Journal of Environmental Science and Management 24-2: 48-53 (December 2021) ISSN 0119-1144

Predicting the Natural Suitability of Teak (*Tectona grandis* L.f.) at Mae Yom National Park, Phrae Province, Thailand Using Logistic Regression Model

ABSTRACT

The natural suitability of teak (Tectona grandis L.f.) at Mae Yom National Park, Phrae Province, Thailand was investigated using logistic regression. The study identified the relationship between key physical and soil properties with the presence of teak at natural sites, and applied a geographic information system platform to identify potential natural sites of teak. Fourteen variables were included in the model, five of these were found to have a significant effect (p < 0.05). Elevation was the most significant topographic variable, whereas magnesium, organic matter, potassium, and calcium were the most significant soil variables. The site prediction had 91.8% accuracy for identifying areas with high, moderate, and low probabilities of being suitable natural habitats for teak. These areas covered 87.63, 244.33, and 163.54 km², respectively.

Torlarp Kamyo[1*]
Sura Pattanakiat[2]
Lamthai Asanok[1]
Khanitin Samanmit[3]
Arnon Cherpaiboon[4]
Sathid Thinkamphaeng[5]
Dokrak Marod[5,6]

[1] Department of Agroforestry, Maejo University Phrae Campus, Phrae, Thailand, 54140
[2] Faculty of Environment and Resource Studies, Mahidol University, Nakhonpathom, Thailand, 73170
[3] Department of Basic Sciences, Maejo University Phrae Campus, Phrae, Thailand, 54140
[4] Department of National Parks, Wildlife and Plants Conservation, Bangkok, Thailand, 10900
[5] Cooperation Centre of Thai Forest Ecological Research Network, Kasetsart University, Bangkok, Thailand, 10900
[6] Faculty of Forestry, Kasetsart University, Bangkok, Thailand, 10900

* corresponding author: torlarp66@yahoo.com

Keywords: *logistic regression model, Mae Yom National Park, suitability prediction, Tectona grandis L.f., Thailand*

INTRODUCTION

Teak (*Tectona grandis* L.f., family Lamiaceae) is one of the world's most sought-after timbers due to its distinctive qualities and wide variety of applications. (*Narong et al. 2007*; *Warner et al. 2016*). It originates in India and has spread naturally to several tropical countries, such as Myanmar and Thailand (*White 1991*). In addition, teak has been grown in about 36 tropical countries and covers an area of 57,000 km² (*Narong et al. 2007*). Teak grows best in tropical deciduous forests and on well-drained alluvial soils (*Nidavani and Mahalakshmi 2014*). The species prefers subtropical climate with average minimum temperature of 9 °C and average maximum temperatures of 41°C. It needs 1,300–3,800 mm of rainfall per year, and three to five months dry season (*White 1991*). Teak is a light-demanding tree species and it is highly resistant to fire. Naturally, teak can be extensively found throughout northern Thailand (16–20°N, 97–101°E), and it occurs in mixed deciduous forest associated with *Pterocarpus macrocarpus*, *Xylia kerrii*, *Afzelia xylocarpa*, *Lagerstroemia calyculata*, and bamboos (*Kaosa-ard 1992*). Among Thailand'smNational Parks, Mae Yom National Park is the last virgin "golden" teak forest in Thailand's national parks.

Surveying the natural distribution of teak is difficult and expensive. However, alternative approaches are mathematical or statistical modelling, which can be used to indicate and predict the most appropriate site by considering various environmental conditions. The species distribution model (SDM) is an important tool for predicting how species are distributed throughout a study area, and it can be applied to determine the most suitable habitats. SDMs have been used in various fields of biology as well as other several purposes, such as conservation planning, surveying, evolutionary studies, and climate-change impact assessments (*Engler et al. 2004; Guisan and Thuiller 2005; Thuiller et al. 2005; Marini et al. 2009; Nanglae and Nilthong 2015*). SDMs are generally used to identify appropriate environmental factors (e.g., topographic, climate, and geological factors) and identify the areas containing these (*Pearson 2007*).

