Modern Applied Science
www.ccsenet.org/journal.html
Vol. 3, No. 1
January 2009

Mapping and Quantification of Land Area and Cover Types with Landsat^{TM} in Carey Island, Selangor, Malaysia

Hj. Kamaruzaman, J (Corresponding author)
Yale University, Yale's School of Forestry & Environmental Studies
205 Prospect St, New Haven, CT 06511
Tel: 20-3-432-1384 E-mail: jusoff.kamaruzaman@yale.edu
Mohd Hasmadi, I
Department of Forest Production, Faculty of Forestry
Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia
Tel: 60-3-8946-7220 E-mail: mhasmadi@hotmail.com

Abstract

Information about current land cover type is essential at a certain level to ensure the optimum use of the land resources. Several approaches can be used to estimate land cover area, where remote sensing and Geographic Information System (GIS) is among the method. Therefore, this study was undertaken to evaluate how reliable these technologies in preparing information about land cover in Carey Island, Selangor of Peninsular Malaysia. Erdas Imagine 9.1 was used in digital image processing. A primary data of Landsat TM, with spatial resolution of 30 m was acquired from scene 127/58 on July 2007. Area estimate was calculated using direct expansion method from samples proportion of each segments of land cover type (1 km by 1 km sample size). In this study, four classes of land cover type have been identified and the areas were oil palm, mangrove, water bodies and urban/bare land area. The area estimate for all classes are 11039.28 ha (oil palm), 5242.86 ha (mangrove), 4894.92 ha (water bodies), and 4751.96 ha (urban/bare land), respectively. The overall classification accuracy obtained for this study is 96%. The results showed that the use of direct expansion method for estimating land cover type area is practical to be used with remote sensing approaches.

Keywords: Land covers mapping, Satellite-based sensing, Direct expansion method

1. Introduction

Knowledge of land cover and land use is significant for management and planning activities concerned with the parcel of the earth. Land use is one of the essential tools for nearly all land development. Information about land cover type is essential to be gathered due to it important to understand the landscape process and generally refer to the functional character of an area with value to the occupations and culture. In traditional way, classification of land cover type and estimate their area requires a relatively time consuming and very costly. However, area estimation survey is very important in managing our land parcel and natural resources because it can facilitate land manager in land use planning. Most of time it is difficult to estimate land cover types when dealing with large area and result is requires in short time. Nowadays, remote sensing is one of the technology tools to provide advance complete information about land circumstance and human activities. The importance of remote sensing has been proved through the time serve as a useful tool in management of many types of natural resources. Land use managers planners are increasingly using remote sensing demanding that information for the decision making process (Junior and Simonett, 1976). In the beginning, the use of panchromatic, medium scale aerial photographs to map land use and land cover has been accepted practice since 1940s. Later on, small scale aerial photographs and satellite images have been utilized for land use and cover mapping of large areas (Lillesand and Kiefer, 1994).

Malaysia has started using remote sensing data such as SPOT and Landsat TM imageries for land cover mapping and land use change detection since early 1980's (Fook *et. al.,* 1994). These space borne data mainly serve as the basis for further improvement of existing procedures for land use data collection in the department of agriculture. In fact, according to Kamaruzaman and Souza (1997) the use of remote sensing data was started in 1961 by using extensive aerial photography for demarcated the upper altitudinal limit of forest land suitability for inclusion in forest state production. A black and white aerial photograph was use widely in forest resource survey in Peninsular Malaysia in 1962 which involved the systematic assessment for management purposes. Remote sensing is a less costly over large area and a more rapid means of data collection than field observation and some detail verification data are essential to its effective use. On the other hand, remote sensing is a practical solution when area estimation is conducted using direct expansion method. Crucial needs for land cover inventory in many developing countries have already prompted the use of remote sensing and GIS to provide up-to-date information (Luney and Dill, 1970). Thus, remote sensing is often the most commercial and valuable source of current land cover and land use planning. Previously land cover area estimate by integration of sampling techniques and satellite images has been performed with high accuracy (Taylor et al., 1997; Ferencz et al.(2004).

