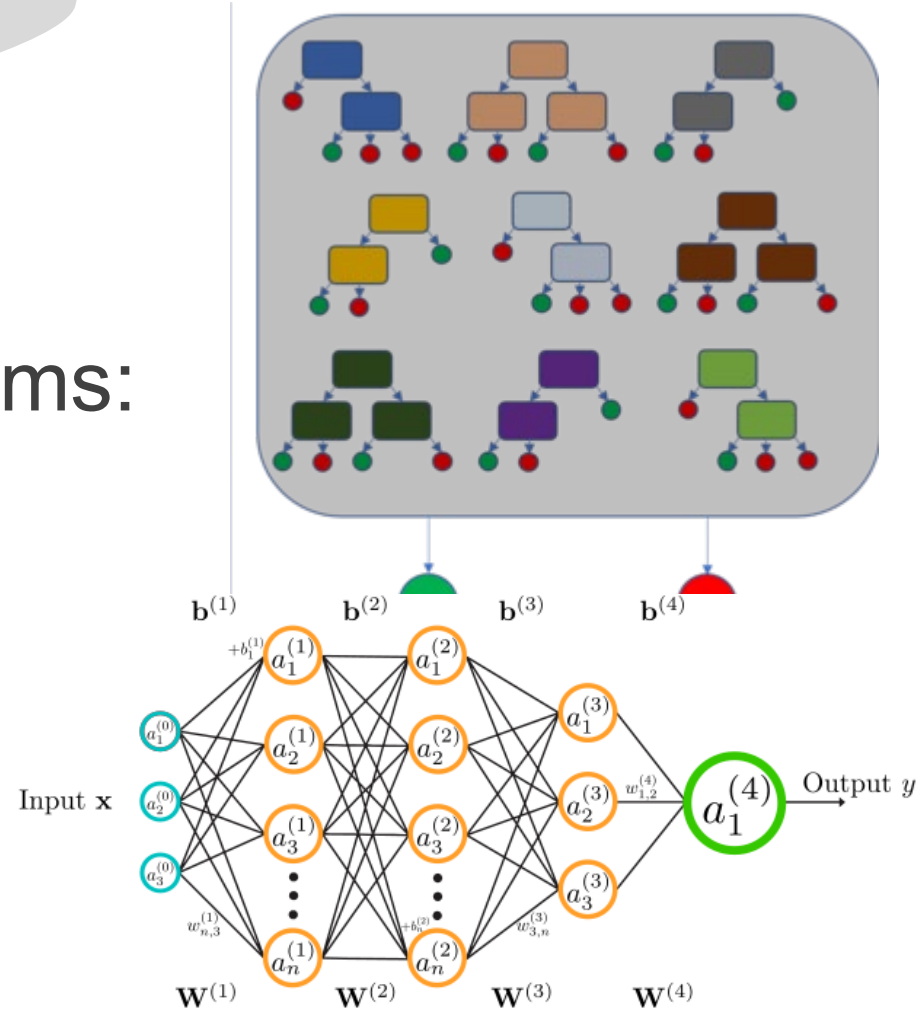
- Vegetation Layer from CC introduces 14 categorical vegetation alliances used in modeling efforts
- Some reduction of points due to smaller footprint of available data
- Can lead to different predictions

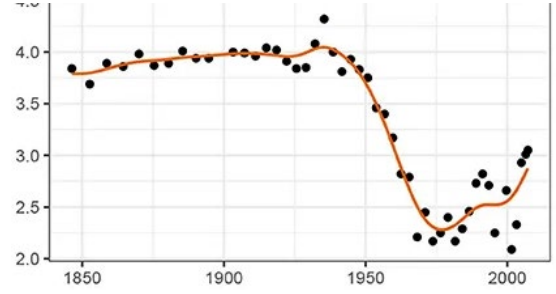


Modeling Approach

- Ensemble method using four algorithms:
  - GAM (Generalized Additive Models)
  - Random Forest (RF)
  - Generalized Boosted Models (GBM)
  - MaxEnt (Maximum Entropy)

- All implemented in R
  - with parallelization

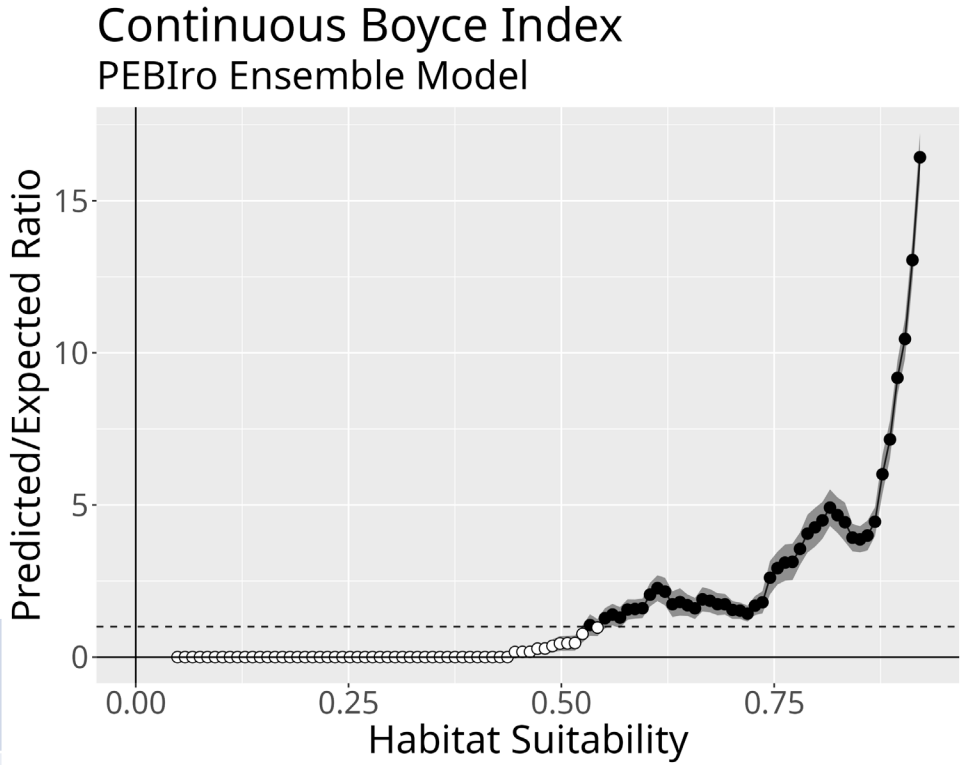




Performance Evaluation

Rosy Two-toned beardtongue

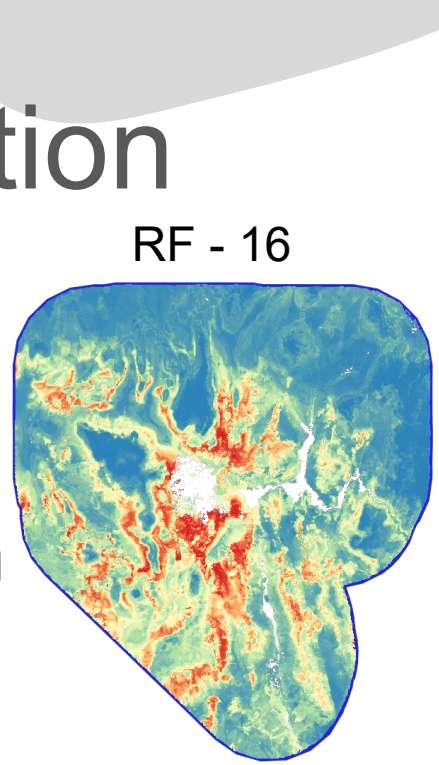
- Metrics used:
- • AUC: Area Under ROC Curve
- • TSS: True Skill Statistic
- • BI: Boyce Index (P/E Curve)
- Evaluation on withheld (20%) test data.

