

# MAPPING 10 YEARS OF LAND COVER CHANGE IN THE MEKONG DELTA

Ta Hoang Trung, Nguyen Van Tuan

Center of Survey and Mapping Data - Viet Nam  
Department of Survey, mapping and geographic information

Received: 9 July 2020; Accepted: 12 August 2020

**Abstract:** *In the last decade, the Mekong delta has seen a significant decrease in the total area of the wetland due to the expansion of the paddy field, and aquaculture farms. The expansion negatively affected the biodiversity of the region, threaten the environment. To conserve the biodiversity of the region and reach sustainable development, authorities have to monitor the transformation process effectively, thus an updated land cover map is necessary. This research used free satellite data, open-source library, and machine learning algorithm to create 10-meter spatial resolution land cover map of the region 2007 and 2017... In this study, Landsat 5 Thematic Mapper (TM), Landsat 8 Operational Land Imager (OLI), Sentinel-2, ALOS PALSAR, and PALSAR 2 mosaic, ALOS Global Digital Surface Model were employed to produce land cover map of the region by using Kernel Density Estimation classifier. Other ancillary data sources such as Open Street Map, regional geographical database was used as information supplement. To creating and validating the maps, 60,000 reference data points were created based on the field GPS photos as well as visual interpretation on Google Earth images. The overall accuracy of the maps is 82% and 84% in 2007 and 2017, respectively. The maps reveal the rapid loss of mangrove, mainly due to the expansion of aquaculture farms and the urban area between the two periods. The result also demonstrates the potentiality of automatically producing high-accuracy land cover maps in the large area.*

**Keywords:** *Land cover change, multi-temporal, machine learning.*

## 1. Introduction

The Mekong delta is intensely vulnerable to climate change and sea level rising (Birkmann et al., 2012). According to the Climate change and sea level rise scenarios for Viet Nam, if the sea level rises 100 centimeters, nearly 40% of the total area of the delta would be sub-merge beneath the sea (Ministry of Natural Resources and Environment, 2016), negatively affected to the livelihood of the farmers (Tuan and Chinvano, 2011). Sea level rising also expedites the salinity intrusion (Smajgl et al., 2015) which caused a loss of 360 million US dollars in 2016 (Nguyen, 2017), and a lack of clean water for more than 200,000 households. As the largest and one of the most important

agricultural areas, which produce half of Viet Nam rice export of the country, the delta has an important role in the food security of Viet Nam, requiring authorities have the plans to mitigate the impact of climate change, developing the region sustainably.

Human activities are among the major factor behind climate change (Trenberth, 2018). Human's need for food, raw materials are considered as the factor that accelerate deforestation which leads to soil erosion, loss of biodiversity, and climate change. For example, the high demand for shrimp and catfish has led to the rapid expansion of aquaculture farms in the Mekong delta, accompanying with the loss of mangrove (Tran et al., 2015) and wetland loss (Huu Nguyen et al., 2016), which is considered as one of the factors made the effect of climate change to the region more serious. Besides, intensification of agriculture degraded water

---

Corresponding author: Ta Hoang Trung  
E-mail: tahoangtrung@gmail.com

quality of the region (Anh et al., 2010; Guong and Hoa, 2012).

To mitigate the impact of climate change, land cover maps are widely used to monitor the land cover transformation process. However, the current process to produce land cover maps is long, limit the effort of authorities to mitigate the impact of climate change. To produce land cover products, aerial photos were widely used as the main data source of land cover change detection. The photos usually were taken at altitude of 1,130 meters, at a scale of 1/15,000 thus it has the advantages of high spatial resolution. However aerial photo has a narrow field of view, thus to create maps in a large area, a big amount of images are required, requiring many times and effort to producing.