The methodology utilized in this study is based on the use of remote sensing data and direct expansion method. Remote sensing can be valuable tools for land cover area estimate when area frame are used. Typically remote sensing helps in the definition of sampling unit and can be exploited to optimize the sample allocation and size of sampling units. In area estimation, classified remotely sensed images are used as auxiliary variables in a direct expansion estimator. This effort was demonstrated for classification and land cover statistics (Mohd Hasmadi and Kamaruzaman, 2007).

The aim of the study is to investigate the usefulness of remote sensing data for estimating land cover types at Carey Island, Peninsular Malaysia using direct expansion method, therefore the specific objectives of the study are to classify and map land cover types of the island using Landsat TM image.

2. Methodology

2.1 Description of study area

<<Figure 1. Location of the study area>>

2.2 Data acquisition and equipments

Data acquisition involved the satellite data and secondary data. Primary data was Landsat TM which taken on July 2007 for 127/58 (path / row) and acquired with spatial resolution of 30 m. Satellite imagery was obtained from Malaysia Center for Remote Sensing (MACRES). The secondary data such as 1:50000 scaled topographical map from Department of Mapping Malaysia, dated 1990 was used as reference to conduct ground verification and image interpretation process. Other equipments were used such as the Digital Plannimeter (Koizumi KP 90N) to measure the area on topographical map, digital camera, and Global Positioning System (GPSmap 60CSx) for navigate point location during ground verification. The image processing and analysis was carried out using Erdas Imagine software version 9.1, while ArcView Version 3.3 was used for the GIS analysis.

2.3 Methods

Basically several steps were carried out in order to prepare satellite imagery to be used for classify and estimate area coverage of land cover types. Figure 2 illustrated the flow chart of the research process.

<<Figure 2: The flow chart of the study>>

Preprocessing involved on the imagery was geometric and radiometric correction. In geometric correction the Malaysia Rectified Skew Orthomorphic(MRSO) projection type (Spheroid name:Modified Everest and Datum name:Kertau 1948) with 30m pixels using a cubic convolution was used as a grid base system. Utilizing 8 ground control points, the final projected imagery accurate to RMSE of the registration was less than 1.0 m. De-stripping technique was done to remove radiometric problem on the image, thus improved the quality of the image for further interpretation and analysis.

Image enhancement was performed to edit the original image data by improving quality of the data for visual interpretation for creating 'new' image. Technically, image enhancement operations distort the original digital values of image data to more effectively display. Modifying contrast and brightness level to 55 and 58 percent was adequate to enhance the quality of the data for ease information extraction. However, there is no ideal or best image enhancement because the results are ultimately evaluated by humans, who make subjective judgments as to whether a given images enhancement is useful (Jensen, 1996). The original spectral band and artificial spectral band are used to create another value to the image to create different composite effects and increased information on land cover. For this study, enhancement image from 5-4-3 (R-G-B) band was the best to be used for visual interpretation and digital image classification. Image classification is the process of assigning the individual pixels of various spectral signatures to categories based on the reflectance of spectral characteristic. In this study unsupervised classification was chosen instead of supervised classification technique. At the beginning, the numbers of classes assigned for the unsupervised classification were 7 classes. Amalgamation was done to reduce the class number to 4. During the process, the maximum iterations were set to 60 times.

Ground verification was performed to verify features retrieved from the satellite images to the features on the ground. It was conducted for 2 days on 12 -13 September 2007. A total of 52 locations were visited randomly. GPS was used to assist in navigate the point location. Photographs were taken and parameters related to land cover types were recorded. All collected data obtain from ground verification were used to determine the accuracy of mapping. The data were organized in a confusion or error matrix from which both the overall classification accuracy on the individual classes can be calculated. Kappa coefficient (K) measurement was used to assess the accuracy of the classification. Kappa analysis is discrete multivariate technique of use in accuracy assessment (Congalton and Mead, 1983). It was calculated by multiplying the total number of pixels in all the ground verification classes (N) by the sum of the confusion matrix diagonals (X_{kk}), subtracting the sum of the ground verification pixels in a class time the sum of the classified pixels in that class summed over all classes ($\sum X_{k\sum} X_{k\sum}$), and dividing by the total number of pixels squared minus the sum of the ground verification pixels in that class times the sum of the classified pixels in that class summed over all class. The calculation of kappa statistic is as follow.

