ISSN 2250-0987
Chonlanet Preechacharoensri et al, UNIASCIT, Vol 3 (2), 2013, 394-399

EXPERT SYSTEM APPROACH TO LAND COVER CLASSIFICATION AT SAKAERAT ENVIRONMENTAL RESEARCH STATION, THAILAND

Chonlanet Preechacharoensri[1,2], Xiaoli Zhang[1]

[1]*College of Forestry, Beijing Forestry University, No.35 Qinghua East Rd., Haidian, Beijing, 100083, P.R.China*

***Abstract*— This research aims to apply the expert system to classify the land cover of Sakaerat Environmental Research Station, Thailand and compares the classification results with the maximum likelihood classification, the most common supervised classification. Expert system is a hierarchy of rules in which the variables can be defined extensively by users for describing hypotheses or classes. This study applied digital number of satellite image, principle component analysis, tasseled cap transformation, ferrous minerals, and the digital elevation model of the study area to the Land sat 5 TM image which was acquired on January 18, 2011 to classify the land cover types of the study area. The land cover types of the study area were classified into seven classes: agricultural area, water body, grass and abandon land, bamboo, deciduous forest, plantation, and evergreen and restoration forest. The results show that the overall accuracy and overall Kappa statistics of the expert classification were 88.10% and 0.7579, respectively, which were greater than the overall accuracy and overall Kappa statistics, 82.38% and 0.6425, respectively, of the maximum likelihood classification. The expert system could be applied to helping obtain greater accuracy of the classification. A higher accuracy of classification shows a higher of map accuracy, which is more reliable for further use and study.**

***Keywords*— Expert system, Land cover classification, Remote sensing, Landsat 5 TM**

1. INTRODUCTION

Land cover has been defined by [7] as the observation of (bio)physical cover on the earth's surface or the surface cover of the ground [10]. Land cover and land use are multidisciplinary in nature. It can be used to manage the natural resource, to facilitate sustainable of the land [2].

The classification is the arrangement of objects into groups on the basis of their relationships and describes systematic framework, with the name of the classes and the criteria used to discriminate them, and the relation between classes. The classification involves class boundaries, and therefore, it should be clear, precise, possibly quantitative, and based upon objective criteria [7].

The classification of remote sensing imagery has long been of interest by the remote sensing community because the results of the classification are the basis for many environmental and socioeconomic applications [11]. Further, it is a basic component of the planning and decision-making processes for area management [3]. The achievement of satellite image classification to the thematic maps is a complex process because it may be affected by many factors, such as the complexity of the landscape, type of remotely sensed data, and image pre-processing and classification approaches. In recent years, many advanced classification methods, one of which is the expert system, have been widely applied for image classification [11].

The expert classification is a hierarchy of rules, or a decision tree, that describes the condition under which a set of low level constituent information, which consists of user-defined variables, raster imagery, vector coverage's, spatial models, external programs, and simple scalars, becomes abstracted into a set of high level information classes [6]. As different kinds of ancillary data, such as the digital elevation model, soil map, housing and population density, road network, temperature, and precipitation, may be incorporated into the classification procedure, it has proven effective in improving classification accuracy. A critical step is to develop rules that can be used in an expert system or a knowledge-based classification approach. It is now becoming attractive because of its capability of accommodating multiple sources of data [11]. It has been reported in previous study related to expert classification that the accuracy of expert classification is greater than the accuracy of the traditional supervised classification using the maximum likelihood decision rule [4], [8], [9], [18], [20].

Sakaerat Environmental Research Station (SERS) is very important location for the study and research on natural science, environment and ecology of dry dipterocarp and dry evergreen forest. The thematic map derived from land cover classification of SERS is the basic information of the area. It is very important to the managers, the researchers and the interested people to plan their work to reach the highest achievement.

The objective of this study was to design an expert classification system, a knowledge-based classification approach, to classify land cover types of the Sakaerat Environmental Research Station in Nakhon Ratchasima Province, Thailand by Landsat 5TM of January 18, 2011 compared with the maximum likelihood classification, the traditional classification technique.

