EVALUATING PERFORMANCE OF MAXENT AND EFFECTS OF SAMPLING STRATEGIES ON MODELING FOREST TREE SPECIES—*Schima superba*

Hou-Chang Chen[a], *Nan-Jang Lo*[b], *Wei-I Chang*[c], *and Kai-Yi Huang*[*d]

[a]*Master student, Dept. of Forestry, National Chung-Hsing University, Taiwan E-mail: zkchris@hotmail.com*

[b]*Specialist, E. F. M. O., National Chung-Hsing University, Taiwan E-mail: njl@dragon.nchu.edu.tw*

[c]*Director, HsinChu F. D. O., Forest Bureau, Council of Agriculture, Taiwan E-mail: weii@forest.gov.tw*

[d*]*Professor, same as with author-*a *E-mail:* *kyhuang@dragon.nchu.edu.tw* ***(corresponding author)***

250 Kuo-Kuang Road, Taichung, Taiwan 402, Tel: +886-4-22854663; Fax: +886-4-22854663

ABSTRACT: Predictive model for species distribution has been the core of ecological research since late 20th century with the development of statistical techniques and 3S tools, which not only can be applied to biodiversity conservation and management, but also enhance the ability of predicting species habitat distribution. The sample points of Chinese guger-tree (*Schima superba*, CGT) in the Huisun study area were obtained by GPS, and GIS technique was used to overlay five environmental factors (including terrain factors and vegetation index derived from SPOT-5 satellite images). Besides, we designed different ratios of background to target to evaluate different sampling designs for modeling individual species' distribution. The study applied MAXENT, DT and DA models to predict the suitable habitat of CGT in Huisun. The results showed that the accuracy of DT was only slightly better than that of MAXENT and accuracies of the two models were much better than that of DA. Implementation of model creation and validation was efficient, but it needed a cross-platform operation for modeling and mapping CGTs' suitable habitats. More importantly, MAXENT and DT can greatly reduce the area of field survey to 6% of the entire study area at the first stage, and thus saving both cost and labor. However, increasing background samples was not always beneficial to model accuracy. Especially when the ratio of background to target became too large, the species prediction did not correspond with real distribution, thereby reducing model accuracy. This ratio falling within the range from one to five was good for species distribution modeling, but may not be optimal. Next studies will attempt to incorporate predictor variables with species spectral information extracted from high spatial, spectral resolution remotely sensed data into predictive models so that newly developed models can be applied at a larger spatial scale.

KEY WORDS: Maximum Entropy (MAXENT), Sampling Design (SD), Decision Tree (DT), Discriminant Analysis (DA), Species Distribution Model.

1. INTRODUCTION

Species distribution modeling can provide a measure of a species' occupancy potential in areas not covered by biological surveys and consequently is becoming an indispensable tool to conservation planning (Guisan and Zimmermann, 2000). These models combine points of known occurrence with spatially continuous environmental layers to infer ecological requirements of a species, generally using a statistical algorithm to build model. A variety of modeling methods reviews of some of these techniques can be found in Guisan and Zimmermann (2000), which offered complete suggestions. We used maximum entropy (MAXENT), decision tree (DT) and discriminant analysis (DA) to build model since they have been successful in prediction in many studies (De'ath and Fabricius, 2000; Bourg *et al*., 2005; Ordo´n˜ ez *et al*., 2005; Phillps *et al*., 2006; Elith *et al*., 2006; Phillps *et al*., 2008; Riordan *et al*., 2009).

Despite the frequent use of distribution models, the number of occurrence data available for individual species from which to generate prediction is often quite limited. Because some species are difficult to take sample or available data are not yet recorded in coordinates (Graham *et al*., 2004). Hence, sampling issue has become an important aspect of ecological studies. But only a limited number of papers examined how to set up an optimal sampling strategy for species habitat suitability modeling (Guisan and Hirzel, 2002). Previous studies evaluated sample size effects on distribution model for only a few algorithms each and most did not test with data collected independently of the training data (Wisze *et al*., 2008). Hence, the goal of this study improved sampling design deficiency, and we could understand how background sample size affected model prediction. In addition we used independent data for model evaluation. Thus we could objectively evaluate the effects of sample size on the model performance.