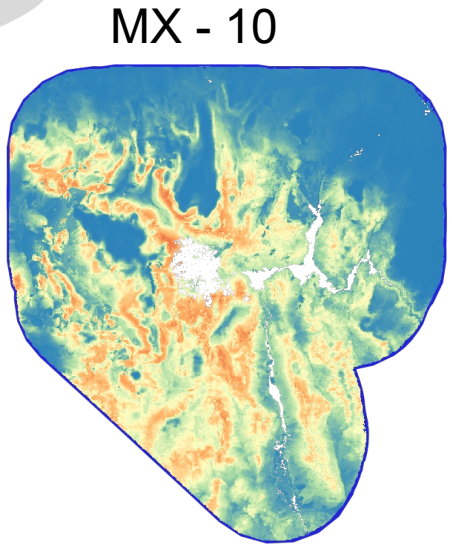


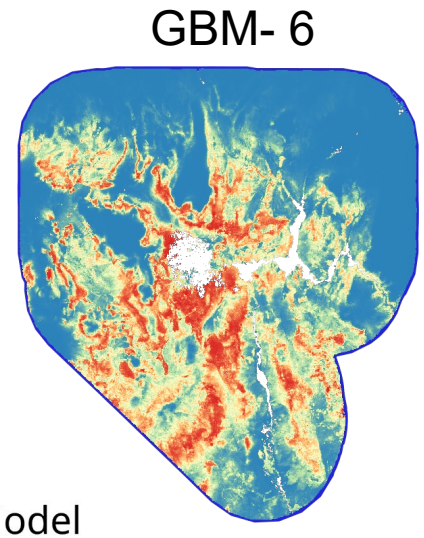
Stat	Training	TrainingSD	Testing	TestingSD
AUC	1.00	0.00	1.00	0.00
TSS	0.94	0.01	0.95	0.02
BI	0.99	0.01	0.94	0.04

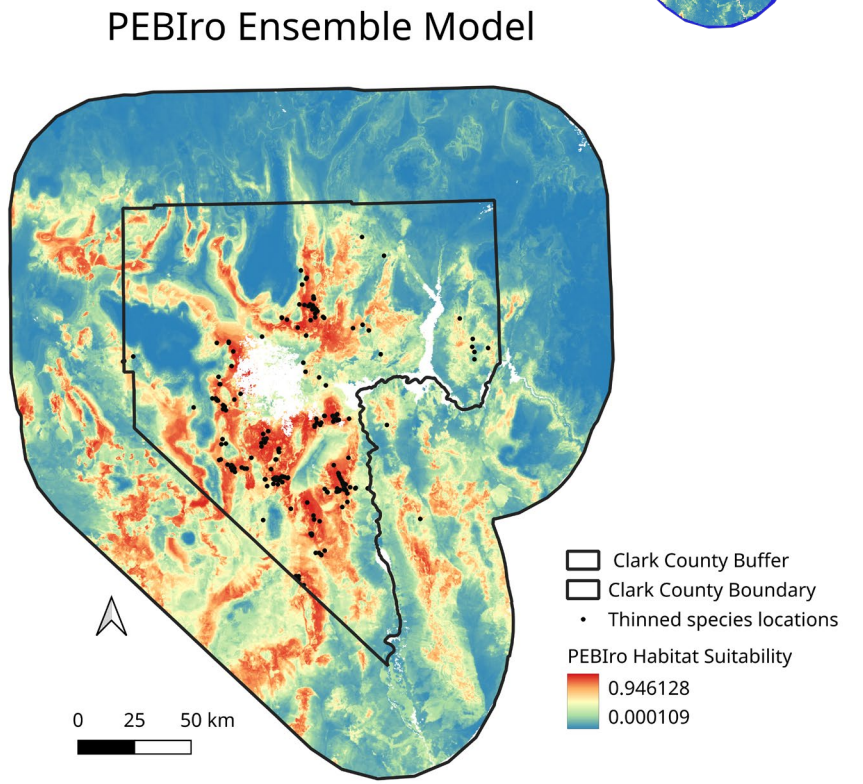
Ensemble Construction

- 20-fold cross-validation for model robustness.
- Performance-based selection using True Skill Statistic to identify the best models among all algorithms
- Models are combined using weighted averaging by TSS scores.



