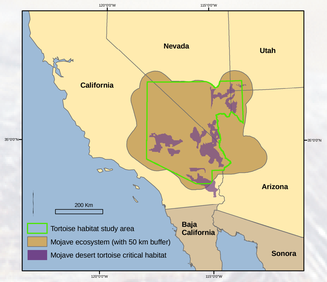
Choices and Challenges for Spatial Habitat Models: Predicting Desert Tortoise Habitat

Kathryn Thomas, Leila Gass, Todd Esque, Robert Webb, and David Miller

The Desert Tortoise

Desert tortoises (*Gopherus agassizii*) are land-dwelling reptiles that occur in the Mojave and Colorado/Sonoran deserts of California, southern Nevada, Arizona, southwestern Utah, and northern Mexico. Two populations exist; the Mojave population, living north and west of the Colorado River-Grand Canyon complex, was listed in 1980 and is considered 'Threatened'. We defined the study area of this project to incorporate the known range of the Mojave population of desert Tortoise.



Project Goal

Multiple-species habitat conservation plans associated with rapidly urbanizing areas, such as the Las Vegas Valley, often rely on translocation of displaced animals or mitigation through purchase and preservation of habitat elsewhere. The efficiency, and ultimately the success, of these management plans depend on the ability to define and delineate potential habitat for the species of concern.

Our goal is to define and delineate suitable habitat for the Mojave Desert population of the desert tortoise. In order to accomplish this we are:

- developing statistically-based spatial models of desert tortoise distribution and habitat,
- comparing different modeling approaches, and
- using the best models to develop habitat suitability maps

Step 1: Species Occurrence Data

We compiled seven existing datasets of desert tortoise occurrence. These data are not systematic, and they are not consistent in collection methods and sampling of the study area. In filtering these occurrence data we:

1) defined a positive occurrence as the observation of a live tortoise, presence of shell remains, or presence of a carcass, and
2) used the requirements of the modeling applications to mitigate the effects of the non-random and 'clumped' distribution of the occurrence records.

Both GARP and Biomapper require raster input data, but they differ slightly in the process of converting presence points to occurrence rasters. GARP assigns a positive value to any raster cell in which one or more occurrence points may occur. Biomapper requires the modeler to create the raster data a-priori, so we used the same logic as is applied by GARP. In either case, raster cell size influences the number of positive raster cells created from presence points.

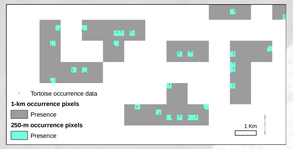


Table 1. Results of filtering presence point data into occurrence rasters.

	GARP	Biomapper
Number of presence points, point dataset	3610	3610
Number of presence raster cells, 250 m raster	2782	2778
Number of presence raster cells, 1 km raster	2028	2022

Step 2: Environmental Predictor Maps

We assembled spatial datasets that include existing data and new datasets developed for this project. These include:

Topography: elevation, slope, and aspect
Climate: seasonal precipitation and variation in precipitation
Substrate: soils (bulk density depth, and rock size) and bedrock geology
Vegetation: perennial cover

Except for bedrock geology, which was used only in GARP, we used the same environmental data for both GARP and Biomapper modeling.

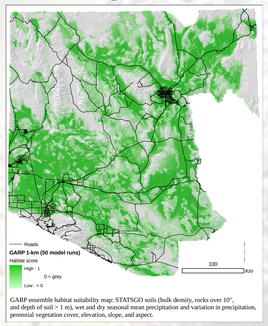
Step 3: Predictive Modeling

We are currently comparing modeling approaches using two existing applications: GARP (Genetic Algorithm for Rule-set Production, Schachetti-Pereira 2002) and Biomapper (Hirzel, A., Hausser, J., Perrin, N., 2002.). Both applications use species occurrences and environmental predictors as input data, but they differ in both modeling algorithms and input data requirements.

GARP

GARP produces models that map the environmental conditions under which a species should be able to maintain populations by searching iteratively for non-random correlations between species occurrence and environmental attribute values. Pseudo-absence locations are automatically generated by random selection from "background" (non-presence) cells. GARP employs four search rules: atomic, logistic regression, bioclimatic envelope, and negated bioclimatic envelope rules. Since the algorithm is stochastic, the model output is not static: each model run produces a somewhat different predicted distribution.

The modeler chooses: 1) the number of model runs (50), 2) the convergence limit (0.01) and maximum number of iterations per run (1000), and 3) the proportions of occurrence points to be used for training and for validation (70%/30%). We chose 50 model runs as a compromise between computation time and number of models run.



The output of each model run is a) a grid, with each cell categorized either 1 (yes) or 0 (no) for predicted desert tortoise occurrence, and b) statistics on the performance of the model run. GARP also automatically selects the ten best performing distributions for each 50 runs. We created cell values (habitat scores) that ranged between 0 and 1 by summing all 50 model output rasters, then dividing by 50. We used the same procedure to create one map from the ten best model output rasters. Higher values mean more models predicted habitat in that location.

Table 2. Percentage of raster cells in each of three habitat score categories for various GARP model outcomes with different model parameterizations. Maps created from the best ten model outputs resulted in the highest percentage of raster cells with high scores, and maps with 250 m resolution had fewer cells in the middle category.