Species ecological characteristics have been shown to affect model performance (*Elith et al*., 2006). Generally, models for species with broad geographic ranges and environmental tolerances tend to be less accurate than those for species with smaller geographic ranges and limited environmental tolerance (Elith *et al*., 2006). According to species characteristic, the target species chosen for this study was Chinese guger trees (*Schima superba*, CGTs), which are widespread with elevation ranging from 300 to 2,300 m in central Taiwan, is one of the fine broad-leaf tree species and good for fitment. CGTs have high water content and dense crown closure, and high dispersal ability; therefore, they have excellent fire resistance characteristics and can grow to form a fire line (Liu *et al*., 1994).

In sum, the study consisted of the following five steps. (1) *In-situ* CGT data were collected from the Tong-Feng watersheds by using a GPS. (2) GIS technique was used to overlay five environmental factors, including four topographic factors and vegetation index derived from SPOT-5 satellite images. (3) MAXENT, DT and DA were applied to build models for predicting the potential habitat of the species in the study area. (4) Comparison of three models using a common dataset. (5) Creation of twelve sample sizes to evaluate the relative influence of both species ecology and sample size on model performance. Further, we use Cohen's (1960) *kappa* value to assess those sampling designs for modeling individual species' distribution to determine what is the effect of sample size on the predictive ability of a model and if there exists a threshold in the ratio of background to target.

2. STUDY AREA

We chose the study area with rectangular shape, which encompasses the Huisun Forest Station and has the total area of 17,136 ha. The Huisun Forest Station is in central Taiwan, situated within 24˚2´–24˚5´ N latitude and 121˚3´–121˚7´ E longitude (Figure 1). The station is the property of National Chun-Hsing University, and study area ranges in elevation from 454 m to 3,419 m, and its climate is temperate and humid. Hence, the study area has nourished many different plant species and is a representative forest in central Taiwan. It comprises five watersheds, including two larger watersheds, Kuan-Dau at west and Tong-Feng at east. At the present stage, all of the CGT samples were collected from the Tong-Feng watershed by using a GPS.

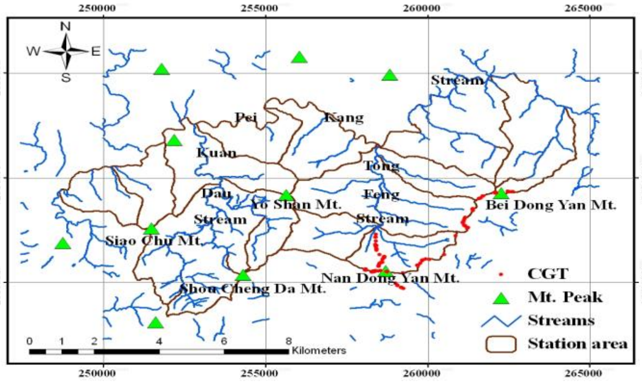

Figure 1 Location map of the study area

3. MATERIALS AND METHODS

3.1 Data collection

Digital elevation model (DEM) with grid size 5 × 5 m, orthophoto base maps (1:10,000), and nine-date SPOT images were collected. *In situ* CGT data were also acquired by using a GPS linked with a laser range. Two-date SPOT-5 images (07/10/2004 and 11/11/2005) were chosen because the two-date images (two out of nine) have the best quality with the least amount of clouds among the nine-date SPOT images.