$$\kappa = \frac{\theta_1 - \theta_2}{1 - \theta_2}$$

where,

$$\theta_1 = \frac{\sum_{i=1}^{r} x_{ii}}{N}$$

$$\theta_2 = \frac{\sum_{i=1}^{r} x_{i+} x_{+i}}{N^2}$$

xij = number of counts in the ijth cell of the confusion matrix

N = total number of counts in the confusion matrix

xi+ = marginal total of row i

x+i = marginal total of column i

Calculation of the variance of Kappa:

$$\mathrm{var}(\kappa) = \frac{1}{N}\left[\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_2)^4}\right]$$

where,

$$\theta_1 = \frac{\sum_{i=1}^{r} x_{ii}}{N}$$

$$\theta_2 = \frac{\sum_{i=1}^{r} x_{i+} + x_{+i}}{N^2}$$

$$\theta_3 = \frac{\sum_{i=1}^{r} x_{ii}(x_{i+} + x_{+i})}{N^2}$$

$$\theta_4 = \frac{\sum_{i=1, j=1}^{r} x_{ij}(x_{j+} + x_{+i})^2}{N^3}$$

xj+ = marginal total of row j

3. Sampling design and direct expansion method

Systematic random sampling was selected for sampling design where sample plots were distributes evenly to all parts of the target area. A total of 39 sample segments were selected using systematic sampling with 2 km by 2 km of size and the size of sample is about 1 km by 1 km. Later on, from sample segments a total of 39 sample unit were selected randomly. Each sample point represented an area corresponding to the size of the grid cell of the sample layout. For example, when sample points were selected randomly from a square systematic grid with 1000 meters distance between the points, each sample point represents an area of 1 km by 1 km. The percentage of sample proportion obtained from samples unit was 4.6 percent. Figure 2 illustrated the sample segment and sample unit of the study area. The value from each digitized of sample unit in each sample segment were then calculated using direct expansion equation to produce the estimate of the land cover area.

<<Figure 3. A systematic random design for 39 samples and unit segments over Carey Island.>>

Calculation of the area estimates has been done by proportions rather than as absolute areas because the proportion is automatically takes account of errors resulting from small localized variations in the segment scale and digitizing errors (Taylor et al., 1997; Mohd Hasmadi and Kamaruzaman, 2007). The unbiased estimate of land cover proportion of land area covered by class c is given by the equation:

$$\overline{y}_c = \frac{1}{n}\sum_{i=1}^{n} y_i$$

with variance v $Var(\overline{y}_c) = (1-\frac{n}{N})\frac{1}{n(n-1)}\sum_{i=1}^{n}(y_i - \overline{y}_c)^2$

where: y_i is the proportion of segment *i* covered by class *c*; *N* = total number of segments in the region, *n* = number of segments in the sample. The proportion of the study region sampled (n/N) is referred to as the sample fraction. When this is less than 5%, the correction factor for a finite population $(1-\frac{n}{N})$ can be omitted from the above formula (Cochran, 1977). The estimate of the class area is:

$$\hat{Z}_c = D\overline{y}_c$$

with variance: $Var(\hat{Z}_c) = D^2 Var(\overline{y}_c)$

where, *D* is the area of the region.

The standard error or accuracy of $\hat{Z}_c$ is estimated by calculating the 95% confidence interval as follows:

$$\hat{Z}_c \pm 1.96\sqrt{\mathrm{var}(\hat{Z}_c)}$$

4. Result and discussion

4.1 Image classification

In this classification, the land cover type was classified and divided into several classes based on natural grouping of the original Landsat TM data. The result of unsupervised classification technique from composite image of 5-4-3 (R-G-B) band produced four land cover classes namely oil palm, mangrove, water bodies and urban/bare land, respectively. The unsupervised classification image was shown in Figure 4.