2. MATERIALS AND METHODS

2.1 Study area

Sakaerat Environmental Research Station (SERS) is situated on the Korat plateau, which covers latitude 14° 26′ to 14° 32′ N and longitude 101° 51′ to 101° 57′ E. The altitude of SERS ranges from 250 – 762 meters above mean sea level. It is a part of Khao Phu Luang National Reserved Forest. The area of SERS is approximately 78.08 km^2. It is located in Pakthongchai and Wangnamkhiao Districts of Nakhon Ratchasima Province in Northeastern, Thailand (fig. 1). SERS was founded under the Thai cabinet resolution in September 1967. Afterwards, Man and Biosphere Program (MAB) of the United Nations Education, Scientific and Cultural Organization (UNESCO) has established SERS as a Biosphere Reserve of Thailand since 1976 for educational and research on natural science, environment and ecology of dry dipterocarp and dry evergreen forest, and other fields [12], [13].

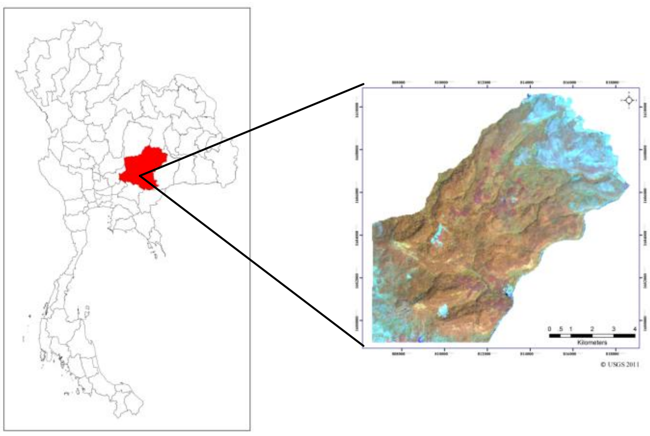

Fig. 1 Maps show Nakhon Ratchasima Province and the study area

Forest resources of SERS can be classified to two types by species composition. There are dry evergreen forest and dry dipterocarp forest. The dry evergreen forest consists of broadleaved species without shedding their leaves during dry season. The important species are *Hopea ferrea* Laness., *H. odorata* Roxb., *Shorea henryana* Pierre, and *Irvingia malayana* Oliv. Ex A.W.Benn. [15]. The dry evergreen forest which is situated on a higher elevation often found mixed with bamboo [21]. While the species composition of dry dipterocarp forest are the species of Dipterocapaceae Family which almost shed their leaves during dry season. The important species are *Shorea obtusa* Wall. Ex Blume and *S. siamensis* Miq. [15]. Moreover, the fast growing tree of the Legumes species and economic native species, such as *Dalbergia cochinchinensis* Pierre, *Dipterocarpus alatus* Roxb. ex G.Don., *Hopea odorata* Roxb, were planted as plantations in the encroached areas [5].

2.1 Methods

The satellite data used in this study was Landsat 5 TM image, path 128 row 50, which was acquired on January 18th, 2011. The software used for preparing and analyzing the data was ERDAS Imagine and ArcGIS remote sensing software.

The geometric correction using image-to-map rectification was performed for the satellite image, where topographic maps scale 1:50,000 were used as the reference maps. The satellite image was geo-referenced to WGS84 datum and UTM Zone 47 North coordinate system. The nearest neighbor re-sampling method was generated because it can maintain the properties of pixels reflectance [16]. RMSe was less than 1 pixel. Six spectral bands, band 1-5 and band 7, were used in the classification procedure.

Land cover was divided into seven classes: agricultural area, water body, grass and abandon land, bamboo, deciduous forest, plantation, and evergreen and restoration forest. These were divided using a preliminary survey of the satellite image, field survey and modification of land use classification system of the Land Development Department, Thailand [17].

