

# CLASSIFICATION OF LAND COVER IN COASTAL ZONE AND ISLANDS OF THE SOUTHWEST REGION OF VIETNAM USING LANDSAT 8 IMAGERY

Tran Anh Tuan<sup>1\*</sup>, Pham Viet Hong<sup>1</sup>, Le Dinh Nam<sup>1</sup>, Tran Thi Tam<sup>2</sup>,  
Vu Le Phuong<sup>1</sup>, Nguyen Thi Anh Nguyet<sup>1</sup>

<sup>1</sup> Institute of Marine Geology and Geophysics - Vietnam Academy of Science and Technology,  
18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam

\*Email: tatuan@imgg.vast.vn

<sup>2</sup>Vietnam Institute of Meteorology, Hydrology and Climate change,  
No.23 - 62 Alley, Nguyen Chi Thanh Road, Dong Da District, Hanoi, Vietnam.

## ABSTRACT

*The study represented the result of land cover classification on the coastal zone and islands of Southwest region of Vietnam using Landsat 8-OLI imagery and sampling sites during 2017. The land cover had been classified by using the combination of unsupervised and supervised classification algorithms (ISODATA and Maximum Likelihood, respectively). Land cover pattern from 147 sampling sites were initially analyzed and used to divided the study area into 6 isolated regions for classification. Land cover complexity was used to define appropriate algorithm for each single region, as the ISODATA algorithm was applied for homogenous region and Maximum Likelihood for much complex region. Synthetic classification result from 6 isolated regions had shown 8 land cover types including mangrove forest, natural forest, plantation forest, aquaculture water, paddy rice field, perennial plant farm, residential area and water body. Assessment on 147 sampling sites showed the overall accuracy was 86,4% and Kappa coefficient was 0.838. The accuracy was generally acceptable for the application of this study to establish 1:50.000 scaled map.*

## 1. INTRODUCTION

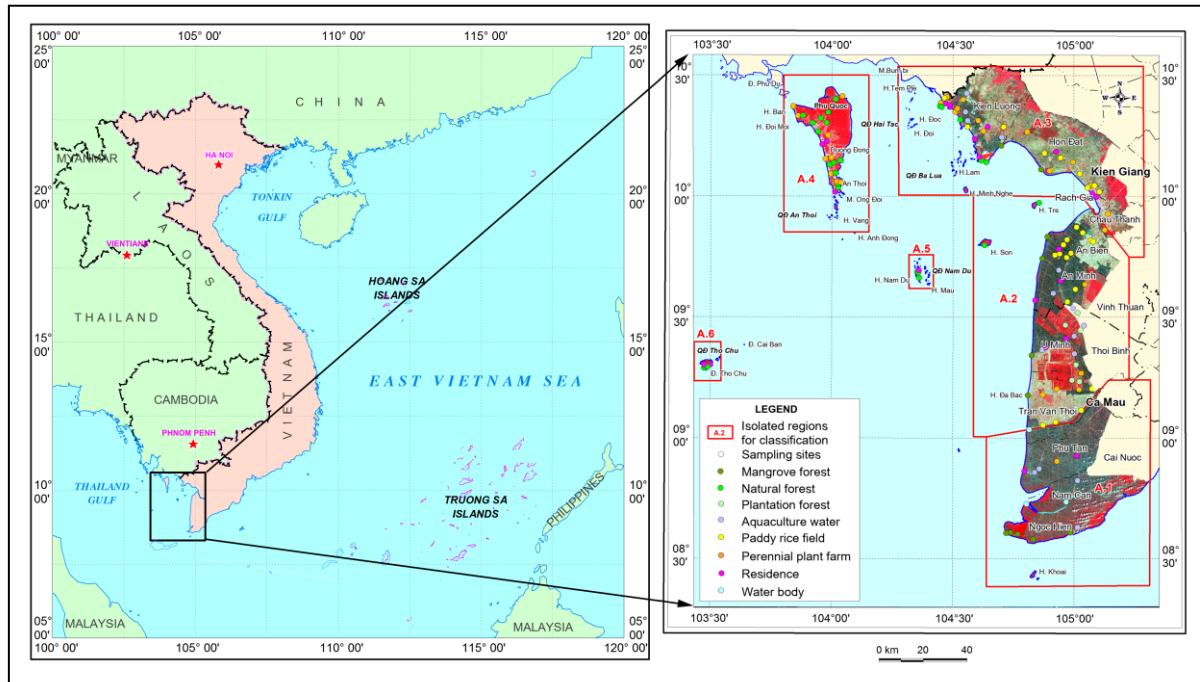
Land cover is the “observed (bio)physical cover of the Earth’s surface” (Di Gregorio and Jansen, 2005), as seen from the ground or through remote sensing, including the vegetation (natural or planted) and human constructions (buildings, roads, etc.) which cover the Earth's surface. Water, ice, bare rock or sand surfaces count as land cover (FAO, 1997). Land cover classification using remote sensing is based spectral reflectance measurement and assessment when electromagnetic spectrum interacts to various land cover patterns. Such approaches using the pixel-based classification methods have been applied globally on coarse spatial resolution satellite images, e.g. AVHRR for 1-degree (DeFries and Townshend, 1994) to 1km (Loveland et al., 2000) resolution, or MODIS for 1km resolution (Friedl et al., 2002). Apart from it, generations of SPOT and Landsat satellite series are applied extensively for mapping land cover pattern of large areas (Roberts et al., 2002; Marceau et al., 1990), while finer spatial resolution images such as QuickBird, IKONOS,... are generally handled on small scale region using object-based approaches in recent years (Chen et al, 2007; Kux and Araújo, 2008).