Recently, giant leaps of satellite image technology provide better quality, quantity, and easier to access data for users. As a result, it becomes popular, providing user variety of choices for their specific purpose. For instance, the U.S. Geological Survey (USGS) is providing Landsat satellite data that has been collected spectral information of the Earth's surface from 40 years ago. The Landsat data is a valuable data source for the studies which focus on the history of biophysical changes on the surface of the Earth. If the 30-meter spatial resolution of Landsat is not enough, users can use Sentinel-2 images, provided by ESA (European Space Agency) or ALOS-2 images, provided by JAXA (Japan Aerospace Exploration Agency). Both data have spatial resolution of 10 meters. The advantages of these data are the wide field of view and high temporal resolution (the revisit time fluctuates from 5-10 days). Besides, users can freely access the data. The easy access and adequate quality of satellite data are suitable for produce updated land cover products in a relatively large area. Also, the open-access data is suitable for countries that have limited financial resources for environment protection like Viet Nam where the low-cost land cover map is preferred.

Optical and Synthetic Aperture Radar (SAR) data are popular among data sources to create land cover maps. Optical sensors are widely

used to create land cover mapping due to the valuable information it provides (Clerici et al., 2017). The sensor collects information of the earth's surface in several spectral bands, from visible to infrared, provide useful information for users. However, optical data, as a passive sensor, only observe the earth's surface on daylight, narrowing the chance to observe the Earth's surface. In contrast, (SAR) data, an active sensor is not limited by daylight. Unlike optical sensors, the SAR sensor, an active sensor, uses longer wavelength range, in comparison with the optical sensor, hence it is not affected by cloud, and can observe Earth's surface regardless of time and weather condition. Furthermore, SAR data collects information about the structure of the objects, provide useful information for the analysis required structure of the objects. The limitation of the SAR data is speckle which can reduce the classification accuracy (Joshi et al., 2016). Utilizing the advantages of both optical and SAR data, several studies combine the optical sensor with SAR sensor to create land cover maps (Bagan, Kinoshita, & Yamagata, 2012; Clerici, Valbuena Calderón, & Posada, 2017; Potapov, Hansen, Stehman, Loveland, & Pittman, 2008). These studies used SAR data to fill the gap that optical data left, for example, seasons that surface is covered by cloud, led to the missing data of the region. In This study, we also used the combination of optical and SAR data such as Sentinel-2, Landsat 5-8, and The Advanced Land Observing Satellite (ALOS) PALSAR mosaic to create 10-meter spatial resolution land cover maps of the study area.

## **2. Methodology**

### **2.1. Study area**

The Mekong delta is located in the southern of Vietnam, covers over 40,000km<sup>2</sup>, spread from latitude 8° to 11° North, longitude 103° to 107°, including 13 provinces. The region has tropical monsoon climate condition, the average temperature is about 30 Celcius degree. Creating by the alluvium of the Mekong River, the region has natural advantages in producing agriculture products (i.e., rice, tropical fruits,

catfish). The region also has rich biodiversity. It is the place where thousands of species call home and have some national wetland parks such as Cat Tien national park. The two conditions require authorities must have a suitable plan for the development of the region to balance development and conservation.

## 2.2. Data pre-processing

Optical data always contaminated by cloud and cloud shadow (Hughes and Hayes, 2014), thus requires methods for masking cloud and cloud shadow. Based on the characteristic of the spectral bands, different approaches for cloud removal are considered. For instance, (Zhu and Woodcock, 2012) utilized the difference between high reflectance value on the visible band and low value in thermal bands of cloud pixels to recognize contaminated pixels. In this study, Landsat data is cloud masking based on quality assessment bands of the data. The cloud of Sentinel-2 data is removed based on the reflectance value of each pixel. A threshold is

set, if the reflectance value of Red, Blue, and Green band is bigger than this number, this pixel is considered as cloud, and will be ignored in the classification process. The cloud masking process is done by using open source software GRASS GIS.

Detecting objects using spectral features may be confusing because of the similarity in spectral behavior of different objects, thus in many studies, spectral indices are used for achieving higher accuracy (Ghasemian and Akhoondzadeh, 2018). Spectral indices are the arithmetic combination of different spectral bands, which are usually used as a classification enhancing technique because of its targeting on a specific land cover type (Mongus and Žalik, 2018). Furthermore, spectral indices help reduce the feature space size, by combining the spectral value of 2 bands into 1 index, thus increasing the speed of the classification process. In this study, three spectral indices NDVI, NDBI, NDWI are used (Table 1).

