Ecological Niche Modeling As a Tool for Conservation Planning: Suitable Habitat for Hypericum sinaicum in South Sinai, Egypt

O. Khafagi a, E.E. Hatab b, K.Omar c

a Botany Department, Faculty of Science (Girls branch), Al-Azhar University, Cairo, Egypt.

b Nature Conservation Sector, Egyptian Environmental Affairs Agency.

c Saint Katherine Protectorate, the Nature Conservation Sector, Egyptian Environmental Affairs Agency.

Corresponding author: kariemomar@gmail.com

Abstract:
Prediction and mapping of potential suitable habitat for threatened and endangered species is critical for monitoring and restoration of their declining native populations in their natural habitat, artificial introductions, or selecting conservation sites, and conservation and management of their native habitat. We used technique called maximum entropy distribution modeling or Maxent for predicting potential suitable habitat for Hypericum sinaicum, a threatened and endangered species in Saint Katherine Protectorate (SKP), South Sinai, Egypt, using small number of occurrence records. Our objectives were to: (1) predict suitable habitat distribution for threatened herb H. sinaicum using a small number of occurrence records to inform conservation planning in Saint Katherine Protectorate; and (2) identify the environmental factors associated with H. sinaicum habitat distribution. Results showed that the environmental variable with highest gain when used in isolation is bio9 (Mean Temperature of Driest Quarter). The approach presented here appears to be quite promising in predicting suitable habitat for threatened and endangered species with small sample records and can be an effective tool for biodiversity conservation planning, monitoring and management.

Keywords: Biodiversity conservation; Ecological Niche Modeling; geographical distribution; Hypericum sinaicum; Maxent, Saint Katherine Protectorate; suitable habitat;

1.0 Introduction:
Preservation of species requires not only detailed knowledge of their natural history and biology, but also information on the availability of suitable areas where species can survive; such knowledge can greatly aid in conservation planning. Recent developments in ecological niche modelling (ENM) have explored applications to diverse conservation issues, including suitable habitat and species range estimates (Chefaoui et al., 2005; Gaubert et al., 2006; Guisan et al., 2006), protected area prioritization and network design (Margules and Austin, 1994; Rondinini et al., 2005; Sánchez-Cordero et al., 2005a), and effects of habitat disturbance on species distributions (Banks et al., 2005; Sánchez-Cordero et al., 2005b; Rhodes et al., 2006).

The ecological niche can be defined as the set of environmental conditions (abiotic factors) under which a species is able to maintain viable populations without immigration (Grinnell, 1917 and 1924). The challenge of identifying distributional areas for species requires two conditions to be met: favourable abiotic conditions and favourable biotic factors (e.g. presence of symbionts and mutualists, absence of serious parasites and predators); a third condition, geographical accessibility (landscape configuration, dispersal abilities of species), both at present and through history, is necessary for the actual presence of species (Soberón and Peterson, 2005). A growing literature deals with methodological challenges specific to best ENM-based predictions of suitable areas (Peterson and Kluza, 2003; Guisan and Thuiller, 2005; Elith et al., 2006) and identification of conservation priorities (Loiselle et al., 2003).

Biodiversity conservation is one of the major concerns in biogeography and ecology. Species richness is distributed non-uniformly across the biosphere (Sechrest et al., 2002) and nature conservation is often based on the concept of biodiversity hotspots (Myers et al., 2000; Brooks et al., 2002; Roberts et al., 2002). Many studies have discussed the factors determining the spatial
distribution of species. Their results also depend on the spatial scale of the study (Mackey and Lindenmayer, 2001; Quist et al., 2004; Trivedi et al., 2008; Blach-Overgaard et al., 2010; Soberón, 2010). Species distribution models (SDMs) that use environmental factors based on historical collections are increasingly being used to not only analyze species distributions, but also to predict the presence or absence of species or their habitats in unrecorded areas (Guisan and Hofer, 2003; Araújo et al., 2005; Wintle et al., 2005; Elith et al., 2006; Elith and Leathwick, 2009). Notably, SDMs have been used to predict potentially suitable areas for the preservation of endangered and rare species (Papes and Gaubert, 2007; Solano and Feria, 2007; Ko et al., 2009; Thorn et al., 2009; Gallagher et al., 2010; Rebelo and Jones, 2010), for the identification of potential sites for reintroduction or restoration (Klar et al., 2008; Kumar and Stohlgren, 2009) and for assessing potential future climate change on species distributions as well as on local species diversity (Pearson and Dawson, 2003; Hole et al., 2009). To enable the analysis of the impacts of climate change on species, it is essential to quantify the relative importance of climate relative to other descriptors of the environment (Morueta-Holme et al., 2010; Newbold, 2010).