The study area is usually referred as part of the Mekong Delta (Leinenkugel et al., 2014; Perera et al., 2010) and particularly emphasized for the mangrove coverage on the Cape of Ca Mau (Tong et al., 2004; Van et al., 2015). The study highlighted the details on evaluate the land cover on the coastal zone and island of the Vietnam southwest territory using fusion of Landsat 8 OLI and sampling sites.

## 2. DATA USED AND METHODS

### 2.1 Data used

The study area is situated in the southern part of Vietnam. It is located between longitudes 103° 26' E and 105° 21' E, and latitudes 08° 17' N and 10° 35' N (Figure. 1). The data used included:



**Figure 1. The location of the study area and the sampling sites located on the false color composite (bands 5, 4, 3) of Landsat 8 OLI imagery.**

- The Landsat 8 OLI imagery (path 126 and row 53, 54) downloaded from USGS via <http://glovis.usgs.gov/>. Scenes were selected at April 21, 2015, with acquired local time from 10:19:47am to 10:20:11am. The obtained Landsat data are Level 1 Terrain-corrected (L1T) product were pre-georeferenced to UTM zone 48 North projection using WGS-84 datum.

- Ground truthing information from 147 sampling sites implemented during March - April and October - November of 2017 by observe current condition of local land cover and positioning using handheld GPS devices within the scope of state project coded VT-UD.01/16-20 of the Vietnam Aerospace Science and Technology Program (2016-2020). Each sampling site is assigned for individual emblem, coordinates positioning, describe present surrounding environment and photographed (Figure. 1).

### 2.2 Methods

#### 2.2.1 Data pre-processing methods

Pre-processing methods are applied to create a database for image classification procedure. General framework includes value conversation from original digital number (DN) to surface reflectance values. The Landsat 8 data user handbook (2016) provides corresponding equations for the value conversation with the entire band spectrum. Finally, converted images were clipped and mosaicked to fulfill the study area.

### 2.2.2 Image classification methods

The study applied two classification approaches, i.e. unsupervised and supervised classification for each isolated region depend on the homogeneous or heterogeneous pattern of its land cover. The blending pattern of land cover in each isolated region determines the number of applied class. The algorithm for the unsupervised and supervised classification is ISODATA and Maximum Likelihood, respectively.

