

## TWO-WAY SPATIAL EXTRAPOLATION AND VALIDATION ON ECOLOGICAL PATTERNS OF *ELAEOCARPUS JAPONICUS* BETWEEN MAIN WATERSHEDS IN HUISUN OF CENTRAL TAIWAN

S. Y. Su<sup>a</sup>, N. J. Lo<sup>b</sup>, W. I. Chang<sup>c</sup>, K. Y. Huang<sup>d,\*</sup>

<sup>a</sup>Graduate student, Dept. of Forestry, Chung-Hsing University, Taiwan, E-mail: j82831079@hotmail.com

<sup>b</sup>Specialist, EPMO, Chung-Hsing University, Taiwan, E-mail: njl@dragon.nchu.edu.tw

<sup>c</sup> Director, Hsinchu Forest District Office, Taiwan, E-mail: weii@forest.gov.tw

<sup>d</sup>Professor, Dept. of Forestry, Chung-Hsing University, Taiwan, E-mail: kyhuang@dragon.nchu.edu.tw  
250 Kuo-Kuang Road, Taichung 402, Taiwan, R. O. C.

Commissions: VIII/6

**KEY WORDS:** Forestry, GIS, Modeling, Pattern, SPOT, Prediction, Accuracy, Performance.

### ABSTRACT:

Spatial extrapolation has become a *sine qua non* and an *ad hoc* major research focus in applied ecology in the latter half 20<sup>th</sup> century. Progressive innovations in data acquisition and processing technologies over the last few decades, especially in the fields of 3S (RS, GIS and GPS) and statistical modeling method, have greatly enhanced ecologists' capacity to face the challenge by enabling them to describe patterns in nature over larger spatial scales and a greater level of details than ever before. *Elaeocarpus japonicus* (Japanese *Elaeocarpus* tree, JET) was selected for applying in the concurrent developed technology, such as ecological distribution modeling and ecological extrapolation. The GPS-located JET samples were introduced in a GIS for overlaying with five environmental layers (elevation, slope, aspect, terrain position and vegetation index derived from two-date SPOT-5 images) for ecological information extraction and model building. We created three sampling designs (SD), Tong-Feng samples for SD1, Kuan-Dau samples for SD2, and the merge of the two former datasets for SD3, according to watersheds, and the three SDs were used individually to test the extrapolation ability of predictive models. The results of the two-way extrapolation indicated it is hard to extend the predicted distribution patterns through different watersheds. The main reasons resulting in this outcome were the difference in microclimate and micro-terrain between these two watersheds. Consequently, the models built with SD3 were the more robust. The information of vegetation index in this study poorly improved the models, so we will adopt the hyperspectral data to overcome the shortage of the SPOT-5 images.

### 1. INTRODUCTION

To plant right tree at right place is the most critical concept in plantation project and forest management. Different tree species need different habitat conditions, which are as the same as the concept of Odum (1997) proposed, ecological niche. Different environmental conditions result in different tree species composition. The niche breadth of each species is not the same, but equally means the species with wider niche breadth could adapt border environmental conditions. Presence or absence of a tree species will mainly decide by the interaction of numerous environmental factors, those usually contain direct factors and indirect factors. Generally, direct factors are referred climate, soil and biotic factors, as well as indirect factors are composted with topographic factors (including elevation, slope, aspect and terrain position). To obtain broad-extent and high accuracy data of direct factors is really difficult because of that the field data collecting stations are fragmentary result in introducing serious error when performing spatial interpolation (Prudhomme and Reed, 1999; Marquezine *et al.*, 2003). In the contrast, by introducing the digital elevation model (DEM) to a geographic information system (GIS), we can derive the high accuracy and broad-extent data of indirect factors, such as elevation, slope, aspect, and terrain position.

