INT. J. REMOTE SENSING, 2001, VOL. 22, NO. 10, 1909–1918

Mapping land use/cover distribution on a mountainous tropical island using remote sensing and GIS

SERWAN M. J. BABAN

Department of Surveying and Land Information, The University of the West Indies, St Augustine, Trinidad, West Indies

and KAMARUZAMAN WAN YUSOF

School of Engineering, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia

(*Received 13 May 1999; in final form 22 March 2000*)

Abstract. The land use/cover distribution on Langkawi Island, Malaysia was mapped using remote sensing and a Geographic Information System (GIS). A Landsat Thematic Mapper (TM) satellite image taken in March 1995 was processed, geocorrected and analysed using IDRISI, raster-based GIS software. An unsupervised classification was performed based on spectral data from a composite image of the bands TM3, TM4 and TM5. Using this output, field data together with available secondary data consisting of topography, land use and soil maps were used to perform a maximum likelihood supervised classification. The overall accuracy of the output image was 90% and individual class accuracy ranged from 74% for rubber to 100% for paddy fields. The classified areas on the image were mainly confined to the mountainous and hilly regions on the island. A shaded relief map, simulating sunshine conditions, showed that the unclassified areas are located in the shadowed slopes, i.e. the slopes facing west. Consequently, the imagery was subdivided on the basis of slope aspect and a stratified classification was performed. As a result of this procedure, the overall accuracy increased to 92% and the individual class accuracy for the inland forest class increased by 9% to 90%. Using IDRISI, individual class areas as well as percentages were calculated. The kappa coefficient for the classified image was 0.90. Qualitative analysis indicates that topography is the main control on the spatial distribution of land use/cover types on the island. As Langkawi Island has been developing rapidly over the last decade, successful planning will require reliable information about land use/cover distribution and change. This study illustrates that remote sensing and GIS techniques are capable of providing such information.

1. Introduction

Economic prosperity in Malaysia over the past three decades has generated a significant demand for all forms of recreational activities (Draper 1994). This and the declaration of the Langkawi Island as a duty free port in 1987 have made the island the focus for recreational activities in Malaysia. Langkawi Island is currently among the most popular tourist destinations: from 1988 to 1989 the number of tourists visiting the island increased by about 31% (Jusoff and Hassan 1996).

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2001 Taylor & Francis Ltd
http://www.tandf.co.uk/journals

Subsequently, the island has been undergoing rapid development in terms of expanding road networks and the construction of hotels and holiday complexes.

Land use/cover classification and evaluation surveys using remote sensing have been conducted successfully for several areas in the tropics (Wu *et al.* 1985, Kachhwaha 1989, Ratanasermpong *et al.* 1995). In Malaysia, Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM) data have been used for land use surveys and the results have showed that it is possible to map various natural land cover types and man-made features, including terrain, forest, soil, dams and urban centres (Salleh 1976, Mahmood *et al.* 1983). The resolution of the MSS, however, was found to be unsuitable for mapping Malaysian agricultural land utilization due to small farm sizes and irregular cropping patterns (Darus 1989). Landsat imagery and GISs have also been used to detect land use changes and to identify suitable areas for tourism-related development in Malaysia (Azmi and Shattri 1989, Kam and Foo 1989, Ibrahim *et al.* 1993).

This paper aims to:

(i) map land use/cover distribution in a mountainous tropical environment, the Langkawi Island, using remote sensing and GIS;
(ii) qualitatively and quantitatively evaluate the outcome and consequently the methodology for providing the necessary information for accommodating development with minimum impact on the environment.

