Indian J. Plant Genet. Resour. 29(1): 16-21 (2016)
DOI 10.5958/0976-1926.2016.00004.8

Maximum Entropy (Maxent) Approach to Sorghum Landraces Distribution Modelling

N Sivaraj[1], M Elangovan[2], V Kamala[1], SR Pandravada[1], P Pranusha[1] and SK Chakrabarty[1]
[1]*ICAR-National Bureau of Plant Genetic Resources, Regional Station, Hyderabad-500030, Telangana, India*
[2]*ICAR-Indian Institute of Millets Research, Rajendranagar, Hyderabad-500030, Telangana, India*

(Received: 02 September 2014; Revised: 21 October 2015; Accepted: 27 October 2015)

Maximum entropy method (Maxent) for modelling of selected sorghum landraces (*Errajonna, Kondajonna, Pachchajonna* and *Tellajonna*) geographic distributions with presence-only data from Khammam district of Telangana state, India for future climate was attempted. Variability occurs among the models generated for specific landraces. Specific potential districts in Telangana state were identified for further cultivation and *on-farm* conservation of sorghum landraces in the light of climate change regime.

Key Words: DIVA-GIS, Maxent, Niche modelling, Sorghum

Introduction

Sorghum is the fifth most important cereal crop and is the dietary staple of more than 500 million people in 30 countries. It is grown on 40 million ha in 105 countries of Africa, Asia, Oceania and the Americas. Sorghum [*Sorghum bicolor* (L.) Moench] landraces have been cultivated in parts of Khammam district of Telangana state by the tribal communities since ancient times providing nutritional security. Early introduction of *Sorghum* into India from African region has a profound effect on the genetic variability available in the country, and India thus becoming a secondary centre of diversity. Landraces of sorghum from this centre of diversity harbour genes for resistance to diseases, insect pest and other abiotic stresses *viz*. drought and high temperature. They also act as sources of traits to improve food, animal feed and nutritional security among the tribal population in this region. However, sorghum landrace diversity is under severe threat due to the loss of habitats, industrial activities and large-scale adoption of cash crops. As sorghum is the staple crop of many tribal communities of this region, the genetic diversity needs to be collected and conserved for ensuring food security. These landraces are rich sources of novel genes that may be utilized for enhancing quality, resistance to biotic and abiotic stresses. This variability and genetic diversity needs to be collected and conserved. With the changing cropping pattern and farmers' preference and the ensuing climate change, these characters are at the verge of extinction. Hence, it is imperative to develop suitable strategies for the conservation of sorghum landraces both *ex-situ* and *in-situ*. In the present study, an attempt has been made for the niche modelling of sorghum landraces of Khammam district of Telangana state, India in the light of climate change using MaxEnt method.

Characteristics of the Study Area

The study zone includes the Khammam district of Telangana state, India. It occupies an area of approximately 16,029 square kilometres (Geographical coordinates: 17.2500°N and 81.1500°E). Summer season is hot and the temperatures can rise rapidly during the day. Monsoon season brings certain amount of rainfall and the temperatures gradually reduce during this period. After the onset of the monsoon day temperatures are much lower and as the winter approaches they reduce further. Summer season commence in March and lasts till the end of May. During this time day temperatures are high and can reach 50°C to 55°C. Humidity is low as it is not located near the ocean. Conditions are generally dry during this period and the day temperatures range from a minimum of 40°C and can rise up to a maximum of 45°C to 50°C. Monsoon season brings much needed relief from the heat. Monsoon seasons are from the months of June to September. Temperatures average around 30°C during this period. The place gets rain from the South West Monsoon. Some amount of rainfall can be experienced in October as well. Winter season is from December to February. January is usually the coldest part

*Author for Correspondence: E-mail: sivarajn@gmail.com

of the year. Temperatures range around 28 °C to 34 °C during this time. The annual average rainfall is 900-1150 mm mostly from the south west monsoon. Red soils are predominant in the district which includes chalkas, red sandy and deep red loamy. Very deep black cotton soils are also seen in some parts of the district. Khammam district is endowed with agro-climatic and soil conditions in which a wide range of agri-horticulture crops like paddy, sorghum, cotton, mango, banana, cashew, coconut, oil palm, cocoa, pepper etc. are grown.