SDMs techniques were classified into two types: profile techniques and group discrimination techniques. Profile techniques, such as environmental envelope (BIOCLIM) and environmental distance (DOMAIN) rely on presence-only data that include locations where associations among species occurrences, or other environmental factors (*Franklin 2010*). Group discrimination techniques was the methods that require both presence and absence data for example, generalized linear models (GLMs) and generalized additive models (GAMs). Both groups are based on statistical models(*Stokland et al. 2011*) Predicting and generating statistical functions or discriminative rules. In general, group discrimination techniques better results than profile techniques. (*Mateo et al. 2010*) for predicting species distributions because they incorporate absence data or background data, which are important for considering the relationship between species occurrence and environmental variables (*Brotons et al. 2004*). However, in case of lacking accurate absence data, pseudo-absence data or artificial absence data can be generated. While pseudo-absence data are based on statistical theory, background data are random sampled, based on machine learning model (e.g., maximum entropy) (*Mateo et al. 2010*). Contemporary statistical analysis, including SDMs, have been considerably used in ecological studies. Approximately 70-80% of articles published in the Journal of Wildlife Management and the Journal of Landscape Ecology used regression models for model predictions occurred from 1998 to 2007. (*Drew et al. 2011*). Such methods also require absence or pseudo-absence data. Logistic regression models that include both presence and absence data have been evaluated and widely accepted as a great statistical tool for forecasting the probability of occurrence of categorical variables (*James et al. 2013*).

This study aims to determine the factors affecting the distribution of teak in Mae Yom National Park, Phrae Province, Thailand, and to identify suitable environmental factors for teak using logistic regression analysis, a GLM technique, to remedy sample size limitations.

MATERIALS AND METHODS

Study Area and Occurrence Data

Mae Yom National Park is the 2nd national park in Phrae Province, Thailand. It was named after the Yom River and was declared the 51st National Park of Thailand in 1986. The area is situated between 18°36′8″–18°49′57″N and 100°4′1″–100°23′33″E and covers 454.75 km^2 area in Ngao District, Lampang Province and Song District, Phrae Province (**Figure 1**). The park's boundary is dominated by teak forest and rugged mountain ranges. The Yom River flows through the park all year round. The rainy season starts from June to September. Winter season is from October to January and summer season is from February to May. A group of teaks is densely grown along the banks of the Yom River, and bright yellow teak flowers are in bloom during October - November, giving delightful scenery of the park. The teak forest is located about 13 km from the park's headquarters and can only be accessed by foot (*Department of National Parks, Wildlife and Plant Conservation 2016*).

Environmental Data

Environmental variables were divided into two groups: topographic variables and soil variables (**Table 1**). The topographic variables included elevation (DEM), slope aspect, and distance to water source. The DEM data were derived from the Land Development Department of Thailand. An interpolation technique was used to calculate the distance to water source. Aspect was calculated from the DEM data. The soil variables obtained from soil sampling were analysed at the Department of Soil Science, Faculty of Agriculture, Kasetsart University, Bangkok, Thailand.

Sampling Points

Sampling points were generated using a two-step method. First, the presence area (50 points) was built using random sampling in the study area. Second, absence points (50 points) were randomly selected outside the presence site that did not contain teak.

Logistic Regression

Logistic regression is a type of linear model (GLM) *Nelder and Wedderburn* (*1972*). In this research, the