3.2 Data processing

Slope and aspect data layers were generated from 5 × 5 m DEM by using ERDAS Imagine software module. The ridges and valleys in the study area were used together with DEM to derive terrain position layer. The main ridges and valleys over the study area were directly interpreted from the contour lines shown on the orthophoto base maps; these lines were then digitized to establish the data layer of main ridges and valleys by using ARC/INFO software for later use. The data layer of main ridges and valleys in a vector format was converted into a new data layer in a raster format by ERDAS Imagine software, and then combined with DEM to generate terrain position layer (Skidmore, 1990). Vegetation indices were derived from the two-date SPOT-5 images, one in autumn, the other in summer, by using Spatial Modeler of ERDAS Imagine. CGT samples obtained by a GPS were corrected by using post-processed differential correction and converted into ArcView shapefile format for later use.

3.3 Database building and sampling

The GIS database used in the study was constructed by using ERDAS Imagine software module Layer Stack to overlay elevation, slope, aspect, terrain position, vegetation index, and CGT data layers. All of these layers were rectified and projected onto TWD67 (Taiwan Datum) Transverse Mercator map projection. The CGT sample layer was overlaid with five data layers, and those pixels of the five layers lying at the same position with tree sample pixels were clipped out. Besides, *In-situ* 122 CGT samples were obtained from the Tong-Feng watershed and two-thirds of Tong-Feng samples (82) were used as training data for base model development, and the remaining, one-third of Tong-Feng samples (40), was used as test data for model validation.

We also tested several sampling designs for modeling individual species' distribution to determine: (1) what is the effect of sample size on the predictive ability of a model and (2) if there exists a threshold in the ratio of background to target. And then we created twelve sampling approaches were set up with a fixed target size and with different background size. The training background sets of S (100, 500, 1000, 2000, 3000…and 10000 sample points) were drawn completely at random across the study area (sampling was computed on 3 replicates). Composed of twelve sampling sample to test different background to target ratio for affecting species' distribution, which the background to target ratio are: 1, 6, 12, 24, 37, 49, 61, 73, 85, 98, 110, 122 respectively (see Table 3). Owing to the need of validation, the test background sets of S (50, 250, 500, 1000, 1500… and 5000 sample points) to keep similar ratio. MAXENT, DT and DA models were developed based on these sampling designs and used to compare the effect of different background sample size on the accuracy of predictive models.

3.4 Model development

The predictive models for predicting potential habitat of CGT were created using three approaches: (1) Maximum entropy, (2) Decision tree and (3) Discriminant analysis.

3.4.1 Maximum entropy

Maximum entropy (MAXENT) is to find the probability distribution of maximum entropy–that which is closet to uniform–subject to constraints imposed by the information available regarding the observed distribution of the species and environmental conditions across the study area. MAXENT represents features as binary functions known as contextual predicates in the form (1):

$$f_{cp,y'}(x,y)=\begin{cases}1 & if\ \ y=y'\ and\ cp(x)=true\\ 0 & otherwise\end{cases} \tag{1}$$

where *cp* is the contextual predicate that maps a pair of outcome y and context x to {true; false} (Le, 2004). The requirement for the model given such feature constraints are that expected value of *f* according to our model is equal the expected value of *f* with respect to the empirical distribution such as formula (2):

$$\sum_{x,y}\tilde{P}(x)p(y/x)f(x,y)=\tilde{P}(x,y)f(x,y) \tag{2}$$

where $\tilde{P}$ is an empirical distribution and *p* characterizes the model. These requirements, however, do not generally specify a unique model. A set of models will satisfy the given constraints and among these there will be one with the largest entropy, which according to MAXENT principle should be chosen as optimal. The entropy of a model, *H(p)*, is calculated using the following formula (3):

$$H(P)=-\sum_{i=1}^{N}P(X_i)log_2 P(X_i) \tag{3}$$

intuitively, models with high entropy are more uniform and therefore they assume less about the world. The maximum entropy model can be interpreted as the model that assumes only the knowledge that is represented by the features derived from the training data and nothing else (Le, 2004; Phillips *et al*., 2006; Phillips *et al*., 2008).