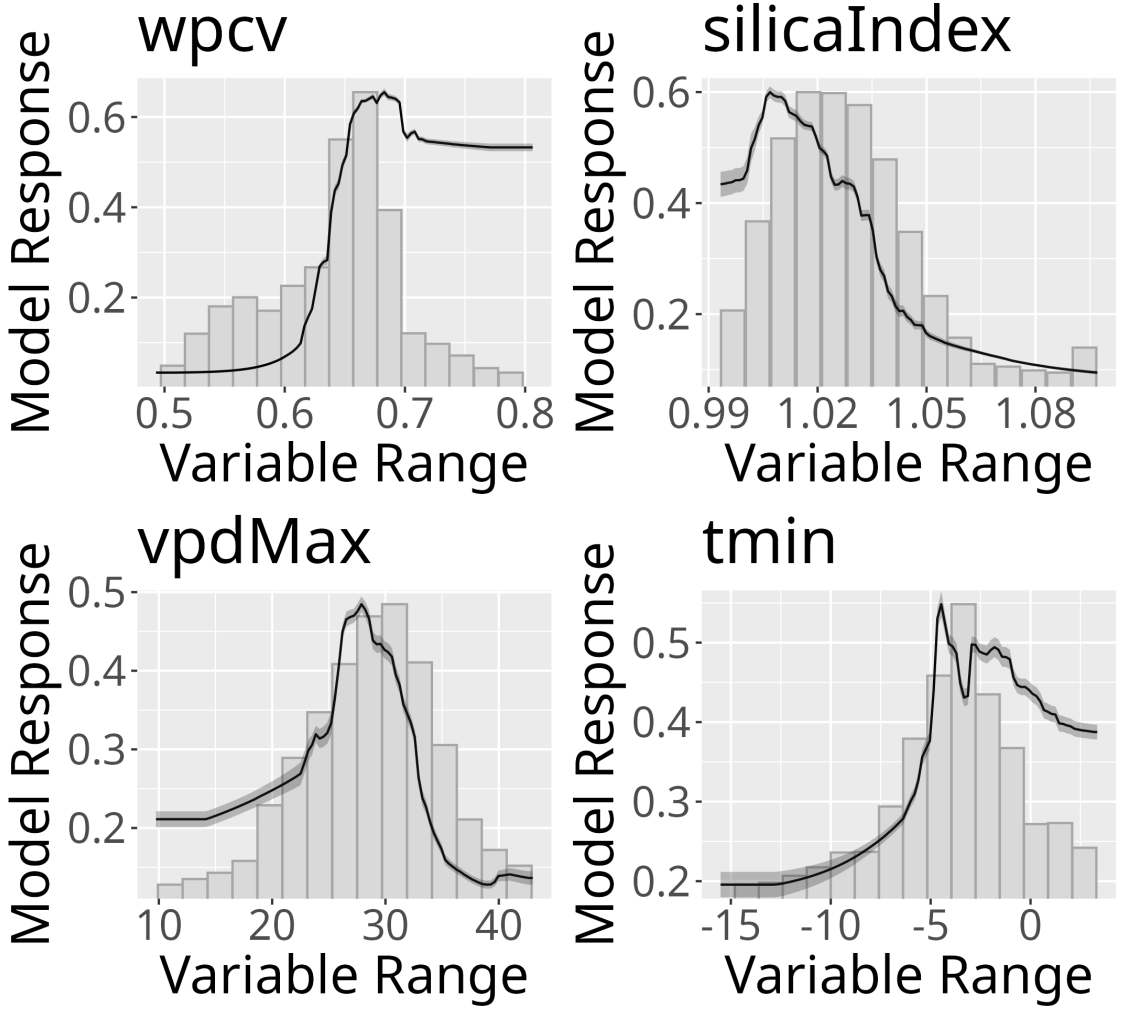






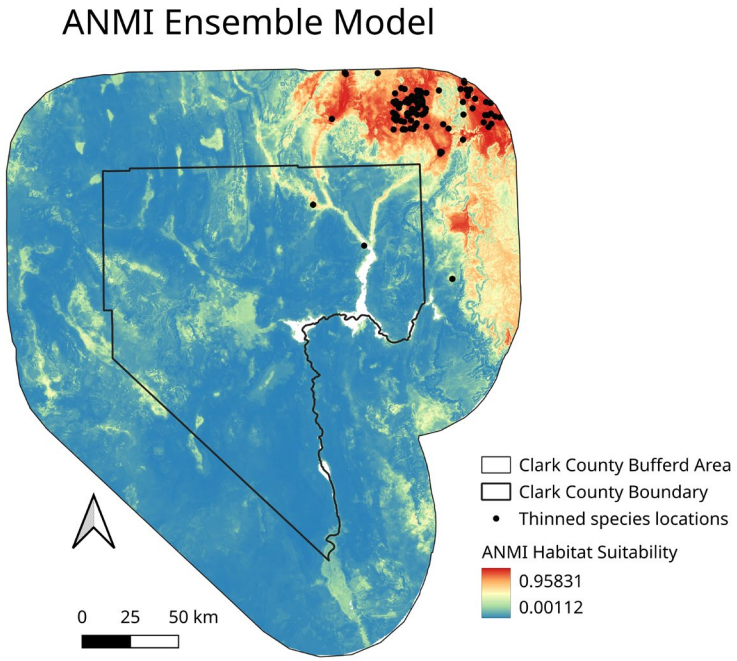
Model Interpretation

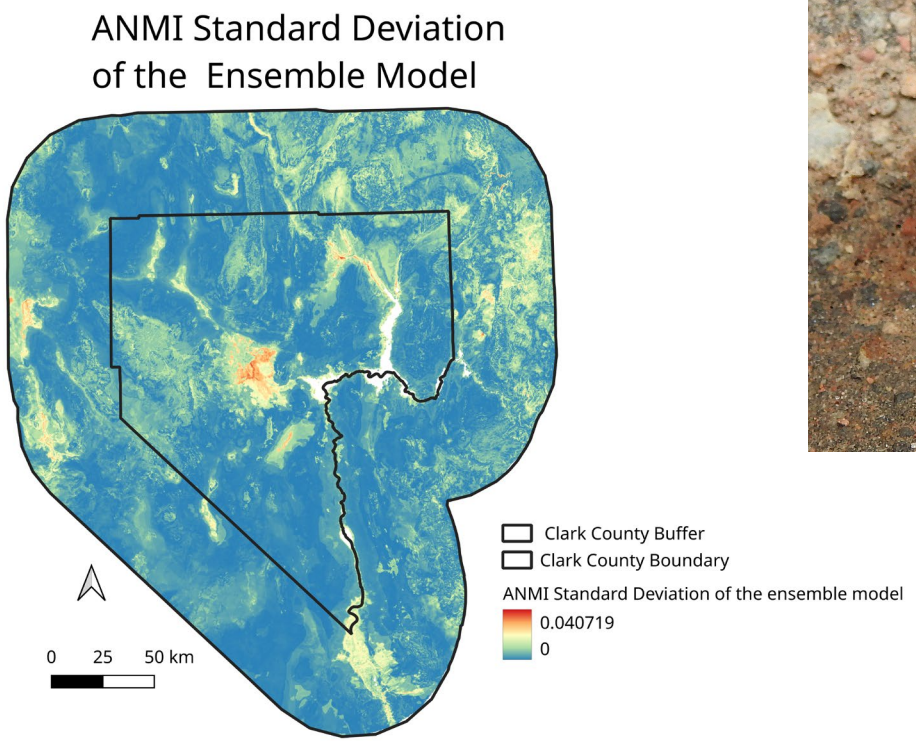
- Relative importance of predictors computed per model.
- Partial response curves created for top 4 variables.
- Histograms overlay to show covariate distribution.



Arizona Toad - *Anaxyrus microscaphus*

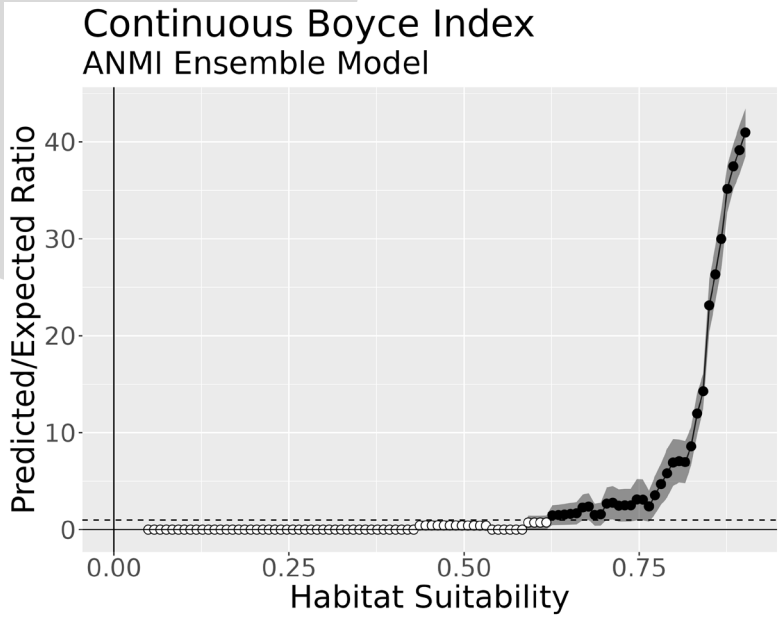
- Ensemble model map and uncertainty visualization.