Table 1. Spectral indices used (Taufik and Ahmad, 2016)

Name	Definition
Normalized difference vegetation index	$NDVI = \frac{NIR - Red}{NIR + Red}$
Normalized difference water index	$NDWI = \frac{Green - NIR}{Green + NIR}$
Normalized difference built-up index	$NDBI = \frac{NIR - SWIR}{NIR + SWIR}$

Texture information has been used in many studies to enhance the accuracy of the classification result. (Zhang et al., 2017) proposed using 4 types of statistics-energy for remote sensing images classification which is an Angular second moment (ASM), contrast, correlation, and entropy. To calculate the ASM, firstly, a grey level co-occurrence matrix (GLCM) is built based on the averages of the results of 4 GLCM over 4 orientations: 0, 45, 90 and 135 degrees with 9x9 pixel sliding window sizes and the distance between 2 samples is 1 pixel. The angles are counter-clockwise from the East.

After that, ASM is calculated based on the equation proposed by (Ghasemian and Akhoondzadeh, 2018) where  $P(i,j)$  is the mean GLCM value in position  $(i,j)$ ,  $L$  is the number of levels

$$ASM = \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} P(i,j)^2 \quad (1)$$

In this study, the satellite images of the region of interest are collected at different times of the year (table 2), thus the phenological change of vegetation is detected more accurately. Furthermore, the date of capture was considered as one of the

elements of the feature space. The time information is very important in the case of the Mekong delta due to the harvest of paddy

field in each 3 to 6 months, thus time information could enhance the discrimination between paddy field and barren.

Table 2. Feature space size and data used

Data	Feature space	No of scenes	Resolution (m)	Year	Provider
Sentinel 2 TOA Product	Band: 2,3,4,8a,11,12, NDVI, NDBI,NDWI, ASM	138	10	2017	ESA
ALOS DSM	Slope	1	30	2017	JAXA
Landsat 5 TM SR product	Band: 1,2,3,4,5,7 NDVI, NDBI,NDWI, ASM	93	30	2007	USGS
Landsat 8 OLI SR product	Band: 2,3,4,5,7, NDVI, NDBI,NDWI, ASM	42	30	2017	USGS
ALOS PALSAR Mosaic	HH,HV	2	25	2007,2017	JAXA

### 2.3. Classification scheme

A classification scheme is a method for mapping objects on the Earth’s surface into groups based on their characteristics in common (Congalton et al., 2014). Each current land cover product used its classification scheme based on the target of the products, however, Land cover Classification system (LCCS) which is developed by FAO (Food and Agriculture Organization) and

UNEP (United Nation Environment Programme) has been used widely (Congalton et al., 2014). In this study, 9 categories were chosen base on the reference of LCCS and “Circular of Technical Specifications for the Content, Structure model of the Geographic Database Scale 1:50,000” to enhance the integration of the maps to current products. Details of the categories are shown in Table 3.

Table 3. Description of the classification scheme

No.	Categories	Code	Definition
1	Water	WB	Can be either natural or artificial water bodies.
2	Urban	UB	Land covered by buildings and other man-made structures.
3	Paddy	PD	The cover type is rice paddy influenced by the presence of water.
4	Crop	CR	Lands covered with crops followed by harvest and a bare soil period (e.g., single and multiple cropping systems).
5	Grass	GR	Lands with herbaceous types of cover. Tree and shrub cover is less than 10 %.
6	Orchard	OR	An orchard is an intentional planting of trees or shrubs that is maintained for food production.
7	Mangrove	MG	Mangroves are a group of trees and shrubs that live in the coastal inter-tidal zones.
8	Forest	FR	Lands dominated by woody vegetation with a percent cover > 60% and height exceeding 2m. Almost all trees and shrubs remain green year-round. Canopy is never without green foliage.
9	Bare land	BR	Lands with exposed soil, sand, rocks, or snow and never has more than 10 % vegetated cover during any time of the year.

## 2.4. Classification method

To produce a land cover map of the Mekong region, firstly satellite data is collected, then pre-processing. Optical data such as Landsat and Sentinel are cloud removed and transformed into WGS-84 coordinate system. SAR data such as PALSAR and PALSAR-2 mosaic data is calculated backscatter value from digital number value. After that, each image is classified using the probabilistic approach of the Kernel Density Estimation classifier.