### 2.2.3 Accuracy evaluation method

In the remote sensing classification, the Kappa coefficient was recommended as a standard measurement of classification accuracy. The Kappa coefficient was developed by Cohen (1960). To interpret the formula of the Kappa coefficient, the following formulation is more useful:

$$K = \frac{P_o - P_c}{1 - P_c} \quad \text{with:} \quad P_o = \sum_{i=1}^M P_{ii} = \frac{1}{N} \sum_{i=1}^M n_{ii} \quad \text{and} \quad P_c = \sum_{i=1}^M P_{i+} P_{+i} = \frac{1}{N^2} \sum_{i=1}^M n_{i+} n_{+i}$$

Where:

$K$ : The Kappa Coefficient

$P_o$ : Accuracy of observed agreement

$P_c$ : Estimate of chance agreement

$N$ : Total number of observations

$M$ : Number of columns (rows) in the error matrix (number of land cover types)

$n_{ii}$ : Observation in column  $i$  and row  $i$  (on the main diagonal of the error matrix)

$n_{i+}$ : Marginal total of row  $i$

$n_{+i}$ : Marginal total of column  $i$

The Kappa coefficient is always between 0 and 1 and it is divided into 3 groups:

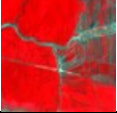


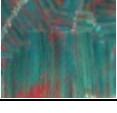
$K \geq 0.8$ : Good agreement;  $0.4 \leq K < 0.8$ : Moderate agreement;  $K < 0.4$ : Poor agreement



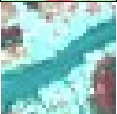

## 3. RESULTS

### 3.1 Building the interpreted keys

Pre-processed data of 147 sampling sites with corresponding land cover classification types of the study area, including main objectives and key criteria, is shown in Table 1.

**Table 1. Interpreted keys of study area.**

ID	Land cover types	False color composite (bands 5, 4, 3)	Description
1	Mangrove forest		Mostly distributed on the Cape of Ca Mau and other coastal regions.
2	Natural forest		Scattering on hilly slope and mountain, islands.
3	Plantation forest		Mostly distributed in National Parks and other region, clear grid architecture with defined boundaries.
4	Aquaculture water		Generally distributed in Ca Mau province, parts of Kien Giang province. Intersected by mangroves.

5	Paddy rice field		Visualized as blue-gray color on false-color composited images, usually bordered by perennial vegetation and/or residential areas.
6	Perennial plant farm		Visualized as red color on false-color composited images, distributed as elongated bands on both two banks of stream channels.
7	Residential area		Visualized as bright blue color on false-color composited images, condensed on both two banks of river channels.
8	Water body		Other sea, lakes, ponds, streams and channels.

### 3.2 Classification and accuracy evaluation

Due to large extension, the study area is divided into 6 isolated region (Figure 1) based on pre-processing data analysis of 147 sampling sites, in order to reduce memory size of input image for later classification process. Moreover, in addition to each object has its corresponding characteristics as identified in the Table 1, the segmentation into isolated regions is a supportive technique to boost the progress and minimize error during classification process. Land cover blending pattern and characteristics are the key criteria for picking a proper classification approach. Consequently, isolated regions tagged A4, A5, and A6 have the considerably homogeneous vegetation covers, including natural forest and perennial plant; isolated region A1 is mostly mangrove and fish farms, therefore they are classified by the unsupervised classification using ISODATA algorithm in order to produce result at higher accuracy. The supervised classification method using Maximum Likelihood algorithm is applied for isolated regions tagged A2 and A4 due to their complicated heterogeneous pattern. Finally, the classified types of each isolated region are composed into a synthetic classification map of land cover for the study area.

In the Table 2, land cover types including mangrove, natural forest, plantation forest and perennial plant have the relatively similar spectrum characteristics as red color in the false-color composited images and it might affect the accuracy of the classification process. Direct and indirect image interpretation using image architecture and geographical distribution laws, respectively - are implemented to distinguish them, e.g. thick mangrove forest usually has finer texture and distributed only on the coastal zone, natural forest disperses on hilly/mountain area, and perennial plant appears as paired bands along stream banks. Another misidentification error is fish farms and harvested paddy fields which share the fairly similar spectrum characteristics. At the image acquiring time (April 2015), harvested paddy fields on high and dry terrain were visualize as color shown in Table 2, while lower and flooded fields were completely identical to fish farms. This identification issue was corrected and enhanced using the ground truthing information collected from the fieldtrip in October - November 2017 during vegetative phase of paddy rice and literature maps.