Nowadays, ecologists especially value the ecological modelling techniques. The specialists can apply the 3S technology to

extract the point data and related data for ecological model building, and the potential distribution map can be produced. According to predicted accurate distribution maps, ecologists can reduce the field survey tasks to save labor and fund spending. The predicted map also can be used to evaluate the ability of model extrapolation, and help the ecologist to evaluate the area inaccessible but we are interested in.

It is extraordinarily necessary to acquire the spatial information for parametric or non-parametric algorithm to build species distribution models. We can compare the distribution maps of different algorithms to realize the performance of different models. Felicísimo (2004) applied discriminant analysis (DA) and decision tree (DT) along with GIS to predict the suitable habitat of tree species. Maximum entropy (MAXENT), with increasing application in ecology field, is a promising tool in many domains. MAXENT doesn't suffer the statistical assumption and limitation, and it can use only fewer point data and incomplete information to build robust predictions (Phillips *et al.*, 2006; Kumer and Stohlgren, 2009). The advantages of MAXENT modelling are very indispensable in ecological related field because it is unusual to collect abundant and representative point data in field survey. Maximum likelihood (ML) algorithm is commonly used in multispectral image classification (Mu and Shao, 2002; McIver and Friedl, 2002). Carpenter (1993) and Hernandez (2006) used DOMAIN to modeling species potential distribution. Carpenter (1993) also proposed DOMAIN is variable sensitivity, and perform well

---

\* Corresponding author.

with limited site data.

The target tree species of this study is *Elaeocarpus japonicas* (Japanese *Elaeocarpus* tree, JET), a kind of evergreen tree species. It widely spread in whole Taiwan from low upland to mountainous areas with elevation 2200 m above sea level. JET is also founded in Japan and China. JET is a kind of dominant tree species in the Huisun forest station. It is usually founded on the ridge with thinner soil layer, direct sunlight and water stress. JET is a kind of pioneer tree species in second succession, and therefore it plays a necessary role in ecosystem.

We aimed at applying 3S (GIS, GPS and RS) technology to derive elevation, slope, aspect and terrain position from DEM and vegetation index (derived from the two-date SPOT-5 images), and using these five environmental layers to build predictive models. In this study, we adopted five methods (DA, DT, MAXENT, ML and DOMAIN) and three sampling designs (SD) to build “Tong-Feng (SD1)” model, “Kuan-Dau (SD2)” model and “two watersheds (SD3)”, eventually we totally built 15 models. The models’ reliability and performance were evaluated, and used as the criteria of model comparison.

## 2. STUDY AREA

We chose the study area with rectangular shape, which covers 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. The station is the property of National Chung-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. All of the JET samples were collected from the two watersheds by using a GPS (Figure 1.).

## 3. METHODS AND MATERIAL

### 3.1 Data Collection

The collected data contained DEM with 5 m × 5 m resolution, orthophoto maps with 1: 10,000 scale and two-date SPOT-5 images taken in 2004/07/10 and 2005/11/11. The JET samples were acquired by field survey with Trimble PRO XR series GPS system. Furthermore, an expandable antenna rod with 5m in length and a laser ranging were adopted with GPS for enhancing the capacity of the system. All of the JET point data were field-collected from Tong-Feng and Kuan-Dau watersheds (© SPOT Image Copyright 2004 and 2005, CSRSR, NCU).

### 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). The equation is expressed as follows. 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. JET samples obtained by a GPS were corrected by using post-processed differential correction and converted into ArcView shapefile format for later use.

$$P_{ij} = PV / (PV + PR)$$

Where PV = the Euclidean distance between a certain pixel P and the nearest valley pixel;

PR = the Euclidean distance between a certain pixel P and the nearest ridge pixel;

When  $P_{ij} = 0.0$ , it is referred to valley;  $P_{ij} = 1.0$ , it is referred to ridge. The  $P_{ij}$  from 0.0 to 1.0 is partitioned into eight equal intervals.