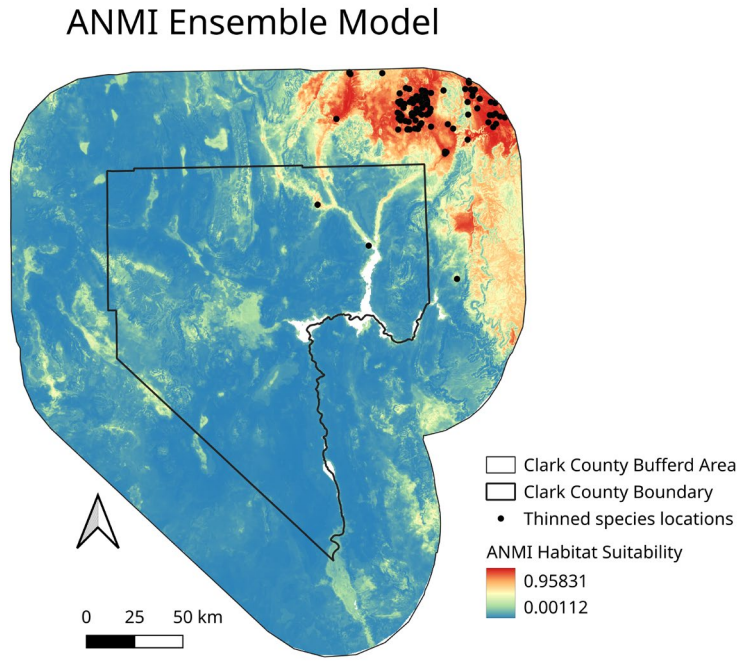


- Model performance

Arizona Toad

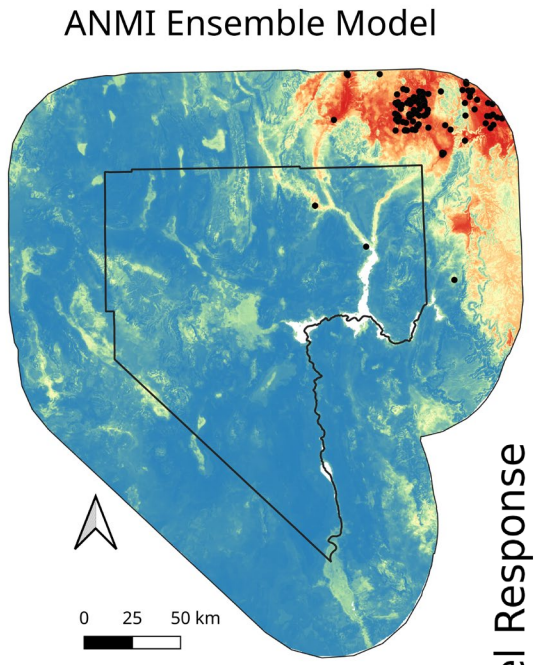


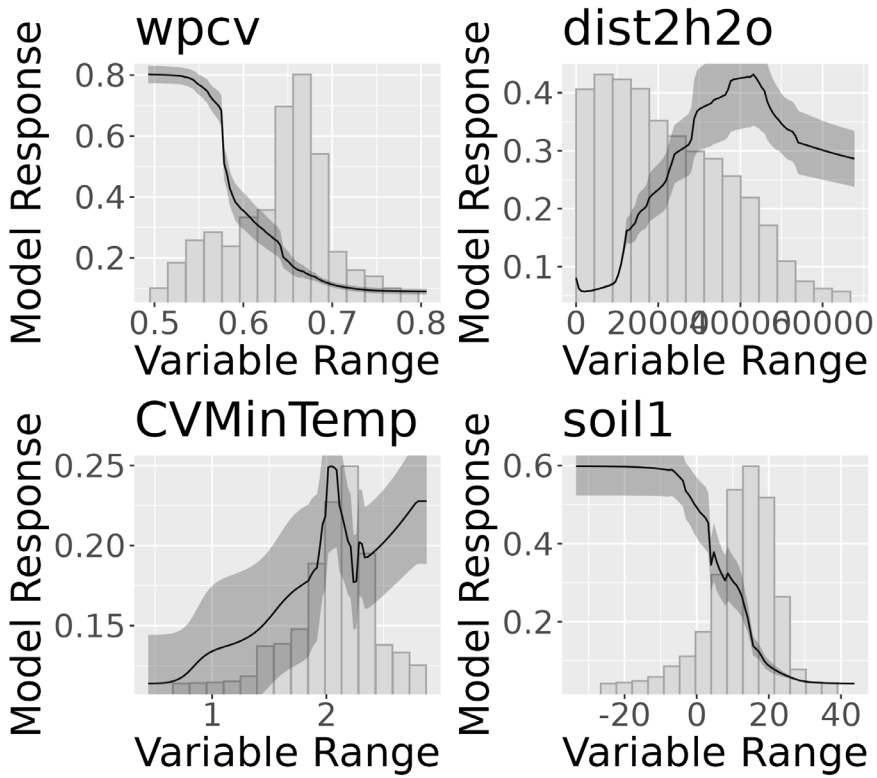
Stat	Training	Training SD	Testing	Testing SD
AUC	1.00	0.00	1.00	0.00
TSS	0.99	0.01	0.99	0.02
BI	0.93	0.04	0.83	0.13



- Variable Importance

Arizona Toad

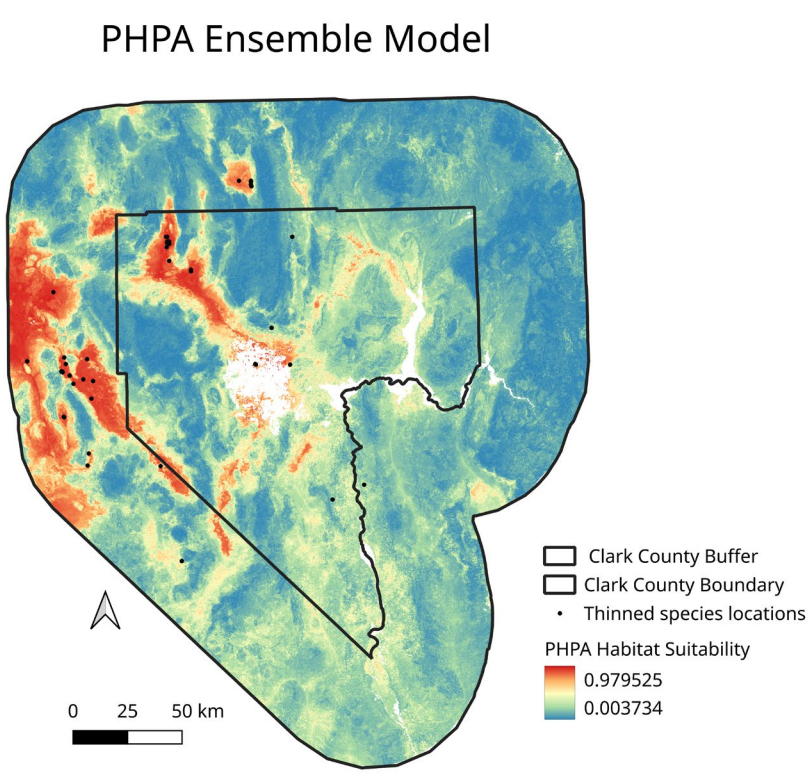


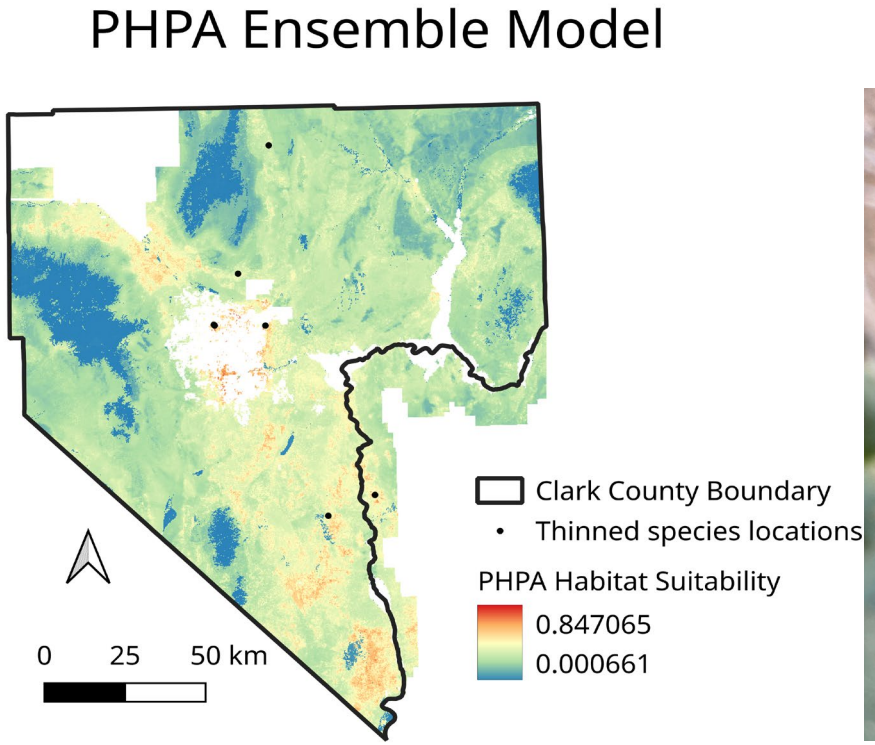


Variable	gam	gbm	mx	rf
aspect.r	6	0	0	1.9
bare.r	0	0	0	2.4
bedrock.r	0	0.9	0	4.8
CVMinTemp.r	8	0	9.4	8.4
slope.r	12	0.2	0.3	4.2
soil1.r	15	0.9	0.2	10.5
tmin.r	7	0	1.2	9.6
vpdMax.r	19	2.2	0.4	6.9
wpcv.r	20	93.3	88.3	37.4
dist2h2o.r	13	2.5	0	13.9

Parish phacelia - Phacelia parishii

- Ensemble model maps





Calflora.org

Parish phacelia - Phacelia parishii

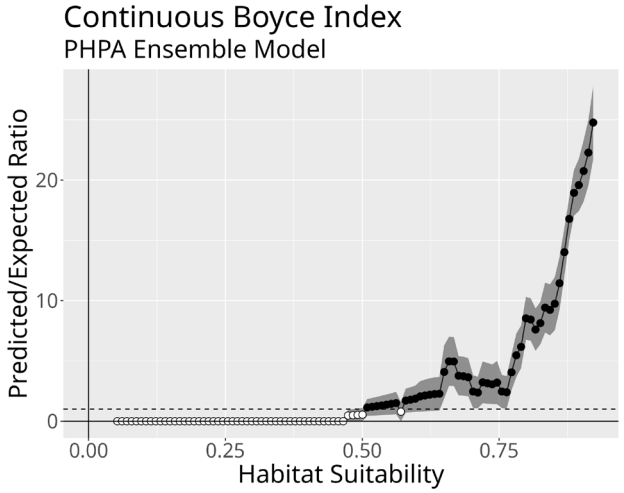
- Model Performance

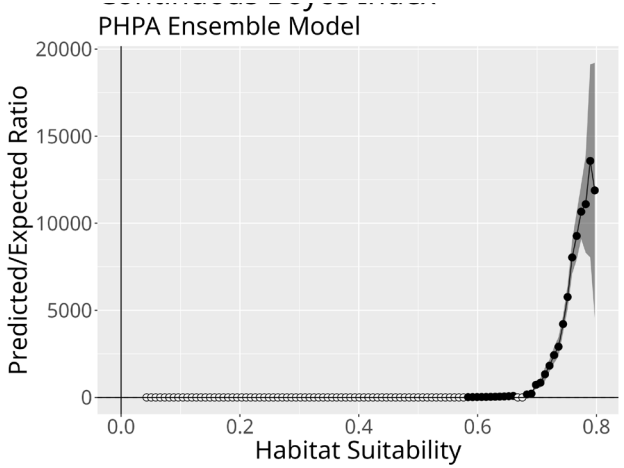
Global Model N=31

Stat	Training	Training SD	Testing	Testing SD
AUC	1.0	0.0	1.0	0.0
TSS	1.0	0.0	1.0	0.0
BI	1.0	0.0	0.9	0.1