<<Figure 4. Land covers type within the Carey Island using unsupervised technique>>

4.2 Area estimate calculation

Using the direct expansion equation, the land cover area estimates were calculated. Area proportion and total area estimate of each land cover class was shown in Table 1. The table includes standard error, coefficient of variation and 95% confidence interval. From the table it showed that area estimation results are as follows; oil palm (11039.28 ha.), mangrove (5242.86 ha), water bodies (4894.92 ha), and urban/bare land (4751.96 ha). Meanwhile, Figure 5 showing the extent of the area estimates in hectare and percent.

<<Table 1. Area proportion and total area estimate of each class by direct expansion method>>

<<Figure 5. Estimated land cover area (in ha and %) of Carey Island>>

4.3 Accuracy assessment

During accuracy assessment process a total of 52 random sampling points were selected in order to determine the accuracy for image classification output (Figure 6). The result of confusion matrix was expressed in tabular form as shown in Table 2. The overall accuracy achieve was 96.97% which is higher than qualification accuracy of a digital classified image of 80% using optical data suggested by Paul (1995). From the table most of the land cover types were classified correctly, but oil palm was confused with mangrove and urban/bare land. In Kappa analysis, kappa coefficient result is 0.95. Table 3 showed the producers and users accuracy of classified image of Carey Island, where the highest producer's accuracy was oil palm (100.00%) and followed by other land cover class with water bodies (96.30%), mangrove (95.65%), and urban/bare land (87.50%). Meanwhile, mangrove and water bodies classes were the highest of the user's accuracy (100.00%), oil palm (94.74%) and urban/bare land had the lowest accuracy (93.33%) due to false classification during classification process.

<<Figure 6: Selected ground verification points in Carey Island (fillet black)>>

<< Table 2. Confusion matrix of the classification image of Carey Island>>

<< Table 3. Producer's and user's accuracy of the classification image of Carey Island>>

5. Conclusion

Applying remote sensing data and for land cover classification and direct expansion method for area estimation revealed the benefit for managing land resources. This study concluded that remote sensing and GIS technique is a powerful tool for generating land cover types and estimate land cover area in Carey Island. False color composite of Landsat TM band 5-4-3 (R-G-B) has showed the best combination for image interpretation. Image classification result by unsupervised technique showed that 96% of the land cover mapping is correct. The result of using direct expansion method for area estimation is reliable and produced result in short time. The highest land cover proportion was oil palm (11039.28 ha or 43%), followed by mangrove (5242.86 ha or 20%), water bodies (4894.92 ha or 19%) and the lowest percentage was urban/bare land type (4751.96 ha or 18%), respectively. Further studies are recommended to integrate remote sensing data and GIS in developing land cover data base system for efficient management of natural resources and it changes over time. Area estimation studies should be tested with different resolution of optical data and cost effectiveness of the methodologies also should be investigated.

Acknowledgements

The authors would like to thank Mr. Khairudin Hashim, the Director of Golden Hope Research Sdn. Bhd for his permission to conduct this research in Carey Island. Special thanks also go to Malaysian Centre for Remote Sensing (MACRES) for providing satellite data for this work.

References

Cochran, W.G., (1997). *Sampling Techniques.* 3rd Edition. John Wiley and Sons, 413p.

Congalton, R.G. and Mead, R.A., (1983). A Quantitative Method to Test for Consistency and Correctness in Photointerepretation. *Photogrammetric Engineering and Remote Sensing*, 49 (1): 60-74.

Ference,C.,Bognar,P,Lichtenberger,J.Hamar,Tarcs,I.G.,Timar.G.,Molnar,G.,Pasztor,R.S., Steinbach,P.,Szekel,Y.B.,Ference,O.E., and Ference-Arkos.I., (2004). Crop yield estimation by remote sensing. *International Journal of Remote sensing*, 25: 4113-4149.

