WOULD ILLEGAL LOGGING BE PREDICTED ACCURATELY AND PREVENTED BEFOREHAND BY USING REMOTE SENSING AND GIS?

Tien-An Lin[1] Sheng-Hua Tseng[2], Nan-Chang Lo[3], and Kai-Yi Huang[4]
[12] Graduate Student, Dept. of Forestry, Chung-Hsing University, Taiwan, R.O.C., tammy0841@yahoo.com.tw
[3]Specialist, Experimental Forest Management Office, Chung-Hsing University, Taiwan, R.O.C., njl@nchu.edu.tw
[4]Professor, Dept. of Forestry, Chung-Hsing University, Taiwan, R.O.C., kyhuang@dragon.nchu.edu.tw
[1234]250, Kuo-Kuang Rd., Taichung 402, Taiwan, R.O.C.

KEY WORDS: Normalized Difference Vegetation Index (NDVI), Species Distribution Model (SDM), Maximum Entropy (MAXENT), Support Vector Machine (SVM), DOMAIN

ABSTRACT: Taiwan is rich in forest physiognomy and has lots of valuable forest resources. However, they disappeared gradually owing to rampant illegal logging in recent years. The target species are Taiwan red cypress and Taiwan cypress. The study aimed to predict the hot zone of illegal logging (HZIL) through species distribution model (SDM) coupled with 3S technology. In order to simulate different situations of HZIL, four sampling designs (SD) were created, including random sampling (T1), equal sampling (T2), inclined sampling (T3) based on the preference of illegal loggers for trees near a road and on gentle slopes, and the three SDs plus NDVI (normalized difference vegetation index) values above 0.5 respectively (T4, T5, and T6), which represents healthy and big trees. The SDMs were built based on predictor variables, including elevation, slope, aspect, cost distance, easting, northing, and change in NDVI, by using DOMAIN, maximum entropy (MAXENT) and support vector machine (SVM). Regardless of which algorithm being used, T3 was better than T1 and T2 owing to its inclined sampling. And T6 was the best; in contrast, T2 was the worst one among them. MAXENT was the best, followed by SVM, and DOMAIN was the last on T6. The greater the difference in NDVI between before and after logging, the more easily and accurately the HZIL can be estimated. However, several other factors such as windfall, forest fire, and landslide may still cause erroneous forecasting. Hence, a follow-up study will use high spatial resolution remotely sensed imagery and create new variables and sampling designs to tackle different situations in illegal logging so that the HZIL can be estimated more accurately. This approach will help the government to reverse the situation from run-after to preemptive strike on illegal loggers via 3S technology and protect the valuable forest resource.

1. INTRODUCTION

Taiwan is rich in forest physiognomy because of the large gradient in altitude. Inside the narrow 36,000 km^2, the elevation sharply declines about 4,000 m, breeding abundant species because of the complex climate. According to the third forest resources survey, Taiwan's forest area is 2,102,400 ha, dominating 58.53% of the entire island area (Taiwan Forestry Bureau, 2010). In recent years, the consciousness of conserving and protesting forest has risen up for people because of extreme weathers accompanied by the climate change. How to keep forest away from destruction is the important issue for forest management. Since 1992, the policy of prohibiting on cutting the trees has passed to improve the ability of forest to conserve the soil and water. But the illegal logging events have become more frequent for the past several years under the impact of economical depression. It is still a big problem to forestry about how to protest the forest resources effectively to fulfill the ability of public welfare. Both the system of forest protection and the operation of the forest police system are not effective to deter illegal logging (Hsu, 2010). In the era of science and technology changing rapidly, it is worth thinking how to use high technologies to supervise forest areas and suppress illegal logging effectively (Leu, 2010).

Species distribution modeling (SDM) plays an important role in the ecology and biogeography in recent years. SDM uses species habitat environmental factors with GIS technique, which integrates remote sensing imagery and GPS field measurement data. Using different kinds of multivariate statistics in conjunction with 3S to build predictive models and predict the spatial distribution of species may assist in ecological investigation on a large spatial scale and decision-making (Guisan and Zimmermann, 2010). This study attempted to use the predictor variables characterizing inclined behavior of the illegal logger to model the hot zone of illegal logging (HZIL) by SDM.

The objectives of this study were: (1) to use different sampling designs and different variables to simulate the multiple situations of illegal loggers' inclination; (2) to compare the predictive accuracies of different models; (3) to determine the HZIL and to identify high-risk CF trees of illegal logging which enables related government agencies to take effective measures readily deterring illegal logging.