The Khammam district has a tribal population of 5,58,958 which is about 13.29% of the total tribal population of the Telangana state. *Koyas, Guttikoyas, Konda Reddis, Lambadas* are the major tribal groups of this region cultivating sorghum landraces. Tribal women belonging to the above groups are responsible to conserve the sorghum germplasm for growing in the next season and their contribution is more in farm activities apart from their routine household activities.

Materials and Methods

MaxEnt software version 3.3.3k downloaded from www.cs.princeton.edu/~schapire/maxent was used for sorghum landraces habitat modelling. Maximum entropy (Maxent) is a niche modelling method that has been developed involving species distribution information based only on known presences and is a general-purpose method for making predictions or inferences from incomplete information. Using 27 selected sorghum landraces evidence points (Table 1) which were collected during two exploration missions jointly organised by the ICAR-National Bureau of Plant Genetic Resources (NBPGR) and Directorate of Sorghum Research (DSR), MaxEnt runs were performed for each of the sorghum landraces (*Errajonna, Kondajonna, Pachchajonna* and *Tellajonna*). The sorghum landrace evidence points (27) were recorded using Global Positioning System (Garmin GPS-12). WorldClim (WC) database (http://www.worldclim.org) was used for the study and it provides interpolated global climate surfaces using latitude, longitude, and elevation as independent variables and represents long-term (1950–2000) monthly means of maximum, minimum, mean temperatures and total rainfall. For future climates (2020), we used the Global circulation model outputs from the Intergovernmental Panel on Climate Change, Fourth Assessment Report (Solomon *et al*., 2007). Default settings were used in Maxent so that the complexity of the model varied depending upon the number of data

Table 1. Evidence points of *Sorghum* landraces sourced from Khammam district, Telangana state, India for Maxent modelling

Accession identity	Vernacular name	Longitude (East)	Latitude (North)	Altitude (feet)
IC0283753	*Errajonna*	80.9105	17.7210	428
IC0347580	*Errajonna*	81.5100	17.7417	400
IC0596013	*Errajonna*	80.9000	17.6667	435
IC0596016	*Errajonna*	80.5500	18.4833	311
IC0596017	*Errajonna*	80.4000	18.5500	315
IC0596011	*Kondajonna*	81.3833	17.7500	432
IC0306486	*Kondajonna*	81.3833	17.5833	320
IC0596021	*Kondajonna*	81.4000	17.6000	310
IC0596015	*Kondajonna*	80.5900	18.2890	287
IC0596009	*Kondajonna*	80.3667	18.6000	315
IC0596018	*Kondajonna*	80.9105	17.7210	428
IC0347592	*Kondajonna*	81.0400	17.7133	207
IC0596014	*Kondajonna*	81.2467	17.6845	311
IC0596008	*Pachchajonna*	81.1500	17.7333	445
IC0596010	*Pachchajonna*	81.0333	17.7000	447
IC0596026	*Pachchajonna*	81.0667	17.7333	432
IC0596029	*Pachchajonna*	81.2667	17.5833	310
IC0596012	*Pachchajonna*	80.9000	17.6667	435
IC0596020	*Pachchajonna*	81.2267	17.6033	307
IC0596021	*Pachchajonna*	82.2267	18.6033	307
IC0596022	*Pachchajonna*	81.1500	17.7333	326
IC0249063	*Tellajonna*	81.3717	17.7533	214
IC0347590	*Tellajonna*	80.6717	18.1983	222
IC0347585	*Tellajonna*	81.5033	17.7483	500
IC0347591	*Tellajonna*	80.4233	18.5417	307
IC0347588	*Tellajonna*	81.0400	17.7117	206
IC0249067	*Tellajonna*	81.0200	17.3167	700

points used for model fitting. The *ASCI* file generated by the maxent run for each of the sorghum landraces was imported to grid file using DIVA-GIS software version 7.5 (Hijmans *et al*., 2012). The grid layer generated for each run was overlayed on India shapefile using DIVA-GIS and analysed.

Results and Discussion

The important landraces in sorghum germplasm studied include *Errajonna, Kondajonna, Pachchajonna* and *Tellajonna*. The panicles are compact, loose or semi-compact. The seed colour varied from shades of white/yellow/red and the seed size ranging from medium-bold to bold (Fig. 1). Characteristics of the sorghum landraces selected for niche modelling are presented in Table 2.