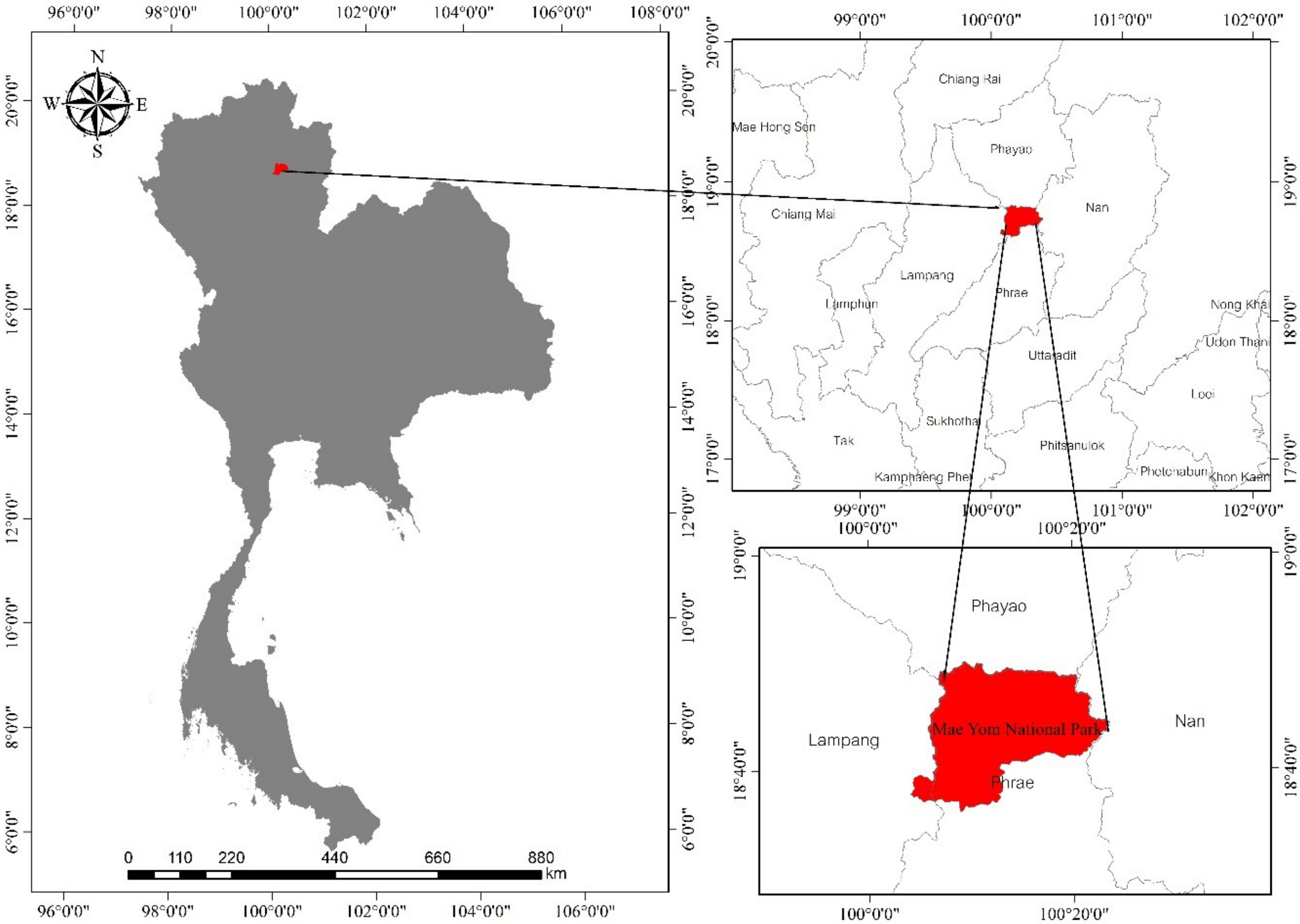

Figure 1. The area in Mae Yom National Park, Thailand where the study in predicting the natural suitability of Teak (*Tectona grandis* L.f.) was conducted.

Table 1. Variables in the species distribution models used in predicting the natural suitability of teak in Mae Yom National Park, Phrae Province, Thailand (2017).