2.1.1) Supervised classification

Landsat 5 TM imagery year 2011 was performed using the maximum likelihood decision rule, which is the most common supervised classification used with remote sensing data [14]. The classification method uses the training data as a means of estimating the mean and variance of the class, which are used to estimate probability. It considers not only the mean or average values in assigning classification, but also the variability of each class. Its decision rule is based on the probability that a pixel belongs to a particular class [1]. In this classification method, the training data sets for each land cover classes were assigned using a land use map of SERS, bands combination, and field data.

2.1.2) Expert classification

Expert classification was performed using the IMAGINE Expert classifier. The Expert classifier is composed of two parts: the Knowledge Engineer and the Knowledge Classifier. The Knowledge Engineer provides the interface for an expert with first-hand knowledge of the data and the application to identify the variables, rules, and output classes of interest and to create a hierarchy decision tree. The Knowledge Classifier provides an interface for non-experts to apply the knowledge base and to create the output classification [6].

The conditions used in the expert classification are as follows:

- Digital number (DN) of an image
- Band ratio, which highlights the specific features that have high reflectance in one band and relatively low in another. The band ratio that was used in this study was ferrous minerals (FM), which was calculated from TM5/TM4.
- The principle component analysis (PCA), which changes the values of pixels based on multivariate statistical relation derived from the spectral reflectance values for all pixels in all bands. It also compresses redundant data values into fewer bands, which are often more interpretable than the source data.

- Tasseled cap transformation (TC), which rotates the data structure axes to optimize data viewing for vegetation studies. It can be used to divide brightness, greenness, and wetness [6].
- Ancillary data, the digital elevation model (DEM), as also used as a condition in the study.

Table 1 shows the hypotheses, rules, and variables used in the Knowledge Engineer module.

TABLE 1 THE HYPOTHESES, RULES, AND VARIABLES USED IN THE KNOWLEDGE ENGINEER MODULE

Class	Rule	Variable
Agricultural area	Rule base I	DN3 (≥67.2, ≤188.76) DN4 (≥77.14, ≤149.14) DN5 (≥ 123.03, ≤232.58) DEM (>300)
Water body	Rule base II	DN 4 (≥1.69, ≤25.28)
Grass and abandon land	Rule base III	TC2 (≥-34.83, ≤39.89) TC3 (≥-66.29, ≤-7.3) DEM (≥540)
Bamboo	Rule base IV	DN3 (≥37.08, ≤58.99) DN4 (≥ 171.91, ≤222.47) DN5 (≥96.07, ≤139.89)
Deciduous forest	Rule base V	TC1 (≥157.87, ≤279.78) TC2 (≥-26.97, ≤47.75) TC3 (≥-38.76, ≤16.29) PCA2 (≥-103.37, ≤6.47)
Plantation	Rule base VI	DN3 (≥23.6, ≤40.45) DN5 (≥45.51, ≤79.21) TC1 (≥122.47, ≤201.12) TC2 (≥47.75, ≤118.54) TC3 (≥8.43, ≤39.89) PCA1 (≥ 136.52, ≤223.03) PCA2 (≥10.61, ≤69.66) PCA3 (≥-28.65, ≤6.47) FM (≥0.31, ≤0.47)
Evergreen and restoration forest	Rule base VII	PCA1 (≥73.6, ≤234.83) PCA2 (≥-20.79, ≤53.93) PCA3 (≥-20.79, ≤22.47)

3. RESULTS AND DISCUSSION

3.1 Maximum likelihood classification

Land cover classification, performed by the maximum likelihood decision rule, is presented in fig. 2, which shows the results after the 3 x 3 majority filtering. The result of the overall accuracy is 82.38% and the overall *k* statistic is 0.6425. The producer's accuracy and user's accuracy of grass and abandon land are 0.00% and 0.00%, deciduous forest are 94.74% and 73.47%, plantation are 30.77% and 61.54%, and evergreen and restoration forest are 92.14% and 89.58%, respectively. These accuracy results are shown in table 2.