This study used the kernel-based probabilistic classification method, which based on Bayesian theory using a generative model of classification scheme and data. The model is built using kernel density estimation, which is a non-parametric method, estimating the probability density function of a set of random variables. Each image is classified independently, after that they are combined to produce the single land cover product. The flow chart of the process can be seen in Figure 1.

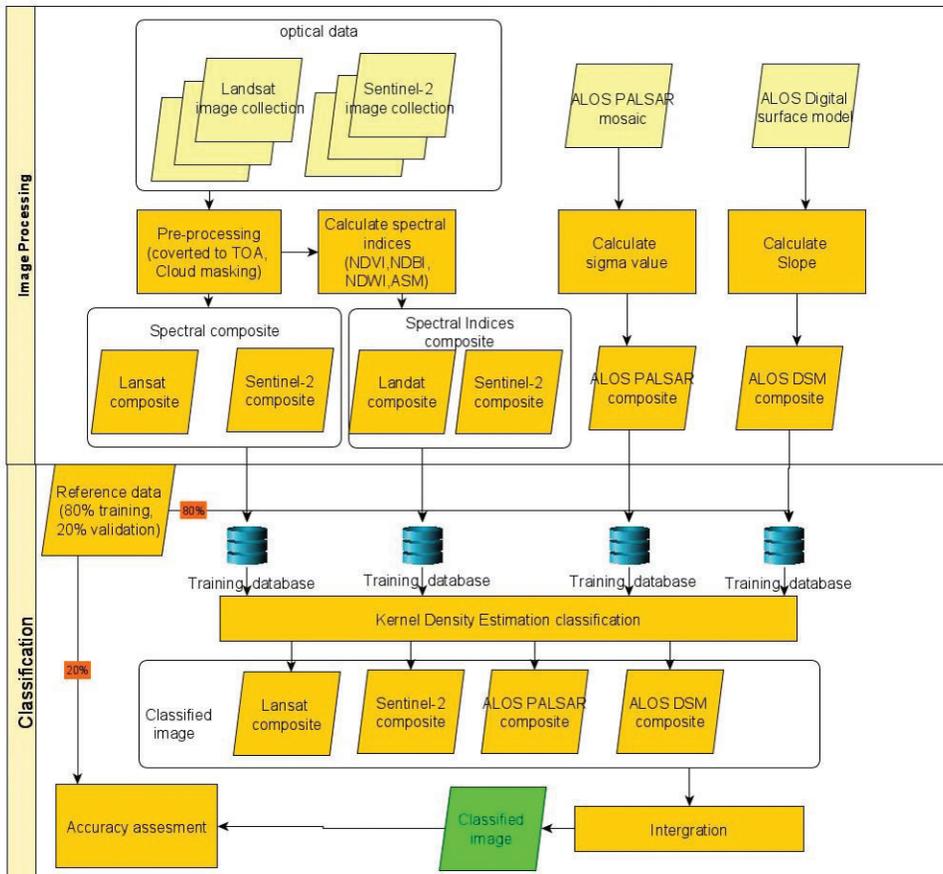


Figure 1. Flow chart of the classification process

Firstly, two dimensions of observation date are calculated by equation 1, where DOY is the Julian day of the year when the image is captured.

$$[t_1, t_2] = \left[ \cos\left(2\pi \frac{DOY}{DOY_{max}}\right), \sin\left(2\pi \frac{DOY}{DOY_{max}}\right) \right] \quad (2)$$

KDE classifier approach is based on the Bayesian rule where  $p(C_k)$  is prior probability of class  $k$ ,  $p(C_k|x)$  is the probability of class  $k$  when

given feature vector  $x$ , so-called posterior probability,  $p(x|C_k)$  is the probability of feature vector  $x$  when given class  $k$ , which can be derived by using kernel density estimation to training data,  $p(x)$  can be considered as constant, in this study, it is equal to  $1/(\text{Number of categories})$

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p(x)} \quad (3)$$

Each pixel of the resulted image is integrated by (Equation 4)

$$p' = \prod_{i=1}^M ap_i + \frac{1-a}{M} \quad (4)$$

$a$  is constant, in this study  $a = 0.7$ ,  $M$  is the number of scenes.