Maxent model for *Errajonna* generated for future climate (2020) is presented in Fig. 2. Warmer colours show areas with better predicted conditions. The graph represented in Fig. 2 shows the omission rate

Table 2. Characteristics of selected landraces of Sorghum occurring in Khammam district, Telangana State, India

Characters/Landraces	Errajonna	Kondajonna	Pachchajonna	Tellajonna
Race	Durra-bicolor	Guinea-caudatum	Durra-bicolor	Guinea-caudatum
Plant height (cm)	155-250	160-390	230-440	240-330
Plant pigmentation	Pigmented	Pigmented	Pigmented	Pigmented
Basal tillers number	2	2	2	2
Midrib colour	White	White	White	White
Days to flowering	68-106	92-143	62-118	69-128
Panicle length (cm)	26	16	16	16
Panicle width (cm)	12	7	7.5	11
Panicle compactness and shape	Semi-compact elliptic	Semi-compact elliptic	Semi-loose stiff branches	Semi-compact elliptic
Glume colour	Purple	Straw	Straw	Straw
Glume covering	Grain fully covered	1/4th grain covered	3/4th grain covered	½ grain covered
Seed colour	Reddish brown	White	Yellow	White
Seed lustre	Lustrous	Lustrous	Lustrous	Lustrous
Seed size (mm)	3	2.5	3	2.8
100 Seed weight (g)	2.9	2.8	2.8	2.75

and predicted area as a function of the cumulative threshold. The omission rate is calculated both on the training presence records, and (if test data are used) on the test records. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.The image uses colours to indicate predicted probability that conditions are suitable, with red indicating high probability (0.77 to 1.0) of suitable conditions for the Errajonna, green indicating conditions typical of those where the species is found and lighter shades of green indicating low predicted probability of suitable conditions. Adilabad, Karimanagar and Nalgonda districts of Telangana state are the potential regions for introducing and cultivating the *Errajonna* landrace and for planning *in-situ on-farm* conservation sites in the light of climate change scenario.

MaxEnt models for *Kondajonna*, *Pachchajonna* and *Tellajonna* and respective omission and predicted areas are provided in Figs. 3-5. The highest probability (0.7 to 1.0) of distribution of these landraces is represented by red colour. Of all the landraces, *Pachchajonna* may be affected due to climate change. Most of the high probable areas of distribution for *Pachchajonna* shifted to coastal region of Andhra Pradesh state. Highest probable areas of distribution of sorghum landraces in the light of climate change as indicated in the present study is as follows: *Tellajonna*>*Errajonna*>*Kondajonna*> *Pachchajonna* respectively.

The probability distribution of selected sorghum landraces is the sum of each weighted variable divided by a scaling constant to ensure that the probability value ranges from 0 to 1. The information available about the target distribution of sorghum landraces often presents itself as a set of real-valued variables, called 'features', and the constraints are that the expected value of each feature should match its empirical average (Phillips *et al*., 2006). The program starts with a uniform probability distribution and works in cycles adjusting the probabilities to maximum entropy. It iteratively alters one weight at a time to maximize the likelihood of reaching the optimum probability distribution. Maxent is considered as the most accurate model performing extremely well in predicting

A. *Errajonna* **B. *Kondajonna*** **C. *Pachchajonna*** **D. *Tellajonna***

Fig. 1. Select landraces of sorghum from Khammam district, Telangana State, India