CC Vegetation Model N=5

Stat	Training	Training SD	Testing	Testing SD
AUC	1	0	1	0
TSS	1	0	1	0
BI	1	0	NA	NA



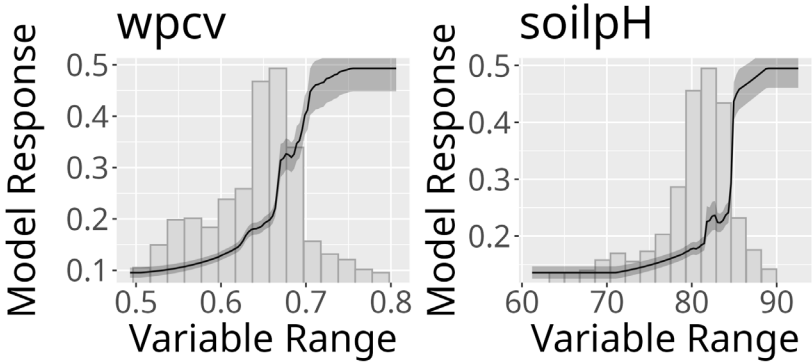


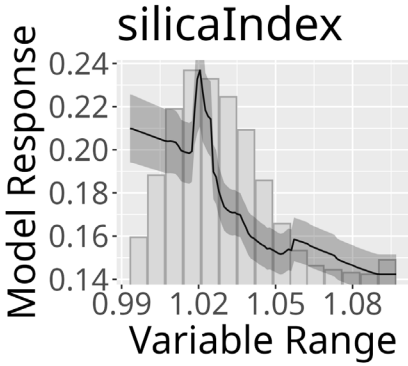
Calflora.org

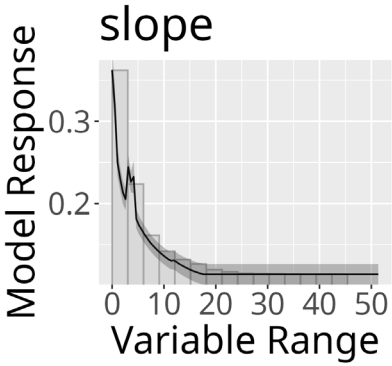
Parish phacelia - *Phacelia parishii*

- Variable Importance

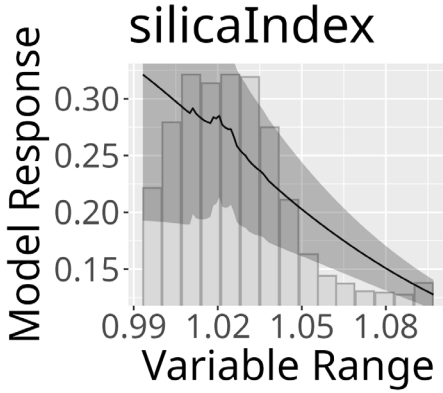
Global Model N=31

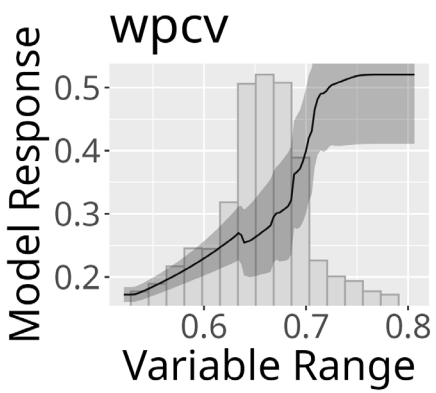


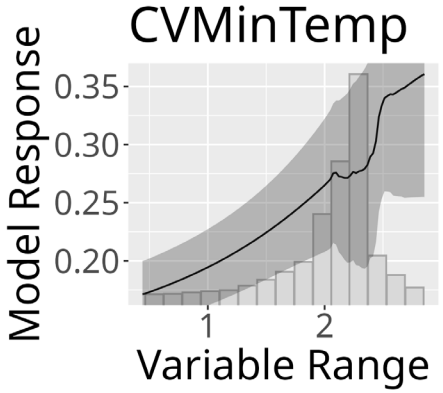


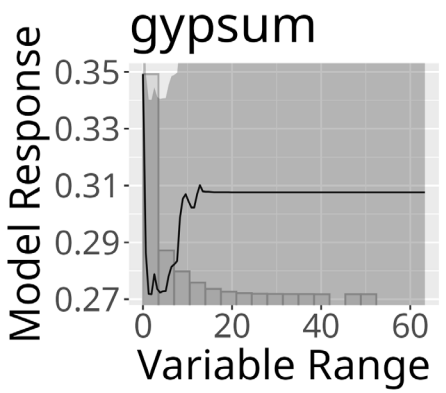


CC Vegetation Model N=5







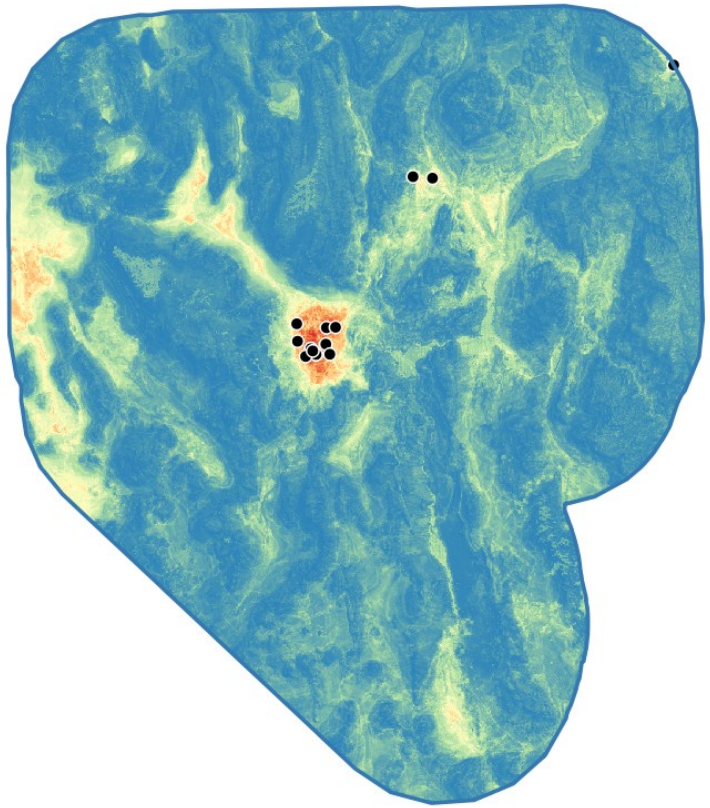


Calflora.org

Spotted bat - *Euderma maculatum*

The problem of too few points

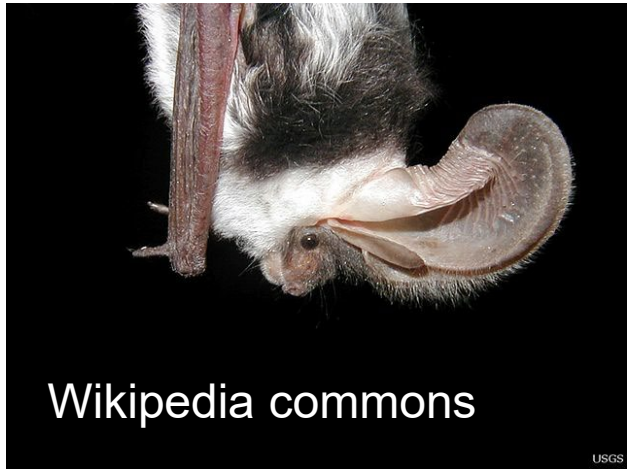
- Ensemble model map – Ensemble of Small Models