Fook, L.K., Ku Mohd and Laili, N., (1994). Complimentary Nature of SAR and Optical Data for Land Cover Mapping. Paper Presented at *The National Seminar for Decision Makers on 'Potential Aplications of ERS-1 in Malaysia'*, 1 September 1994, Kuala Lumpur. 8 p.

Jensen, J.R., (1996). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Prentice-Hall, Inc. U.S.A. 318 p.

Junior, J.L and Simonett, D.S., (1976). *Remote Sensing of Environment*. Addison-Wesley Publishing Company, Inc. USA. 694 p.

Kamaruzaman Jusoff and Souza,G.D. (1997). Use of satellite remote sensing in Malaysia and its potential. *International Journal of Remote Sensing*, 18,(1): 57-70.

Lillesand, T.M. and Kiefer, R.W., (1994). *Remote Sensing and Image Interpretation*. 3rd Edition. John Wiley and Sons, Inc. USA. 750 p.

Luney, P.R and Dill, H.W., (1970). Uses, Potentialities, and Needs in Agriculture and Forestry. In *Remote Sensing with Special Reference to Agriculture and Forestry.* Nat. Acad. Of Sci. 1-34 p.

Mohd Hasmadi,I., and Kamaruzaman,J., (2007). Estimating Forest Area Using Remote Sensing and Regression Estimator, *WSES Transaction on Signal Processing.* Greece. Issue 1, Volume 3, January 2007. ISSNN 1790-5022 : 88-94

Paul,M. (1995). *Guideline for the Use of Digital Imagery for Vegetation Mapping*, United State Dept. of Agriculture, Washington, D. C., U.S.A. 125 p.

Taylor,J.C.,Sannier,C., Delince,J. and Gallego,F.J. (1997). *Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988-1993*. Published by European Commission (EUR 1719EN). 71p.

Table 1. Area proportion and total area estimate of each class by direct expansion method

No. of segment:39	Oil palm	Mangrove	Water bodies	Urban/bare land
$\sum y_{ic}$	23.154	6.077	4.593	5.508
Proportion ($\overline{y}_c$)	0.6093	0.2894	0.2702	0.2623
Area (Z_c)[ha]	11039.28	5242.86	4894.92	4751.96
Area (Z_c)[%]	43	20	19	18
Std.Error [ha]	0.0587	0.0754	0.0572	0.0943
C.V (%)	59.33	119.36	87.27	164.59
Z_c+1.96SE[ha]	0.71	0.41	0.36	0.42
Z_c-1.96SE [ha]	0.51	0.17	0.18	0.11

*C.V=Coefficient of Variation (Std.Error/Expanded area)*100, $\sum y_{ic}$ * 100*

Table 2. Confusion matrix of the classification image of Carey Island

	REFERENCE DATA (Ground verification)					
		Urban/bare land	Oil palm	Mangrove	Water bodies	Total
CLASSIFIED	Urban/Bare Land	2	0	0	0	2
	Oil Palm	2	25	1	0	28
	Mangrove	0	0	4	0	4
	Water Bodies	0	0	0	18	18
	Total	4	25	5	18	52
	Overall accuracy 96.97%					

Table 3. Producer's and user's accuracy of the classification image of Carey Island

	Producer's Accuracy (%)	**User's Accuracy (%)**	**Kappa**
Oil Palm	100.00	94.74	0.9043
Mangrove	95.65	100.00	1.0000
Urban / Bare Land	87.50	93.33	0.9231
Water Bodies	96.30	100.00	1.000
Overall Kappa Statistics			**0.9514**