Prediction and mapping of potential suitable habitat for threatened and endangered species is critical for monitoring and restoration of their declining native populations in their natural habitat, artificial introductions, or selecting conservation sites, and conservation and management of their native habitat (Gaston, 1996). But distribution data on threatened and endangered species are often sparse (Ferrier et al., 2002; Engler et al., 2004) and clustered making commonly used habitat modeling approaches difficult. Species distribution modeling tools are becoming increasingly popular in ecology and are being widely used in many ecological applications (Elith et al., 2006; Peterson, 2006). These models establish relationships between occurrences of species and biophysical and environmental conditions in the study area. Most species distribution modeling methods are sensitive to sample size (Wisz et al., 2008) and may not accurately predict habitat distribution patterns for threatened and endangered species.

The Saint Katherine region is situated in the southern part of Sinai and is a part of the upper Sinai massif. It is located between 33˚ 55' to 34˚ 30' East and 28˚ 30' to 28˚ 35' North. The Saint Katherine Protectorate (SKP) is one of Egypt’s largest protected areas and includes the country’s highest mountains. This arid, mountainous ecosystem supports a surprising biodiversity and a high proportion of plant endemics and rare plants. The flora of the mountains differs from the other areas, due to its unique geology, morphology and climatic aspects. The soil is formed mainly from mountains weathering, thus it is mainly granitic in origin. The soil layer is generally shallow were the bed rock is close to the surface. Annual rainfall is less than 50 mm. However, rainfall is not of annual character, rather 2 to 3 consecutive years without rainfall is common. Rain takes the form of sporadic flash floods or limited local showers, thus highly spatial heterogeneity in received moisture is also common (Hatab, 2009).

Our objectives were to: (1) predict suitable habitat distribution for threatened herb *H. sinaicum* using a small number of occurrence records to inform conservation planning in Saint Katherine Protectorate; and (2) identify the environmental factors associated with *H. sinaicum* habitat distribution. We used species occurrence records, GIS (geographical information system) environmental layers (bioclimatic and topographic), and the maximum entropy distribution modeling approach (Phillips et al., 2006) to predict potential suitable habitat for *H. sinaicum*.

### 2.0 Material and Methods:

#### 2.1 Target species and occurrence data

We recorded 116 sites of *H. sinaicum* (Hypericaceae) species in Saint Katherine Protectorate during the period between March to August 2011; these records represent the total known distribution of the species. *H. sinaicum* is one of the near endemic species in SKP only found in Sinai and North West Saudi Arabia (Boulos, 2002). *H. sinaicum* recorded as rare species (IUCN, 1994), this species has a highly medicinal importance value, extraction from aerial parts give substances like Hypericin, protohypericin, pseudohypericin, protopseudohypericin, and hyperforin which showed effect to inhibit the growth of retroviruses including HIV, the AIDS virus) in animals beside the treatment of depression (Rezanka and Sigler, 2007).