The result indicated that the producer's accuracy of plantation was very low, at 30.77%. This shows an omission error of plantation class. The pixels of plantation class were classified to other classes; in this case it was classified as grass

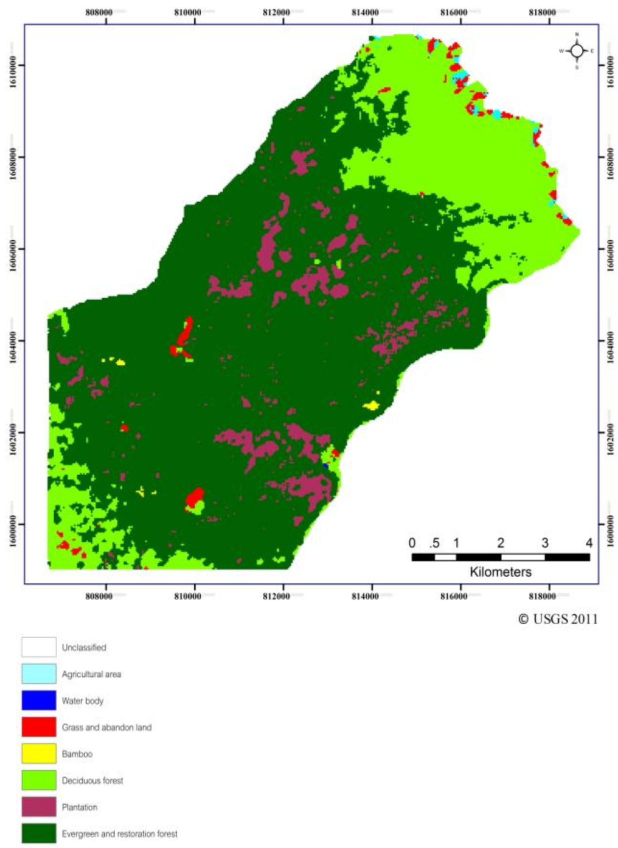

Fig.2 Land cover classification map by the maximum likelihood classification

TABLE 2 ACCURACY ASSESSMENT OF THE MAXIMUM LIKELIHOOD CLASSIFICATION

Class	Ref-erence total	Class-ified total	No. Correct	PA (%)	UA (%)	*k*
Agr	4	0	0	-	-	0.0000
Wa	0	0	0	-	-	0.0000
Gr	1	4	0	0.00	0.00	-0.0048
Bb	1	0	0	-	-	0.0000
De	38	49	36	94.74	73.47	0.6761
Pl	26	13	8	30.77	61.54	0.5610
Ev	140	144	129	92.14	89.58	0.6875
total	210	210	173			
Overall accuracy = 82.38%, Overall Kappa statistic = 0.6425						

Note:Agr:agricultural area, Wa: water body, Gr: grass and abandon land, Bb: bamboo, De: deciduous forest, Pl: plantation, Ev: evergreen and restoration forest, PA: Producer's accuracy, UA: User's accuracy, *k*: Kappa statistic

and abandon land, deciduous forest, and evergreen and restoration forest. It also shows that the spectral signature of these classes is similar to one another; therefore, it was calculated to the class of highest probability [19].

The estimated area in km^2 of each class which was calculated by the maximum likelihood classification showed that agricultural area covered 0.29% (0.2315 km^2), water body covered 0.01% (0.0086 km^2), grass and abandon land covered 1.15% (0.9200 km^2), bamboo covered 0.13% (0.1066 km^2), deciduous forest covered 22.58% (18.0786 km^2), plantation covered 5.58% (4.6854 km^2), and evergreen and restoration forest covered 69.98% (56.2017 km^2). The details are shown in table 3.

TABLE 3 THE ESTIMATED AREA CALCULATED BY THE MAXIMUM LIKELIHOOD CLASSIFICATION

Land cover type	**Area (hectare)**	**Percent cover (%)**
Agricultural area	23.15	0.29
Water body	0.86	0.01
Grass and abandon land	92.00	1.15
Bamboo	10.66	0.13
Deciduous forest	1,807.76	22.58
Plantation	468.54	5.58
Evergreen and restoration forest	5,602.17	69.98
Total	8,005.24	100.00

3.2 Expert classification

Land cover classification conducted by the expert system is presented in figure 3 which shows the results after the 3 x 3 majority filtering. The result of the overall accuracy is 88.10% and the overall *k* statistic is 0.7595. The producer's accuracy and user's accuracy of agricultural area are 75% and 100.00%, grass and abandon land are 100.00% and 50.00%, deciduous forest are 100.00% and 73.08%, plantation are 34.62% and 90.00%, and evergreen and restoration forest are 95.71% and 93.71%, respectively. These accuracy results are shown in table 4.

