Journal of Arid Environments 135 (2016) 140–146

Contents lists available at ScienceDirect

Journal of Arid Environments

journal homepage: www.elsevier.com/locate/jaridenv

Using species distribution models to assess the importance of Egypt's protected areas for the conservation of medicinal plants

Emad Kaky*, Francis Gilbert

School of Life Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

A R T I C L E I N F O

Article history:
Received 6 July 2016
Received in revised form
1 September 2016
Accepted 13 September 2016

Keywords:
MaxEnt
Egypt
Ecosystem services
Conservation

A B S T R A C T

Human activities affect the distribution and abundance of plants, with impacts on ecosystem services and human well-being; it is thus vital that a network of Protected Areas is capable of conserving plants that are useful. Using the species distribution (SDM) model algorithm MaxEnt, we tested whether Egypt's network of Protected Areas performs well in conserving the region's important medicinal plant species. We constructed individual SDMs for each species, and then combined the models into a single 'species-richness' layer, which we then compared to the distribution of the existing Protected Areas. Temperature was the most important of eleven predictor variables used to build the SDMs. Assuming the SDM's prediction of suitable habitat was accurate and corresponded to the occurrence of the medicinal plant species, then on average species richness was significantly higher within than outside the Protected Areas. Based on our findings, Egypt's Protected Areas are effective at conserving its medicinal plants.

1. Introduction

Human activities are having a strong impact on plant abundance and distribution, with consequent effects on ecosystem services and human well-being (Klein et al., 2008). This growing effect of human activities on biodiversity (Chapin et al., 2000) creates an urgent need to understand the elements that determine the distribution and abundance of plants in order to enhance their conservation (Dubuis et al., 2011). The identification of species-rich regions and those where geographically limited species co-occur can optimise the creation of Protected Areas (Bojórquez-Tapia et al., 1995).

Medicinal plants are one of the most important elements of biodiversity around the world (Klein et al., 2008; Okigbo et al., 2008) because of their role in ecosystem services such as healthcare, cultural value and heritage, local economics and human well-being, especially in poor areas (Klein et al., 2008; Okigbo et al., 2008). Conserving and protecting these kinds of species is vital, including improving knowledge about the important ecological requirements of medicinal plants, and raising awareness among all stakeholders to protect this heritage. Consequently, conservation planning and effective management is important in protecting the most threatened species in order to avoid declines in the diversity of medicinal plants.

Species distribution models (SDMs) can be used to predict the geographic distribution of individual species using locality data and ecological variables as predictors (Franklin, 2009: 41–45). While occurrence records can be harvested from museums/herbaria, published reports, and original fieldwork, accurately identifying whether a species is truly absent is exceedingly difficult. To address this challenge, several SDM algorithms have been designed to employ only positive presence data (Phillips et al., 2006). One such SDM algorithm, MaxEnt, has been shown to be one of the most effective tools for accurately predicting species distributions (Elith et al., 2006). SDMs using MaxEnt offer a valuable tool for creating general patterns of species richness without needing to analyse the specific quality or precision of the predictions for every individual species (Pineda and Lobo, 2009). Several studies have added together the models for individual species to create maps of species richness, the approach we adopt here. For example, Ortega-Huerta and Peterson (2004) added the individual maps of 285 bird and 114 mammal species of part of Mexico to create a map of species richness; Newbold et al. (2009) and Pineda and Lobo (2009) used the same approach for Egyptian mammals and butterflies, and Mexican amphibians respectively, as did de Pous et al. (2011) on Moroccan reptiles. It is exciting that the same approach can be used to project into the future under climate change (Distler et al., 2015), as we have also done (Kaky & Gilbert, in prep.). Ideally the maps of

---

* Corresponding author.
E-mail address: plxedk@nottingham.ac.uk (E. Kaky).

http://dx.doi.org/10.1016/j.jaridenv.2016.09.001

predicted species richness should be validated using independent data (Pineda and Lobo, 2009). Such species-richness maps make it possible to distinguish hotspots of species richness (Newbold et al., 2010), and hence to select feasible regions for conservation relatively objectively (Pressey et al., 1993). This is a powerful tool to help build conservation efforts or anticipate the future of biodiversity under worldwide climate change (Algar et al., 2009).

The climatic predictors used in our SDMs should be very suitable for plants. The physiological toleration hypothesis suggests that plant species richness is most elevated in warm and/or wet environments because a more extensive range of functions can persevere under such circumstances (Spasojevic et al., 2014). For instance, Hawkins et al. (2003) found that a measure of the balance between energy and water nearly always described spatial differences in species richness better than other environmental variables. In warm regions of the tropics and subtropics, the most robust predictors are typically water variables, while water/energy variables (for plants) or energy predictors (for animals) predominate in high latitudes (Hawkins et al., 2003).