Perennial herbs, 10-25 cm, woody at the base: stems branched; ascending; leaves 0.3-1.2 x 0.3- 0.8 cm. white tomentose, sessile, ovate to elliptic, with scattered black gland-dots; flowers 1-1.5 cm diam., in
vlew-flowered terminal cymes; sepals 3-4 x 1-1.25 mm, acute, with black glandular dots, the margins with stalked black glands; petals 6-8 x 3.5 mm, yellow, persistent; stamens in 3 bundles; styles 3, free; capsule 5x3 mm, ovoid, 3-valved (Boulos, 1999). No specific microhabitat preference for *H. sinaicum*, this species located into most of the micro-habitats, included Cliffs, Slope, Terraces, Gorge and Farsh, but showed much better growth in cliffs and gorges microhabitats (Omar, 2010). Observation also found that there is no grazing on *H. sinaicum*. Most of the *H. sinaicum* populations are small and the plants occurred sporadically in space, as little groups or as individuals. Special micro-habitat (Mountainous sheltered moist crevices), over-collection for scientific research and overgazing from feral donkeys put this species in a critical conservation condition. In order to develop an efficient and effective conservation strategy using complementary in situ and ex situ techniques, we must have a clear understanding of *H. sinaicum* geographical distribution.

### 2.2 Environmental Variables

We considered twenty three environmental variables as potential predictors of the *H. sinaicum* habitat distribution (Table 1). These variables were chosen based on their biological relevance to plant species distributions and other habitat modeling studies (for example, Kumar et al., 2006; Guisan et al., 2007a,b; Pearson et al., 2007; Murienne et al., 2009). Nineteen bioclimatic variables (Nix, 1986), biologically more meaningful to define eco-physiological tolerances of a species (Graham and Hijmans 2006; Murienne et al., 2009), were obtained from WorldClim dataset (Hijmans et al., 2005; http://www.worldclim.org/bioclim.htm). Altitude (Digital Elevation Model; DEM) data were also obtained from the WorldClim website; 1 km spatial resolution. The DEM data were used to generate slope and aspect (both in degrees) using (ESRI) Environmental Systems Research Institute’s ARC GIS version 9.2 and ‘Surface Analysis’ function. All environmental variables were resampled to 1 km spatial resolution. Maxent’s predictions are ‘cumulative values’, representing, as a percentage, the probability value for the current analysis pixel and all other pixels with equal or lower probability values. The algorithm is implemented in a stand-alone, freely available application. In this study we considered each environmental variable (linear features) and its square (quadratic features) and this because Maxent utilize pseudo-absence.

### 2.3 Modeling Procedure

We used the modeling technique maximum entropy distribution or Maxent which has been found to perform best among many different modeling methods (Elith et al., 2006; Ortega-Huerta and Peterson, 2008), and may remain effective despite small sample sizes (Hernandez et al., 2006; Pearson et al., 2007; Papes and Gaubert, 2007; Wisz et al., 2008; Benito et al., 2009). Maxent is a maximum entropy based machine learning program that estimates the probability distribution for a species’ occurrence based on environmental constraints (Phillips et al., 2006). It requires only species presence data (not absence) and environmental variable (continuous or categorical) layers for the study area. We used the freely available Maxent software, version 3.1 (http://www.cs.princeton.edu/~schapire/maxent/), which generates an estimate of probability of presence of the species that varies from 0 to 1, where 0 being the lowest and 1 the highest probability. The 116 occurrence records and 10 environmental predictors were used in Maxent to model potential habitat distribution for *H. sinaicum*. Testing or validation is required to assess the predictive performance of the model. Ideally an independent data set should be used for testing the model performance, however, in many cases this will not be available, a situation particular prevalent for threatened and endangered species. Therefore, the most commonly used approach is to partition the data randomly into ‘training’ and ‘test’ sets, thus creating quasi-independent data for model testing (Fielding and Bell, 1997).

However, this approach may not work with a small number of samples because the ‘training’ and ‘test’ datasets will be very small (Pearson et al., 2007). Therefore, we explicitly followed Pearson et al. (2007) and used a jackknife procedure, in which model performance is assessed based on its ability to predict the single locality that is excluded from the ‘training’ dataset. Ninety one different predictions were thus made with one of the occurrence records excluded in each prediction and the final potential habitat map was generated using all records (Fig. 1). We used the *P* value program provided by Pearson et al. (2007) to test the significance of the model. The jackknife validation test required the use of a threshold to define ‘suitable’ and ‘unsuitable’ areas. We used two different thresholds, the ‘lowest presence threshold’ (LPT, equal to the lowest probability at the species...
presence locations), and a fixed threshold of 0.10; for more details see Pearson et al. (2007).