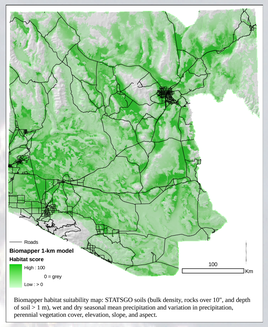
	Model parameterization							
Habitat Score	1 km g 50	1 km g bs	1 km ng 50	1 km ng bs	250 m g 50	250 m g bs	250 m ng 50	250 m ng bs
0-0.33	52.4%	45.9%	57.7%	49.0%	53.0%	53.4%	54.3%	51.1%
0.34-0.66	13.2%	11.8%	11.0%	12.8%	9.0%	7.2%	9.2%	7.5%
0.67-1	34.4%	42.2%	31.3%	38.2%	38.0%	39.4%	36.4%	41.3%

1 km = 1 km pixel input data, 250 m = 250 meter pixel input data
g = bedrock geology map included, ng = bedrock geology not included
50 = habitat score based on all 50 runs, bs = habitat score based on best 10 of 50 runs

Biomapper

Biomapper is a package of GIS and statistical tools based on Ecological Niche Factor Analysis (ENFA) (Hirzel, et al., 2002) and designed to build habitat suitability (HS) models and maps, without using absence data. The modeler chooses the HS algorithm (medians, geometric mean, distance harmonic mean, or minimum distance), and all other HS calculations are done automatically. We used the geometric mean algorithm: good generalization power, no assumption made on species distribution.

Biomapper functions include: using descriptive statistics to explore environmental factor maps, computing the ENFA, exploring model output, and computing a habitat suitability grid map, with values between 1 and 100 (low to high suitability). The HS map is created by computing factors (similar to principal components analysis [PCA]) that explain the ecological distribution of the species. As in PCA, the factors are uncorrelated; however, ENFA factors have biological significance. The marginality factor describes how far the species optimum is from the mean habitat in the study area. The specialization factors describe how specialized the species is compared to the available range of habitat in the study area.



Comparison of GARP and Biomapper

Table 3. Percentage of output cells in each category for Biomapper and GARP (1 km cells). Based on our categorical comparison, Biomapper has notably fewer raster cells with scores in the "high" category.

Habitat Score			Modeling Approach		
GARP	Biomapper		GARP ng 50	GARP ng bs	Biomapper
0-0.33	0-33	Low	57.7%	49.0%	70.7%
0.34-0.66	34-66	Med	11.0%	12.8%	19.8%
0.67-1	67-100	High	31.3%	38.2%	9.5%

Table 4. Comparison of number of cells in each category for Biomapper and GARP for two subset areas (at center and edge of overall study area, 25 km2) with 1-km cells. In the center subset area, Biomapper and GARP predicted approximately 66% of all raster cells similarly. No other patterns are apparent. For the edge subset area, about 82% of all raster cells were predicted similarly, and GARP consistently predicted higher habitat scores for raster cells that are different.

Center		GARP (ng 50)				
	Category	Low	Med	High	Total	Total %
Biomapper	Low	128	28	24	180	28.8%
	Med	50	27	51	128	20.5%
	High	60	2	255	317	50.7%
	Total	238	57	330	625	100.0%
	Total %	38.1%	9.1%	52.8%	100.0%	

Edge						
		Low	Med	High	Total	Total %
Biomapper	Low	505	45	50	600	96.0%
	Med		6	19	25	4.0%
	High			0		0.0%
	Total	505	51	69	625	100.0%
	Total %	80.8%	8.2%	11.0%	100.0%	

We will use additional methods to compare the two model approaches, such as receiver operating characteristic curves, percent of known occurrences correctly classified, and visual inspection by tortoise experts.

Step 4. Model verification

We will verify the model output by 1) withholding part of the occurrence dataset or using jackknife methods to test model performance, 2) using iterative runs of single model strategies (for stochastic models), 3) comparing different models' performance and map output, and 4) field verification.

Step 5. Suitable habitat map(s)

The final map of predicted suitable habitat for the Mojave population of desert tortoise will incorporate the best distribution model(s) developed in the previously described steps with maps of constraining anthropogenic factors (e.g., urban areas, roads, and rural infrastructure). We anticipate that planners, land managers, and agencies will use these maps to inform decisions for conservation activities, including translocations of desert tortoise. The suitable habitat maps will provide a meso-scale view of suitable habitat that can be combined with ground-level expert evaluation of site-specific micro-habitat.

References

Schachetti-Perelta, R. 2002. Desktop GARP. University of Kansas Natural History Museum, Lawrence. URL http://www.lifemapper.org/desktopgarp/ (accessed July 2005).

Hirzel A.H., Hausser J., Chessel D. & Perrin N. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? Ecology, 83, 2027-2036.

Biomapper 3.1. Lausanne, Lab. for Conservation Biology. URL: http://www.unil.ch/biomapper (accessed October 2005).

Contact information

For more information contact Kathryn Thomas at Kathryn_A_Thomas@usgs.gov or Leila Gass at lgass@usgs.gov.

Acknowledgements

This project would not have been possible without the tortoise occurrence data provided by U.S. Fish and Wildlife Service, Mojave Desert Ecosystem Program, National Park Service, Utah Division of Wildlife Resources, Bureau of Land Management, and EG & G Energy Measurements. Joan Blainey produced climate data and Cynthia Wallace produced perennial vegetation cover data used in our models. Other environmental data were created by the Mojave Desert Ecosystem Program and USGS. This project is funded by the U.S. Geological Survey Dept of the Interior Science Initiatives Program.