### 3. Result and discussion

#### 3.1. Result

Figure 2 shows the result of the classification process, which is the land cover map of the Mekong Delta in 2007 and 2017, respectively. Generally, the 2017 result has higher accuracy than 2007 result in overall accuracy, similarity;

almost categories of 2017 result have user's accuracy higher than 2007 result whereas the fluctuation is seen on producer's accuracy. The overall accuracy of the 2017 result is 84%, around 1% higher than the 2007 result. User's accuracy of other crops and grassland is the two lowest accuracy numbers in both 2007 and 2017. In 2007, user's accuracy of other crops and grassland is 32% and 51.8%, respectively while in 2017 is 50.9% and 50.6% correspondingly. In contrast, water and evergreen broadleaf trees are two categories that have the highest user's accuracy, more than 95% for both categories in two years.

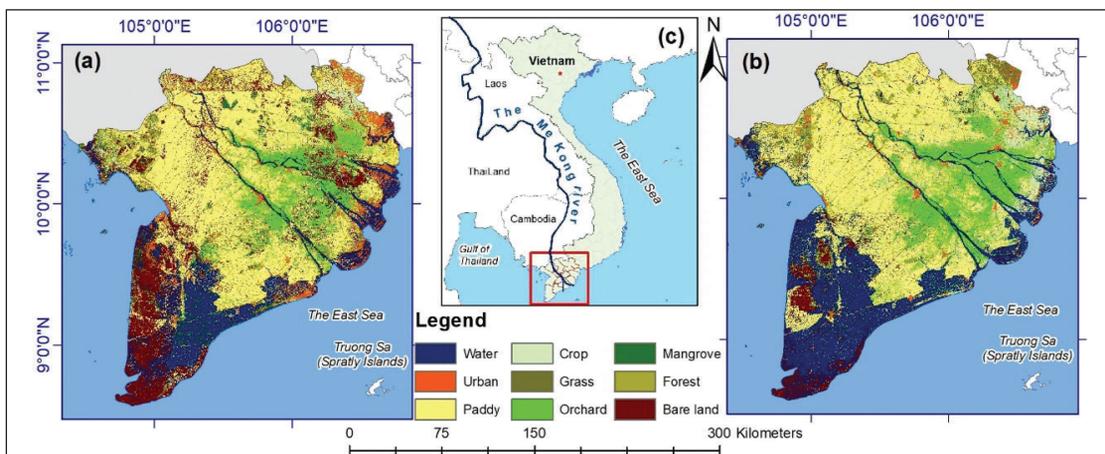


Figure 2. Land cover map of the Mekong Delta in 2007 (a), 2017 (b), and location of the study area in the red box (c)

Table 4. Confusion matrix of 2007 land cover map

		Predicted										*PA (%)
		WB	UB	PD	CR	GR	OR	BR	FR	MG	All	
Actual	WB	<b>338</b>	21	18	2	4	0	13	6	14	416	81.3
	UB	2	<b>205</b>	6	5	0	2	4	1	3	228	90
	PD	0	10	<b>390</b>	2	2	3	14	4	22	447	87.3
	CR	0	0	1	<b>39</b>	0	0	0	1	0	41	95.2
	GR	0	0	0	2	<b>45</b>	2	2	2	1	54	83.4
	OR	0	5	4	21	3	<b>65</b>	4	15	15	132	49.3
	BR	2	36	3	3	7	2	<b>97</b>	3	4	157	61.8
	FR	0	7	5	47	26	11	21	<b>640</b>	16	773	82.8
	MG	3	2	0	1	0	1	0	2	<b>269</b>	278	96.8
	All	345	286	427	122	87	86	155	674	344	<b>2526</b>	80.88
*UA (%)	98	71.7	91.4	32.0	51.8	75.6	62.6	95	78.2	72.9	<b>82.70</b>	