The change in water content and pigment composition in plant owing to the season or stress can be detected by using multi-date imagery. These two phenomena could result in changing plant’s spectral reflectance of different bands in multi-band image (Jensen, 2005). The concept of the vegetation index adopted in this study is explained in Hoffer (1978). The following equation is used to derive the vegetation index data layer.

$$\text{Vegetation Index} = \frac{NIR_{\text{autumn}} - MIR_{\text{autumn}}}{NIR_{\text{summer}} - MIR_{\text{summer}}}$$

Where  $NIR_{\text{summer/autumn}}$  is the reflectance of near infrared band during summer and autumn, and the reflectance of middle infrared is denoted as  $MIR_{\text{summer/autumn}}$ . The output value is scaled in 8-bits data type.

### 3.3 Overlaying the Environmental Layers

The layers of elevation, slope, aspect, terrain position, vegetation index, and JET sample data were overlaid by ERDAS Imagine software. We used the function “AOI (area of interest)” in ERDAS imagine software to clip the concurrent environment factor value of JET locations. These clipped-out data were used as independent variable for building predictive model.

### 3.4 Target and Background Samples

Target sample is the GPS-located JET point sample and the concurrent environment factor value. The ratio of background to target we adopted was followed the criteria Sperduto and Congalton (1996) proposed that the ratio should be more than 3. The sampling strategy is randomly selected following Pereira and Itami (1991) suggested avoiding spatial autocorrelation.

### 3.5 Sampling Designs and Model Building

We designed three sampling designs (SD) for the comparison of model reliability, “Tong-Feng (SD1)”, “Kuan-Dau (SD2)” and “merged samples of two watersheds (SD3)”. SD1 had 104 individual JET samples, and SD2 had 80. SD3 had all of the 184 JET samples. For each of these three SDs, the dataset was split into two subsets, 2/3 and 1/3 of all, used for

split-sample evaluation. We used 2/3 of all as the training dataset for modeling, and used the remaining 1/3 as test dataset for model evaluation. To build the predictive models for each SD, we used five methods, DA, DT, MAXENT, ML, DOMAIN.

**3.5.1 Discriminant Analysis (DA):** DA is an algorithm that tries to find the most robust boundary within variables for group participation. A grouping variable and few discriminant variables are implemented in DA to establish the discriminant function to participate the original samples into few categories (Lowell, 1991). The following equation is the typical structure of discriminant function.

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_i X_i + \dots + b_k X_k$$

Where Y = the grouping variable  
X<sub>k</sub> = discriminant variables

**3.5.2 Decision Tree (DT):** DT (also called Classification and Regression Trees, CART) is a non-parametric classification algorithm for data mining with both classifying and predicting capability. DT could build classified rules from observations or some experiences (Guisan and Zimmermann, 2000). Decision tree algorithm sequentially partitions the dataset with some important predictors in order to maximize differences on a dependent variable. The decision pathways originate from a starting node (root) that contains all observations, then classify step by step into binary subsets based on the important predictors, and so on. Finally, it will end at multiple nodes containing unique subsets of observations. Terminal nodes are assigned a final outcome based on group membership of the majority of observations (De'ath and Fabricius, 2000; Bourg *et al.*, 2005; O'Brien *et al.*, 2005).

**3.5.3 Maximum Entropy (MAXENT):** One of novel methods used in ecology field is MAXENT. It can build robust and stable prediction models by applying incomplete information and small sample size (Kumar and Stohlgren, 2009; Phillips *et al.*, 2006). Entropy means the uniform condition in thermodynamic. The axiom of MAXENT is to searching the maximum entropy of species distribution under limited conditions. When reaching the maximum entropy, the species distribution is similar to the natural condition.

$$P(x) = \exp \left[ \sum_{n=1}^p \lambda_n \frac{f_n(x) - \min_n}{\max_n - \min_n} - \text{LinearPredictorNormalizer} \right] / Z$$

Where  $\frac{f_n(x) - \min_n}{\max_n - \min_n}$ : hinge feature;

$\lambda_n$ : weight coefficient

Linear Predictor Normalizer: a constant for numerical stability

Z: a scaling constant that ensures that P sums to 1 over all grid cells

The MAXENT software is free and online available (<http://www.cs.princeton.edu/~schapire/MAXENT>).