Variable	Description
Elevation	Elevation (m)
Slope	Slope (%)
Distance	Distance from a stream (m)
Aspect	Direction of the slope face (0–360°)
pH	Potential of Hydrogen ions (pH)
$CaCO_3$	Calcium carbonate
Mg	Magnesium
Ca	Calcium
P	Phosphorus
K	Potassium
OM	Organic matter (%)
Sand	Sand (%)
Silt	Silt (%)
Clay	Clay (%)

model has been fit for better extraction of suitable areas in Mae Yom National Park. Meanwhile, dependent variables have 0 and 1 binary values (presence as 1 and absence as 0). This model used a logit link (i.e., logit transformation) to describe the relationships between the response probabilities and the 14 predictor variables (**Table 1**). The logit link is modelled as a linear function (Equation 1):

$$logit(p) = b_0 + \sum_{i=1}^{n} b_i\, x_i\, logit(p) = b_0 + \sum_{i=1}^{n} b_i\, x_i \tag{1}$$

where p is the associated function of causative factors, n is the number of predictor variables, and xi are trail conditioning factors and b0, b1, ..., bn are the coefficient. The model was fitted using the maximum likelihood method. The model prediction was formed as an exponential function (Equation 2):

$$P(Y) = \frac{e^{logit(p)}}{1 + e^{logit(p)}} \tag{2}$$

where P(Y) represents the probability (i.e., probability of presence area). This was transformed into a continuous probability Y ranging from 0 to 1. In this study, the GLM function in the program R was used for logistic regression and the logit function was chosen as the link function.

Prediction Map Creation and Evaluation

The parameters were used in interpolation procedures to create the prediction map based on elevation, pH, $CaCO_3$, sand, silt, clay, phosphorus (P), potassium (K), magnesium (Mg), calcium (Ca), organic matter (OM), slope, aspect, and distance from water based on GIS. The suitability weights were obtained using the logistic regression model. Teak suitability was classified into one of three intervals—high, moderate, and low classification accuracy—based on a model assessment, which was performed to compare the classification results with ground survey results. The number of sample points was calculated using binomial probability theory (*Congalton et al. 1983*).

Table 2. Factors selected in the generalised linear models (GLM) using the logistic regression method in the prediction of the natural suitability of teak at Mae Yom National Park, Phrae Province, Thailand.

Factor	Estimate	Sig.
Elevation	−0.0171	**
Slope	−0.0621	
Distance from water	0.0007	
Aspect	0.0037	
pH	2.9232	
$CaCO_3$	−0.0027	
Mg	0.0084	*
Ca	−0.0065	**
P	0.0232	
K	−0.0363	***
OM	1.4808	**
Sand	−0.1512	
Silt	−0.0207	
Clay	0.0125	

* 0.1 significance level
** 0.01 significance level
*** 0.001 significance level

RESULTS AND DISCUSSION

Factors Affecting the Model

Logistic regression was used to select the variables affecting the model, and Akaike's information criterion (AIC) to select the optimal model ($R^2 = 0.81$). Fourteen variables were selected, of which four were significant ($p < 0.05$) (**Table 2**). Although some variables (pH, $CaCO_3$, sand, silt, clay, K, Mg, OM, P, Ca, slope, aspect, and distance from water) in the model were not statistically significant, these were included in the model due to their presumed effects on teak distribution. Elevation was the most significant topographic variable, while OM, K, Mg, and Ca were the most significant soil variables.

Of the topographic variables, aspect and distance from water had positive effects, whereas slope and elevation had negative effects. Thus, suitable areas for teak should have high solar radiation, as teak is a light-demanding and fire-resistant species (*Kaosa-ard 1992*); have slopes of no more than 30°; and be located far from streams or water sources.

Among the soil variables, pH, clay, OM, P, and Mg had positive effects. However, since the experimental area is dominated by a limestone mountain, $CaCO_3$, sand, silt, K, and Ca had less significant effects. Therefore, suitable areas for teak should have higher OM and clay, which is consistent with the failure of some teak plantations established on low-lying, poorly drained land with clay soil (*Seth and Yadav 1959 ;Fernández-Moya et al. 2014*). In addition, teak is considered a species with high soil nutrient requirements, and deep, well-drained soils with high chemical fertility and low acidity are usually necessary for the habitat of this species (*Fernández-Moya et al. 2015*). Moreover, teak is predominantly found in soils with pH values ranging from 6.5–7.5 (*Kulkarni 1951; Yoshinori et al. 2010*). The OM content of soil also has an important role in the ecological distribution of teak, and Ca is the major constituent of teak wood (*Kaosa-ard 1989*).