3.0 Results and Discussion:

The Maxent model predicted potential suitable habitat for *H. sinaicum* with high success rates (that is, low omission rates), 98% at LPT. Most suitable habitat for *H. sinaicum* was predicted in the northern parts of the SKP in South Sinai (Fig. 1), and its distribution is quite fragmented. The Maxent model’s internal jackknife test of variable importance showed that 'Mean Temperature of Driest Quarter (degree C)', and 'Precipitation of Driest Quarter (degree C)' were the two most important predictors of *H. sinaicum*'s habitat distribution (Fig. 2; Table 1). These variables presented the higher gain (that is, contained most information) compared to other variables (Fig. 2; Table 1). Using four arbitrarily defined probability classes, the high suitability class had an area of 60.8 km² (1.4%); medium-60.2 km² (1.3%); low- 130.6 km² (3%); and very low-4098 km² (Fig. 1).

Figure 1: Predicted potential suitable habitat for *H. sinaicum* species on Saint Katherine Protectorate, South Sinai, Egypt. 1- actual size for habitat with high probability, 2- actual size for habitat with low probability and 3- actual size for habitat with medium probability.

<table>
<thead>
<tr>
<th>No.</th>
<th>Environmental variable</th>
<th>Percent contribution</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean Temperature of Driest Quarter (Bio9, degree C)</td>
<td>28.1</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>2</td>
<td>Precipitation of Driest Quarter (Bio17, degree C)</td>
<td>19.7</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>3</td>
<td>Altitude (m)</td>
<td>7.9</td>
<td>Generated in GIS</td>
</tr>
<tr>
<td>4</td>
<td>Habitat (degree)</td>
<td>7.6</td>
<td>Generated in GIS</td>
</tr>
<tr>
<td>5</td>
<td>Mean Temperature of Coldest Quarter (Bio11, degree C)</td>
<td>5.7</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>6</td>
<td>Precipitation of Driest Period (Bio14, degree C)</td>
<td>5.2</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>7</td>
<td>Precipitation of Warmest Quarter (Bio 18, degree C)</td>
<td>5.1</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>8</td>
<td>Slope (degree)</td>
<td>4.6</td>
<td>Generated in GIS</td>
</tr>
<tr>
<td>9</td>
<td>Min Temperature of Coldest Period (Bio 6, degree C)</td>
<td>4.2</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>10</td>
<td>Max Temperature of Warmest Period (Bio 5, degree C)</td>
<td>3</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>11</td>
<td>Precipitation of Wettest Period (Bio13, degree C)</td>
<td>2.2</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>12</td>
<td>Precipitation of Coldest Quarter (Bio19, degree C)</td>
<td>1.5</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
</tbody>
</table>

Table 1: Selected environmental variables and their percent contribution in Maxent model for *H. sinaicum* species in Saint Katherine Protectorate.
<table>
<thead>
<tr>
<th></th>
<th>Predictor Variable</th>
<th>Importance</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Temperature Annual Range (Bio7, degree C)</td>
<td>1.5</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>14</td>
<td>Aspect (degree)</td>
<td>1.3</td>
<td>Generated in GIS</td>
</tr>
<tr>
<td>15</td>
<td>Mean Temperature of Wettest Quarter (Bio 8, degree C)</td>
<td>0.7</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>16</td>
<td>Precipitation of Wettest Quarter (Bio16, degree C)</td>
<td>0.6</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>17</td>
<td>Mean Diurnal Range (Bio2, degree C)</td>
<td>0.6</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>18</td>
<td>Isothermality (Bio3, degree C)</td>
<td>0.2</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>19</td>
<td>Precipitation Seasonality (Bio15, degree C)</td>
<td>0.2</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>20</td>
<td>Mean Temperature of Warmest Quarter (Bio10, degree C)</td>
<td>0</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>21</td>
<td>Annual Mean Temperature (Bio1, degree C)</td>
<td>0</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>22</td>
<td>Temperature Seasonality (C of V) (Bio 4, degree C)</td>
<td>0</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
<tr>
<td>23</td>
<td>Annual Precipitation (Bio 12, degree C)</td>
<td>0</td>
<td>WorldClim; Hijmans et al. 2005</td>
</tr>
</tbody>
</table>