3.4.2 Decision tree

Decision Tree (DT) is machine learning and data mining technique, we used classification and regression tree (CART) algorithms in this study. Classification tree builds the rule by recursive binary partitioning into regions that are increasingly homogeneous with respect to the class variable. The homogeneous regions are called nodes. At each step in fitting a classification tree, an optimization is carried out to select a node, a predictor variable, and a cut-off or group of codes that result in the most homogenous subgroups for the data, as measured by the Gini index (Breiman *et al*., 1984). This criterion could set the optimum tree as a trade-off between goodness of fit on training data and size of the tree. Such a classification tree is said to be full grown, and the final regions are called terminal node. Terminal nodes are assigned a final outcome based on group membership of the majority of observations (De'ath and Fabricius, 2000). The lower branches of a fully grown classification tree model sampling error, so algorithms for pruning the lower branches on the basis of split-sample validation error have been developed (Breiman *et al*., 1984).

3.4.3 Discriminant analysis

Discriminant analysis (DA) is a technique, which discriminates among *k* classes (objects) based on a set of independent or predictor variables. The objectives of DA are to (1) find linear composites of *n* independent variables which maximize among-groups to within-groups variability; (2) test if the group centroids of the *k* dependent classes are different; (3) determine which of the *n* independent variables contribute significantly to class discrimination; and (4) assign unclassified or "new" observations to one of *k* classes (Johnson and Wichern, 2007). The variates for a discriminant analysis, also known as the discriminant function takes the following formula (4):

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \ldots + W_nX_{nk} \tag{4}$$

Where Z_{jk} = discriminant *Z* score of discriminant; function *j* for object (class) *k; a* = intercept; W_i = discriminant weight for independent variable *i;* X_{ik} = independent variable *i* object (class) *k*.

3.5 Model validation

Evaluation methods of the different samplings, we used split-sample validation. The first one (training dataset) is used to build model; the other one (test dataset) is used to validat the model. For each model, predict the response of the remaining data, and calculate the error matrix (De'ath and Fabricius, 2000). Some common statistical measurements include producer's accuracy (omission error), user's accuracy (commission error), overall accuracy and *kappa* value (Jensen, 2005; Sim and Wright, 2005; Lillesand *et al*., 2008).

4. RESULTS AND DISCUSSION

Owing to very large amount of calculation, it was necessary to reduce the dimension of a model to improve its building efficiency. All three modules can evaluate the relative importance of predictor variables in the models for predicting the potential habitat of CGTs. The evaluation showed that elevation, slope, and terrain position were the variables having the first three highest relative importance. Hence, we used these three predictor variables to build predictive models. Finally, in order to get more reliable results, the study only used "test sample" to assess the accuracy of model.

As shown table1, both the overall accuracy and *kappa* values with DT (91%, 0.63) was only slightly better than that of MAXENT (90%, 0.61) and accuracies of the two models were much better than that of DA (86%, 0.56), respectively. According to the PA and UA with three models in the test case, for both MAXENT and DT, the PA percents of non-habitat (91%, 96%) were close to the UA percents of non-habitat (97%, 96%), whereas the PA percents of habitat (82%, 62%) were slightly greater than the UA percents of habitat (56%, 62%). This indicated that the model modeling result meet with real condition. As shown in the results of DA, the PA percents of non-habitat in the test case (83%) were fewer than the UA percents of non-habitat (99%), and the PA percents of habitat in the test case (94%) were obviously greater than the UA percents of habitat (36%), thereby raising the PA percent of habitat by trading off the PA percent of non-habitat. It made model relaxed restriction in habitat prediction and mitigated omission problem, but caused non-habitat to be erroneously assigned to habitat area (commission error). This is a major shortcoming of DA, thereby decreasing non-habitat accuracy and overall accuracy. Table 2 shows the distribution statistics of CGT potential habitat predicted by three models. Clearly, three models greatly reduced the area of field survey to less than 6% of the entire study area at the first stage, and thus saving both cost and labor.

