Ecological Informatics 29 (2015) 57–65

Contents lists available at ScienceDirect

Ecological Informatics

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

Should species distribution models use only native or exotic records of existence or both?

Farzin Shabani *, Lalit Kumar

Ecosystem Management, School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351, Australia

ARTICLE INFO

Article history:
Received 14 November 2014
Received in revised form 20 July 2015
Accepted 27 July 2015
Available online 1 August 2015

Keywords:
Species distribution records
Climate change
MaxEnt
CLIMEX
Phoenix dactylifera L.

ABSTRACT

This study investigated the importance of the use of appropriate species distribution records in projecting potential distributions under climate change using comparative bioclimatic models and alternative sets of data (native and exotic) to project a species in a new environment. We built bioclimatic models for date palm (*Phoenix dactylifera* L.), using the MaxEnt correlative model and the CLIMEX mechanistic niche model, and fitted the models using three training data sets: native data only, exotic data only and entire data. We compared the ability of the different data sets using the different modelling approaches to project suitable climate envelope for independent records of the species at a global scale. We found that the output of projected species distributions was closely related to the modelling approach as well as the specific categorized distribution of species data used (native data only, exotic data only and entire data).

1. Introduction

There has been a growing awareness of the value of plant and animal species distribution models (SDMs) over the last two decades. The latest techniques in this form of modelling facilitate the forecasting of anthropogenic impact on biodiversity patterns at a variety of spatial scales. The output of predictive modelling of species distribution can be used as a tool in addressing many ecological, biogeographical and evolutionary questions and has been used most recently in the fields of conservation biology and climate change. To be valid, SDMs must by necessity be robust. Concepts such as mobile species in unsuitable habitats, species not observed in suitable habitats, low-detectability species and over-dispersion may be explored through SDMs, provided robustness is properly assessed, to provide as realistic an estimate as possible of the impact of climate change (Guisan and Thuiller, 2005). In this regard, Peterson et al. (2007) documented that the effects of input data gaps and bias in the use of MaxEnt software necessitates careful interpretation of model results. Shabani et al. (2014c) have shown that refinement of the CLIMEX output by non-climatic parameters, such as suitability in soil physicochemical properties, soil taxonomy, slope and land use, are essential and achieve a more realistic result than those based purely on climate. Readers interested in problems associated with the analysis of present distribution of species may refer to Dormann (2007).

Habitat suitability or niche-based model techniques utilize data on distribution of a species, together with other environmental factors, in order to generate statistical functions to be used as the basis of predictions regarding habitat distribution of the species (Guisan and Zimmermann, 2000). In maximum entropy modelling (MaxEnt), species distribution records play a major role in selecting those environmental variables most appropriate for estimating the distribution of species (Grendar and Grendár, 2000; Phillips, 2005). Mechanistic bioclimatic models, such as CLIMEX, also require time and species occurrence records, linking ecophysiological responses to environmental covariates (Kearney and Porter, 2009; Kriticos and Randall, 2001; Webber et al., 2011). Since species distribution records play a major role in the accuracy of model predictions (Zaniewski et al., 2002), knowledge of the predictive abilities of a particular methodology and its valid applications become paramount in the initial stages of developing surveys mapping species distributions.

Species distribution records can be categorized into two groups, native or exotic. The mechanisms governing both native and exotic establishment are similar (Guisan et al., 2014; Seabloom et al., 2003a). To establish itself successfully, whether native or exotic, a species must survive and achieve a positive rate of increase, while living on unconsumed resources left by a resident species (Davis and Pelsor, 2001; Knops et al., 1999; Seabloom et al., 2003b). In terms of a definition, native species occur naturally in a locality, or have existed for many years, in that area (Sellmer, 2010), while exotic species, also termed introduced, non-indigenous or non-native species, are those living outside the species' native distribution, having arrived at the location through human activity, either deliberately or accidentally (Sax and Gaines, 2008). It should be noted that available georeferenced occurrences at global or continent scales can be found in data sets such as Global Biodiversity Information Facility (2015), Atlas of Living Australia (2014) and Discover Life (2014).

* Corresponding author.
E-mail address: fshaban2@une.edu.au (F. Shabani).

http://dx.doi.org/10.1016/j.ecoinf.2015.07.006

In a study by Wharton and Kriticos (2004), CLIMEX software and all distribution records of *Essigella californica* (an invasive pest) were utilized to predict the Australian and global distributions of *E. californica* under current climate conditions. In the study of Pearson et al. (2002), species couples and all distribution records of *Rhynchospora alba* (white-beaked sedge) and *Salix herbacea* (dwarf willow) were utilized to estimate current species distributions. Shabani et al. (2012) utilized the CLIMEX package and all known records of *Phoenix dactylifera* L. (date palm) to project conducive areas for the present and the years 2030, 2050, 2070 and 2100. Taylor et al. (2012) also made use of 1740 available records of *Lantana camara* L., an invasive shrub, to provide a risk map for countries identified as becoming climatically suitable for *L. camara* under future climate scenarios. Shabani et al. (2014a) employed distribution records of *Fusarium oxysporum* f. spp. to map future distributions of this dangerous fungus in terms of climate change, in agricultural regions of Europe, the Middle East and North Africa. Pattison and Mack (2008a) also used CLIMEX together with known records of *Triadica sebifera*, an invasive tree in the United States, to evaluate CLIMEX predictions in field trials. Furthermore, CLIMEX has been utilized to quantify the effects of potential alterations in climatic factors on localities for *Triticum aestivum* L. (wheat) and *Gossypium* (cotton) production, which are two crops important to Australia's economy (Shabani and Kotey, 2015).