2. MATERIALS AND METHODS

2.1 Target Species

Target species in the study are *Chamaecyparis formosensis* (Taiwan red cypress, TRC) and *Chamaecyparis obtuse* (Taiwan Cypress, TC), and these two species in the forest stand are referred to as Cypress forest (CF). They both grow at medium to high altitude range of about 1,500–2,500 meters where is called fog-forest belt in the mountains of Taiwan and often form large area of pure stands or above 2,500 m two species mixed together. They are kinds of valuable commercial timber in Taiwan.

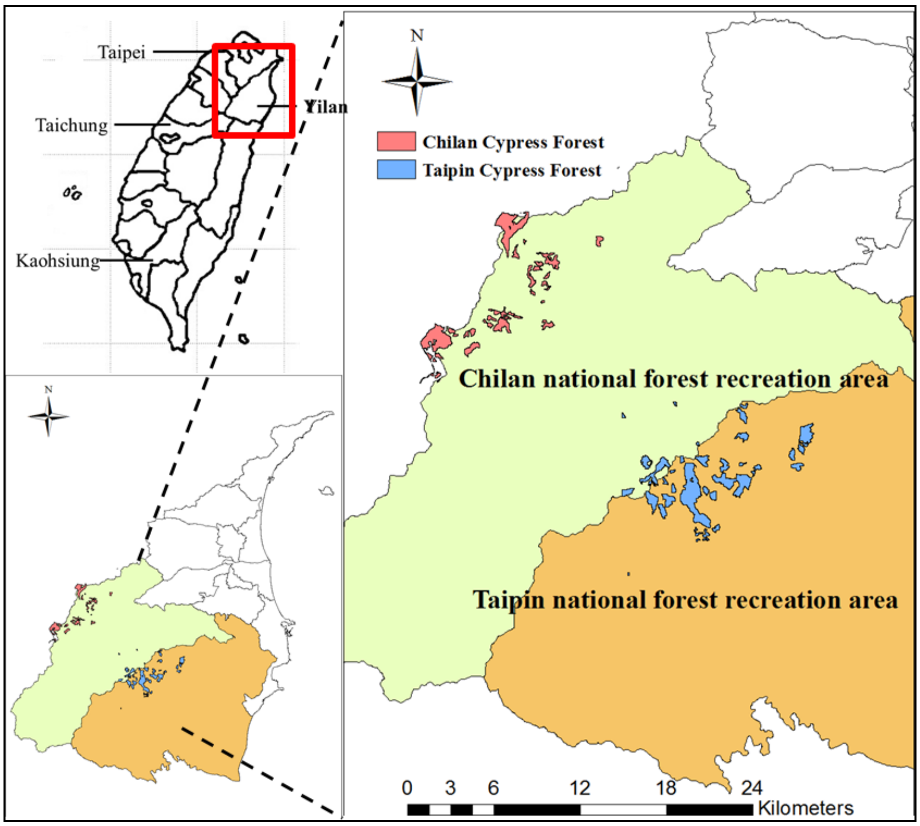

Figure 1 Location map of the study area.

2.2 Study Area

This study includes two areas both in Yilan prefecture, the northeastern of Taiwan: (1) Chilan national forest recreation area and (2) Taipin national forest recreation area (figure 1). This study clipped the two areas with the same size about 22,522.5 ha. Both areas raining nearly throughout the year provide a mild and humid climate conditions, thereby being rich in TRC and TC species. But here also has a serious problem of illegal logging.

2.3 Sampling Design

CF samples were taken directly from digital forest-type map of the study area as dummy data of illegal logging since the real data of illegal logging cases could not be obtained. In order to simulate the tree samples of illegal logging, the study created six sampling designs. (1) The study randomly took the samples from the forest-type of CF (Type 1, T1).

(2) We specified two grades of distance from a road, within and beyond 1,500 m, and specified three grades of slope including: below 20°, 20°–30°, and above 30°. The study divided CF into six parts ($2 \times 3 = 6$ combinations) and chose an equal number of samples from each part (Type 2, T2). (3) To simulate the theft behavior of a "mountain rat" (illegal logger), the study chose the samples associated with mountain rat's inclination (e.g. trees near a road and on gentle slopes) from six parts (Type 3, T3). (4) To create sampling designs involving pronounced tendency of "mountain rats", we took the samples in the same way as in T1, T2, and T3 but plus one more preference for big trees (i.e. tall trees with large diameter at breast height, d. b. h.) having NDVI values over 0.5 (representing healthy, big trees) from each part, (T4, T5, and T6). There were a total of six sampling designs created for building predictive models of illegal logging.