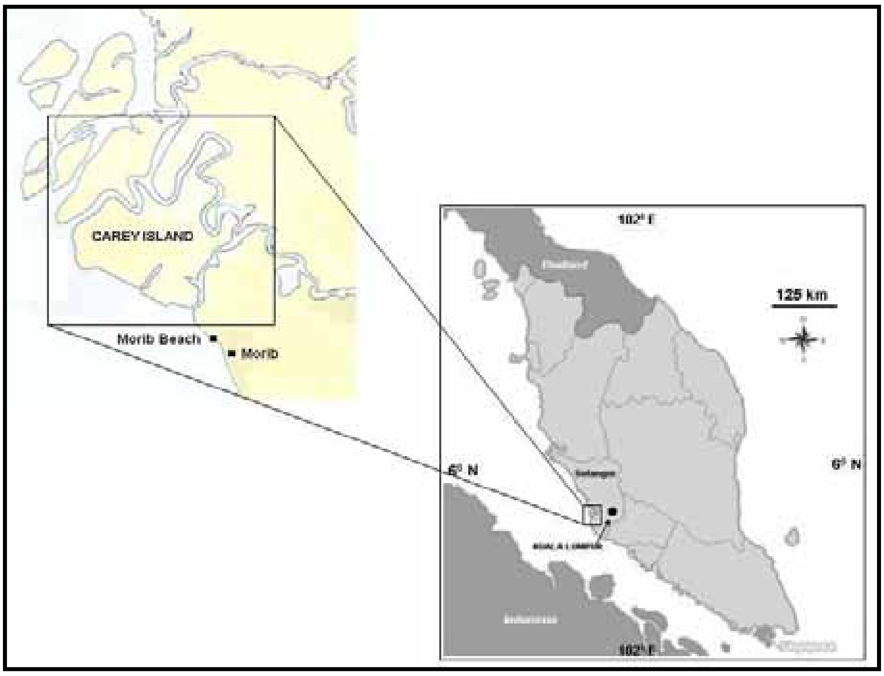

Figure 1. Location of the study area

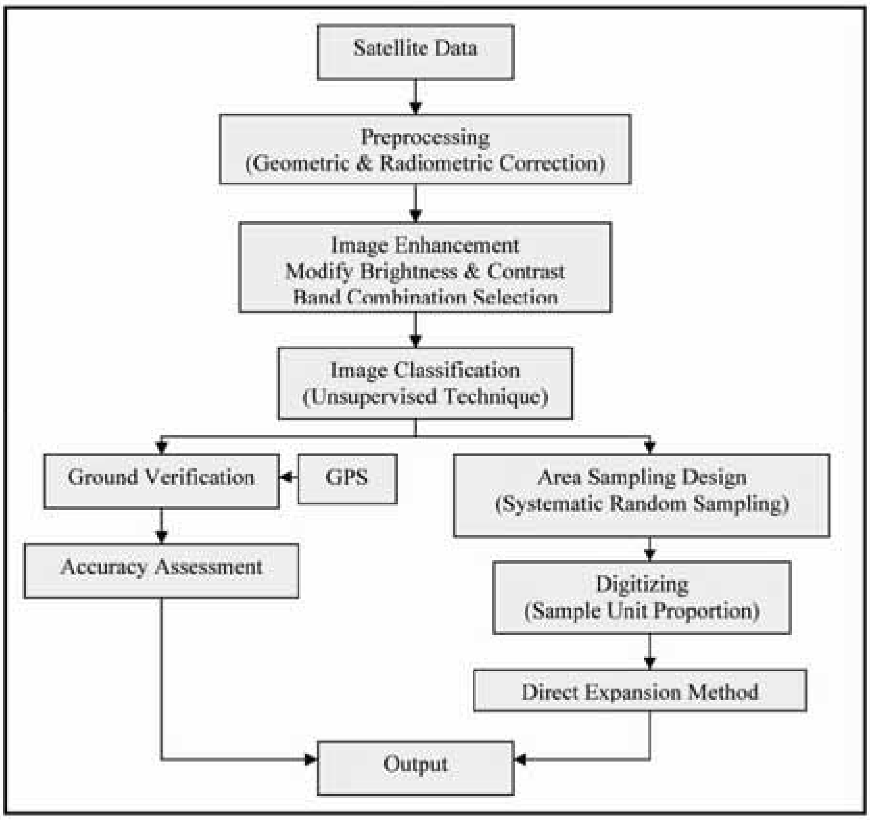

Figure 2. The flow chart of the study