Prediction Map Creation and Evaluation

To predict the most suitable area for teak, thresholds were used for transformation into a probability map using GIS (**Figure 2**). Site prediction had 91.8% accuracy for areas with high, moderate, and low probabilities of being suitable natural habitats for teak; these areas covered 87.63, 244.33, and 163.54 km^2, respectively (**Table 3**).

Table 3. Area of potential natural sites for teak in Mae Yom National Park, Phrae Province, Thailand, 2017.

Prediction Level	Area (km^2)	Area (%)
Low (0 - 0.33)	163.54	33.01
Moderate (0.34 - 0.66)	244.33	49.31
High (0.67 - 1)	87.54	17.68

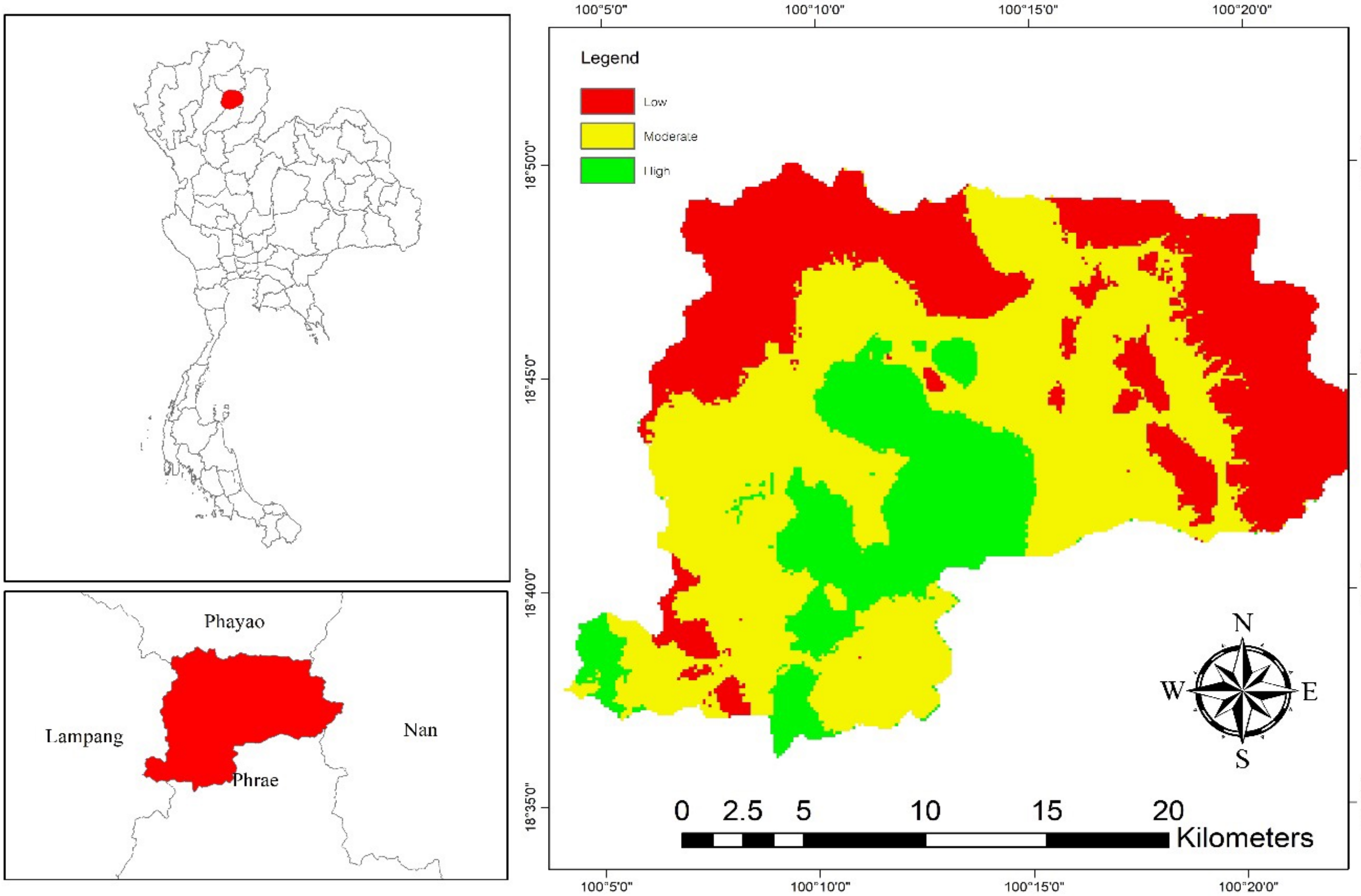

Figure 2. Areas of potential natural sites for teak in Mae Yom National Park, Thailand (2017).

CONCLUSIONS AND RECOMMENDATIONS

A logistic regression model was successfully used to estimate the potential distribution of teak in Mae Yom National Park, Thailand. The potential distribution of teak depended upon topographic and soil variables, and indicated that natural teak habitat should have a slope of no more than 30° and should not be located near rivers or water sources. Elevation was the most important topographic predictor of teak distribution, whereas K, OM, and Ca were the important soil predictors. Moreover, teak prefers a good soil concentrations of K, OM, and Ca. Ca is the major constituent of teak wood. Site prediction using GIS had 91.8% accuracy for areas with high, moderate, and low probabilities of being suitable natural habitats for teak; these areas covered 87.63, 244.33, and 163.54 km^2, respectively. This study of the potential distribution of teak can support conservation efforts and park management in and around teak forests, although more research is needed to determine whether the available protected areas sufficiently cover the most suitable areas for teak distribution. Finally, the methodology applied in this study for teak can be applied to other plant species and in other regions.