Statistic of confidential level and calculated Kappa coefficient value for 147 sampling sites showed as Table 2. Evaluated result shows the overall accuracy is 86.4% and Kappa coefficient value of 0.838 which is sufficient for land cover mapping for the study area (Figure 2).

**Table 2. Statistical result and calculated Kappa coefficient.**

	Classified Data								Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Mangrove forest (1)	9	0	0	1	0	1	0	0	11
Natural forest (2)	0	21	0	0	0	6	0	0	27
Plantation forest (3)	0	0	6	0	0	0	0	0	6
Aquaculture water (4)	1	0	0	14	2	0	0	0	17
Paddy rice field (5)	0	0	0	0	25	2	0	0	27
Perennial plant farm (6)	0	2	1	0	0	23	4	0	30
Residential area (7)	0	0	0	0	0	0	25	0	25
Water body (8)	0	0	0	0	0	0	0	4	4
<b>Total</b>	<b>10</b>	<b>23</b>	<b>7</b>	<b>15</b>	<b>27</b>	<b>32</b>	<b>29</b>	<b>4</b>	<b>147</b>

Reference Data

Where N = 147

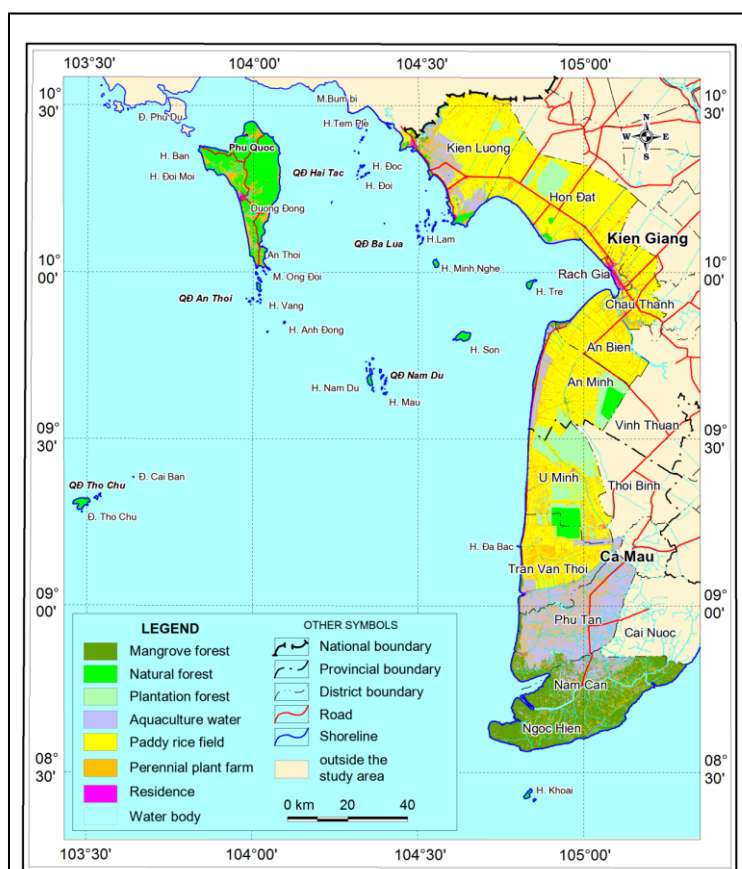
$$P_0 = \frac{9 + 21 + 6 + 14 + 25 + 23 + 25 + 4}{147} = 0.864$$

$$P_c = \frac{(10*11) + (23*27) + (7*6) + (15*17) + (27*27) + (32*30) + (29*25) + (4*4)}{147^2} = \frac{3458}{21609} = 0.16$$

