Species Distribution Modeling

Matt Woodstock

1/31/2020

Objectives

The main aim of this study was to develop a species distribution model for the saguaro cactus using climatological data as a predictor.

Methods

Species Information

The saguaro cactus Carnegia gigantea is a long-lived (>150 years) cactus species taht is native to the southwestern United States and northwestern Mexico. (C. gigantea) can survive in an arid climate because of its ability to store large amounts of water. Since the desert biome is largely devoid of livable structure and resources are scarce, C. gigantea is an important species for others in the ecosystem. The ecological importance of C. gigantea has prompted conservation efforts to protect the species in its native range. As deserts become more susceptible to climate effects, C. gigantea populations are threatened. The occurace data used in this study originate from the Global Biodiversity Information Facility (GBIF) database (www.gbif.org), which handles occurance data of various species worldwide.

Statistical Analyses

Species distribtuion modeling is a method to estimate where a species resides within a defined range. Typically these models require a suite of environmental variables that combine with either presence/absense or abundance data to determine what environmental variables the species occupies the most. In this exercies, we use climatological data from WorldClim (www.worldclim.org), which is available in the dismo R package. One species distribution modeling method, Bioclim, can be used with just occurance and environmental data. The use of absence data is valuable and produces better estimates, but these data are rarely available. As a proxy, a subset of locations within the model domain where C. gigantea were not observed were randomly chosen to serve as absence data in this exercise. With presence and absence data, the probability that a C. gigantea occupies each area within the model domain is estimated based on the environmental data conditions.

Results

Based on the Bioclim model, the greatest proportion of the C. gigantea distribution is in Mexico (Fig. 1). The placement of absence information, which was randomly chosen, has an impact on the overall outcome, but not enough to extend the distribution much further beyond the observed data.
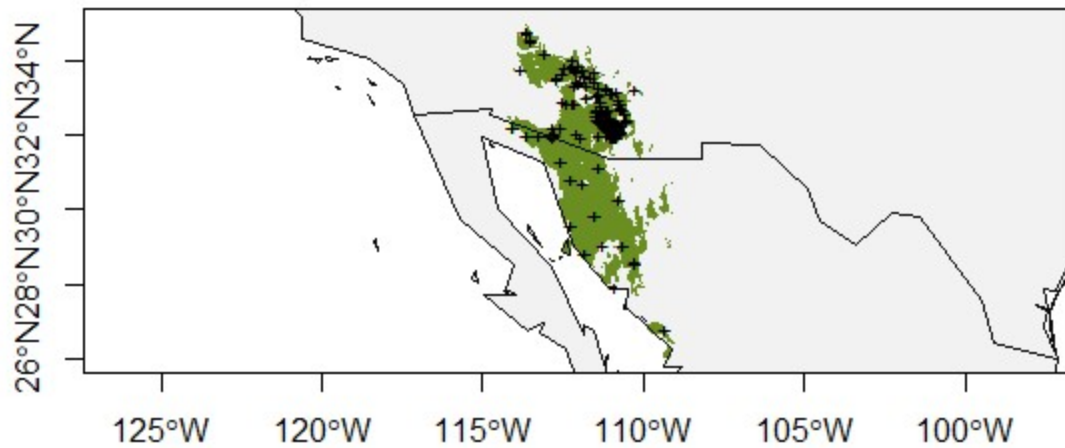
Figure 1. Presence of C. gigantea in the southwestern United States and northwestern Mexico. Black Plus signs are observed presence data and the green shading is modeled distribution.

The observed distribution of C. gigantea extends from 34.80 degrees north to 26.78 degrees north (Table 1). Much of this distribution is likely caused by the location of the observed values, as the predicted distribution does not show habitat beyond the observed coordinate maxima and minima.

Table 1. The geographic extent of observations of C. gigantea.

	Maximum	Minimum
Longitude	-114.01	-109.30
Latitude	34.80	26.78

Discussion

The distribution of C. gigantea extends further north-south than east-west. Despite most of the observations being in the United States, the majority of the probable species' distribution is in Mexico, suggesting climate conditions are more favorable there. A large density of observations occur in one location, which may be a product conservation efforts that are ongoing in the United States. However, this model only considers climate conditions. C. gigantea are harvested by native american tribes that primarily exist in the northwestern Mexican state of Sonora, who eat the fruits and seeds. Removal of seeds from the population over time may limit the distribution of C. gigantea in locations where it is harvested.

The modeled distribution is limited by the total survey area, and would exclude any potential range beyond the most extreme points. Additions of C. gigantea occurance data points beyond the GBIF dataset would likely expand the modeled range of this species. For instance, the greatest density of occurances is near the edge of the current distribution. Hypothetically, the distribution extends further eastward into the New Mexico desert, but this will not be shown with the current dataset.