

BIEN Range Modeling Meeting

7 June 2012

1:00 PM PST

Invited

Martha Narro, Nirav Merchant, Edwin Skidmore, Jim Regetz, Mark Schildhauer, John Donoghue

Discussion Topics

1. Current range modeling status
2. Proposed range mapping deliverables
 - See page 2-3
3. Range modeling decision tree
 - See page 4
4. Script changes and need for retesting new scripts in the cluster
 - Update R packages for Dismo and Raster
 - Add R packages needed for latest clipping method (????)
5. Raster to vector conversion issues
 - Need a way to generate a shapefile from a raster map
 - Tried gdal_polygonize.py, but it hangs with complex rasters (perimeter/area ratio)
6. Scheduling a model run in the cluster
7. Need to archive maxent results that are stored in temp path
 - We currently use a temp path for each model so all maxent model outputs are stored in this temp location
 - Need to change temp path to species name or keep only the maxentResults.csv file from each run, rename it to the species name and copy it to an output location
8. Scheduling subsequent analysis runs in the cluster
 - Desire to run subsequent analyses on all species
 - Range specific climatic attributes (mean temp, precip, variance, etc)
 - Maxent variable importance (this is why we need 5)
9. Range mapping API
 - iPlant wants to build an API for range model deliverables. Discuss what API should do

Proposed Species Range Modeling Products

As of 28 October 2011

This page outlines the species-specific products that we propose to generate and retain from the range modeling effort. This list does not include additional aggregate products, such as diversity maps which will also be generated and retained. The list below only pertains to products specifically generated during the massive species-by-species range modeling effort. Comments from BIEN members are welcome and encouraged so we can ensure that we retain and generate the species range products needed by all members for their science questions.

Goal: Achieve balance between developing products that are needed by group members while minimizing storage space and access/downloading availability.

We propose to generate and retain the following products as part of our range modeling effort:

- **Convex Hull range** - polygon (**SHAPEFILE**)
- **Bounding box range** - polygon (**SHAPEFILE**) - can we just do this as a CSV file (simple file with 4 values)
- **Latitude extent range** - polygon (**SHAPEFILE**)
Is this in degrees or band across the continent in area (I prefer the first) - in which case this would just be two values (vector)
- **Area of occupancy** - polygon (**SHAPEFILE**)
Not sure this is useful as a shapefile? Raster, list of grid cells, or ???. Probably redundant with occurrence CSV file
- **Maxent raw probability surface** - environmental suitability grid - raster (**GEOTIFF**)
Raster layer generated from the Maxent model runs of each species, using a combination of 19 bioclim layers and 19 spatial filters and various thresholds.
- **Maxent binary range map** - from 1 percent training presence threshold - polygon (**GEOTIFF**)
Raster layer generated from the best modeling method as defined in our range area study.
- **Maxent binary range map** - from 1 percent training presence threshold - polygon (**SHAPEFILE**)
Polygon map generated from the Maxent raster defined above, perhaps smoothed for better cartographic representation.
- **Species occurrences** - points (**CSV FILE**)
- **Range size estimates** - text (**CSV FILE**), contains the following measures of range size:
 - Number of occurrence points
 - Convex hull area
 - Clipped convex hull area
 - Bounding box area
 - Clipped bounding box area

- Latitudinal extent (linear dimension without area)
- Latitude extent area
- Clipped Latitude extent area
- Area of occupancy (area of occupied cells)
- Areas from the three Maxent model types (bioclim, spatial filters, bioclim + spatial filters) and various thresholds defined below:
 - Maxent minimum training presence threshold area
 - Maxent 1 percent training presence threshold area
 - Maxent 5 percent training presence threshold area
 - Maxent max Kappa threshold area
 - Maxent balanced sensitivity and specificity area
 - Maxent maximum sensitivity and specificity area
- **Maxent all thresholds** - text (**CSV FILE**), contains a list of all thresholds generated by Maxent for each species
- **MaxentResults.csv** - text file (**CSV FILE**)
File named maxentResults.csv that is generated in the temp directory by Dismo for each Maxent model run of each species

Range Modeling Decision Tree

The following decision tree will be used to determine how each species is modeled so that we are able to calculate a range area for each species – including those for which we are unable to model using Maxent because they have too few occurrence records in distinct geographic locations (lat/long)

- A. ≥ 5 unique¹ points..... Model range area using Maxent
- A. < 5 unique points, or failed Maxent..... B

- B. 3-4 unique points or failed Maxent Convex hull
- B. < 3 unique points..... C

- C. 2 unique points..... Two point algorithm
- C. 1 unique point..... Single point algorithm

Two point algorithm:

Create one box around both points or a separate box around each point?
Calculate a box around both points and compare area of box with Gentry area.

- Box area $>$ Gentry area (75,000 km²)..... Fit two boxes²
- Box area \leq Gentry area (75,000 km²)..... Fit single box³

One point algorithm:

Calculate one box around each point 2

* Fit one Gentry box

Notes:

- ¹ Unique defined as having distinctly separate geographic locations (lat/long)
- ² Two point boxes will be constructed using the bounding latitudes and longitudes of both points.
- ³ Single point boxes will be constructed by setting the point as the center of the box and extending 136.9 km in each direction to create a box of area approx. 75,000 km²

Gentry cutoff for small range species:

“We defined small-ranged species as those that have never been collected outside the Peruvian department of Madre de Dios, and large-ranged species as those that have. The rationale is that the size of the department of Madre de Dios, 78,415 km², fits the most widely used criterion of tropical plant species endemism, 50–75 000 km² (Gentry 1986, 1992).” - From Pitman 1999

Pitman, N. C. a., Terborgh, J., Silman, M. R., & Nuñez V., P. (1999). Tree Species Distributions in an Upper Amazonian Forest. *Ecology*, 80(8), 2651-2661. doi:10.1890/0012-9658(1999)080[2651:TSDIAU]2.0.CO;2Gentry. 1986. Endemism in tropical vs. temperate plant communities. Pages 153–181 in M. E. Soule', editor. *Conservation biology: the science of scarcity and diversity*. Sinauer Associates, Sunderland, Massachusetts, USA.

Gentry. 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden* 75(1): 1–34.