The question as to which data set or combination is most valid, as well as assessing differences in projections, remains unanswered. Should species distribution models use only native or exotic records of existence or both? In this study, attempts were made to investigate the value of the use of (1) native data only, (2) exotic data only and (3) entire data to model species distributions. In regard to the question, this study, which primarily used current distributions of *P. dactylifera* L. for modelling, incorporates the three different sets of distribution records for this species utilizing both MaxEnt and CLIMEX software packages to assess the validity and projection differences of the use of this species distribution records in projecting potential distributions under climate change.

2. Methodology

2.1. Date palm (P. dactylifera) distribution

The Global Biodiversity Information Facility (GBIF) (Global Biodiversity Information Facility, 2015), Missouri Botanical Gardens' database (Missouri Botanical Garden, 2015) and additional literature on the species in CAB Abstracts databases (CAB Direct, 2015) formed the basis for the collection of data on *P. dactylifera* distributions. Supplementary information was drawn from additional sources (Eshraghi et al., 2005; Mahmoudi and Hosseininia, 2008; Shayesteh and Marouf, 2010; Tengberg, 2011). In total, 508 records of *P. dactylifera* were collected, as shown in Fig. 1 (395 records in areas where this species is considered a native species and 113 records where date palm is considered an exotic species; Jain et al., 2011).

2.2. MaxEnt

MaxEnt is a machine learning software technique used for creating a maximum entropy model of geographical distribution of a species. The level of suitability of each cell of the grid is expressed by the environmental variables' functions of the specific cell (Phillips et al., 2006). Absence data are not a requirement, and conversely the model incorporates background environmental variables of the complete study area, with a regularization function preventing predictions from over-fitting of data (Nazeri et al., 2012; Peterson et al., 2007; Phillips et al., 2006). MaxEnt allows the incorporation of continuous and categorical variables, estimating a resource's selection probability, by establishing

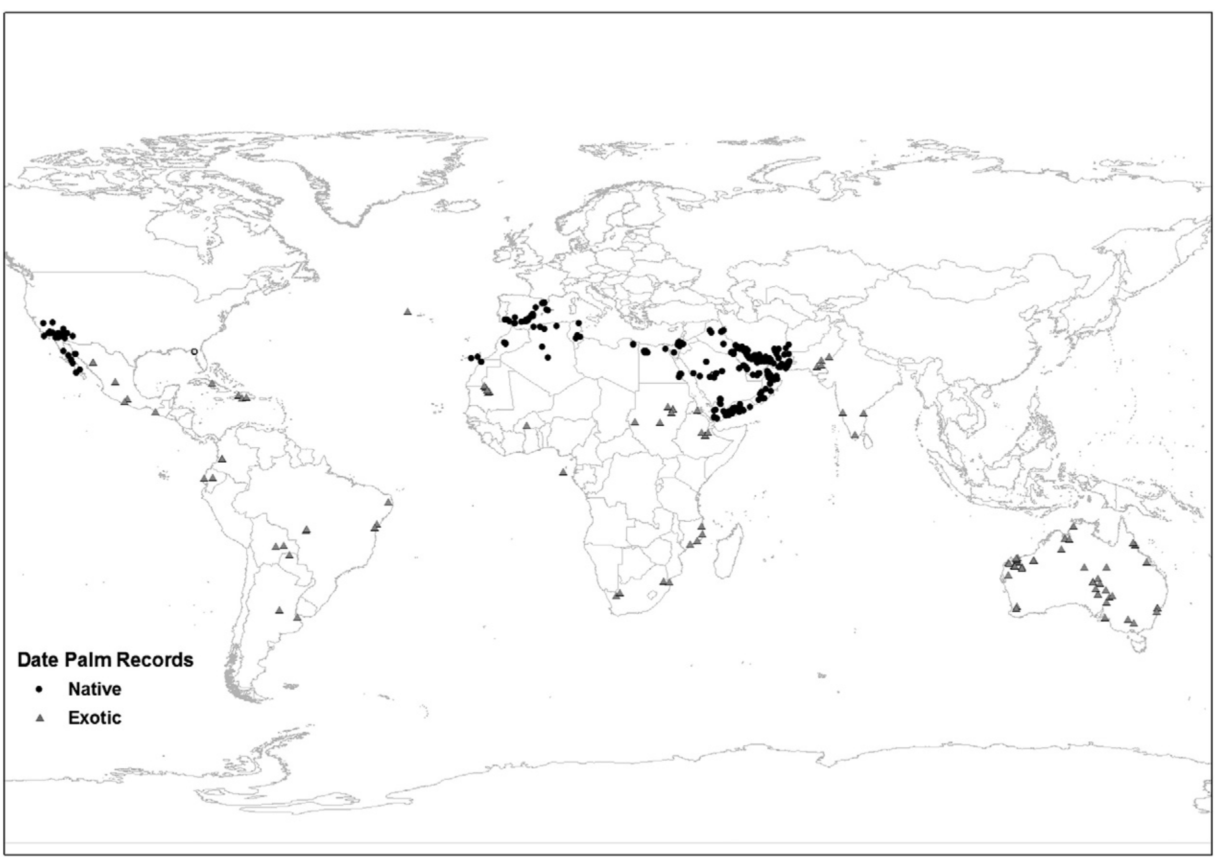

Fig. 1. *P. dactylifera* current global distribution taken from Global Biodiversity Information Facility, Missouri Botanical Gardens' database and other date palm literature in CAB Abstracts databases at global scale.