2. The study area

The group of Langkawi Islands are made up of 104 islands which lie off the north-west coast of Peninsular Malaysia and are located approximately at latitude 6° 20′ N and longitude 99° 40′ E. The largest of the group is Langkawi Island, which has an area of 390 km^2 and houses 93% of the total population of the islands (figure 1). Its climate is influenced by the two wet monsoons occurring in April/May, when the monthly rainfall average is over 180 mm, and during August/October when the average reaches over 220 mm. The annual average rainfall is between 1600 and 2000 mm. The daily average temperature is about 31° C. Langkawi Island is mainly covered by forest, mangroves, shrub, agricultural land, sandy beaches and newly expanding developed urban areas. The topography varies from flat coastal plains, through hilly areas, to rugged mountains rising to 870 m above sea level at Gunung Raya (Jusoff and Hassan 1996).

The geology of Langkawi Island consists of quartzite sandstone, the oldest strata, in the north-western part overlain by limestone (marble) which outcrops in the eastern parts. This formation is overlain by shale or silty shale which outcrops in the east of Gunung Raya. The youngest rock formation is a massive dolomitic limestone. In addition, granite forms the highest peak (Gunung Raya) and sediments, mainly of marine origin, cover the south-western part, around Kuah town (figure 1). Alluvium was developed for residential or agricultural land. However, the alluvium area in the north-east of the island is relatively untouched and covered by thick mangrove (Abdullah and Akhir 1997).

3. Data and methodology

A Landsat TM satellite image for Langkawi Island, acquired on March 1995, was analysed using IDRISI, a raster-based GIS software program. The image was atmospherically corrected using the darkest pixel approach and geocorrected with a

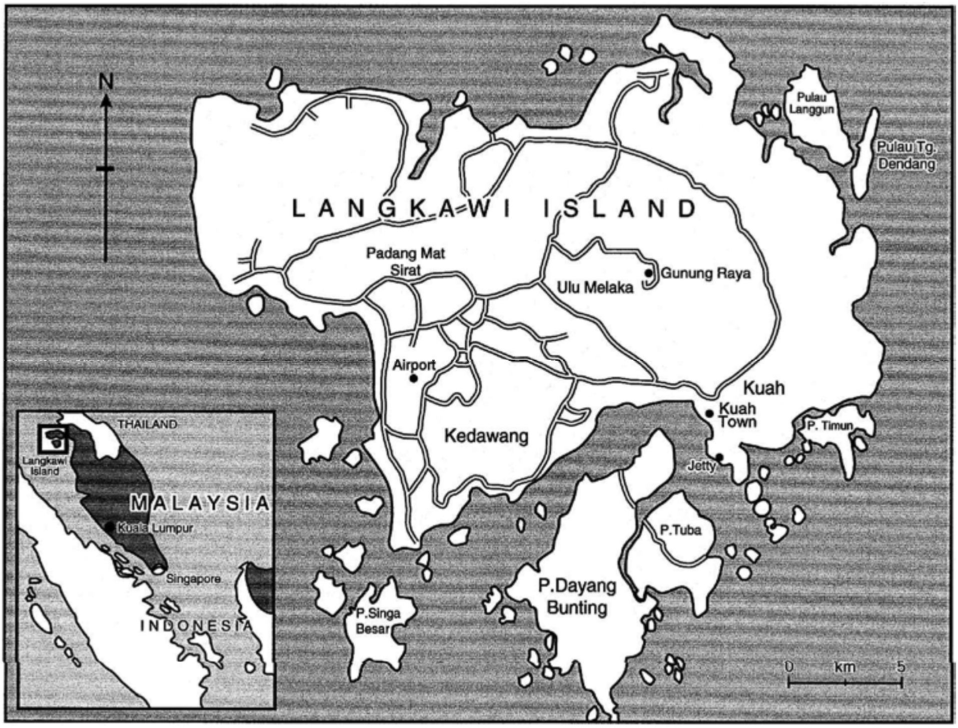

Figure 1. Study area.

residual mean square (RMS) error of less than half a pixel. Using IDRISI with a composite image of the bands TM3, TM4 and TM5, an unsupervised classification of land use/cover was performed on the basis of spectral signatures for nine clustering areas (figure 2).