Table 1 The accuracies of three models for predicting the potential habitat of CGTs

Model	OA (%)	*Kappa*	Habitat PA (%)	Habitat UA (%)	Non-Habitat PA (%)	Non-Habitat UA (%)
MAXENT	90	0.61	82	56	91	97
DT	91	0.63	62	62	96	96
DA	86	0.56	94	36	83	99

* The results are computed on 3 replications; OA = overall accuracy; PA = producer's accuracy; UA = user's accuracy.

Table 2 Distribution statistics of CGT potential habitat predicted by using three approaches

Class	MAXENT		DT		DA	
	Area (ha)	%	Area (ha)	%	Area (ha)	%
Habitat	874.11	5	1,046.39	6	1,058.15	6
Non-habitat	16,261.89	95	16,0.89.61	94	16,077.85	94
Sum	17,136.00	100	17,136.00	100	17,136.00	100

Table 3 shows the comparison of the predictive accuracy resulting from the different sample size. On the whole, results showed that the *kappa* values of all models declined largely as background sample size (BSZ) rose from 250 to 500 (see Figure 2a), which means BSZ obviously affected models performance. DT declined slowly with the increase in BSZ, whereas MAXENT and DA dropped sharply with the increase in BSZ. DT had better reliability to resist change in B/T ratio and tease apart complex relationship. Hence, model performance with DT was better than MAXENT and DA, and the effect of BSZ decreased the accuracy of model predictions of habitat suitability.

Table 3 Comparison of the prediction accuracy resulting from different background sample size

BSZ	B/T ratio	MAXENT *kappa*		DT *kappa*		DA *kappa*		MAXENT OA (%)		DT OA (%)		DA OA (%)	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
50	1	0.650	0.000	0.743	0.031	0.710	0.035	84	0.000	87	1.735	85	1.963
250	**6**	**0.610**	**0.030**	**0.633**	**0.055**	**0.563**	**0.046**	**90**	**0.577**	**91**	**1.514**	**86**	**2.021**
500	13	0.443	0.031	0.597	0.047	0.417	0.035	90	1.000	95	0.557	85	1.504
1000	25	0.303	0.035	0.520	0.046	0.227	0.012	89	1.000	97	0.252	81	1.015
1500	38	0.243	0.012	0.457	0.074	0.173	0.015	89	0.577	98	0.473	82	2.060
2000	50	0.197	0.012	0.477	0.121	0.143	0.006	89	0.000	99	0.200	83	0.656
2500	63	0.170	0.010	0.437	0.076	0.107	0.006	89	0.577	99	0.200	82	0.586
3000	75	0.143	0.015	0.427	0.080	0.093	0.006	89	1.000	99	0.100	83	0.850
3500	88	0.130	0.000	0.320	0.201	0.083	0.006	89	0.000	99	0.100	83	0.153
4000	100	0.117	0.015	0.300	0.078	0.070	0.000	89	1.000	99	0.000	82	0.289
4500	113	0.107	0.006	0.277	0.087	0.067	0.006	89	0.577	99	0.058	83	0.577
5000	125	0.100	0.010	0.107	0.095	0.060	0.000	89	0.577	99	0.058	83	0.551

*Means and S. D. are computed on 3 replications; BSZ = background sample size; B/T ratio = background to target ratio; OA = overall accuracy.

Table 4 summarizes producer's and user's accuracy in different sample size. Results showed that non-habitat (background) of three models had higher producer's accuracy and were more representative as the background to target ratio increased. In other words, three models traded off the accuracy of habitat (target) to improve non-habitat prediction. This reduced omission errors in the background prediction, but increased omission error in the target prediction. Thus it was hard to balance commission and omission errors between target and background distribution.