**3.5.4 Maximum Likelihood (ML):** ML is a widely used method in classification algorithm (Wu and Shao, 2002; McIver and Fridel, 2002). ML algorithm is based on the probability to assign the pixel to one of the predefined k class with

maximum likelihood (Atkinson and Lewis, 2000; Lo and Yeung, 2002).

**3.5.5 DOMAIN:** This method assigns a classification value to the candidate area according to a point-to-point similarity metric, and also base on this criterion to find the area where environment is similar with the sample data. Sum of the standardized distance between two points of each environment variable is used to quantify the similarity. And equalization of variable contribution is achieved by standardizing the environment variables. The classification value of each pixel in the study area is decided by the maximum similarity between each pixel a set of data points. It is necessary to set a similarity threshold to converge the predicted distribution pattern (Carpebter *et al.*, 1993; Hernandez *et al.*, 2006). In this study, the similarity threshold was set in 0.97, in which the kappa coefficient was reasonable.

### 3.6 Model Evaluation and Assessment

The test and training data sets are used to evaluate the model performance and reliability. In each data set, the evaluation indices contain producer's accuracy (PA), user's accuracy (UA) and overall accuracy (OA). Kappa agreement coefficient is extremely important to assess the agreement between predicted map and reference test dataset. The kappa coefficient compares the marginal and diagonal value in matrix fairly due to the calculation containing not only PA and UA but also OA (Referred and Fitzpatrick-Lins, 1986; Congalton, 1991; Paine and Kiser, 2003). Furthermore, in the model evaluation of "Tong-Feng (SD1)" model, the test dataset of "Kuan-Dau" JET samples were used as independent samples to evaluate the ability of extrapolating predicting model through space. Again, we treat the same process evaluating "Kuan-Dau (SD2)" model with the test dataset of "Tong-Feng" samples. In the evaluation in "merged samples of two watersheds (SD3)" model, we split the test set into two subsets according to the watersheds' boundary. The two subsets of SD3's test sample were used as two independent sample sets to demonstrate that the model performance was still reasonable when using these two subsets solely.

## 4. RESULTS AND DISCUSSION

We calculated the statistics of five environmental factors corresponding to the entire study area and all of the JET samples in two watersheds and compared the difference in statistics between them, as shown in Table 1. The elevation range of the "Tong-Feng" and "Kuan-Dau" JET samples (1,122–2,027 m and 1,076–1,559, respectively) were within the nature distribution range, from low elevation to 2,200 m above sea level. The means of slope statistics in "Tong-Feng" and "Kuan-Dou" samples were 22° and 27°, respectively. The mean slope of all JET samples is obviously lower than that of the entire study area; consequently, this result is due to the nature behavior of JET. JETs prefer to grow on the flat areas beside ridges with unclosing canopy structure, where they are illuminated by abundant solar radiation. This behavior could be demonstrated by the mean of terrain position statistics.

The predicted distribution maps of SD3 used to represent overall prediction showed in Figure 2. At the earlier stage of this result, each method eliminated vegetation index from the effective variables because the contribution of vegetation index in model performance is less than 1 percent. The most important effective variables were slope, aspect and terrain

position, followed by aspect. So, we used these four effective variables to build and evaluate each model in three SDs (Table 2). Overall, in three SDs, the best method for model building was DOMAIN ( $\kappa = 0.83-0.87$ ), followed by DT (0.72–0.80), MAXENT (0.63–0.68), ML (0.51–0.54) and DA (0.23–0.47) in order. SDs didn't affect the model performance. It is clear to realize that DOMAIN, DT and MAXENT is efficient in converging the predicted distribution pattern. Convergence of prediction helps ecologist to reduce the consuming of field survey.