The results indicate the very low value of the producer's accuracy of the plantation class, at 34.62%. Even though, the percentage of the producer's accuracy of the plantation class by the expert classification is higher than that of the maximum likelihood classification, it shows that the omission error is still high. In this case, the pixels of the plantation class were classified as grass and abandon land, deciduous forest, and evergreen and restoration forest. To solve this problem, others variables can be added to in a hierarchical rules to improve the accuracy of the producer of the plantation class.

The estimated area in km^2 of each class which was calculated by the expert classification showed that agricultural area occupied 1.32% (1.0523 km^2), water body occupied 0.01% (0.0086 km^2), grass and abandon land occupied 0.49% (0.3871 km^2), bamboo occupied 0.26% (0.2070 km^2), deciduous forest occupied 22.39% (17.8164 km^2), plantation occupied 4.31% (3.4288 km^2), and evergreen and restoration forest occupied 71.22% (56.6832 km^2). The details are shown in table 5.

When comparing the results of the two classifications, maximum likelihood and expert classification, it was found that the overall accuracy, the overall *k* statistic, the producer's and user's accuracy, and the *k* statistic of every land cover class of the expert classification were greater than those of the maximum likelihood classification. However, this excluded the deciduous forest class where the user's accuracy of the maximum likelihood classification was slightly higher than the expert classification. The details of the overall accuracy and overall Kappa statistic of the expert and maximum likelihood classification are shown in table 6.

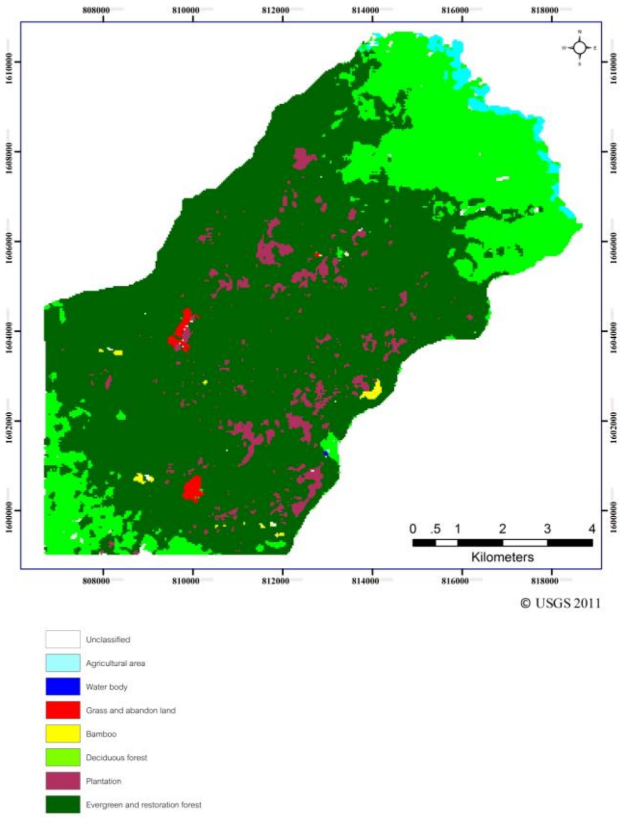

Fig.3 Land cover classification map by the expert classification

TABLE 4 ACCURACY ASSESSMENT OF THE EXPERT CLASSIFICATION

Class	**Reference total**	**Classified total**	**No. Correct**	**PA (%)**	**UA (%)**	***k***
Agr	4	3	3	75.00	100.00	1.0000
Wa	0	0	0	-	-	0.0000
Gr	1	2	1	100.00	50.00	0.4976
Bb	1	0	0	-	-	0.0000
De	38	52	38	100.00	73.08	0.6713
Pl	26	10	9	34.62	90.00	0.8859
Ev	140	143	134	95.71	93.17	0.8112
total	210	210	185			
Overall accuracy = 88.10%, Overall Kappa statistic = 0.7595						