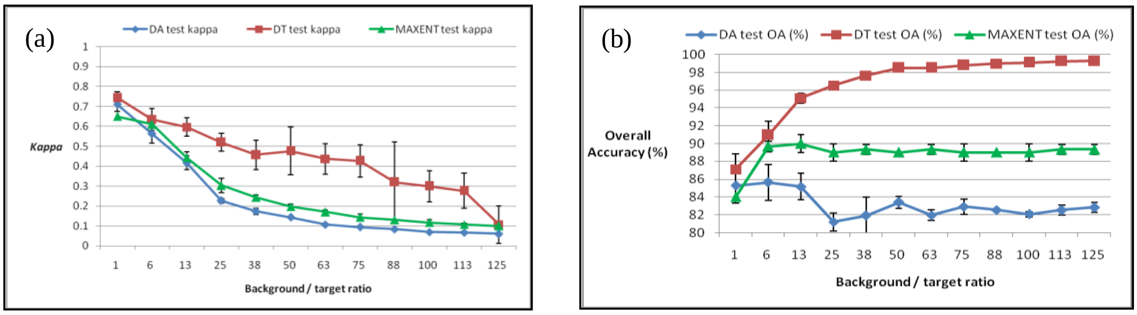

Figure 2 (a) Comparison of *kappa* values of three models (± standard deviation);
(b) Comparison of overall accuracy of three models (± standard deviation)

Table 4 Comparison of the producer's and user's accuracy in different sample size

BSZ	B/T ratio	MAXENT Habitat		MAXENT Non-Habitat		DT Habitat		DT Non-Habitat		DA Habitat		DA Non-Habitat	
		PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
50	1	72	85	92	83	96	79	81	96	97	76	76	97
250	**6**	**82**	**56**	**91**	**97**	**62**	**62**	**96**	**96**	**94**	**36**	**83**	**99**
500	13	85	34	90	99	40	70	99	98	96	32	84	100
1000	25	84	21	89	99	54	53	98	98	97	16	81	100
1500	38	87	16	89	100	43	54	99	99	96	12	82	100
2000	50	90	13	89	100	39	67	100	99	93	10	83	100
2500	63	87	11	89	100	40	50	99	99	93	7	82	100
3000	75	87	9	89	100	44	46	100	99	93	6	83	100
3500	88	87	8	89	100	25	58	100	99	93	5	82	100
4000	100	92	7	89	100	21	58	100	99	93	5	82	100
4500	113	87	7	89	100	18	68	100	99	93	4	82	100
5000	125	90	6	89	100	6	52	100	99	92	4	83	100

*PA and UA are computed on 3 replications; PA = producer's accuracy; UA = user's accuracy;
BSZ = background sample size; B/T ratio = background to target ratio.

This interesting finding inspired research team to further understand if there exists a threshold in the ratio of background to target. Table 5 indicated how the model performance in different sample size. The results showed that three models have different model performance at the same ratio. As the ratio was larger than 25, DA model had a poor agreement but MAXENT and DT kept moderate agreement. This means MAXENT and DT had a better reliability than DA as the ratio changed. The ratio of background to target falling within the range from one to six is good for species distribution modeling, but may not be optimal. There may have a critical ratio with the model. Especially when the ratio of background to target became too large, the species prediction did not correspond with real distribution, thereby reducing model accuracy. This indicates that more background samples are not always better for model accuracy.

Table 5 Model performance in different sample sizes

Strength of agreement	*Kappa* value	MAXENT B/T ratio	DT B/T ratio	DA B/T ratio
Very good	0.8–1.0	—	—	—
Good	0.6–0.8	< 6	< 6	< 1
Moderate	0.4–0.6	6 < x < 13	6 < x < 75	1 < x <13
Fair	0.2–0.4	13 < x < 38	75 < x < 113	13 < x < 25
Poor	0.0–0.2	x > 38	x > 113	x > 25

* B/T ratio = background to target ratio.

5. CONCLUSIONS

The study developed MAXENT, DT and DA models that related the known CGT sites to habitat characteristics (topographic variables) and spatially extrapolated CGT's potential sites in the study area. DT was slightly better than MAXENT, and the two models were much better than DA. Implementation of model creation and validation was efficient, but it needed a cross-platform operation for modeling and mapping CGTs' suitable habitats. Because MAXENT and DT greatly reduced the area of field survey to fewer than 6% of the entire study area, they were better suited for predicting the tree's potential habitat.