2.4 Variables Used in the Study

The study generated elevation, slope and aspect from the DEM of the study area by ERDAS Imagine. It also derived the data layers of normalized difference vegetation index (NDVI) from SPOT image**s** of 2002 (previous image) and 2005 (post-image after illegal logging) by ERDAS Imagine (c SPOT Image Copyright 202 and 2005 CNES). Then the study generated differencing vegetation index (DVI) image from the two-date NDVI images. In order to account for where the NDVI image value of a pixel may drop after a CF tree has been logged within the pixel, the study assigned five likely values to the post NDVI image (0.3, 0.35, 0.4, 0.45, and 0.5). The study produced the data layers of easting and northing coordinates from a geo-referenced imagery by Idrisi and used the VARCOS function to obtain the data layer of cost distance, which is useful when considering the movement across a mountain area. There were a total of seven predictor variables in the study.

2.5. Model Calibration and Evaluation

For each sampling design, the study chose three replicates of 180 training samples for model calibration and one independent dataset of 90 test samples for model validation. In order to avoid spatial autocorrelation, the number of background samples randomly selected from the non-CF area was five times more than that of logging CF (target) (Pereira and Itami, 1991; Sperduto and Congalton, 1996). Finally, three models were built by seven variables (elevation, slope, aspect, cost distance, easting, northing, change in NDVI value) to predict the HZIL of CF by coupling 3S technology with three algorithms, including MAXENT, SVM, and DOMAIN, in the study. The study implemented each model with three replicates and evaluated the models by the mean *Kappa* values obtained from the independent dataset for each area.

3. RESULTS AND DISCUSSION

3.1. Model Fit with the Change in NDVI

As shown in figure 2, the accuracies of the three models in different sampling designs declined with the change in NDVI value as it becomes smaller from 0.3 to 0.5. T1 (random sampling) was used as baseline for comparison. T1, T2, and T3 (solid lines in figure 2 and samples without NDVI inclination) had much lower accuracies, especially T2 (red solid line). Among three designs without NDVI preference, the performance of T3 was best and T2 was the worst because samples in T2 were intentionally taken in an equal number from each part. T4, T5, and T6 simulated the situation in which tree samples were taken with the most inclinations of "mountain rats". Among six sampling designs, the three models on sampling designs with inclination to NDVI value above 0.5 (T4, T5, and T6, dashed lines in figure 2) performed much better than T1, T2, and T3 did as varied with changes in NDVI value. It indicates that there was substantial difference in accuracy between sampling designs with and without preference for a big NDVI value (above 0.5). Hence, the mountain rat's inclination to big and healthy trees (i.e. a large NDVI value above 0.5) was extremely important to prediction of illegal logging.

3.2. Comparisons of Model Performance

As shown in table 1, for example the accuracies of MAXENT in T6 declined gradually with the change in NDVI value as it becomes less from 0.3 to 0.45, whereas those accuracies declined sharply as the NDVI value is above 0.45. In contrast, the accuracies of DOMAIN and SVM in T6 declined gradually with the change in NDVI value as it becomes less from 0.3 to 0.35, whereas they declined sharply as the NDVI value is above 0.35. The mean *Kappa* value of MAXENT was 0.78, SVM was 0.71, and DOMAIN was 0.56. As a result, MAXENT was the best among three models, followed by SVM, and DOMAIN was the worst on T6, and similar results were with T4 and T5.