Note:Agr:agricultural area, Wa: water body, Gr: grass and abandon land, Bb: bamboo, De: deciduous forest, Pl: plantation, Ev: evergreen and restoration forest, PA: Producer's accuracy, UA: User's accuracy, *k*: Kappa statistic

4. CONCLUSION AND RECOMMENDATIONS

Expert system allows users to apply the variables for a hierarchical rule for describing the hypotheses or classes. The various variables can be applied to a hierarchy decision tree. This study applied DN, band ratio (FM), PCA, TC, and ancillary data (DEM) to construct a hierarchy decision tree and to create the defined output land cover classes. The result of the overall accuracy is 82.38% and the overall *k* statistic is 0.6425. The producer's accuracy and user's accuracy of grass and abandon land are 0.00% and 0.00%, deciduous forest are

TABLE 5 THE ESTIMATED AREA CALCULATED BY THE EXPERT CLASSIFICATION

Land cover type	Area (hectare)	Percent cover (%)
Agricultural area	105.23	1.32
Water body	0.86	0.01
Grass and abandon land	38.71	0.49
Bamboo	20.70	0.26
Deciduous forest	1,781.64	22.39
Plantation	342.88	4.31
Evergreen and restoration forest	5,668.32	71.22
Total	7,958.34	100.00

TABLE 6 THE COMPARISON OF ACCURACY OF THE EXPERT AND MAXIMUM LIKELIHOOD CLASSIFICATION

Land cover type	Expert classification		Maximum likelihood classification	
	PA (%)	UA (%)	PA (%)	UA (%)
Agr	75.00	100.00	-	-
Wa	-	-	-	-
Gr	100.00	50.00	0.00	0.00
Bb	-	-	-	-
De	100.00	73.08	94.74	73.47
Pl	34.62	90.00	30.77	61.54
Ev	95.71	93.71	92.14	89.58
OA (%)	88.10		82.38	
k	0.7595		0.6425	

Note:Agr:agricultural area, Wa: water body, Gr: grass and abandon land, Bb: bamboo, De: deciduous forest, Pl: plantation, Ev: evergreen and restoration forest, PA: Producer's accuracy, UA: User's accuracy, *k*: Kappa statistic

94.74% and 73.47%, plantation are 30.77% and 61.54%, and evergreen and restoration forest are 92.14% and 89.58%, respectively.

It was found that the percentage of the producer's accuracy, the user's accuracy, and the *k* statistic of every land cover class of expert classification were higher than that of the maximum likelihood. This excluded the user's accuracy of deciduous forest class of the expert classification, which was slightly lower than the user's accuracy of maximum likelihood classification. In addition, the overall accuracy and overall *k* statistic of the expert classification is greater than that of the maximum likelihood classification. The overall accuracy and overall *k* statistic of the expert and maximum likelihood classification were 88.10%, 0.7579, 82.38% and 0.6425, respectively.

The results from this study show that the performing of land cover classification by the expert system can improve the accuracy of the classification. When the accuracy of the classification is higher it provides a more accurate land cover map, and therefore, the land cover map is more reliable for use in specific purposes or with further study. Moreover, it can be applied in every land use/land cover classification project because the extensive variables that are used in a hierarchy decision rules can improve the accuracy of the classification.

ACKNOWLEDGMENT

The author would like to thank the Geo-informatics and Space Technology Development Agency (GISTDA), Thailand for their kindly supported the satellite data in this study.