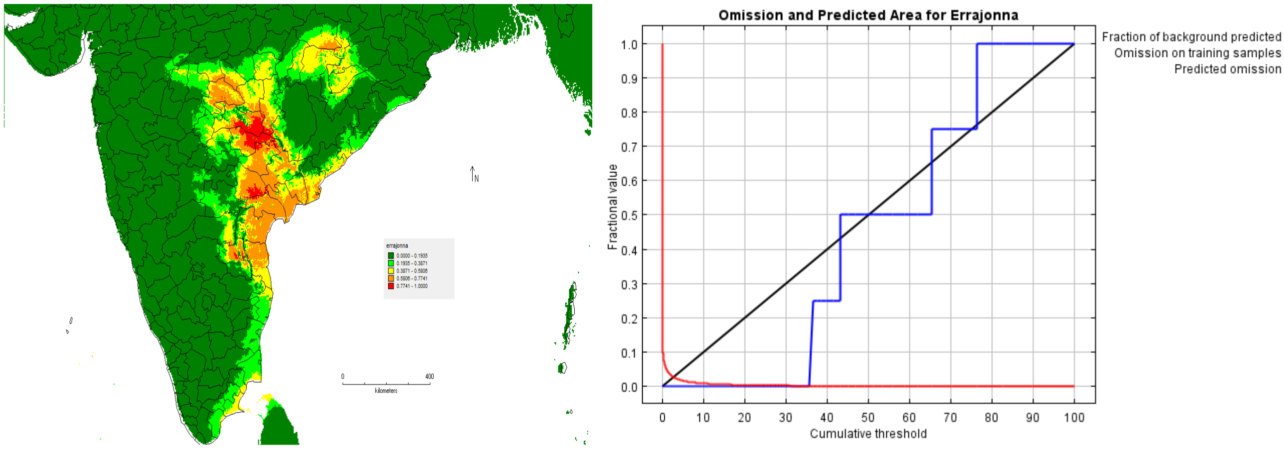

Fig. 2. MaxEnt model for *Errajonna* sorghum landrace and plot showing predicted area

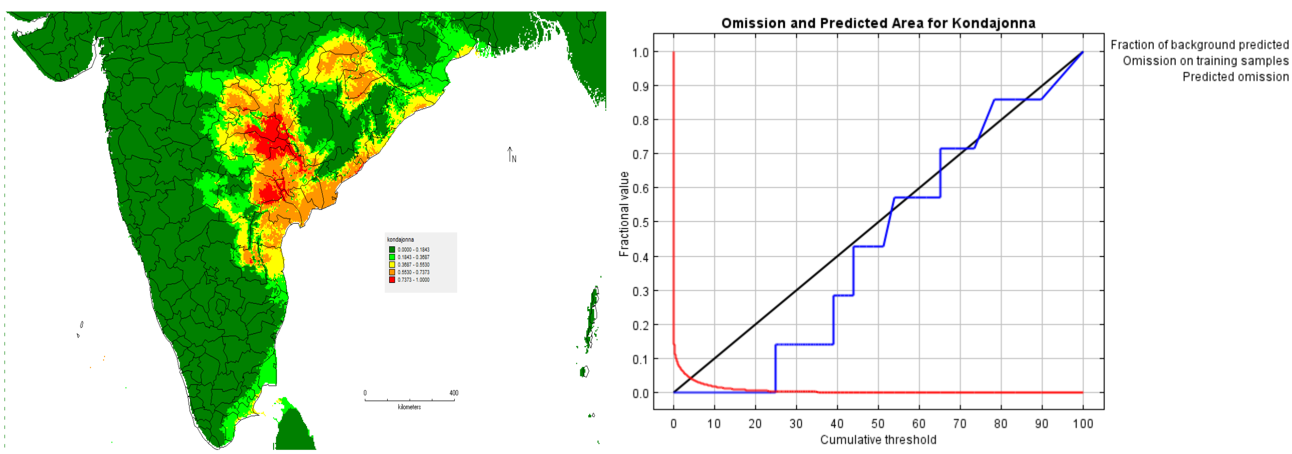

Fig. 3. MaxEnt model for *Kondajonna* sorghum landrace and plot showing predicted area

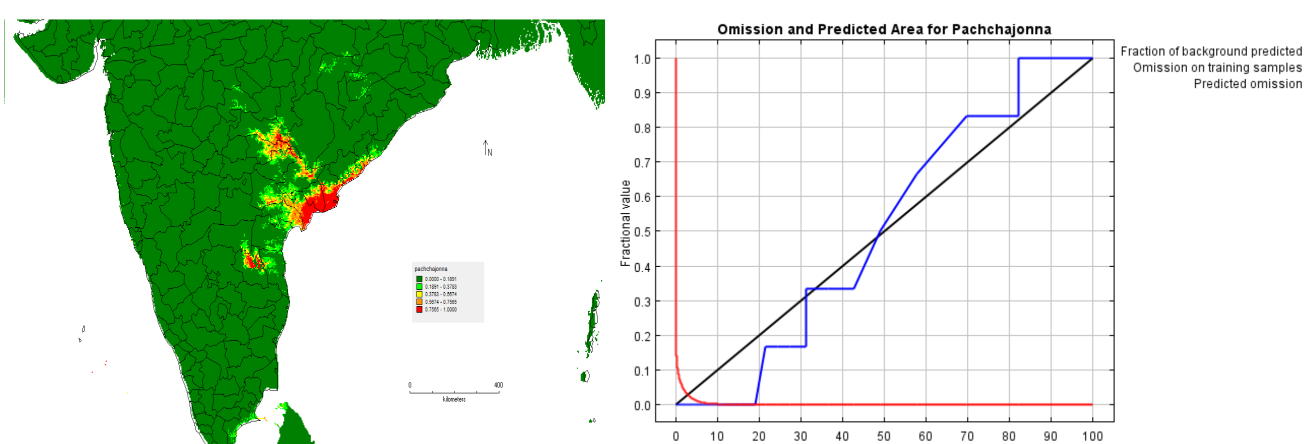

Fig. 4. MaxEnt model for *Pachchajonna* sorghum and plot showing predicted area

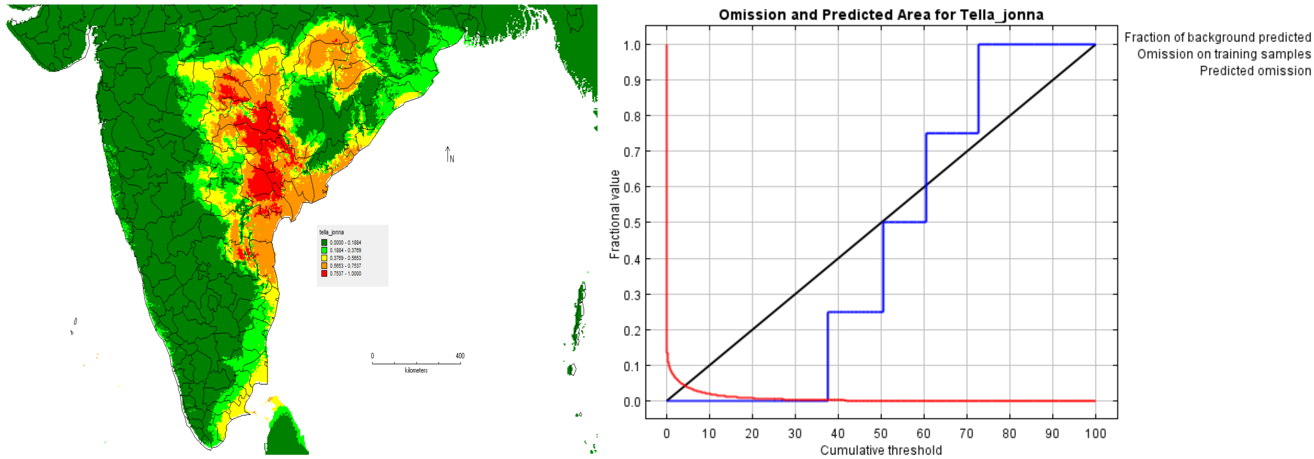

Fig. 5. MaxEnt model for *Tellajonna* sorghum and plot showing predicted area

occurrences in relation to other common approaches (Elith *et al*., 2006; Hijmans and Graham, 2006; Graham and Hijmans, 2006), especially with incomplete information. Predictive modelling of sorghum landraces geographic distributions based on the environmental conditions of sites of known presence establishes an important technique in analytical biology, with applications in conservation and reserve planning, ecology, evolution, epidemiology, invasive-species management and other fields (Corsi *et al*., 1999; Peterson and Shaw, 2003; Peterson *et al*., 1999; Scott *et al*., 2002; Welk *et al*., 2002; Yom-Tov and Kadmon, 1998). The MaxEnt modelling approach can be used in its present form for many applications with presence-only datasets, and merits further research and development (Eitzinger *et al*., 2013). Maxent has been successfully used by many researchers earlier to predict distributions such as stony corals (Tittensor *et al*., 2009); macrofungi (Wollan *et al*., 2008); seaweeds (Verbruggen *et al*., 2009); forests (Carnaval and Moritz, 2008), rare plants (Williams *et al*., 2009) and many other species (Elith *et al*., 2006).

The availability of detailed environmental databases, composed with cheap and efficient computer technology, has driven a rapid increase in predictive modelling of species environmental requirements and geographic distributions. The potential impact of climate changes on agricultural crop production varies spatially and depends on crop specific biophysical constraints (Eitzinger *et al*., 2013). Due to population expansion worldwide, the demand for food sources would increase in the changed climate regime. Because of sorghum's extensive adaptation would prove vital in meeting the food security of the people, use of MaxEnt method is highly warranted in preserving the important sorghum landraces.

Acknowledgements

Authors are thankful to Director, ICAR-NBPGR, New Delhi and Head, Division of Plant Quarantine, ICAR-NBPGR, New Delhi for facilities and help.

References

Carnaval AC and C Moritz (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J. Biogeo.* **35**: 1187-1201.

Corsi F, E Dupr`e, L Boitani (1999) A large-scale model of wolf distribution in Italy for conservation planning. *Conserv. Biol.* **13**: 150-159.

Eitzinger A, P Laderach, S Carmona, C Navarro, L Collet (2013) Prediction of the impact of climate change on coffee and mango growing areas in Haiti. *Full Tech. Report*. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

Elith, J, CH Graham, RP Anderson, M Dudík, S Ferrier, A Guisan, RJ Hijmans, F Huettmann, R Leathwick, A Lehmann, LG Lohmann, BA Loiselle, G Manion, C Moritz, M Nakamura, Y Nakazawa, JMcC Overton, AT Peterson, J Phillips, K Richardson, R Scachetti-Pereira, E Schapire, J Soberon, S Williams, M Wisz and E Zimmermann (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129-151.

Graham CH and RJ Hijmans (2006) A comparison of methods for mapping species ranges and species richness. *Glo. Eco. & Biogeo.* **15**: 578.

Hijmans RJ and C Graham (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Glo.Change Bio.* **12**: 2272-2281.

Hijmans RJ, L Guarino and P Mathur (2012) *DIVA-GIS Version* **7.5**: Manual. Available at: www.diva-gis.org.

Peterson AT and J Shaw (2003) *Lutzomyia* vectors for cutaneous leishmaniasis in southern Brazil: ecological niche models, predicted geographic distribution, and climate change effects. *Int. J. Parasitol*. **33**: 919-931.

Peterson AT, J Sober´on and V S´anchez-Cordero (1999) Conservation of ecological niches in evolutionary time. *Science* **285**: 1265-1267.

Phillips SJ, RP Anderson and RE Schapire (2006) Maximum entropy modeling of species geographic distributions. *Eco. Modelling* **190**: 231-259.

Scott JM, PJ Heglund, ML Morrison, JB Haufler, MG Raphael, WA Wall and FB Samson (eds.) (2002) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Island Press, Washington.

Solomon S, D Qin, M Manning, Z Chen, M Marquis, KB Averyt, M Tignor and HL Miller (2007) IPCC 2007: *Summary for Policy makers*. Climate change 2007: The physical science basis. *Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change.* Available at: http://goo.gl/N4RMw.

Tittensor DP, AR Baco, PE Brewin, MR Clark, M Consalvey, J Hall-Spencer, AA Rowden, T Schlacher, KI Stocks and AD Rogers (2009) Predicting global habitat suitability for stony corals on seamounts. *J. Biogeo*. **36**: 1111-1128.

Verbruggen H, L Tyberghein, K Pauly, C Vlaeminck, K VanNieuwenhuyze, W Kooistra, F Leliaert and O De Clerck (2009) Macroecology meets macroevolution: evolutionary niche dynamics in the seaweed Halimeda. *Glo. Eco. and Biogeo*. **18**: 393-405.

Welk E, K Schubert and MH Hoffmann (2002) Present and potential distribution of invasive mustard (*Alliara petiolata*) in North America*. Divers. Distributions.* **8**: 219-233.

Williams JN, CW Seo, J Thorne, JK Nelson, S Erwin, JM O'Brien and MW Schwartz (2009) Using species distribution models to predict new occurrences for rare plants. *Divers. Distributions*. **15**: 565-576.

Wollan AK, V Bakkestuen, H Kauserud, G Gulden and R Halvorsen (2008) Modelling and predicting fungal distribution patterns using herbarium data. *J. Biogeo.* **35**: 2298-2310.

Yom-Tov Y and R Kadmon (1998) Analysis of the distribution of insectivorous bats in Israel. *Divers. Distributions.* **4**: 63-70.