Available secondary data to support this work consisted of a 1990 topographic map at scale 1550 000 issued by the Department of Survey and Maps, Malaysia, a 1990 land use map at scale 15150 000, and a 1986 soil map at scale 15100 000, both obtained from the Ministry of Agriculture, Malaysia. Based on figure 2 and all available secondary data, a field work programme was organized during August 1998, during which extensive ground truth data, totalling 225 sample points were collected.

The following land use/cover types were selected as representative classes in a supervised classification; inland forest, mangrove forest, rubber, paddy fields, mixed horticulture, grassland, urban areas and water bodies. A minimum of 30 pixels were selected in the training areas for each class. The training areas were selected on the basis of the unsupervised classified image (figure 2), primary field data and the secondary data. Using IDRISI, the polygons around each training area were digitized and assigned a unique identifier for each cover type (Eastman 1997); a spectral signature file for each class was subsequently created. These signature files were used to categorize the continuum of spectral data in the entire image by using a maximum likelihood classifier involving bands TM3, TM4 and TM5 (figure 3). All relevant TM bands had been examined for autocorrelation and bands TM3, TM4 and TM5

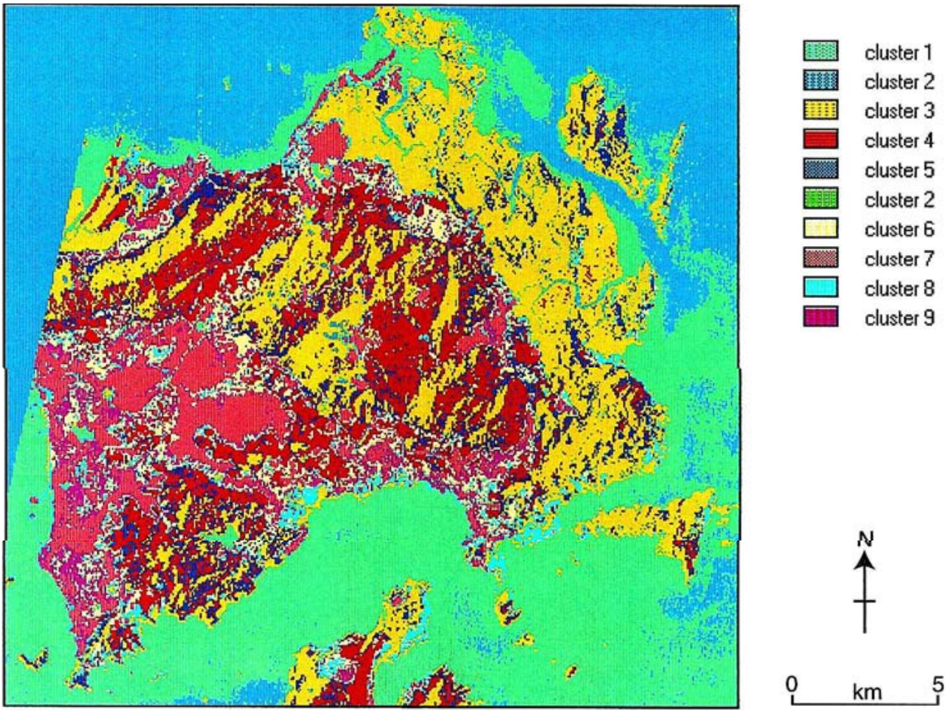

Figure 2. Unsupervised classification based on spectral data.