66 Models for each of 3 algorithms using Pairwise Combinations of 12 Environmental Variables

3 Algorithms – GBM, Maxent and RF

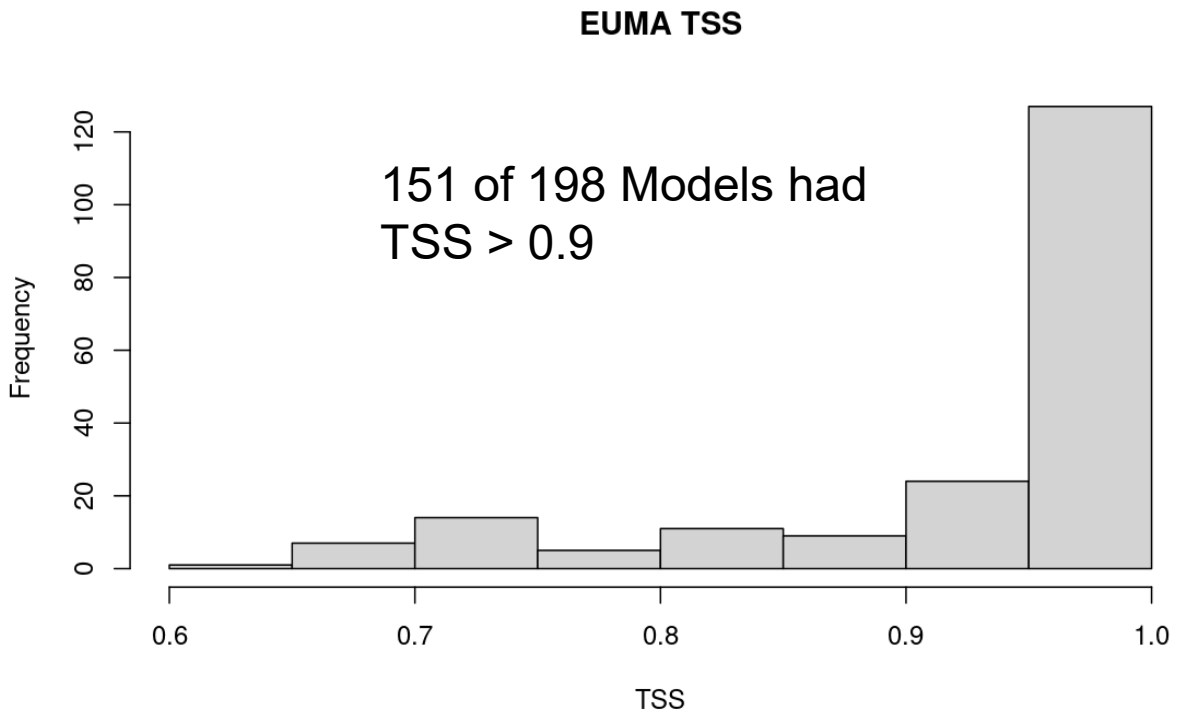
Threshold by best TSS (0.9) and Weighted by individual model TSS

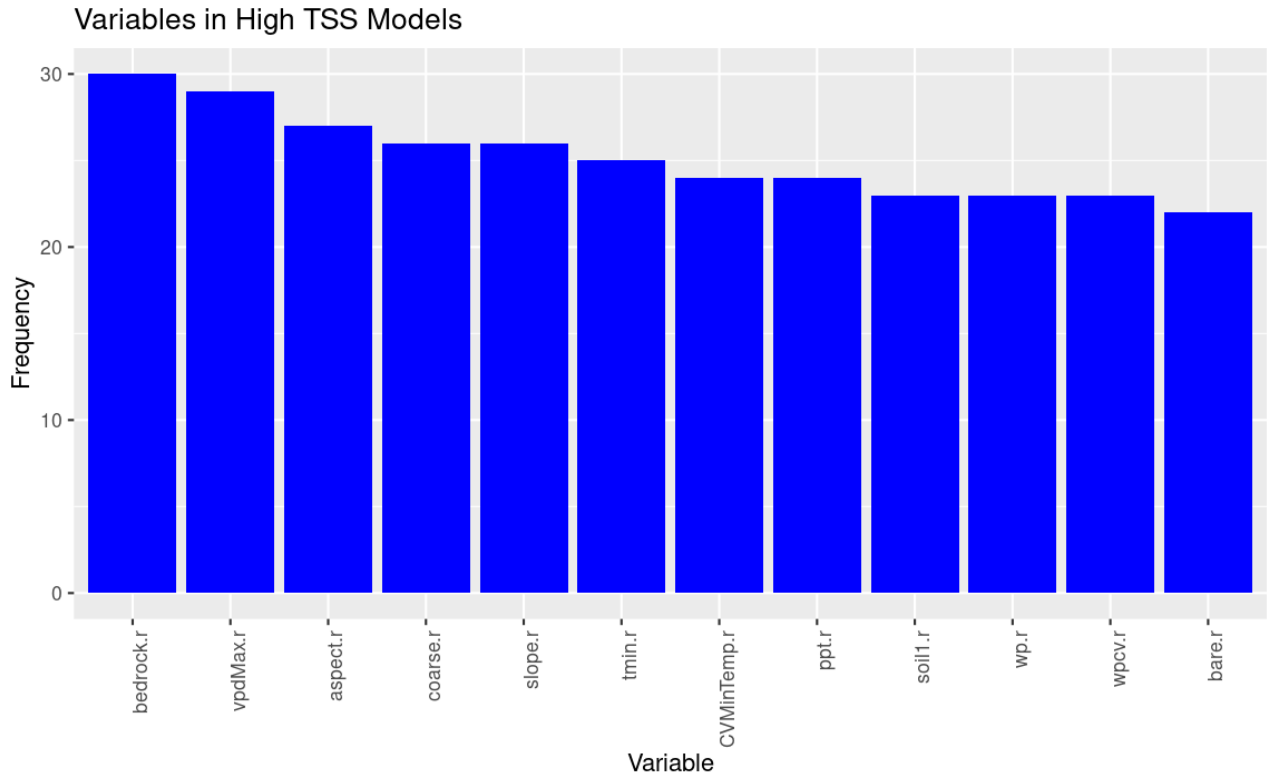


Spotted bat - *Euderma maculatum*

The problem of too few points

- Performance – Ensemble of Small Models



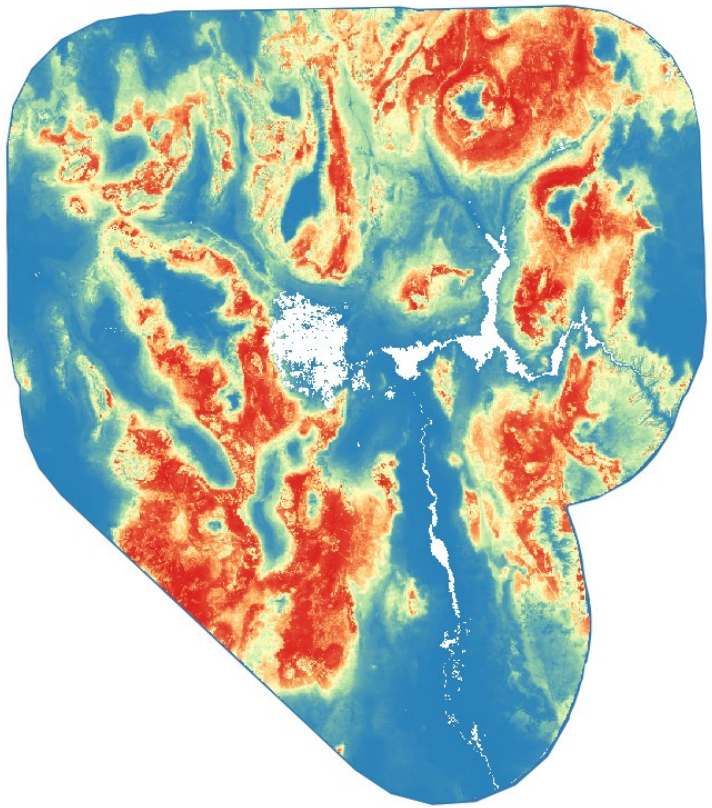


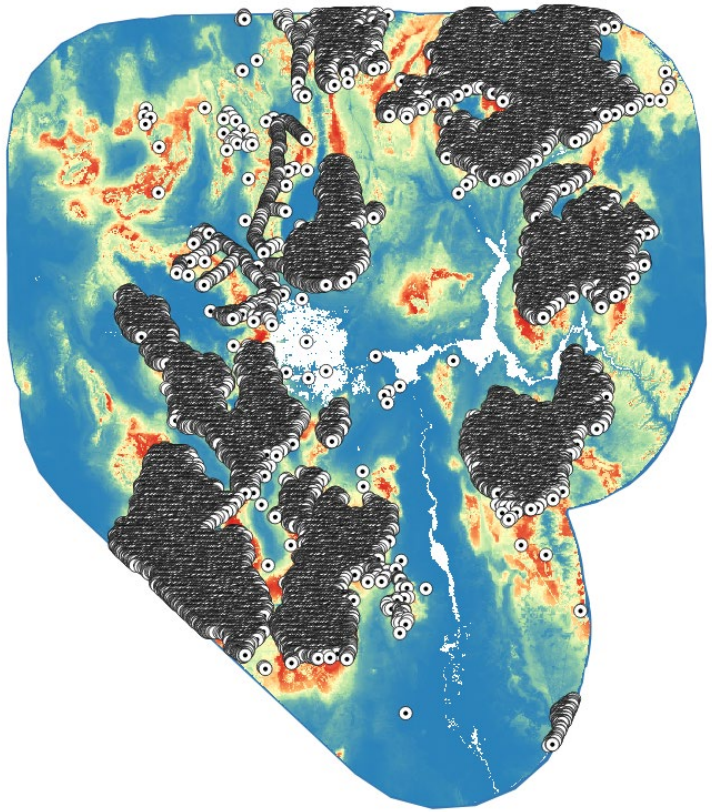
Eastern Joshua tree - *Yucca jaegeriana*