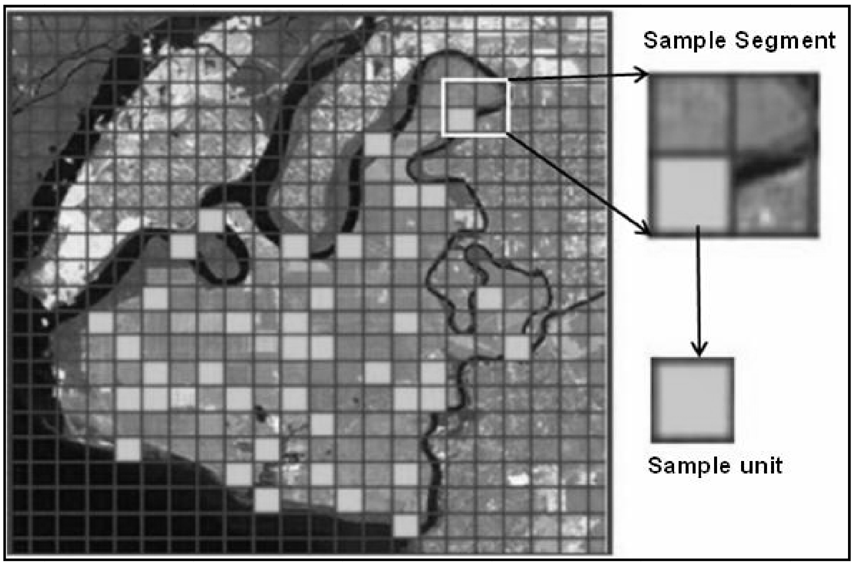

Figure 3. A systematic random design for 39 samples and unit segments over Carey Island