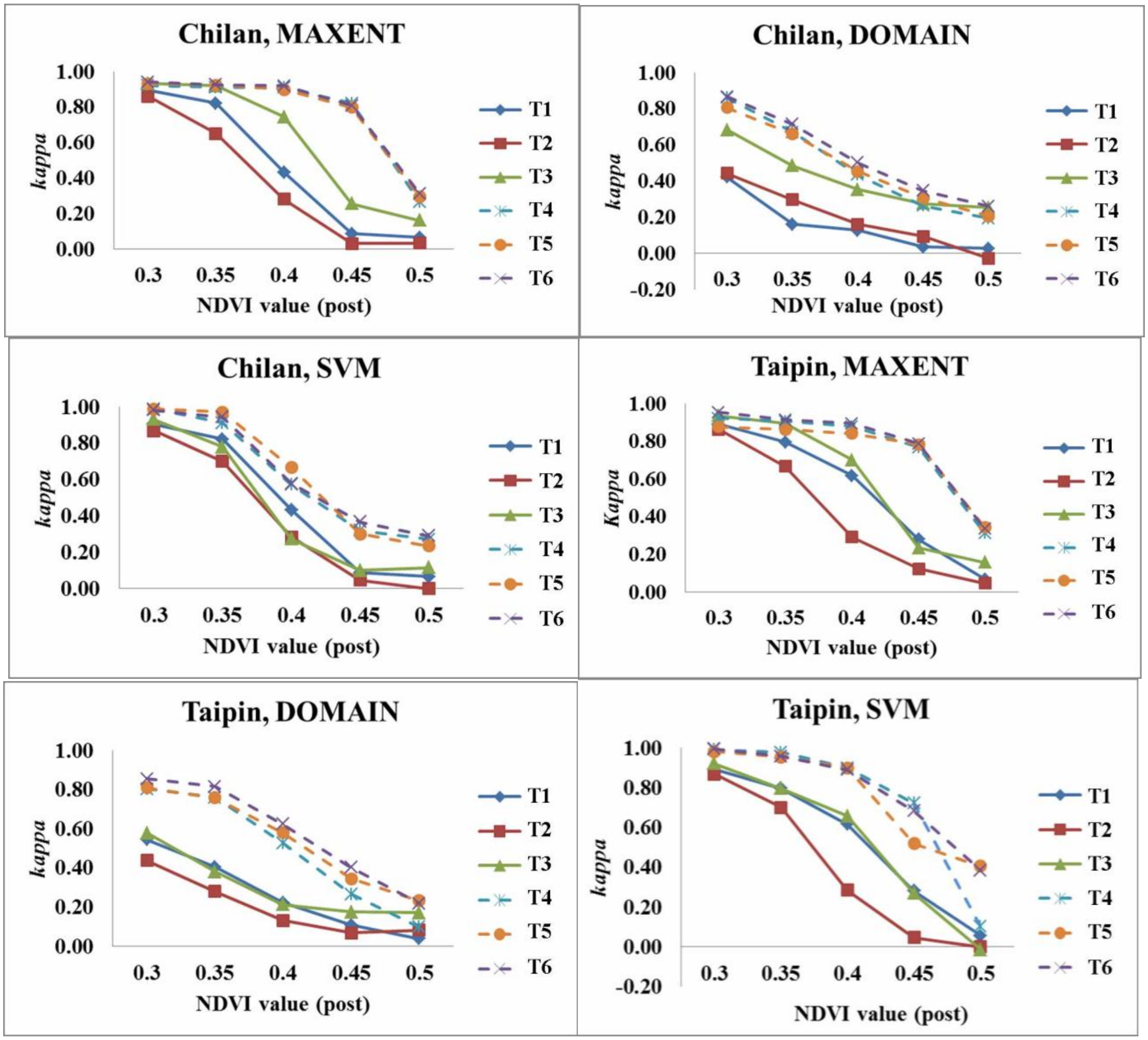

Figure 2 Performances of three models on six sampling designs varied with changes in NDVI over two areas.

Table 1 The accuracies of three models for T6 validated by independent datasets in Chilan and Taipin

Statistic	*Kappa* value										
NDVI (*Δ NDVI)	**0.3 (0.3)**		**0.35 (0.25)**		**0.4 (0.2)**		**0.45 (0.15)**		**0.5 (0.1)**		Mean
Area	Chilan	Taipin	Chilan	Taipin	Chilan	Taipin	Chilan	Taipin	Chilan	Taipin	
MAXENT	0.94	0.95	0.93	0.91	0.92	0.9	0.81	0.79	0.31	0.34	**0.78**
DOMAIN	0.87	0.86	0.71	0.81	0.5	0.62	0.35	0.4	0.26	0.22	**0.56**
SVM	0.98	1	0.94	0.96	0.58	0.89	0.37	0.68	0.29	0.38	**0.71**
Mean	**0.93**	**0.93**	**0.86**	**0.9**	**0.67**	**0.8**	**0.51**	**0.62**	**0.29**	**0.31**	—

***Δ NDVI = the change in NDVI value against the maximum NDVI value for CF (0.6) in this case (0.6 - NDVI).**

3.3. Demonstration for Predicting the HZIL

The predictive result from MAXENT is a map of continuous probability distribution, and it was reclassified into three probability groups. The study overlaid 450 randomly chosen CF tree samples on this map and counted the number of CF trees at each probability interval. As shown in figure 3, the predictive result from MAXENT on T6 (only show T6) seemed to have a consistent distribution across probability intervals for five NDVI values. In contrast, the number of samples widely varied with different probability intervals for five NDVI values when sampled without preference for big NDVI value (T1, T2, and T3 not shown here), especially for T2. Consequently, the prediction from MAXENT on T6 had the most stable probability distribution among five NDVI values. This indicates that MAXENT on T6 had good ability to predict the CF trees with high-risk illegal logging accurately and effectively in different situations (red: high-risk zone of illegal logging; orange: medium-risk; dark-green: low-risk; light blue: non-logging CF trees).