REFERENCES

Brotons, L., Thuiller, W., Araujo, M.B. and Hirzel, A.H. 2004. "Presence–Absence Versus Presence-Only Modelling Methods for Predicting Bird Habitat Suitability". *Ecography* 27:437–448.

Congalton, R., Oderwald, R. and Mead, R. 1983. "Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques". *Photogrammetric Engineering and Remote Sensing* 49 (12):1671–1678

Department of National Parks, Wildlife and Plant Conservation. 2016. Mae Yom National Park. http://portal.dnp.go.th/Content/nationalpark?contentId=4091. Retrieved on 10 July 2017.

Drew, C.A., Wiersma, Y.F. and Huettmann, F. 2011. Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications. Springer New York, Dordrecht Heidelberg, London.

Engler, R., Guisan, A. and Rechsteiner, L. 2004. "An Improved Approach for Predicting the Distribution of Rare and Endangered Species from Occurrence and Pseudo-Absence Data". *Journal of Applied Ecology* 41:263–274.

Fernández-Moya J., Alvarado, A., Mata, R., Thiele, H. and Segura, J.M. Vaides, E., San Miguel-Ayanz, A. and Marchamalo-Sacristán, M. 2015. "Soil Fertility Characterisation of Teak (*Tectona grandis* L.f.) Plantations in Central America". *Soil Research* 53:423–432.

Fernández-Moya, J. Alvarado, A. Miguel-Ayanz, A. S. Marchamalo-Sacristán, M. 2014. "Forest Nutrition and Fertilization in Teak (*Tectona grandis* L.f.) Plantations in Central America". *New Zealand Journal of Forestry Science* 44: 1-8.

Franklin, J. 2010. Mapping Species Distribution: Spatial Inference and Prediction. Cambridge University Press, New York.

Guisan, A. and Thuiller, W. 2005. "Predicting Species Distribution: Offering More Than Simple Habitat Models". *Ecology Letters* 8:993–1009.

James, G., Witten, D., Hastie, T. and Tibshirani, R. 2013. An Introduction to Statistical Learning. Springer Texts in Statistics Springer New York, New York.