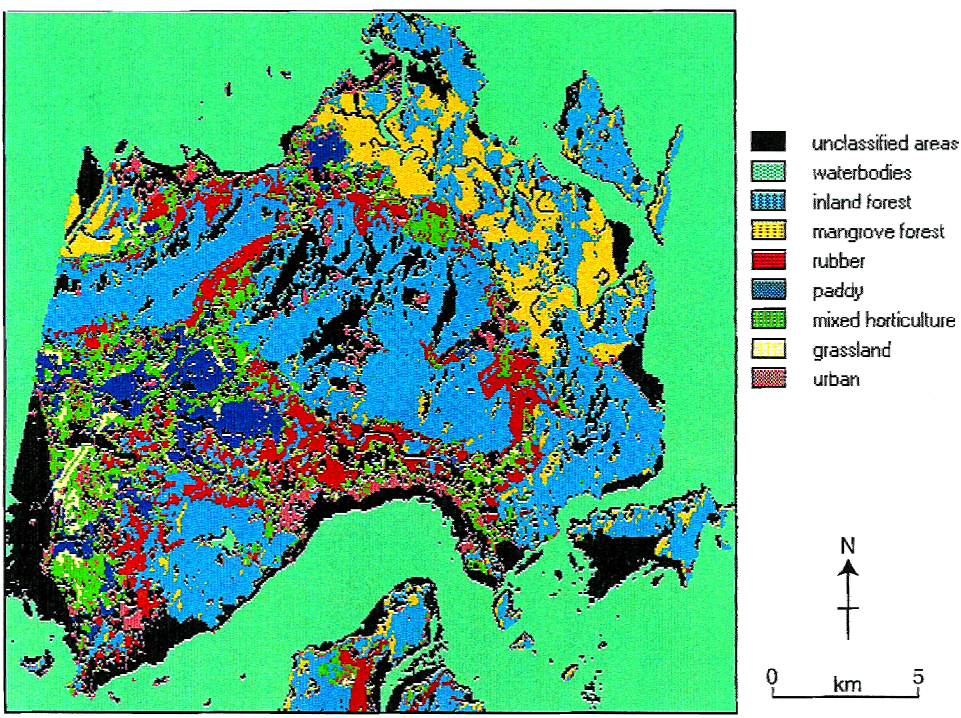

Figure 3. Supervised maximum likelihood classification of the island.

emerged as having the least similarity in terms of information content. These bands were consequently used in the classification process.

The thematic content of the classified image was quantitatively assessed for accuracy by evaluating the correspondence between the class label assigned to a pixel in the image and the 'true' class as measured on the ground (figure 3). This evaluation is often based on a statistical evaluation of a sample population of the classified image (Janssen and Van der val 1994). In this study, a stratified random sampling method was used to generate 1000 sample points on the classified image using IDRISI. Subsequently, a direct comparison of classes was carried out based on geographical location, with the available primary and secondary data (table 1). Overall, most of the classes were easily separated and mapped due to their distinctive spectral reflectance signatures within the TM bands used, and their confinement to different parts of the island was based mainly on topography. The overall accuracy achieved was 90%. Individual class accuracy ranged from 74% for rubber to 100% for paddy and grassland (table 1).

The output image (figure 3) shows a large number of unclassified areas, mainly concentrated in the inland forest area in the mountainous region. This is a familiar problem and is caused by shadowing, which is probably the major constraining factor for modelling vegetation in mountainous and hilly areas using remotely sensed data (Garcia and Murguia 1996, Jusoff and Hassan 1996). In order to confirm that this is the case in this study, a shaded relief map of the island was created with a Sun azimuth of 315° and an altitude 45° to cast a shadow over the DEM landscape (figure 4). This configuration aims to simulate the state of illumination conditions on the imagery at the time of data capture (Benson and Greene 1996). Comparison of figures 3 and 4 confirms that the unclassified areas are located in the shadowed slopes facing the west.

One way of solving this problem is by readjusting the radiance pixel values on the basis of aspect and slope (Howard 1991). In practical terms this can be achieved by conducting a stratified classification by dividing the imagery into subscenes based

Table 1. Landsat TM data accuracy assessment (key: 1, waterbodies; 2, inland forest; 3, mangrove forest; 4, rubber; 5, paddy; 6, mixed horticulture, 7, grassland; 8, urban area; 0, unclassified).