We used the independent samples aforementioned to evaluate the extrapolating ability, and the result is shown in Table 1. The result reveals that when evaluating model performance by independent samples, the  $\kappa$  value of each model decreased sharply in both SD1 and SD2. In contrast, the  $\kappa$  value of each model in SD3 declined slightly. The result in Table 3 indicated that the model prediction of SD1 and SD2 could not extend through watershed.

### 5. CONCLUSIONS

The methods we adopted papered in different ratability, but the performance efficiency of these methods should be at the same grade. DOMAIN had the highest prediction accuracy, and the agreement of model performance between training dataset and test dataset in DOMAIN was also the best. Although the

performance of DT and MAXENT were slightly lower than DOMAIN, but this three method had the same level of kappa coefficient in this study. Hence, we suggest that DOMAIN, DT and MAXENT a high potential in similar research and ecological application.

The evaluation of SDs showed that it is hard to extend the distribution pattern through spatial (e.g. watersheds and mountain). The phenomenon is strongly established by the result of two-way extrapolation we designed. The result indicated that it is hard to extend the spatial patterns of JETs from one watershed to another and vice versa. By comparing the microclimate and micro-terrain of the two watersheds, the humidity and sunlight affected by micro-terrain of these two watersheds are remarkably different. Consequently, the models merely based on topographic variables performed poorly on two-way spatial extrapolation between these two watersheds. Not surprisingly, the  $\kappa$  values of predictive models developed from the merged samples of the two watersheds in SD3 just declined slightly. The results suggested that the vegetation indices derived from SPOT-5 images could not improve model accuracy for a widely distributed tree species due to the limitations of spectral resolution and spatial resolution with SPOT-5 imagery. Follow-up studies will attempt to extract spectral information associated with species from hyperspectral data and LIDAR DEM and use it as variable for model development so that models are applicable on a broader spatial scale.

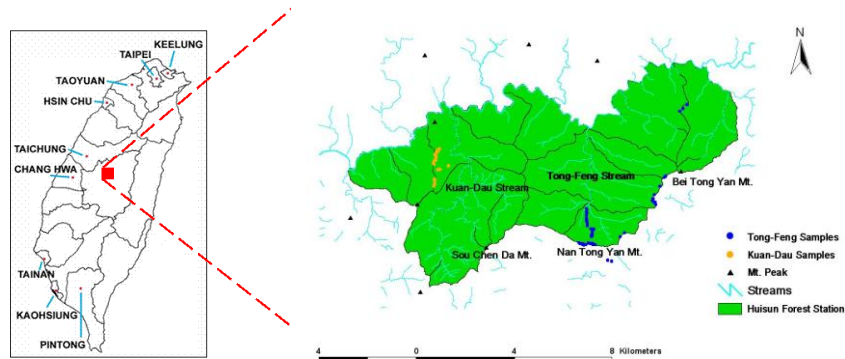


Figure 1. Location map of study area.