Protected Areas currently cover about 12% of the terrestrial surface of the earth (Seiferling et al., 2012), while those that have been declared in Egypt cover 15% of the total land area (El-Gabbas et al., 2016). The 30 Egyptian Protected Areas were all established since 1983, based on the recommendations of experts familiar with Egyptian biodiversity (Newbold et al., 2009). An obvious issue is the extent to which these Protected Areas are capable of conserving Egypt's fauna and flora: a basic requirement is that they contain a high proportion of the biodiversity of the country. Thus ideally there should be higher species richness within the Protected Areas than outside them. Several studies have measured this: for example, Sciberras et al. (2013) showed that the density and biomass of fish and invertebrates inside partially protected areas was higher than in unprotected areas; Newbold et al. (2009) and Lee et al. (2007) found that species richness inside Protected Areas was higher than outside, but others found the reverse (Pawar et al., 2007; Traba et al., 2007). Human activities are one of the main reasons for declines both inside and especially outside Protected Areas: thus forest cover decreased between 1980 and 2001 in areas surrounding most tropical Protected Areas (DeFries et al., 2005), and one might anticipate similar declines in the fauna. The active management of Protected Areas needs many more such comparisons to guide management decisions (Linkie et al., 2006).

Our objective is therefore to assess the role of the network of Egyptian Protected Areas in conserving medicinal plants by comparing their diversity within and just outside each Protected Area, averaging this difference across all the Protected Areas. We did this by predicting the distribution of each species using SDMs, and summing together all the SDMs to create two kinds of species-richness maps (by either using or not using thresholds to binarize the predicted habitat suitabilities). We then use these maps to assess the predicted species richness inside and outside Egypt's Protected Areas.

2. Methods

We used data for 121 medicinal plant species of the Egyptian flora. The occurrence data for these species were collated by the BioMAP project (http://www.biomapegypt.org/), a project run from Cairo in 2004–2008 and funded by Italian Debt Swap. The data are presence-only records collected from different sources (i.e. literature, herbarium, and field work). To avoid inaccurate predictions, we deleted species with fewer than ten records to avoid overfitting (Baldwin, 2009), species with more than ten but spatially very restricted records, and the one species whose SDM had a mean AUC less than 0.7 (Franklin, 2009: 222–223). We ended up with 114 species of Egyptian medicinal plants, with 14396 point records.

The environmental variables used in this study were 23 predictors, 19 of them (Bio layers) downloaded from the WorldClim v1.4 dataset at resolution of 2.5 arc-minutes (http://www.worldclim.org/bioclim) (Hijmans et al., 2005) (Table 1). Normalized Difference Vegetation Index (NDVI) data for seven years (2004–2010) were downloaded from the Spot Vegetation website (http://free.vgt.vito.be/) and used to create two layers: maximum NDVI (Max_NDVI), and the difference between the Minimum and Maximum NDVI values (NDVI_differences). A further environment layer was a habitat layer, derived from the Biomap project, which divided Egypt's terrain into eleven classes (“sea, littoral coastal land, cultivated land, sand dune, wadi, metamorphic rock, igneous rock, gravels, serir sand sheets, sabkhas and sedimentary rocks”) (for more detail, see Newbold et al., 2009). Altitude data were downloaded from http://www.cgiar-csi.org/data/elevation and the resolution rescaled from 90 m to be 2.5 arc-minutes (see (El-Gabbas et al., 2016). Eleven of the 23 environmental variables (see Table 1) remained for use after 12 were removed based on collinearity analysis using the Variance Inflation Factor, implemented in R v2.15 (the 'car' package: R Development Core Team, 2012).

We used Maximum Entropy (MaxEnt) version 3.3.3 k (Phillips et al., 2006) (downloaded from: http://www.cs.princeton.edu/~schapire/maxent/) to run the models, choosing a set of options (i.e. feature classes QPT, 10000 background points, 1000 iterations, cross-validation with 10 replications, 10% training presence threshold, and logistic output format) to create both ‘probability’ (i.e. raw values of habitat suitability) and ‘binary’ (predicted 'suitable'/'unsuitable' via thresholding) maps. MaxEnt performance is good with presence-only data and small numbers of records (Elith et al., 2006; Franklin, 2009: 62–63), and its performance is good in comparison with other algorithms (Elith et al., 2006). The options were chosen after exhaustive runs with different option combinations (of feature classes, number of background points, number of iterations and regularization values) to obtain the best models. Two statistics were used to evaluate the accuracy of each model, the AUC, and the true skill statistic (TSS) (Allouche et al., 2006). TSS values lie between -1 and $+1$: close to $+1$ indicates perfect performance, while close to zero or less than zero point to model performance no better than random (details, see Allouche et al., 2006). (For details of each SDM, see Supplementary Table S1.).

The relative importance of the environmental predictors can be determined in three ways by Maxent (percent contribution, permutation importance, jacknife (Phillips et al., 2006):). Care is needed when there are high correlations between variables, but pre-screening variables for collinearity (as we have done) minimises this problem. Here we used permutation importance to determine the importance of the environmental predictors, calculated by permuting the values of each predictor and calculating the resulting reduction in the training AUC: a large reduction shows that the model is influenced by that predictor. The values are standardized to a percentage (Phillips et al., 2006).

We created two kinds of maps of the distribution of species richness. The first was the ‘probability’ map, made manually by obtaining the average of the replicate ascii files obtained from Maxent for each species, and then adding all the species layers together using the ‘raster calculator’ of ArcGIS10.2.2. This map was then rescaled to fit the same range as the second type, the ‘binary’ map, which is the product of adding together the binary maps for each species. The binary map converts each pixel value of the MaxEnt output (a continuous value between 0 and 1) into binary data (predicted suitable/unsuitable) by choosing a threshold rule (see Liu et al., 2005). We chose the “10% training presence” as our threshold rule (El-Gabbas et al., 2016), which produced a binary map for each of the 10 replicates for each species. Subsequently we