Besides, we also evaluate effects of sample size on the performance of species distribution modeling. Preliminary results indicated that DT had a better stability than MAXENT and DA as the ratio changed. Furthermore, the increase in background samples was not always good for model accuracy. The species prediction did not correspond with real distribution, thereby reducing model accuracy. Again, The ratio of background to target falling within the range from one to six is good for species distribution modeling, but may not be optimal. This point will need to be confirmed in a follow-up study.

Sampling is an expensive, time-consuming task, our results should encourage further, though cautious, use of predictions based on background/target sample ratio. Further study will be needed to evaluate the effect of the different target sample size, and design different spatial scales for model performance. More importantly, we shall attempt to use predictor variables involving species spectral information extracted from high spatial, spectral resolution remotely sensed data, especially hyperspectral image data, for building a model so that it can improve model prediction.

6. REFERENCES

Bourg, N. A., W. J. Mcshea and D. E. Gill, 2005. Putting a CART before search: successful habitat prediction for a rare forest herb. *Ecology*, 86 (10), pp. 2793-2804.

Breiman, L., J. H. Friedman, R. A. Olshen and C. J. Stone, 1984. *Classification and regression trees.* Monterey, Calif., U.S.A.: Wadsworth, Inc.

Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20 (1): 37-46.

Congalton, R. G. and M. Story, 1986. Accuracy assessment: a user's perspective, *Photo-grammetric Engineering and Remote Sensing*, 55 (9): 1303-1309.

Congalton, R. G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37: 35-46.

De'ath, G. and K. E. Fabricius, 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81 (11), pp. 3178-3192.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. Overton, M. Mcc., A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz and N. E. Zimmermann, 2006. Novel methods improve prediction of species' distribution from occurrence data. *Ecography*, 29: 129-151.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz and A. T. Peterson, 2004. New developments in museum-based informatics and application in biodiversity analysis. *Trends Ecol*. Evol, 19:497-503.

Guisan, A. and N. E., Zimmermann, 2000. Predictive habitat distribution models in ecology. *Ecological Modelling,* 135: 147-186.

Hirzel, A. and A. Guisan, 2002. Which is the optimal sampling strategy for habitat suitability modeling. *Ecological Modelling*. 157: 331-341.

Jensen, J. R., 2005. *Introductory Digital Image Processing–A Remote Sensing Perspective, 3rd ed*., Pearson Education, Inc, New Jersey.

Johnson, R. A. and D. W., Wichern, 2007. *Applied Multivariate Statistical Analysis*, Prentice-Hall, Inc.

Le, Z., 2004. Maximum entropy modeling toolkit for python and c++ (version 20041229). Natural Language Processing Lab, Northeastern University, China.

Lillesand, T. M., R. W. Kiefer and J. W., Chipman, 2008. *Remote Sensing and Image Interpretation. 5th Edition.* John Wiley & Sons. Inc., New York.

Liu, Y. C., F. Y. Lu and C. H., Ou, 1994. *Trees of Taiwan. Taichung, Taiwan: College of Agriculture*, National Chung-Shing University, pp. 440.

Ordo´n˜ ez, C., J. Mari´a mati´as and A. Manuel felici´simo, 2005. Learning Machines Applied to Potential Forest Distribution, *Environmental Management,* 35 (1), 109-120.

Phillips S. J., R. P. Anderson and R. E., Schapire, 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190: 231-259.

Phillips, S. J. and M., Dudik, 2008, Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31: 161-175.

Riordan, E. C. and P. W., Rundel, 2009. Modelling the distribution of a threatened habitat: the california sage scrub. *Biogeography*, 36 (11), 2176-2188.

Sim, J. and C. C., Wright, 2005. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85: 257-268.

Skidmore, A. K., 1990. Terrain position as mapped from a grided digital elevation model. *Geographical Information Systems*. 4 (1): 33-49.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan A. and NCEAS working group, 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14: 763-773.