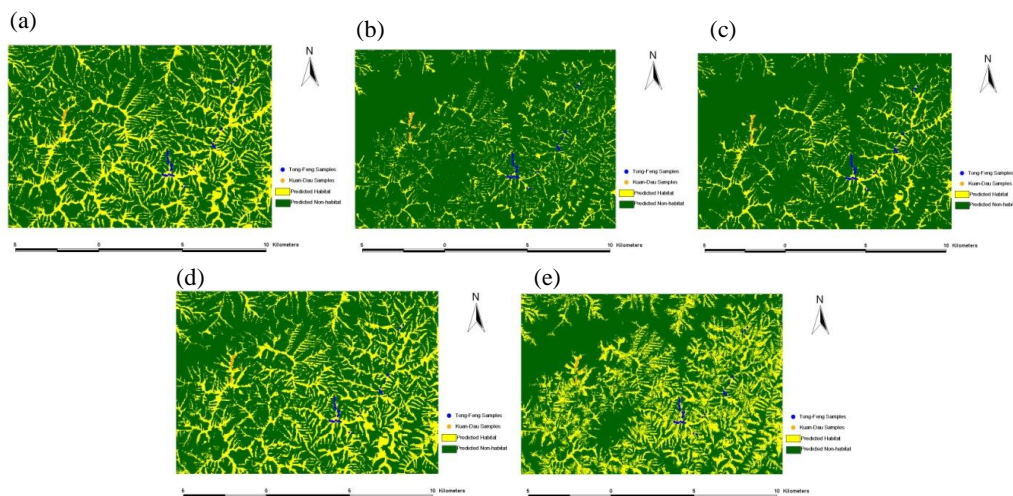


Figure 2. Predicted distribution maps of SD3 included (a) DA, (b) DT, (c) MAXENT, (d) ML, and (e) DOMAIN.

Statistics	Study Area					“Tong-Feng” Samples					“Kuan-Dau” Samples				
	Elevation (m)	Slope (°)	Aspect (°)	TP	VI	Elevation (m)	Slope (°)	Aspect (°)	TP	VI	Elevation (m)	Slope (°)	Aspect (°)	TP	VI
Mean	1314	34	—	5	24	1712	22	—	7	24	1259	27	—	7	30
Mode	1239	37	127	6	22	1227	22	247	7	22	1076	22	7	7	21
Max.	2418	89	361	8	119	2027	46	358	8	51	1559	38	355	8	51
Min.	454	0	0	1	0	1122	2	2	2	20	1076	16	7	4	21

VI: Vegetation Index. TP: Terrain position.

Table 1. The statics of environmental factors with study area and the samples of each watershed.

Method and Dataset		SD1		SD2		SD3	
		OA (%)	<i>kappa</i>	OA (%)	<i>kappa</i>	OA (%)	<i>kappa</i>
DA	Training	81	0.44	72	0.22	76	0.44
	Test	88	0.49	73	0.23	79	0.49
	Average	85	0.47	73	0.23	78	0.47
DT	Training	96	0.79	97	0.81	92	0.76
	Test	93	0.64	96	0.79	91	0.69
	Average	95	0.72	97	0.80	92	0.73
MAXENT	Training	90	0.62	94	0.68	90	0.70
	Test	91	0.63	93	0.66	89	0.66
	Average	91	0.63	94	0.67	90	0.68
ML	Training	84	0.49	88	0.56	79	0.50
	Test	86	0.52	88	0.51	82	0.53
	Average	85	0.51	88	0.54	81	0.52
DOMAIN	Training	97	0.86	98	0.91	95	0.85
	Test	96	0.79	97	0.82	94	0.80
	Average	97	0.83	98	0.87	95	0.83

OA: Overall Accuracy. SD1: “Tong-Feng” samples. SD2: “Kuan-Dau” samples. SD3: “Merged samples of two watersheds”.

Table 2. The model evaluation result of each method in the three SDs.

Method	SD1: Using KDS as IS		SD2: Using TFS as IS		SD-3: Using TFS as IS		SD3: Using KDS as IS	
	OA (%)	<i>kappa</i>	OA (%)	<i>kappa</i>	OA (%)	<i>kappa</i>	OA (%)	<i>kappa</i>
DA	75	0.32	77	0.51	80	0.43	78	0.28
DT	77	0.13	70	-0.00	92	0.63	93	0.64
MAXENT	74	0.08	74	0.22	91	0.61	91	0.53
ML	73	0.21	67	0.07	81	0.43	81	0.37
DOMAIN	78	0.18	72	0.10	98	0.91	97	0.82

OA: Overall Accuracy. SD1: “Tong-Feng” model. SD2: “Kuan-Dau” model. SD3: “Merged samples of two watersheds” model. TFS: Test dataset of Tong-Feng Watershed. KDS: Test dataset in Kuan-Dau watershed. IS: Independent Samples.