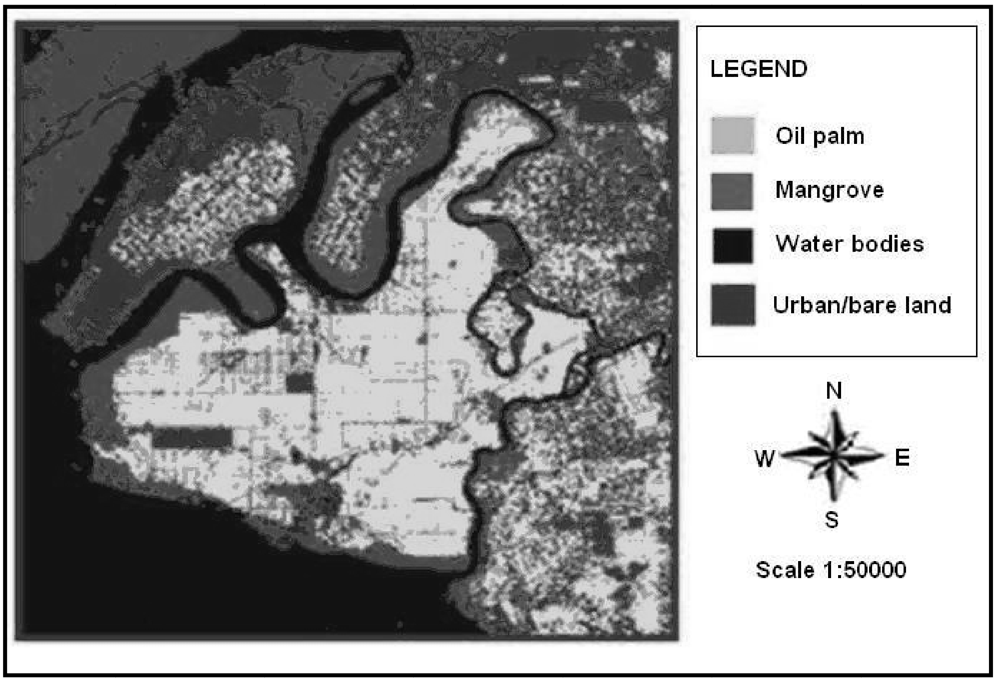

Figure 4. Land cover types within the Carey Island using unsupervised technique

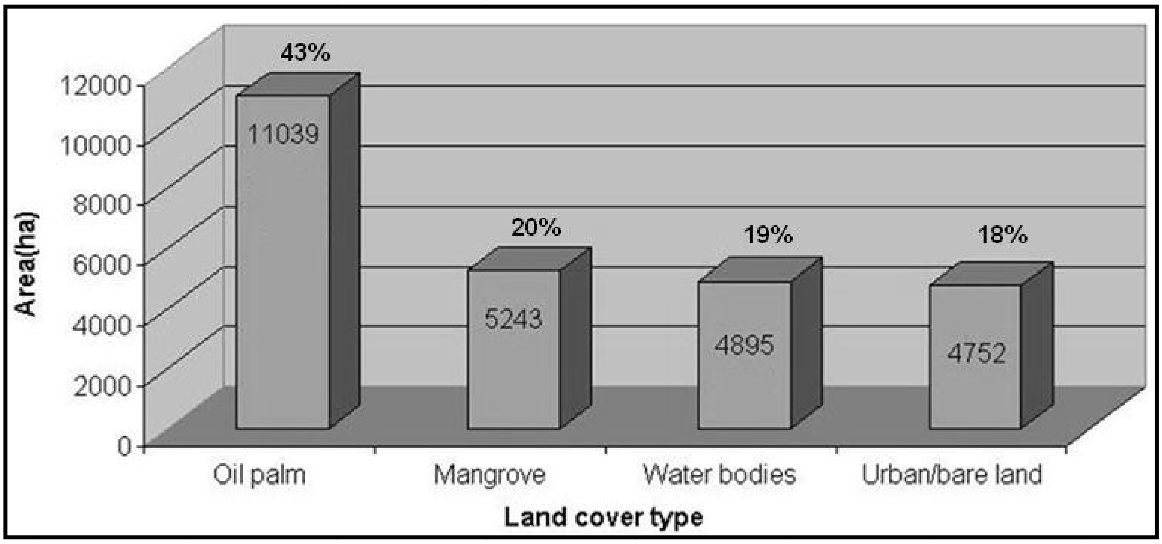

Figure 5. Estimated land cover area (in ha and %) of Carey Island

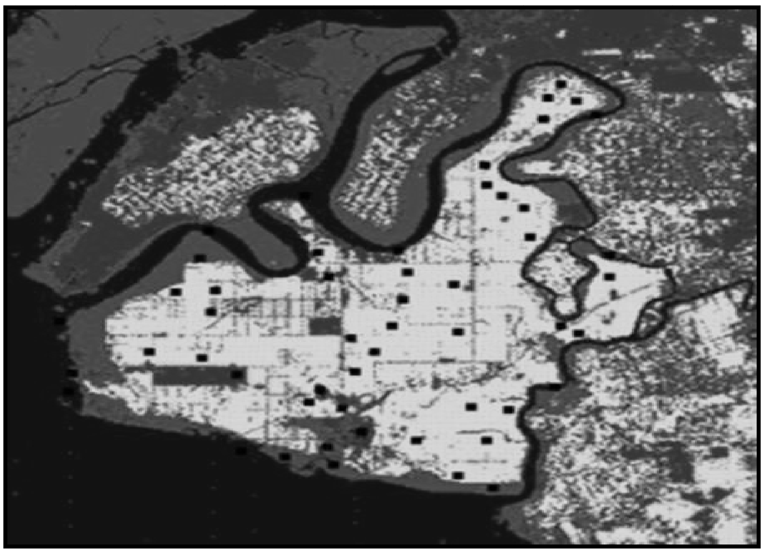

Figure 6. Selected ground verification points in Carey Island (fillet black)