Kaosa-ard, A. 1989. "Teak *Tectona grandis* L.f. its Natural Distribution and Related Factors". *The Journal of the Natural History Society of Siam* 29:55–74

Kaosa-ard, A. 1992. Country Report on Teak in Thailand: Teak in Asia. In: Wood, H. (Ed.). Proceedings of the CHAINA/ESCAP/FAO Regional Seminar on Research and Development of Teak, Guangzhou, China :79–86.

Kulkarni, D.H. 1951. Distribution of Teak (*Tectona grandis* L.) on the Northern Slope of the Satpuras with Special Reference To Geology. In: Proceedings 8th Silviculturists Conference, Dehra Dun 254–263.

Marini, M.A., Barbet-Massin, M., Lopes, L.E. and Jiguet, F. 2009. "Major Current and Future Gaps of Brazilian Reserves to Protect Neotropical Savanna Birds". *Biological Conservation* 142: 3039–3050.

Mateo, RG., Croat, T.B., Felicísimo, A.M. and Muñoz, J. 2010. "Profile or Group Discriminative Techniques? Generating Reliable Species Distribution Models Using Pseudo-Absences and Target-Group Absences From Natural History Collections". *Diversity and Distributions* 16: 84–94.

Narong, K., Katsutoshi, S. and Sota, T. 2007. "Composition and Diversity of Woody Regeneration in a 37-year-old Teak (*Tectona grandis* L.) Plantation in Northern Thailand". *Forest Ecology and Management* 247: 246–254.

Nanglae, S. and Nilthong, R. 2015. "Predicting Assam Tea Distribution in Upper Northern Thailand Using Species Distribution Models". *Applied Ecology and Environmental Research* 13(4): 1085–1096.

Nelder, J.A. and Wedderburn, R.W.M. 1972. "Generalized Linear Models". *Journal of the Royal Statistical Society Series* A 135: 370–384.

Nidavani, R.B. and Mahalakshmi, A. 2014. "Teak (*Tectona grandis* Linn.): A Renowned Timber Plant with Potential Medicinal Values". International Journal of Pharmacy and Pharmaceutical Sciences 6:48–54.

Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. and Prentice, I.C. 2005. "Climate Change Threats to Plant Diversity in Europe". Proceedings of the National Academy of Sciences of the United States of America 102: 8245–8250.

Pearson, R.G. 2007. Species' Distribution Modeling for Conservation Educators and Practitioners: Synthesis. American Museum of Natural History. http://ncep.amnh.org. Retrieved on 10 July 2017.

Seth, S.K. and Yadev, M.A.W. 1959. "Teak soils". Indian Forester 85: 2–16.

Stokland, J.N., Halvorsen, R. and Støa, B. 2011. "Species Distribution Modelling – Effect of Design and Sample Size of Pseudo–Absence Observations". *Ecological Modelling* 222: 1800–1809.

Warner, A.J., Jamroenprucksa, M. and Puangchit, L. 2016. "Development and Evaluation of Teak (*Tectona grandis* L.f.) Taper Equations in Northern Thailand". *Agriculture and Natural Resources* 50: 362–367.

White, K.J. 1991. Teak: Some Aspects of Research and Development. RAPA Publication 1991/17. FAO, Bangkok.

Yoshinori, W., Ebenezer, O., Tsugiyuki, M., Mohammed, M.B., Oladele, O. I. and Toshiyuki, W. 2010. "Teak (*Tectona grandis*) Growth as Influenced by Soil Physicochemical Properties and Other Site Conditions in Ashanti Region, Ghana". *Journal of Food, Agriculture & Environment* 8(2):1040–1045.

ACKNOWLEDGMENT

This research was supported by the Office of Agricultural Research and Extension at Maejo University and the National Research Council of Thailand. We are grateful to the staff of Mae Yom National Park and the agroforestry students for their help during the field data collection. We also thank the Thai Forest Ecological Research Network for kindly supporting the data collection and soil analysis.