TM DATA

		1	2	3	4	5	6	7	8	0	Total	Accuracy (%)
G												
R	1	296	1								297	99
O	2		196	3						421	241	81
U	3		3	65							68	96
N	4		14	5	77	2		6			104	74
D	5					42					42	100
	6		4	1	5		73				83	87
D	7							8			8	100
A	8		2		1	2			46		51	90
T	0											
A												
	Total	296	220	74	83	46	73	14	46	42	894	

Overall accuracy: $(296+196+65+77+42+73+8+46)/894=0.898\%$

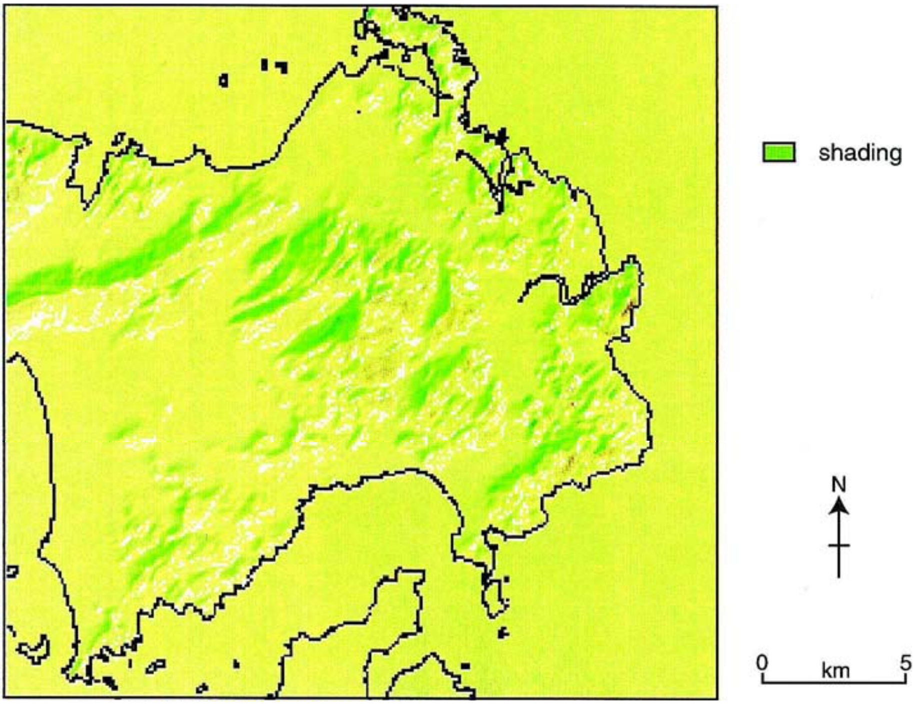

Figure 4. A shaded relief map of the island.