$$K = \frac{0.864 - 0.16}{1 - 0.16} = 0.838$$

#### 4. CONCLUSIONS

The paper presents the land cover classification based on spectrum characteristics of Landsat 8 OLI and preliminary evaluation on vegetation cover distribution of 147 sampling sites for area partition into 6 isolated regions in order to improve classification accuracy and reduce input image size. The unsupervised classification approach using ISODATA algorithm was applied for homogeneous pattern area in the isolated regions tagged A1, A4, A5 and A6, meanwhile the supervised classification approach using Maximum Likelihood algorithm was used for heterogeneous pattern area in isolated region A2 and A3.



**Figure 2. Land cover map of the study area.**

Land cover classification process was performed on each isolated region independently using corresponding image interpretation keys. Single classified land cover type was post-processed and joined together for the study area with 8 main types: mangrove forest, natural forest, plantation forest, aquaculture water, paddy rice field, perennial plant farm, residential area and water body. Confidential level assessment on classification results of 147 sampling sites showed overall accuracy is 86.4% and Kappa coefficient value is 0.838, which is suitable for land cover mapping of the study area.

## 5. ACKNOWLEDGMENT

The authors would like to thank to the project code VT-UD.01/16-20 supported the research team.

## 6. REFERENCES

- Chen Y., Shi P., Fung T., Wang J., Li Y., 2007). Object-oriented classification for urban land cover mapping with ASTER imagery. *International Journal of Remote Sensing* 28 (29), 4645\_4651.
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37-46.
- DeFries, R. S., and Townshend, J. R. G., 1994. NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, 15, 3567-3586.
- Di Gregorio A, LJM Jansen, 2005. Land Cover Classification System Classification concepts and user manual Software version (2). In: 8 *EaNRS* (ed). Food and Agriculture Organization of the United Nations, Rome.
- FAO. 1997. *Africover Land Cover Classification*. Rome.
- Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1-2), 287 - 302.
- Kux H.J.H., Araújo E.H.G. 2008. Object-based Image Analysis using QuickBird satellite images and GIS data, case study Belo Horizonte (Brazil). In: Blaschke T., Lang S., Hay G.J. (eds) *Object-Based Image Analysis. Lecture Notes in Geoinformation and Cartography*. Springer, Berlin, Heidelberg.
- Leinenkugel, Patrick, Natascha Oppelt, and Claudia Kuenzer. 2014. "A new land cover map for the Mekong: Southeast Asia's largest transboundary river basin." *Pacific Geographies* 41: 10 - 14.
- Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21, 1303-1330.
- Marceau DJ, Howarth PJ, Dubois JM, Gratton DJ., 1990. Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 1990;28:513-519.
- Perera, L. K., Herath, S., Apan, A., Samarakoon, L., 2010. Mapping Mekong land cover at 250m resolution without in situ observations, *Asian Journal of GEOINFORMATICS*, Vol. 10 No. 4, pp. 31-41.
- Roberts DA, Numata I, Holmes K, Batista G, Krug T, Monteiro A, Powell B, Chadwick OA., 2002. Large area mapping of land-cover change in Rondonia using decision tree classifiers. *Journal of Geophysical Research*. 2002;107(D20):8073 LBA 40-1-40-18.
- Tong P. H. S., Y. Auda, J. Populus, M. Aizpuru, A. Al Habshi and F. Blasco. 2004. Assessment from space of mangroves evolution in the Mekong delta, in relation with extensive shrimp-farming. *International Journal of Remote Sensing*, 2004, vol. 25, no21, pp. 4795-4812.
- United States Geological Survey, 2016. *Landsat 8 (L8) data users handbook*. Retrieved from <https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf> on September 14, 2018.
- Van, TT, Wilson, N, Thanh-Tung, H, Quisthoudt, K, Quang-Minh, V, Xuan-Tuan, L, Dandouh-Guebas, F & Koedam, N., 2015. Changes in mangrove vegetation area and character in a war and land use change affected region of Vietnam (Mui Ca Mau) over six decades. *Acta Oecologica*, vol 63, pp. 71-81.