![Figure 2. Results of jackknife evaluations of relative importance of predictor variables for H. sinaicum Maxent model.](image)

**Note:** alt is elevation; Bio 1 - Annual Mean Temperature; Bio 2 - Mean Monthly Temperature Range; Bio 3 - Isothermality (2/7) (* 100); Bio 4 - Temperature Seasonality (STD * 100); Bio 5 - Max Temperature of Warmest Month; Bio 6 - Min Temperature of Coldest Month; Bio 7 - Temperature Annual Range; Bio 8 - Mean Temperature of Wettest Quarter; Bio 9 - Mean Temperature of Driest Quarter; Bio 10 - Mean Temperature of Warmest Quarter; Bio 11 - Mean Temperature of Coldest Quarter; Bio 12 - Annual Precipitation; Bio 13 - Precipitation of Wettest Month; Bio 14 - Precipitation of Driest Month; Bio 15 - Precipitation Seasonality (CV); Bio 16 - Precipitation of Wettest Quarter; Bio 17 - Precipitation of Driest Quarter; Bio 18 - Precipitation of Warmest Quarter; Bio 19 - Precipitation of Coldest Quarter.
The distribution of highly and moderately suitable areas appears to follow the distribution of highly elevated areas in SKP (Map7 in Omar, 2010). The parts of the study area predicted in the ‘very low’ suitability class (probability < 0.10) can be interpreted as unsuitable for H. sinaicum (Figure 1). We also calculated total extent of occurrence (EOO, as defined by IUCN, 2001) of H. sinaicum based on the commonly used threshold of 0.4 (That is, the threshold above which the species is more likely to be present; Jimenez-Valverde and Lobo, 2007); it was estimated to be 926 km².

Figure (3) summaries curves represent a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other
variables. They may be easier to interpret if there are strong correlations between variables. Results derived from curves showed that the probability of the presence of *H. sinaicum* increase with the increase in altitude, aspect, Precipitation of Driest Month, Mean Monthly Temperature Range, Temperature Annual Range and Temperature Seasonality (STD * 100); but decrease with the increase of slope, Habitat, Max Temperature of Warmest Month; Min Temperature of Coldest Month; Mean Temperature of Wettest Quarter; Mean Temperature of Driest Quarter; Mean Temperature of Warmest Quarter; Mean Temperature of Coldest Quarter and Annual Mean Temperature.

4.0 Conclusions:
In this study we showed that the habitat distribution patterns for threatened and endangered plant species such as *H. sinaicum* can be modeled using a small number of occurrence records and environmental variables using Maxent. This study provides the first predicted potential habitat distribution map for a plant species (*H. sinaicum*) in SKP. Since Maxent is mapping the fundamental niche (different from occupied niche) of the species using bioclimatic variables the suitable habitat for *H. sinaicum* may be over predicted in some areas (Pearson 2007; Murienne et al., 2009). The potential habitat distribution map for *H. sinaicum* can help in planning land use management around its existing populations, discover new populations, identify top-priority survey sites, or set priorities to restore its natural habitat for more effective conservation. More research is needed to determine whether the existing protected areas adequately cover suitable habitat for *H. sinaicum*. The methodology presented here could be used for quantifying habitat distribution patterns for other threatened and endangered plant and animal species in other areas and may aid field surveys and allocation of conservation and restoration efforts.

5.0 Acknowledgement:
I would like to express deepest grateful to Mr. Mohammed Kotb, General Manager of Saint Katherine protectorate for his support during research steps.

References:


