Strengthening MaxEnt modelling through screening of redundant explanatory Bioclimatic Variables with Variance Inflation Factor analysis

Prakash Pradhan

West Bengal Biodiversity Board, Department of Environment, Government of West Bengal, Poura Bhawan, 4th Floor, Salt Lake City, Sector-III, Kolkata, West Bengal, PIN – 700 106, India
shresthambj@gmail.com

Abstract: Through the past two decades, bioclimatic variables have been utilized as an important tool to understand species distribution and prioritize areas for conservation of target species through the Ecological Niche Modelling. However, the interpolated datasets of bioclimatic variables are known to cause over-fitting of the models mainly due to multicollinearity or redundancy within the variables. In the current work, bioclimatic variables of South and South East Asia region are screened regarding the presence of multicollinearity or redundancy to serve as a convenient reference for investigators of the region.
[Pradhan P. **Strengthening MaxEnt modelling through screening of redundant explanatory Bioclimatic Variables with Variance Inflation Factor analysis.** *Researcher* 2016;8(5):29-34]. ISSN 1553-9865 (print); ISSN 2163-8950 (online). http://www.sciencepub.net/researcher. 5. doi:10.7537/marsrsj08051605.

Keywords: Ecological Niche Modelling, Habitat Suitability Modelling, MaxEnt, Multicollinearity, Over-fitting, South Asia, South-East Asia

Abbreviations: ASCII-American Standard Code for Information Interchange; ENFA-Ecological Niche Factor Analysis; ENM-Ecological Niche Modelling; ESRI-Environmental Systems Research Institute; SDM-Species Distribution Modelling; VIF-Variance Inflation Factor

1. Introduction

Bioclimatic variables represent important explanatory variables to understand species distribution (Busby 1986, Nix 1986). They express spatial variation in annual means, seasonality and extreme or limiting climatic factors and represent biologically meaningful parameters for characterizing species distributions (Saatchi et al., 2008). The advent of ecological niche modeling/ species distribution modeling/ habitat suitability modeling has opened an array of utility of bioclimatic variables. However, being derived from interpolated datasets, these 19 bioclimatic variables are not free from drawbacks, one of them being redundancy/ multicollinearity (Arif, Adams and Wicknick, 2007).

For modeling ENM/SDM if the explanatory variables are used without screening then it may inadvertently lead to inclusion of those variables also which are highly correlated and have basically same set of information. In such instance, these variables may incline statistical weight towards themselves, and when it happens, it may undesirably lead to over-fitting of the model (Anderson and Gonzales, 2011; van Gils et al., 2014).

However, there are methods to counter such anomaly by selecting few explanatory/ predictor variables making the resultant models more 'parsimonius' yet less 'over-fitted'. The methods/ tools to minimize redundancy of explanatory bioclimatic variables include the three-way Mantel test where the relationship between the two variables are evaluated while holding geographic distance constant (Legendre and Legendre, 1998); ENFA analysis with BIOMAPPER 4.0 (Hirzel, Hausser and Perrin, 2007); MaxEnt-based stepwise selection of variables (Parolo, Rossi and Ferrarini, 2008); Correlation analysis through ENM Tools (Warren, Glor and Turelli, 2010); Principal Component Analysis of variables (Rangel, Diniz-Filho and Bini, 2010; Fourcade et al., 2014); Maxent Variable Selection package in R platform (Jueterbock et al., 2016); VIF function in the vegan package in R (Oksanen et al., 2016); SDM toolbox for ArcGIS (ESRI) ver 10, etc.

The current analysis makes use of ENM Tools (Warren, Glor and Turelli, 2010) for correlative screening of bioclimatic variables, principally due to its integrative capacity to MaxEnt program, ease of its graphic user interface (knowledge of programming language not necessary), integrated capacity of criterion-based model selection using AIC, AICc, and BIC (Burnham and Anderson, 2002), and ability to handle and output large data size.

The current work is to screen the redundant bioclimatic variables of South and South East Asia and to suggest a working set of bioclimatic variables for easy reference to those involved in ENM/ SDM/ habitat suitability modeling in the region and to set a protocol to screen explanatory variables for model building elsewhere.

2. Material and Methods

Worldclim (Hijmans et al., 2005) hosts the interpolated climatic records from a global network of 4000 climate stations, with time series of 1950-2000. Bioclimatic variables (Busby 1986; Nix 1986; Hijmans et al. 2005) are one of the output of worldclim which are available in tiles and for South and South East Asia, tiles of 18, 19, 28 and 29 at a spatial resolution of 30 arc seconds (~1 × 1 km resolution) were obtained following Pradhan (2015).The current investigation doesn't incorporate climatic information of tile 110 and 210, hence geographic space of East Asian countries like Taiwan, Japan etc. could not be included in the current analysis.

Each of the 19 downloaded bioclimatic variables of each tile was merged with same variable of the other three tile (i.e. bio 1 of tile 18 with bio 1 of 19, 28 and 29) to obtain variable for larger coverage (South and South East Asia; Latitude 0˚ to 60˚, Longitude 60˚ to 120˚). They were then converted to ESRI ASC (or ESRI ASCII) in DIVA-GIS version 7.5 (Hijmans et al., 2001) for analyzing correlation in ENM Tools (Warren, Glor and Turelli, 2010). Various margin of Pearson correlation coefficient 'r' for screening of variables has been suggested viz. >0.5 by Václavík and Meentemeyer (2009), r>0.7 by Dormann et al. (2013) and Rotllan-Puig and Traveset (2016); >0.9 by Jueterbock et al. (2016). However, the current analysis uses the standardized values of r>0.8, r2>0.8 and VIF value of >10 for screening. After obtaining Pearson correlation coefficient 'r' values for individual set of correlation, coefficient of determination 'r^2' values (r×r) were derived, followed by the derivation of the VIF by the formula [1/ (1-r^2)] (Zuur, Ieno and Elphick, 2010). VIF indicates the degree to which the standard errors are inflated due to the levels of multicollinearity. VIF values of 10 were taken as indicative of problematic collinearity/ redundancy (Montgomery and Peck, 1992).

3. Results

The correlation was performed amongst 19 bioclimatic variables and the values of r are presented in Table 1, values of the corresponding r^2 are presented in Table 2 and values of the corresponding VIF are presented in Table 3. Individual details of correlation analysis of the 19 bioclimatic variables are as follows.

Bio 1 (Annual Mean Temperature): Bio 1 is correlated with bio 5 (Max Temperature of Warmest Month) via r>0.8, while it is correlated with bio 9 (Mean Temperature of Driest Quarter) and bio 10 (Mean Temperature of Warmest Quarter) via r>0.8, r^2>0.8, and it is correlated with bio 6 (Min Temperature of Coldest Month) and bio 11 (Mean Temperature of Coldest Quarter) via r>0.8, r2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 1 is not to be used alongside bio 6 or bio 11 for a modeling procedure.

Bio 2 (Mean Monthly Temperature Range): Bio 2 is not correlated (r<0.8, r2<0.8, VIF<10) with any other variable hence could be used in any combination with other bioclimatic variables.

Bio 3 (Isothermality): Bio 3 is not correlated (r<0.8, r2<0.8, VIF<10) with any other variable hence could be used in any combination with other bioclimatic variables.

Bio 4 (Temperature Seasonality): Bio 4 is correlated with bio 6 (Min Temperature of Coldest Month) and bio 11 (Mean Temperature of Coldest Quarter) via r^2>0.8, while it is correlated with bio 7 (Temperature Annual Range) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 4 is not to be used alongside bio 7 for a modeling procedure.

Bio 5 (Max Temperature of Warmest Month): Bio 5 is correlated with bio 1 (Annual Mean Temperature) via r^2>0.8, while it is correlated with bio 10 (Mean Temperature of Warmest Quarter) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 5 is not to be used alongside bio 10 for a modeling procedure.

Bio 6 (Min Temperature of Coldest Month): Bio 6 is correlated with bio 9 (Mean Temperature of Driest Quarter) via r>0.8, r^2>0.8; with bio 4 (Temperature Seasonality) and bio 7 (Temperature Annual Range) via r^2>0.8, while it is correlated with bio 1 (Annual Mean Temperature) and bio 11 (Mean Temperature of Coldest Quarter) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 6 is not to be used alongside bio 1 or bio 11 for a modeling procedure.

Bio 7 (Temperature Annual Range): Bio 7 is correlated with bio 6 (Min Temperature of Coldest Month) via r^2>0.8, while it is correlated with bio 4 (Temperature Seasonality) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 7 is not to be used alongside bio 4 for a modeling procedure.

Bio 8 (Mean Temperature of Wettest Quarter): Bio 8 is not correlated (r<0.8, r2<0.8, VIF<10) with any other variable hence could be used in any combination with other bioclimatic variables.

Bio 9 (Mean Temperature of Driest Quarter): Bio 9 is correlated with bio 1 (Annual Mean Temperature), bio 6 (Min Temperature of Coldest

Month) and bio 11 (Mean Temperature of Coldest Quarter) via r>0.8, r^2>0.8.

Bio 10 (Mean Temperature of Warmest Quarter): Bio 10 is correlated with bio 1 (Annual Mean Temperature) via r>0.8, r^2>0.8; while it is correlated with bio 5 (Max Temperature of Warmest Month) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 10 is not to be used alongside bio 5 for a modeling procedure.

Bio 11 (Mean Temperature of Coldest Quarter): Bio 11 is correlated with bio 9 (Mean Temperature of Driest Quarter) via r>0.8, r^2>0.8; with bio 4 (Temperature Seasonality) via r^2>0.8, while it is correlated with bio 1 (Annual Mean Temperature) and bio 6 (Min Temperature of Coldest Month) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 11 is not to be used alongside bio 1 or bio 6 for a modeling procedure.

Bio 12 (Annual Precipitation): Bio 12 is correlated with bio 13 (Precipitation of Wettest Month) and bio 16 (Precipitation of Wettest Quarter) via r>0.8, r^2>0.8; while it is correlated with bio 18 via r^2>0.8.

Bio 13 (Precipitation of Wettest Month): Bio 13 is correlated with bio 12 (Annual Precipitation) via r>0.8, r^2>0.8; while it is correlated with bio 16 (Precipitation of Wettest Quarter) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 13 is not to be used alongside bio 16 for a modeling procedure.

Bio 14 (Precipitation of Driest Month): Bio 14 is correlated with bio 17 (Precipitation of Driest Quarter) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 14 is not to be used alongside bio 17 for a modeling procedure.

Bio 15 (Precipitation Seasonality): Bio 15 is not correlated (r<0.8, r2<0.8, VIF<10) with any other variable hence could be used in any combination with other bioclimatic variables.

Bio 16 (Precipitation of Wettest Quarter): Bio 16 is correlated with bio 12 (Annual Precipitation) via r>0.8, r^2>0.8, while it is correlated with bio 13 (Precipitation of Wettest Month) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 16 is not to be used alongside bio 13 for a modeling procedure.

Bio 17 (Precipitation of Driest Quarter): Bio 17 is correlated with bio 14 (Precipitation of Driest Month) via r>0.8, r^2>0.8, VIF>10. Considering negative effects of VIF and notwithstanding any major necessity of inclusion, bio 17 is not to be used alongside bio 14 for a modeling procedure.

Bio 18 (Precipitation of Warmest Quarter): Bio 18 is correlated with bio 12 (Annual Precipitation) via r>0.8.

Bio 19 (Precipitation of Coldest Quarter): Bio 19 is not correlated (r<0.8, r2<0.8, VIF<10) with any other variable hence could be used in any combination with other bioclimatic variables.

4. Discussions

Through the current analysis, an array of multicollinearity has been observed among the bioclimatic variables of South and South East Asia. Whenever there is a necessity of removal from each pair of redundant variables, the first choice should be to remove the variable from pairs having VIF values >10. The second choice should be from the pair of variables with both r and r^2 having value of >0.8 (though they may not have VIF>10); thirdly the variables from the pair having only r>0.8 should be removed (though they may not have r^2>0.8 and VIF>10).

Variable pairs with the value of only r^2>0.8 (r<0.8, VIF<10) should be considered with care as r^2 represent the proportion/ percentage of variation (fluctuation) of one variable that is predictable from another variable. As evident from Table 1 and Table 2, the r values of some of the negatively correlated variables like bio 4 and bio 6 (r=-0.90155831752501); bio 4 and bio 11 (r= -0.909186862266398); bio 6 and bio 7 (r= -0.896997609716365) when multiplied by the power of 2, return their values (i.e. r^2) as 0.81281, 0.82662 and 0.8046 respectively. It may be suggested to keep biologically meaningful yet negatively correlated variables because they may possess some unmeasured information not present in other variables.

The preliminary MaxEnt runs (minimum of triplicate runs), may be helpful to identify the least significant variables which may further be removed based upon the individual response of variable versus the species occurrence/ percentage contribution to the model (Jueterbock et al., 2016).

Even though the explanatory variables are screened, it may be very helpful to take advice from the expert of the species/ genera/ family regarding choice of variables for model building of a particular species, which may add to the incorporation of the variables that are limiting/ extreme events/ have more biological meaning to the distribution of the target species (Mbatudde et al. 2012; Pradhan et al. 2012).
In this regard 'annual average' factors like average temperatures and precipitations may have little meaning. Composite variables based on the precipitation of the warmest or coldest period or temperature of the wettest or driest period could be

avoided as these datasets may be internally flawed. Therefore, the average temperature of the warmest period, the maximum temperature of the warmest month, minimum temperature of the coldest period, temperature variability, precipitation variability, precipitation of the wettest and driest periods etc may be used in combination unless otherwise VIF and r values restrict them.

Table 1. Table showing values of Pearson correlation coefficient 'r' of the pairs of bioclimatic variables. Values more than 0.8 and indicated as bold.

Variables	Bio 19	Bio 1	Bio 2	Bio 3	Bio 4	Bio 5	Bio 6	Bio 7	Bio 8	Bio 9	Bio 10	Bio 11	Bio 12	Bio 13	Bio 14	Bio 15	Bio 16	Bio 17	Bio 18
Bio 19	1.000	0.294	-0.386	0.556	-0.376	0.107	0.379	-0.438	0.129	0.311	0.175	0.342	0.595	0.399	0.753	-0.309	0.401	0.768	0.283
Bio 1			-0.280	0.659	-0.795	**0.839**	**0.971**	-0.773	0.709	**0.907**	**0.906**	**0.975**	0.580	0.587	0.223	0.176	0.590	0.257	0.346
Bio 2				-0.133	0.332	-0.050	-0.395	0.494	-0.266	-0.202	-0.183	-0.317	-0.615	-0.468	-0.457	0.491	-0.493	-0.485	-0.526
Bio 3					-0.861	0.276	0.752	-0.829	0.348	0.643	0.362	0.765	0.664	0.549	0.520	0.159	0.562	0.544	0.406
Bio 4						-0.355	-0.902	**0.980**	-0.407	-0.774	-0.465	-0.909	-0.697	-0.667	-0.306	-0.223	-0.680	-0.344	-0.487
Bio 5							0.700	-0.312	0.678	0.749	**0.981**	0.706	0.231	0.282	0.035	0.108	0.274	0.052	0.005
Bio 6								-0.897	0.626	**0.905**	0.792	**0.995**	0.678	0.650	0.311	0.127	0.659	0.349	0.429
Bio 7									-0.414	-0.740	-0.446	-0.887	-0.759	-0.690	-0.392	-0.102	-0.707	-0.432	-0.567
Bio 8										0.395	0.741	0.631	0.444	0.466	0.176	0.182	0.464	0.196	0.363
Bio 9											0.791	**0.906**	0.484	0.479	0.189	0.106	0.484	0.221	0.215
Bio 10												0.791	0.360	0.389	0.117	0.095	0.385	0.139	0.151
Bio 11													0.651	0.644	0.267	0.195	0.651	0.304	0.410
Bio 12														**0.898**	0.600	-0.077	**0.927**	0.639	**0.811**
Bio 13															0.266	0.210	**0.991**	0.308	0.729
Bio 14																-0.455	0.299	**0.994**	0.407
Bio 15																	0.172	-0.452	-0.017
Bio 16																		0.341	0.766
Bio 17																			0.440
Bio 18																			1.000

Table 2. Table showing values of Pearson coefficient of determination 'r^2' for the bioclimatic variable pairs. Values more than 0.8 and indicated as bold.

Variables	Bio 19	Bio 1	Bio 2	Bio 3	Bio 4	Bio5	Bio 6	Bio 7	Bio 8	Bio 9	Bio 10	Bio 11	Bio 12	Bio 13	Bio 14	Bio 15	Bio 16	Bio 17	Bio 18
Bio 19	1.000	0.086	0.149	0.309	0.141	0.011	0.144	0.192	0.017	0.097	0.030	0.117	0.354	0.159	0.567	0.095	0.161	0.590	0.080
Bio 1			0.079	0.434	0.632	0.704	**0.943**	0.597	0.502	**0.822**	**0.821**	**0.951**	0.337	0.344	0.050	0.031	0.348	0.066	0.120
Bio 2				0.018	0.110	0.003	0.156	0.244	0.071	0.041	0.033	0.100	0.379	0.219	0.209	0.241	0.243	0.235	0.277
Bio 3					0.741	0.076	0.565	0.688	0.121	0.414	0.131	0.585	0.441	0.301	0.271	0.025	0.316	0.296	0.165
Bio 4						0.126	**0.813**	**0.960**	0.166	0.600	0.216	**0.827**	0.486	0.444	0.094	0.050	0.462	0.118	0.238
Bio 5							0.490	0.098	0.460	0.560	**0.963**	0.498	0.053	0.080	0.001	0.012	0.075	0.003	0.000
Bio 6								**0.805**	0.392	**0.818**	0.627	**0.991**	0.459	0.423	0.097	0.016	0.435	0.121	0.184
Bio 7									0.171	0.548	0.199	0.787	0.575	0.477	0.154	0.010	0.500	0.186	0.321
Bio 8										0.156	0.549	0.398	0.197	0.218	0.031	0.033	0.215	0.038	0.132
Bio 9											0.625	**0.821**	0.234	0.230	0.036	0.011	0.234	0.049	0.046
Bio 10												0.626	0.129	0.151	0.014	0.009	0.148	0.019	0.023
Bio 11													0.424	0.414	0.071	0.038	0.424	0.092	0.168
Bio 12														**0.806**	0.360	0.006	**0.859**	0.408	0.658
Bio 13															0.071	0.044	**0.982**	0.095	0.532
Bio 14																0.207	0.089	**0.988**	0.166
Bio 15																	0.030	0.204	0.000
Bio 16																		0.117	0.586
Bio 17																			0.194
Bio 18																			1.000

Table 3. Table showing values of Variance Inflation Factor (VIF) of the pairs of bioclimatic variables. Values more than 10 are indicated in bold.

Variables	Bio 1	Bio 2	Bio 3	Bio 4	Bio 5	Bio 6	Bio 7	Bio 8	Bio 9	Bio 10	Bio 11	Bio 12	Bio 13	Bio 14	Bio 15	Bio 16	Bio 17	Bio 18
Bio 19	1.095	1.175	1.448	1.164	1.012	1.168	1.238	1.017	1.107	1.031	1.132	1.549	1.189	2.308	1.105	1.192	2.441	1.087
Bio 1		1.085	1.767	2.714	3.377	**17.515**	2.480	2.009	5.619	5.581	**20.208**	1.508	1.525	1.052	1.032	1.535	1.071	1.136
Bio 2			1.018	1.124	1.003	1.185	1.323	1.076	1.043	1.035	1.112	1.609	1.281	1.264	1.317	1.321	1.307	1.383
Bio 3				3.863	1.083	2.300	3.202	1.138	1.706	1.151	2.409	1.789	1.431	1.371	1.026	1.463	1.420	1.198
Bio 4					1.144	5.342	**25.092**	1.198	2.498	1.276	5.768	1.945	1.800	1.103	1.052	1.859	1.134	1.312
Bio 5						1.962	1.108	1.851	2.275	**26.877**	1.992	1.056	1.087	1.001	1.012	1.081	1.003	1.000
Bio 6							5.118	1.646	5.503	2.684	**106.662**	1.849	1.733	1.107	1.016	1.769	1.138	1.225
Bio 7								1.206	2.211	1.249	4.702	2.355	1.911	1.181	1.011	2.002	1.229	1.473
Bio 8									1.185	2.218	1.660	1.245	1.278	1.032	1.034	1.275	1.040	1.151
Bio 9										2.667	5.587	1.306	1.298	1.037	1.011	1.306	1.052	1.049
Bio 10											2.674	1.149	1.178	1.014	1.009	1.174	1.020	1.023
Bio 11												1.737	1.707	1.077	1.040	1.735	1.102	1.202
Bio 12													5.164	1.564	1.006	7.078	1.690	2.921
Bio 13														1.076	1.046	**56.443**	1.105	2.136
Bio 14															1.261	1.098	**82.397**	1.199
Bio 15																1.031	1.256	1.000
Bio 16																	1.132	2.415
Bio 17																		1.240

However, working only with non-redundant ones may not always yield good results as some of the redundant variables may act as good ecological descriptors. It should be noted that sometimes over-fitting may be the result of selecting too much aggregate sampling sites (sample/observation bias). In this regard, mantel test with SAM 4.0 (Rangel, Diniz-

Filho and Bini, 2010) may be performed to check spatial autocorrelation, if any.

In any case, if multiple models (randomly subsampled) are built for the same species (with RAW output in ASC format) utilizing multiple sets of non redundant bioclimatic variables, the final model may be selected based upon lowest AICc score, highest AUC value and incorporating lesser number of correlated variables ($r>0.8$, $r^2>0.8$) (Warren, Glor and Turelli, 2010). The methodology of the current work may be extended to other areas of the world and the future climate scenarios as well for the screening of possible redundancy and decreasing over-fitting of ecological niche models.

References

1. Busby JR. A biogeographical analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. Australian Journal of Ecology 1986;11(1):1–7. http://dx.doi.org/10.1111/j.1442-9993.1986.tb00912.x.
2. Nix H. A biogeographic analysis of Australian elapid snakes. Atlas of Elapid Snakes of Australia. Australian Government Publishing Service, Canberra, Australia, 1986;4–15 pp.
3. Saatchi S, Buermann W, ter Steege H, Mori S, Smith TB. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. Remote Sensing and Environment 2008;112:2000–2017. http://dx.doi.org/10.1016/j.rse.2008.01.008.
4. Arif S, Adams DC, Wicknick JA. Bioclimatic modelling, morphology, and behavior reveal alternative mechanisms regulating the distributions of two parapatric salamander species. Evolutionary Ecology Research 2007;9:843–854.
5. Anderson RP, Gonzales I. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling 2011;222:2796–2811. http://dx.doi.org/10.1016/j.ecolmodel.2011.04.011.
6. van Gils H, Westinga E, Carafa M, Antonucci A, Ciaschetti G. Where the bears roam in Majella National Park, Italy. Journal of Nature Conservation 2014;22(1):23–34. http://dx.doi.org/10.1016/j.jnc.2013.08.001.
7. Legendre P, Legendre L. Numerical Ecology (2nd Ed.). Amsterdam: Elsevier. 1998.
8. Hirzel AH, Hausser J, Perrin N. Biomapper 4.0. Laboratory for Conservation Biology, Department of Ecology and Evolution, University of Lausanne,Switzerland. 2007.
9. Parolo G, Rossi G, Ferrarini A. Toward improved species niche modelling: *Arnica Montana* in the Alps as a case study. Journal of Applied Ecology 2008;45:1410–1418. http://dx.doi.org/10.1111/j.1365-2664.2008.01516.x.
10. Warren DL, Glor RE, Turelli M. ENM Tools: a toolbox for comparative studies of environmental niche models. Ecography 2010;33:607–611. http://dx.doi.org/10.1111/j.1600-0587.2009.06142.x.
11. Rangel, TF, Diniz-Filho JAF, Bini LM. SAM: a comprehensive application for Spatial Analysis in Macroecology. Ecography 2010;33:46–50. http://dx.doi.org/10.1111/j.1600-0587.2009.06299.x.
12. Fourcade Y, Engler JO, Rödder D, Secondi J. Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. PLoS ONE 9(5): e97122. 2014. http://dx.doi.org/10.1371/journal.pone.0097122.
13. Jueterbock A, Smolina I, Coyer JA, Hoarau G. The fate of the Arctic seaweed *Fucus distichus* under climate change: an ecological niche modeling approach. Ecology and Evolution 2016;6(6):1712–1724. http://dx.doi.org/10.1002/ece3.2001 R package available at https://cran.fhcrc.org/web/packages/MaxentVariableSelection/index.html.
14. Oksanen J, Blanchet FG, Kindt R, Pierre L, Minchin PR, O'Harra RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. Community Ecology Package (Package 'vegan'). R Package available at https://cran.r-project.org/web/packages/vegan. 2016.
15. Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd Ed. Springer-Verlag. 2002.
16. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 2005;25:1965–1978. http://dx.doi.org/10.1002/joc.1276.
17. Pradhan P. Potential distribution of *Monotropa uniflora* L. as a surrogate for range of Monotropoideae (Ericaceae) in South Asia. Biodiversitas 2015;16(2):109–115. http://dx.doi.org/10.13057/biodiv/d160201.
18. Hijmans RJ, Guarino L, Cruz M, Rojas E. Computer tools for spatial analysis of plant

genetic resources data: 1. DIVA-GIS. Plant Genetic Resources Newsletter 2001;127:15–19.
19. Václavík T, Meentemeyer RK. Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling 2009;220:3248–3258. http://dx.doi.org/10.1016/j.ecolmodel.2009.08.013.
20. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, García Marquéz JR, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013;36(1):27–46. http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x.
21. Rotllan-Puig X, Traveset A. Declining relict plants: Climate effect or seed dispersal disruption? A landscape-scale approach. Basic and Applied Ecology 2016;17(1):81–91. http://dx.doi.org/10.1016/j.baae.2015.08.003.
22. Zuur AF, Ieno EN, Elphick CS. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 2010;1:3–14. http://dx.doi.org/10.1111/j.2041-210X.2009.00001.x.
23. Montgomery DC, Peck EA. Introduction to Linear Regression Analysis. New York: Wiley. 1992.
24. Mbatudde M, Mwanjololo M, Kakudidi EK, Dalitz H. Modelling the potential distribution of endangered *Prunus africana* (Hook.f.) Kalkm. in East Africa. African Journal of Ecology 2012;50:393-403. http://dx.doi.org/10.1111/j.1365-2028.2012.01327.x.
25. Pradhan P, Dutta AK, Roy A, Basu SK, Acharya K. Inventory and Spatial Ecology of Macrofungi in the *Shorea robusta* Forest Ecosystem of Lateritic Region of West Bengal. Biodiversity 2012;13(2):88–99. http://dx.doi.org/10.1080/14888386.2012.690560.

5/18/2016