Table 3. The evaluation results of extrapolation ability.

## 6. REFERENCE

Atkinson, P. M. and P. Lewis, 2000. Geostatistical classification for remote sensing: an introduction. *Computers & Geosciences*, 26, pp.361-371.

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.

Carpenter, G., A. N. Gillison and J. Winter, 1993. Domain: a flexible modelling procedure for mapping potential distribution of plants and animal. *Biodiversity and Conservation*, 2, pp. 667-680.

Congalton, R. G., 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, 37, 99, pp. 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.

Felicísimo, A. M. and A. Gómez-Muñoz, 2004. *GIS and predictive modelling: a comparison of methods applied to forestal management and decision-making*. *Proceedings of GIS Research*. University of East Anglia, UK, pp. 143-144.

Guisan, A. and Zimmermann, N. E., 2000. Predictive habitat distribution models in ecology. *Ecological Modeling*, 135, pp.

147-186.

Hernandez, A. P., C. H. Graham, L. L. Master, and D. L. Albert, 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29, pp. 773-785.

Hoffer, R. M., 1978. Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: Swain, P. H. and S. M. Davis (Eds.), *Remote Sensing: The Quantitative Approach*, McGraw-Hill, Inc., New York, pp. 227-289.

Jensen, J. R. 2005. *Introductory digital image processing: a remote sensing perspective (3rd ed.)*. New Jersey: Pearson Education, Inc.

Kumar, S. and T. J. Stohlgren, 2009. MAXENT modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. *Journal of Ecology and Natural Environment*, 1(4), pp. 94-98.

Lo, C. P. and A. K. Yeung, 2002. *Concept and Techniques of Geographic Information Systems*. Upper Saddle River, New Jersey, pp.429.

Lowell, K., 1991. Utilizing discriminant function analysis with a geographical information system to model ecological succession spatially. *International Journal of Geographical Information System*, 5 (2), pp. 175-191.

Marquinez, J., J. Lastra and P. Garcia, 2003. Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology*, 270, pp. 1-11.

McIver, D. K. and M. A. Friedl, 2002. Using prior probabilities decision-tree classification of remotely sensed data. *Remote Sensing and Environment*, 81, pp.253-261.

Odum, E. P., 1971. *Fundamentals of ecology, (3rd ed.)*, W. B. Saunders Co. Philadelphia.

O'Brien, C. S., S. S. Rosenstock, J. J. Herver, J. L. Bright and S. R. Boe, 2005. Landscape – level models of potential habitat for Sonoran pronghorn. *Wildlife Society Bulletin*, 33 (1), pp. 24-34.

Paine, D. P. and J. D. Kiser, 2003. *Aerial Photography and Image Interpretation, (2nd ed.)*. John Wiley & Sons, New York.

Pereira, J. M. C., and R. M. Itami, 1991. GIS-based habitat modeling using logistic multiple regression: A study of the Mt. Graham red squirrel. *Photogrammetric Engineering and Remote Sensing*, 57 (11), pp. 1475 – 1486.

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

Prudhomme, C and D. W. Reed, 1999. Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *International Journal of Climatology*, 19, pp. 1337-1356.

Rosenfield, G. H. and K. Fitzpatrick-Lins, 1986. A coefficient

of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering & Remote Sensing*, 52(2), pp. 223-227.

Skidmore, A. K., 1990. Terrain position as mapped from a gridded digital elevation model. *International Journal of Geographical Information Science*, 4, pp. 33-49.

Sperduto, M. B., and R. G. Congalton, 1996. Predicting rare orchid (Small Whorled Pogonia) habitat using GIS. *Photogrammetric Engineering and Remote Sensing*, 62 (11), pp. 1269 – 1279.

Wu, W. and G. Shao, 2002. Optimal combinations of data, classifiers, and sampling method for accuracy characterization of deforestation. *Canadian Journal of Remote Sensing*, 28 (4), pp.601-609.