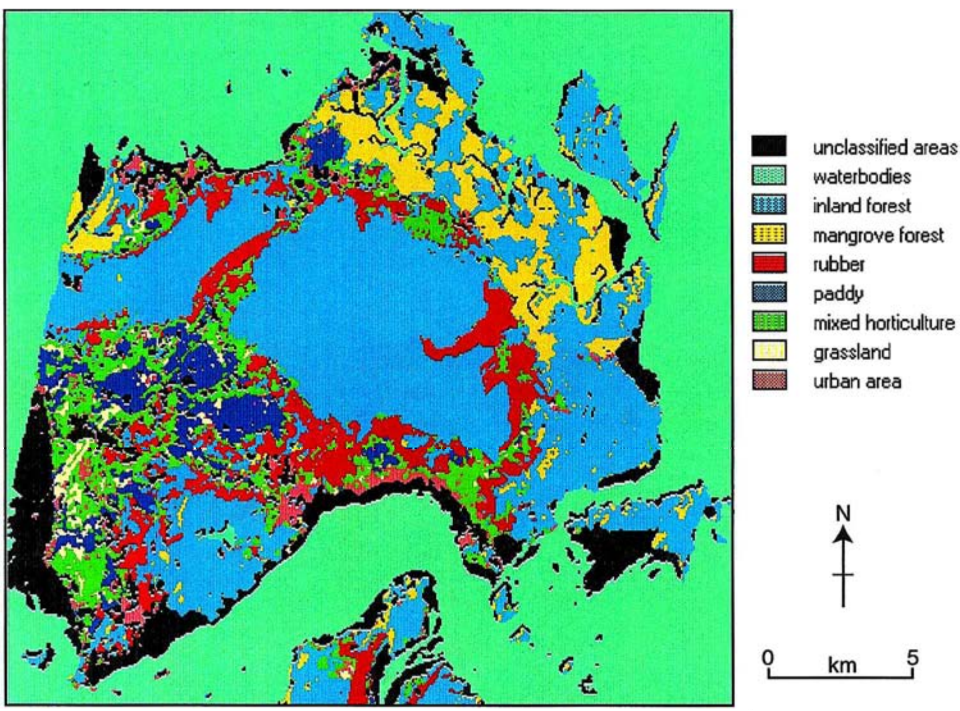

Figure 5. The final outcome of the stratified classification.

on slope aspect, performing classifications on individual subscenes, and combining the results into one image (Garcia and Murguia 1996). This methodology was adopted in this study. Training areas for the inland forest class located on west-facing slopes (based on the aspect map) were taken and the maximum likelihood supervised classification was performed. The output was then overlaid on figure 4 by using the *Cover* option in IDRISI (figure 5). This option allows the pixels of the second image to be covered with those of the first, except where the first image has values of zero. In these cells, the value of the second image shows through (Eastman 1992). A re-count of samples located in the inland forest areas showed that this process had increased the number of points being correctly classified from 196 to 217 and decreased the number of unclassified pixels from 42 to 21. As a result, the classification accuracy for the inland forest class increased from 81% to 90% and the overall image accuracy increased from 90% to 92%. The Kappa coefficient, κ, which is the difference between the observed agreement between the two sets of sample point data and the agreement that might be contributed by chance matching the two datasets was calculated (Campbell 1996) to be 0.90. Using IDRISI, the areas associated with each category in the classified image, as well as their percentages of the total area, were calculated (table 2).

4. Results and discussion

During field work it was evident that topography was the main influence on the distribution of land use/cover on the island. Examining the classified image reinforces this observation. Low-lying areas of alluvium with plentiful amounts of fresh water in the south-western part of the island are dominated by paddy fields. However, a small area of paddy also exists on the alluvium in the north, which is only accessible by salty water and covered by thick mangrove forest. This area is an exception to the general pattern and might be sustained by fresh water draining from the mountains and the edge of the paddy could represent the beginning of brackish water in this area. The lowlands also contain mixed horticulture, beaches and urban settlements. Rubber seems to exist on the hills and hillsides whilst the highlands are mainly occupied by inland forest.

There were some difficulties in distinguishing between inland forest and mangrove in certain locations. This might be attributed to the shadow effect on inland forest at those locations (Jussof and Hassan 1996). Some difficulties were also encountered in the identification of urban areas, for two possible reasons. First, their existence in small spatial units produces mixed class covers with the rubber and paddy units, which exist nearby. Second, traditional urban units are built of the local wood and

Table 2. Land use/cover areas and percentage area of the island based on the classified image.

Land use/cover	Area (km^2)	Area (%)
Inland forest	154.20	52.3
Mangrove forest	30.86	10.5
Rubber	35.55	12.1
Paddy	19.61	6.6
Mixed horticulture	34.81	11.8
Grassland	5.93	2.0
Urban area	13.57	4.6

they do not display great spectral contrast from the surrounding rubber units. Rubber and paddy classes were also difficult to separate in a few locations. Other problems included some misclassification of small units of rubber estates existing in close proximity to inland forest. Some unclassified and misclassified areas could have been caused by land use changes due to developments between the 1995 satellite image and the 1998 field survey.