\*UA: User's accuracy, PA: Producer's accuracy

Table 5. Confusion matrix of 2017 land cover map

		Predicted										
		WB	UB	PD	CR	GR	OR	BR	FR	MG	All	PA (%)
Actual	WB	<b>955</b>	8	17	15	2	14	1	3	7	1022	93.5
	UB	33	<b>2646</b>	11	105	140	63	53	35	1	3087	85.8
	PD	9	36	<b>1646</b>	66	69	68	47	51	4	1996	82.5
	CR	0	12	10	<b>492</b>	17	12	3	14	5	565	87.1
	GR	0	18	9	22	<b>460</b>	28	30	25	4	596	77.2
	OR	0	6	1	8	6	<b>677</b>	1	1	11	711	95.3
	BR	6	105	46	36	135	15	<b>701</b>	27	1	1072	65.4
	FR	3	21	17	219	80	80	44	<b>2582</b>	60	3106	83.2
	MG	9	1	1	4	1	5	0	8	<b>440</b>	469	93.9
	All	1015	2853	1758	967	910	962	880	2746	533	<b>12624</b>	84.88
	UA (%)	94.1	92.8	93.7	50.9	50.6	70.4	79.7	94.1	82.6	78.77	<b>84.00</b>

\*UA: User's accuracy, PA: Producer's accuracy

To check the accuracy of the land cover products, we employed the confusion matrix which is widely used to evaluate the quality of land cover maps. The overall accuracy is 83% and 84% in 2007 and 2017, respectively (Tables 4 and 5). The low user's accuracy or error of commission, which one category does not belong to this category but is predicted as this category of other might be caused by the similarity in spectral (i.e, Crop and Grass). In both 2 years, many crop pixels are predicted as grass. In the classification process, we supposed that the 2 categories have the similarity in spectral bands however, crops usually are planted by human thus it usually appears in satellite images in the form of homogeneous area, while grass usually mixes with another type of plants resulting in texture information differences, hence the texture information is different between two kinds. However, the result shows that texture information might be not enough to clarify two categories.

After 10 years, while the total area of paddy field and orchard remained stable, the significant change can be seen on water body. In provinces along coastal area, a number of areas of other land type have been converted to water body, in this case is aquaculture farms. The expansion is confirmed by the statistic

of the General Statistic Office of Viet Nam (GSO). According to GSO, in the last decade, the total area of aquaculture farm of 13 provinces of the Mekong delta increased from 724,000 ha in 2007 to 796,000 ha in 2017. The rapid conversion is considered to increase the fish production to meet the high demand of the market.

### 3.2. Discussion

This study employed a kernel-based probabilistic classification to create land cover maps for Southern Vietnam in 2007 and 2017, using multi-sensor, multi-temporal remote sensing data. The study proposed a method of fusion many different kinds of data to create a land cover products. This might be necessary to create land cover products in coastal areas where optical data is contaminated by the cloud, causing no-data pixel phenomena. Besides, the method notably increases the frequency of Earth observation, thus useful when monitoring the phenology changes of vegetation.

Our approach proposed an affordable method for creating a land cover map of Viet Nam. The study used open-data sources, which can be downloaded freely via the provider's website, and open-source package to process the data and classify images.

As a result, the product has an affordable price, suitable for countries that have a limited budget.

The study proposed a fully automatic process to produce land cover map. There are two main stages to produce the product, the first is download image, and second one is pre-processing and classifying the images.

In this study, we used shell script which runs on the Linux operation system to automatic all the map-making process. As a result, the time to produce land cover product was significantly reduced. The only concern when produce land cover product is the creation of training data and the re-visit time of satellite data.