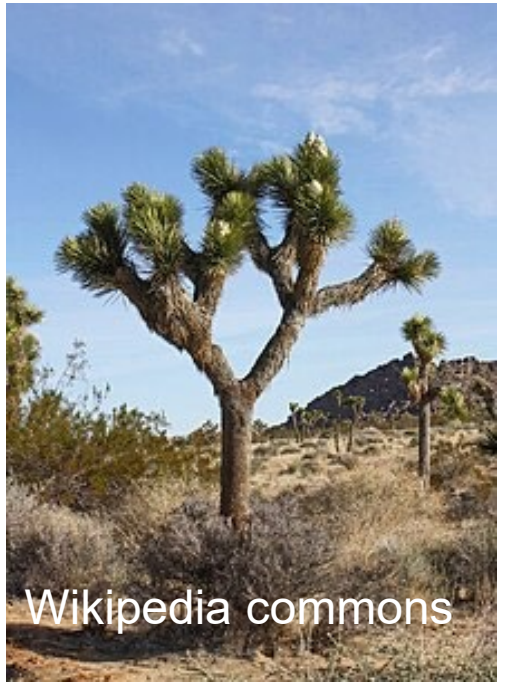
The problem of too many points – but missing a few

- Ensemble model map

109K points! Thinned to 55K for modeling





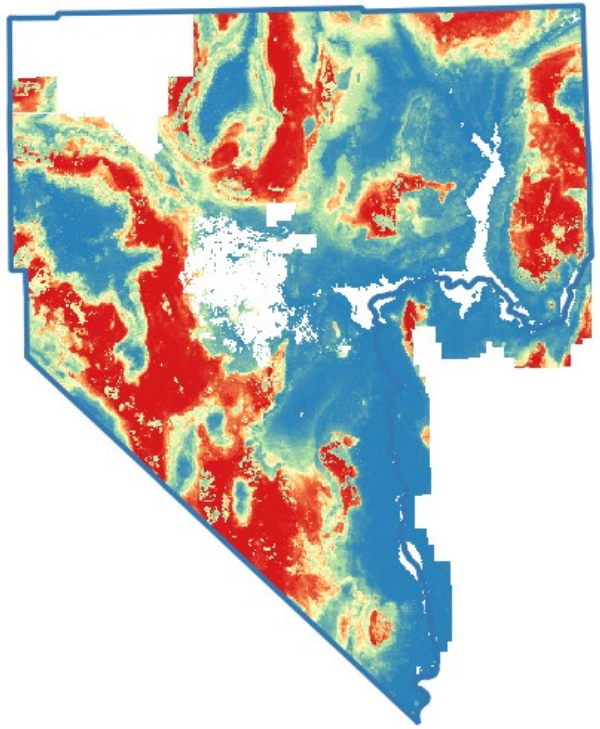


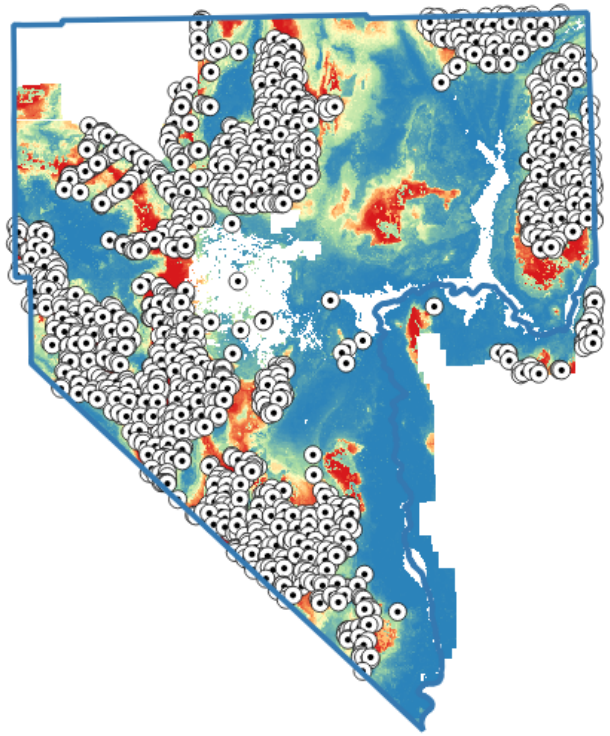
Eastern Joshua tree - *Yucca jaegeriana*

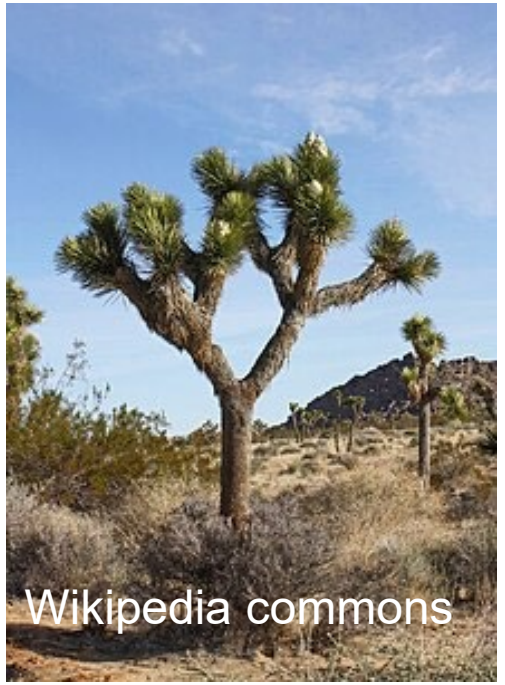
The problem of too many points – but missing a few