The quantitative analysis indicated that the overall accuracy of the classified image is 92%. In terms of areas occupied by each land use/cover and their class accuracy, inland forest represents 52.3% of the area with a class accuracy of 90%, followed by rubber, mixed horticulture, mangrove, paddy, urban and grassland which represent 12.1%, 11.8%, 10.5%, 6.6%, 4.6% and 2.0% in terms of area respectively (the respective individual class accuracies are 74%, 87%, 96%, 100%, 90% and 100% (tables 1 and 2)). The Kappa coefficient, κ, was calculated to be 0.90, which indicates that the classification has achieved an accuracy that is 90% better than what would be expected from random assignment of pixels to land use/cover types on the island.

5. Conclusion

Langkawi's rapid development has created a gap in knowledge regarding the pattern, spatial distribution and actual area occupied by various land use/cover types on the island. Such information is essential for planning and implementing policies to optimize natural resources and accommodate development whilst minimizing the impact on the environment. It is of great importance to gather and provide this dynamic information to decision makers on a regular basis.

In mountainous and hilly areas such as Langkawi Island, shadowed slopes represent a major challenge to digital processing of satellite data as they generate unclassified areas. This problem can be solved by creating a shadow relief map using a GIS and performing a stratified classification based on the slope aspect. This approach improved the overall accuracy of mapping land use/cover based on the statistical evaluation from 90% to 92% whilst the individual accuracy of the inland forest class improved from 81% to 90%. Individual class accuracy for other classes in the classified image ranged from 74% for rubber to 100% for paddy and grassland. The Kappa coefficient for the classification was 0.90. The area occupied by each class was also calculated. Inland forest is the dominant class, covering more than half of the island, followed by rubber, mixed horticulture, mangrove, paddy, urban and grassland.

The main factor controlling land use/cover distribution on the island is topography and the qualitative analysis of the classified image seems to reinforce this observation. Low areas are dominated by paddy, mangrove forest, mixed horticulture, grassland, beaches and urban settlements, hills and hill slopes are covered by rubber, and higher grounds are occupied by inland forest.

These results clearly suggest that the spectral and spatial characteristics of Landsat TM data could serve to identify and map land use/cover types in tropical mountainous environments. The success of the TM data in this application could be due to a number of reasons. Firstly, the spectral characteristics of land use/cover types are distinctive enough to be used for the identification and separation of individual types. Secondly, the spatial distributions of different land use/cover types on the island are confined to separate zones based on topography. Thirdly, the previous two factors result in the production of a small number of mixed pixels.

Finally, the averaging effect on recorded spectral reflectance, provided by the spatial resolution of 30 m by 30 m seems to decrease internal variations within each single class. As a consequence, identifiable classes with uniform radiance values are produced.

The results also suggest that a raster-based GIS can facilitate the necessary digital analysis and manipulation. This includes data integration, geocorrections, introducing information and knowledge from other datasets into the classification process, handling the classification, performing statistical accuracy tests and calculating areas.

In conclusion, satellite remote sensing and GISs can be used to generate the necessary dynamic information for surveying and monitoring land use/cover on Langkawi Island and similar success may be possible in other mountainous tropical environments

Acknowledgments

We thank the Malaysian Centre for Remote Sensing (MACRES) for kindly providing the Landsat TM imagery and Professor Ian Foster for his helpful comments.

References

Abullah, I., and Akhir, J. M., 1997, The correlation of Landsat TM images characteristic with the Geological Information: preliminary results of a case study in Langkawi Islands, Malaysia. *Proceedings of the 18th Asian Conference on Remote Sensing, Kuala Lumpur, Malaysia, 20–24 October 1997* (Kuala Lumpur: MACRES), pp. J.5.1–J.5.7.