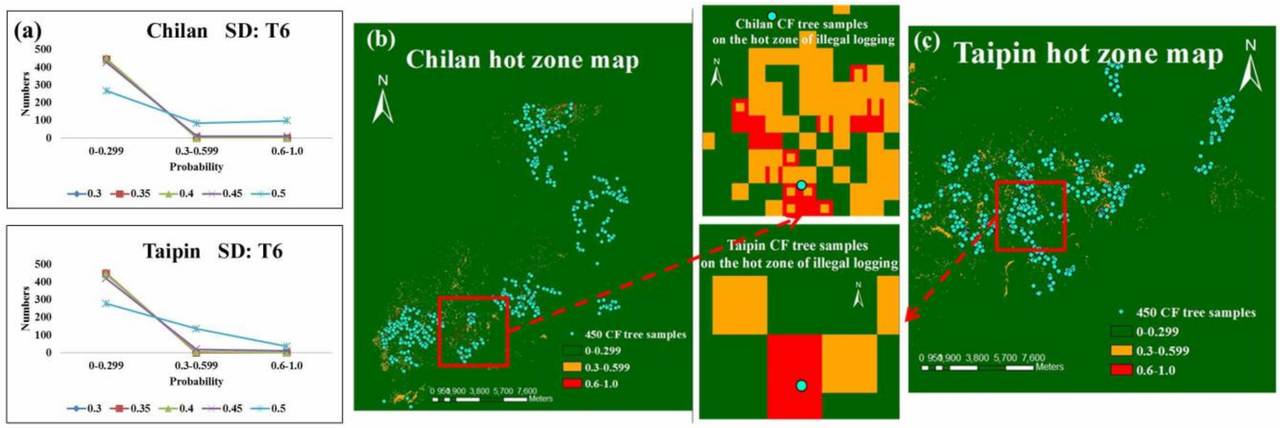

Figure 3 (a) Line graph of frequency at each probability interval for MAXENT on T6 of two areas; (b) map of the hot zone of illegal logging in Chilan (T6 with 0.45 NDVI value) and (c) in Taipin (T6 with 0.45 NDVI value).

4. CONCLUSIONS

The study demonstrated how to predict the HZIL by using 3S technology for an alternative way to forest management. This study simulated different scenarios of illegal logging and generated the distribution map of the HZIL. MAXENT was the best among three models. Among three sampling designs without NDVI preference, T3 had the highest predictive accuracy. There was substantial difference in accuracy between sampling designs with preference and without preference for a big NDVI value, while the difference in accuracy among three sampling designs added with NDVI preference decreased. Hence, the change in NDVI may be a key variable for model predicting the HZIL accurately. Changes in NDVI value after cutting a CF tree within a pixel may vary widely, depending on tree age, tree health, and under-story vegetation. The greater difference in NDVI value, the more easily and accurately the HZIL can be estimated. However, other factors such as windfall, forest fire, forest disease and insect, landslides, and floods may still cause erroneous forecasting. Moreover, the distance from a target tree to a river and reverse operation due to change in mountain rat’s behavior need to be considered as well. Hence, a follow-up study will use high spatial resolution remotely sensed imagery and create new variables and sampling designs, such as the distance from a target tree to stream, to tackle different situations in illegal logging so that the HZIL can be estimated more accurately. This approach will help the government to reverse the situation from run-after to preemptive strike on illegal loggers via 3S technology and protect our valuable forest resource.

5. REFERENCES

Taiwan Forestry Bureau, 2010. The third forest resources survey.
Hsu, M. H., 2010. A study of forest illegal logging in Taitung.
Leu, S. W., 2011. The Organization and operation of illegal logging.
Guisan, A. and N. E. Zimmermann, 2010. Predictive habitat distribution models in ecology. Ecological Modelling 135: pp. 147–186.
Pereira, J. M. C., and Itami, R. M., 1991. GIs-based habitat modeling using logistic multiple regression: A study of the Mt. Graham red squirrel, Photogrammetric Engineering b Remote Sensing, 57(11): pp. 1475–1486.
Sperduto, M. B. and Congalton, R. G., 1996. Predicting rare orchid (S Using species distribution models to predict new occurrences for rare plants. mall Whorled Pogonia) habitat using GIS. Photogrammetric Engineering and Remote Sensing, 62 (11): pp.1269–1279.