## References

1. Anh, P.T., Kroeze, C., Bush, S.R., Mol, A.P.J., (2010), *Water pollution by Pangasius production in the Mekong Delta, Viet Nam: Causes and options for control*. Aquac. Res. <https://doi.org/10.1111/j.1365-2109.2010.02578>.
2. Birkmann, J., Garschagen, M., Van Tuan, V., Binh, N.T., (2012), *Vulnerability, Coping and Adaptation to Water Related Hazards in the Vietnamese Mekong Delta*. [https://doi.org/10.1007/978-94-007-3962-8\\_10](https://doi.org/10.1007/978-94-007-3962-8_10).
3. Clerici, N., Valbuena Calderón, C.A., Posada, J.M., (2017), *Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia*. J. Maps 13, 718–726. <https://doi.org/10.1080/17445647.2017.1372316>.
4. Congalton, R.G., Gu, J., Yadav, K., Thenkabail, P., Ozdogan, M., (2014), *Global land cover mapping: A review and uncertainty analysis*. Remote Sens. 6, 12070–12093. <https://doi.org/10.3390/rs61212070>.
5. Ghasemian, N., Akhoondzadeh, M., (2018), *Introducing two Random Forest based methods for cloud detection in remote sensing images*. Adv. Sp. Res. <https://doi.org/https://doi.org/10.1016/j.asr.2018.04.030>.
6. Guong, V.T., Hoa, N.M., (2012), *Aquaculture and Agricultural Production in the Mekong Delta and its Effects on Nutrient Pollution of Soil and Water*. [https://doi.org/10.1007/978-94-007-3962-8\\_14](https://doi.org/10.1007/978-94-007-3962-8_14).
7. Hughes, M.J., Hayes, D.J., (2014), *Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing*. Remote Sens. 6, 4907–4926. <https://doi.org/10.3390/rs6064907>.
8. Huu Nguyen, H., Dargusch, P., Moss, P., Tran, D.B., (2016), *A review of the drivers of 200 years of wetland degradation in the Mekong Delta of Viet Nam*. Reg. Environ. Chang. 16, 2303–2315. <https://doi.org/10.1007/s10113-016-0941-3>.
9. Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T.A., Reiche, J., Ryan, C.M., Waske, B., (2016), *A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring*. Remote Sens. <https://doi.org/10.3390/rs8010070>.
10. Ministry of Natural Resources and Environment (MONRE), V., (2016), *Climate change and sea level rise scenarios for Viet Nam*.
11. Mongus, D., Žalik, B., (2018), *Segmentation schema for enhancing land cover identification: A case study using Sentinel 2 data*. Int. J. Appl. Earth Obs. Geoinf. 66, 56–68. <https://doi.org/10.1016/J.JAG.2017.11.004>.
12. Nguyen, N.A., (2017), *Historic drought and salinity intrusion in the Mekong Delta in 2016: Lessons learned and response solutions*. Viet Nam J. Sci. Technol. Eng. [https://doi.org/10.31276/vjste.59\(1\).93](https://doi.org/10.31276/vjste.59(1).93).
13. Smajgl, A., Toan, T.Q., Nhan, D.K., Ward, J., Trung, N.H., Tri, L.Q., Tri, V.P.D., Vu, P.T., (2015), *Responding to rising sea levels in the Mekong Delta*. Nat. Clim. Chang. <https://doi.org/10.1038/>

nclimate2469.

14. Taufik, A., Ahmad, S.S.S., (2016), *Land cover classification of Landsat 8 satellite data based on Fuzzy Logic approach*. IOP Conf. Ser. Earth Environ. Sci. 37. <https://doi.org/10.1088/1755-1315/37/1/012062>.
15. Tran, H., Tran, T., Kervyn, M., (2015), *Dynamics of land cover/land use changes in the Mekong Delta, 1973-2011: A Remote sensing analysis of the Tran Van Thoi District, Ca Mau Province, Viet Nam*, *Remote Sens.* 7, 2899–2925. <https://doi.org/10.3390/rs70302899>.
16. Trenberth, K.E., (2018), "Climate change caused by human activities is happening and it already has major consequences", *J. Energy Nat. Resour. Law*. <https://doi.org/10.1080/02646811.2018.1450895>.
17. Tuan, L.A., Chinvano, S., (2011), *Climate Change in the Mekong River Delta and Key Concerns on Future Climate Threats*, in: *Advances in Global Change Research*. [https://doi.org/10.1007/978-94-007-0934-8\\_12](https://doi.org/10.1007/978-94-007-0934-8_12).
18. Zhang, X., Cui, J., Wang, W., Lin, C., (2017), *A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm*. *Sensors (Switzerland)* 17. <https://doi.org/10.3390/s17071474>.
19. Zhu, Z., Woodcock, C.E., (2012), *Object-based cloud and cloud shadow detection in Landsat imagery*. *Remote Sens. Environ.* 118, 83–94. <https://doi.org/10.1016/j.rse.2011.10.028>.