Benson, J., and Greene, B., 1996, Shaded relief images: visualising terrain without contour lines. In *Raster Imagery in Geographic Information Systems*, edited by S. Morian and S. L. Baros (Sante Fe, NM: Onward Press), pp. 192–198.

Campbell, J. B., 1996, *Introduction to Remote Sensing*, 2nd edition (London: Taylor and Francis).

Darus, A., 1989, Application of Landsat MSS data in land use mapping: a Malaysian experience. *Proceedings of the 10th Asian Conference on Remote Sensing, Malaysia, Kuala Lumpur, 23–29 November 1989* (Kuala Lumpur: MACRES), pp. G-8-1–G-8-8.

Draper, W. H. (editor), 1994, *Human Developments Report, United Nations Development Programme* (New York. Oxford University Press).

Eastman, J. R., 1992, *Technical Reference in IDRISI Version 4.0*, Clark University, Worcester, MA, USA, 212 pp.

Eastman, J. R., 1997, *Supervised Classification in IDRISI for Windows Version 2, Tutorial Exercises* (Worcester, MA: Clark University), Ch. 12, pp. 86–94.

Garcia, D. F. L., and Murguia, R. G. G., 1996, Modeling vegetation distribution in mountainous terrain. In *Raster Imagery in Geographic Information Systems*, edited by S. Morian and S. L. Baros (Sante Fe, NM: Onward Press), pp.165–171.

Howard, J. A., 1991, *Remote Sensing of Forest Resources, Theory and Applications* (London: Chapman and Hall).

Ibrahim, S. M., Shamsudin, A., and Adel, A., 1993, Spatial data analysis using a digital image processing system with GIS capabilities: case study location of potential beach tourist sites. *Paper presented at the Regional Remote Sensing Seminar on Tropical Ecosystem Management*, *Kuching, Sarawak, Malaysia*, *8–11 September 1993*.

Janssen, L. L. F., and Van der Wal, J. M., 1994, Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering and Remote Sensing*, **60**, 419–426.

Jusoff, K., and Hassan, H. Md., 1996, Forest recreation planning in Langkawi Island, Malaysia, using Landsat TM. *International Journal of Remote Sensing*, **18**, 3599–3613

Kachhwaha, T. S., 1989, Wildlife habitat assessment and forest mapping in Corbett National Park, U.P. India, using Landsat Thematic Mapper data. *Proceedings of the 10th Asian Conference on Remote Sensing, Kuala Lumpur, 23–29 November 1989* (Kuala Lumpur: MACRES), pp. Q-22-1–Q-22-6.

Kam, S. P., and Foo, L. K., 1989, Detection of land use changes using remote sensing and GIS. *Proceedings of the 10th Asian Conference on Remote Sensing, Kuala Lumpur, 23–29 November 1989* (Kuala Lumpur: MACRES), pp. G-2-1–G-2-7.

Mahmood, N. N., Bruneau, M., and Lemen, H., 1983, A pilot study on the use of satellite remote sensing data for agroecological mapping of peninsular Malaysia. *MARDI Special Report*, **2**, 31.

Ratanasermpong, S., Pornprassertchai, J., and Disbunchong, A., 1995, Natural resources and land use changes of Phuket using remote sensing. *Proceedings of the 16th Asian Conference on Remote Sensing, Thailand, 20–24 November 1995* (Thailand: Nakkon Ratchasima), pp. Q-13-1–Q-13-6.

Salleh, M. W., 1976, Preliminary interpretations of ERTS imagery of Peninsular Malaysia. *Malaysian Forester*, **39**, 189–213.

Wu, J. K., Zheng, Q. F., Wang, W. T., and Li, K. W., 1985, A forest inventory using Landsat Imagery in Mao-Shan area on China. *International Journal of Remote Sensing*, **6**, 1783–1796.