- Ensemble model map - CC Vegetation Included

109K points! Thinned to 55K for modeling







Eastern Joshua tree - *Yucca jaegeriana*

The problem of too many points – but missing a few

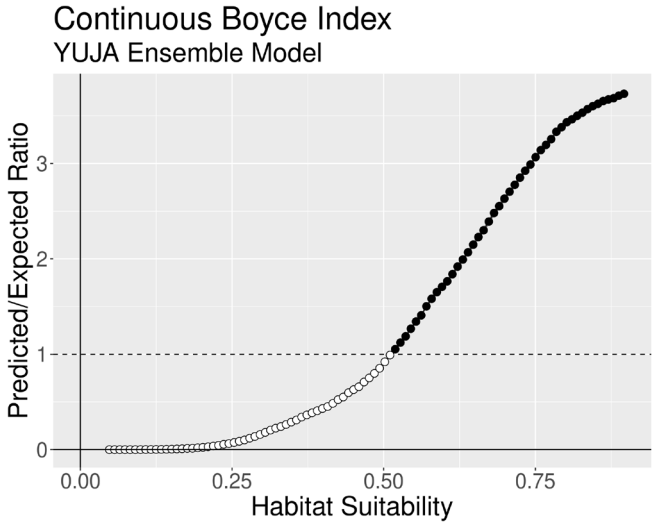
- Model Performance

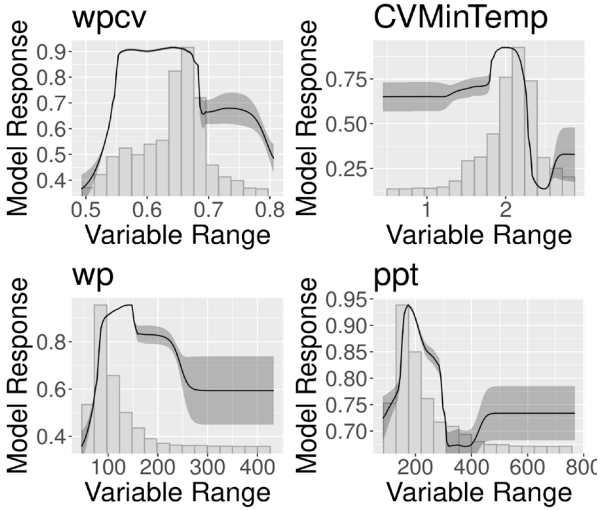
Global Model

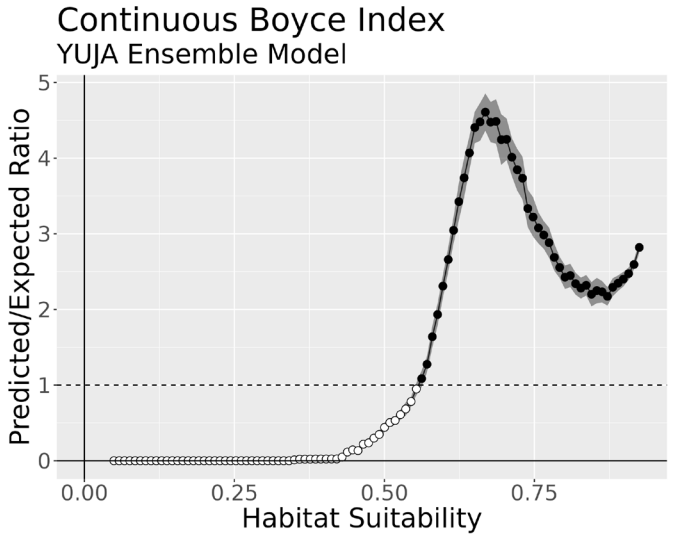
Stat	Training	Training SD	Testing	Testing SD
AUC	0.97	0.00	0.96	0.00
TSS	0.78	0.00	0.77	0.00
BI	1.00	0.00	1.00	0.00

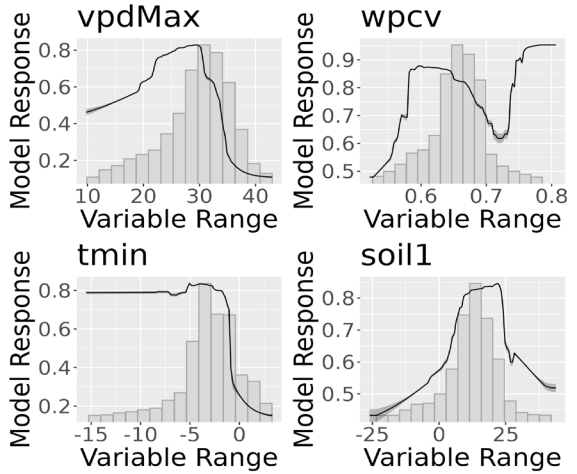
CC Veg Model

Stat	Training	Training SD	Testing	Testing SD
AUC	1.00	0.00	1.00	0.00
TSS	1.00	0.00	1.00	0.00
BI	0.69	0.04	0.59	0.08







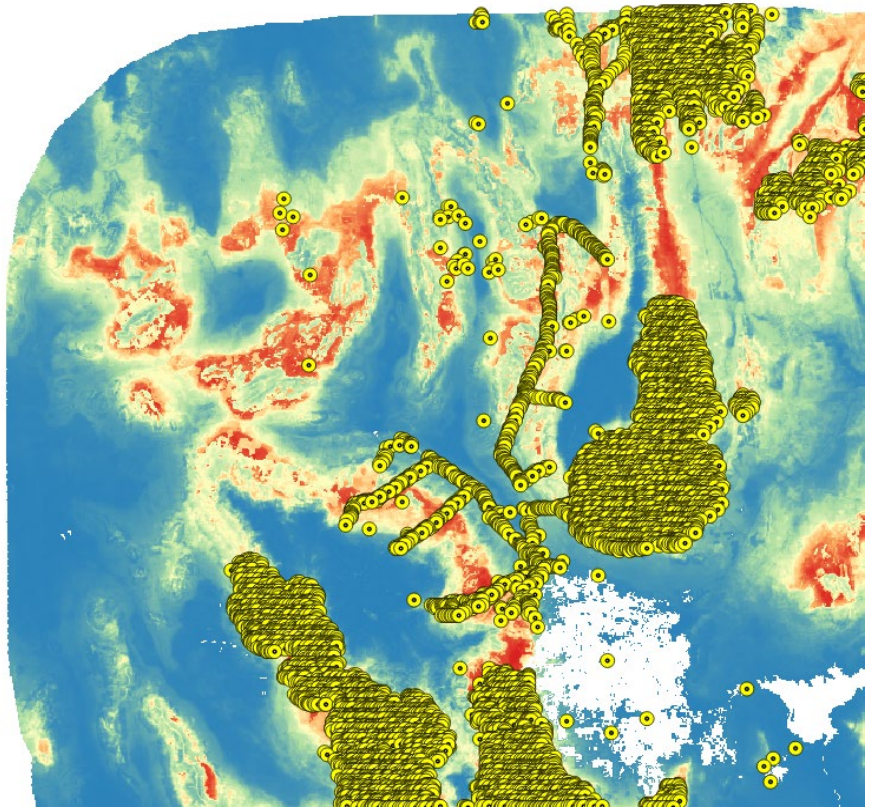


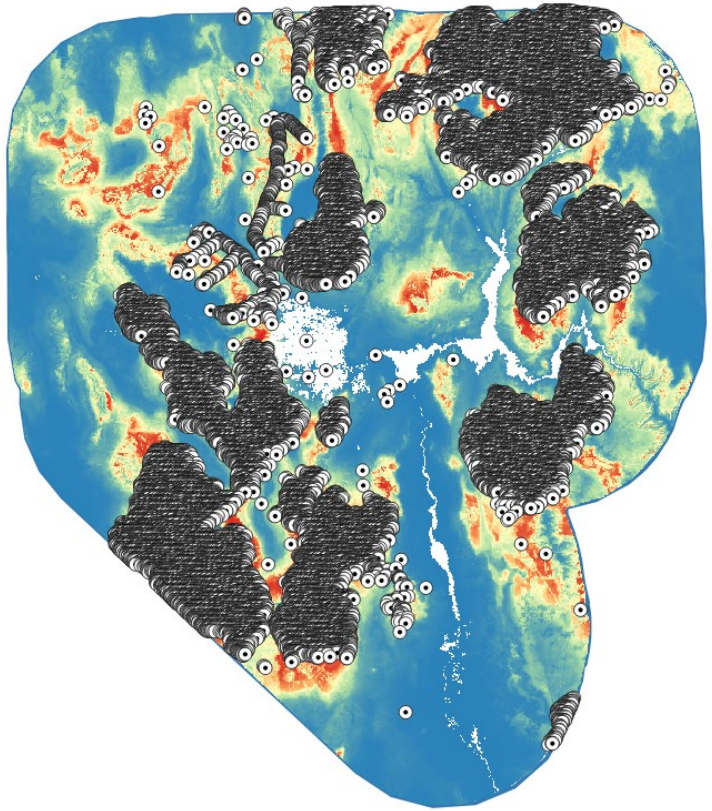
Eastern Joshua tree - *Yucca jaegeriana*

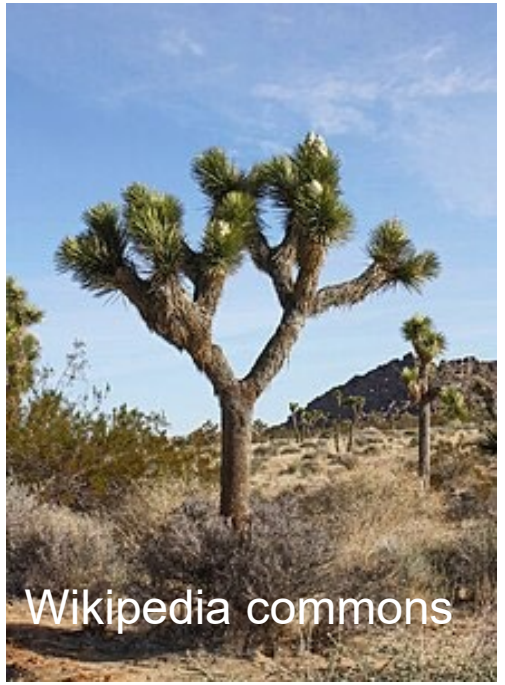
The problem of too many points – but missing a few

- Ensemble model map

109K points! Thinned to 55K for modeling







Summary and Next Steps

- High-performance ensemble SDMs generated for all species.
- Models incorporate multiple data sources and reduce spatial bias.
- Next Steps:
- Finalize models with additional data where needed