REFERENCES

[1] Campbell, J. (2001) *Digital image classification*. [Online] Available: http://www.eng.auburn.edu/users/doughmp/Webfiles/Classification_report.pdf

[2] Canada Centre for Remote Sensing (2007) *Fundamentals of Remote Sensing*. [Online]. Available : http://www.nrcan.gc.ca/earth-sciences/geography-boundary/remote-sensing/fundamentals/1430

[3] Coffey, R. (2013) *The difference between "land use" and "land cover"*. [Online]. Available: http://msue.anr.msu.edu/news/the_difference_between_land_use_and_land_cover

[4] Deeudomchan, K., *Land use classification with expert systems method in Phu Khieo Wildlife Sanctuary*. Master's degree Thesis, Kasetsart University, Thailand, 2003.

[5] Department of Environmental Quality Promotion. (2010) *Sakaerat: plantation model*. [Online]. Available: http://www.deqp.go.th/index.php?option=com_content&view=article&id18349%3A2010-08-10-07-20-28&catid=12%3A2010-02-17-11-32-15&Itemid=50&lang=th

[6] ERDAS, *ERDAS Field Guide*, 5th ed, Georgia, USA: ERDAS® Inc., 1999.

[7] Gregorio, A. D. and L. J. M. Jansen (2000) *Land cover classification system (LCCS): classification concepts and user manual*. [Online]. Available: http://www.fao.org/docrep/003/x0596e/X0596e00.htm#P-1_0

[8] Kahya, O., B. Bayram and S. Reis, Land cover classification with an expert system approach using Landsat ETM imagery: a case study of Trabzon, *Environ Monit Assess*, vol. 160, pp. 431-438, 2010.

[9] Lawawirojwong, S., *Expert classification for land cover mapping for Bang Pakong Watershed*. Master's degree Thesis, Mahidol University. Thaiand, 2002.

[10] Levin, N. (1999) *Fundamentals of remote sensing*. [Online]. Available: http://geography.huji.ac.il/personal/Noam%20Levin/1999-fundamentals-of-remote-sensing.pdf

[11] Lu, D. and Q., Weng, A survey of image classification methods and techniques for improving classification performance, *Int. J. Remote Sens*, vol. 28(5),pp. 823-870, 2007.

[12] Ongsomwang, S., *Application of natural color and color infrared aerial photographs in evaluation of land use: its change and impact of Sakaerat Environmental Research Station, Amphoe Pakthongchai, Nakorn Ratchasima Province*. Master's degree Thesis, Kasetsart University, Thailand, 1986.

[13] Placksanoi, J., *The application of remote sensing and geographic information system for study the boundary change of dry evergreen forest and dry dipterocarp forest in Sakaerat Environmental Research Station*. Master's degree Thesis, Nakhon Ratchasima Rajabhat University, Thailand, 2006.

[14] Richards, J. A. and X. Jia, *Remote sensing digital image analysis: an introduction*. Berlin, Germany: Springer, 1999.
[15] Sahunaru, P. and T. Wisarat, Plant ecology in Sakaerat Environmental Research Station; in: *Sakaerat Environmental Research Conclusion*, Khemnak C. (ed.), pp. 55-79, Thailand, 1983.
[16] Sa-ngawongse, S., *Analysis of digital remotely sensed data: teaching document of GEO 154773*. Chiangmai, Thailand: Chaing Mai University Press, 2007.
[17] Sa-ngawongse, S., *Remote sensing for land-use/land-cover monitoring and application*. Bangkok, Thailand: V Print Ltd., 2009.
[18] Stefanov, W. L., M. S. Ramsey and P. R. Christensen, Monitoring urban land cover change: An expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sens Environ*, vol. 77, pp.173-185, 2001.
[19] Tso, B. and P. M. Mather, *Classification Methods for Remotely Sensed Data*. London, England: Taloy & Francis, 2001.
[20] Taweesuk, S. and P. Thammapala. (2005) Expert classification technique for mapping teak plantation areas in Thailand. [Online]. Available: http://www.asprs.org/a/publications/proceedings/pecora16/Taweesuk_s.pdf.
[21] Watcharakitti, S., P. Intrachundra and S. Mungkorndin, Natural resources and land use study of Sakaerat Environmental Research Station. *Forestry Research Report*, vol.69, 1980.