# Mapbiomas Indonesia
# Algorithm Theoretical Basis Document (ATBD)
Collection 1

**TABLE OF CONTENTS**

# 1. INTRODUCTION

Mapbiomas Indonesia is a development initiative for monitoring annual land cover change in Indonesia that involves a collaborative network of civil society organizations. Spatial information that provides a historical overview of changing land cover and land use in Indonesia is key to formulating plans for equitable and sustainable natural resources management.

Since April 2019, this collaboration has involved ten Indonesian civil society organizations based in nine provinces and in Jakarta, with support from Mapbiomas Brazil and Woods & Wayside International in the development and application of rapid, reliable and low-cost methodologies for producing annual time series maps of land use and land cover (LULC) in Indonesia. All processes were developed from and are operated using the Google Earth Engine (GEE) platform.

The initiative produced annual maps covering the period 2000-2019 showing ten LULC classes together with analyses of their transitions, which are being published as Mapbiomas Indonesia Collection 1. Product development will be continuous, both in terms of updating data with the latest information, and in methods, tools, and knowledge dissemination.

Hopefully, through this Algorithm Theoretical Basis Document (ATBD) data users will be able to understand the methodologies, dataset descriptions, and statistics used in this collection. All Mapbiomas Indonesia maps and datasets are available and accessible through the Map Biomas Indonesia platform at https://Mapbiomas.nusantara.earth/.

## 1.1. Scope

This ATDB document aims to describe the theoretical base, justification and methods applied in producing Mapbiomas Indonesia Collection 1. The scope of this document covers classification methods, imagery management architecture, and approaches for integrating regional maps with cross-sectoral themes, as well as general overviews of satellite imagery datasets, feature inputs, and accuracy assessment methods. Meanwhile, special approaches in cross-sectoral thematic mapping are presented as appendices.

## 1.2. Regional coverage

Indonesia’s terrestrial regions comprise 16,671 islands covering a total area of 1,916,906 square kilometers ($km^2$) based on official data for 2020 reported to and listed by the United Nations (UN). Mapbiomas Indonesia maps LULC in the whole of Indonesia by dividing it into large island and archipelagic areas, hereinafter referred to as regions. These regions are: (1) Sumatra, (2) Java and Bali, (3) Kalimantan, (4) Sulawesi, (5) the Lesser Sundas and Moluccas, and (6) Papua.

These six regions are then divided further into 33 sub-regions to optimize the analysis processes based on a regional homogeneity approach in terms of landscape, vegetation structure and composition, and land use activities.

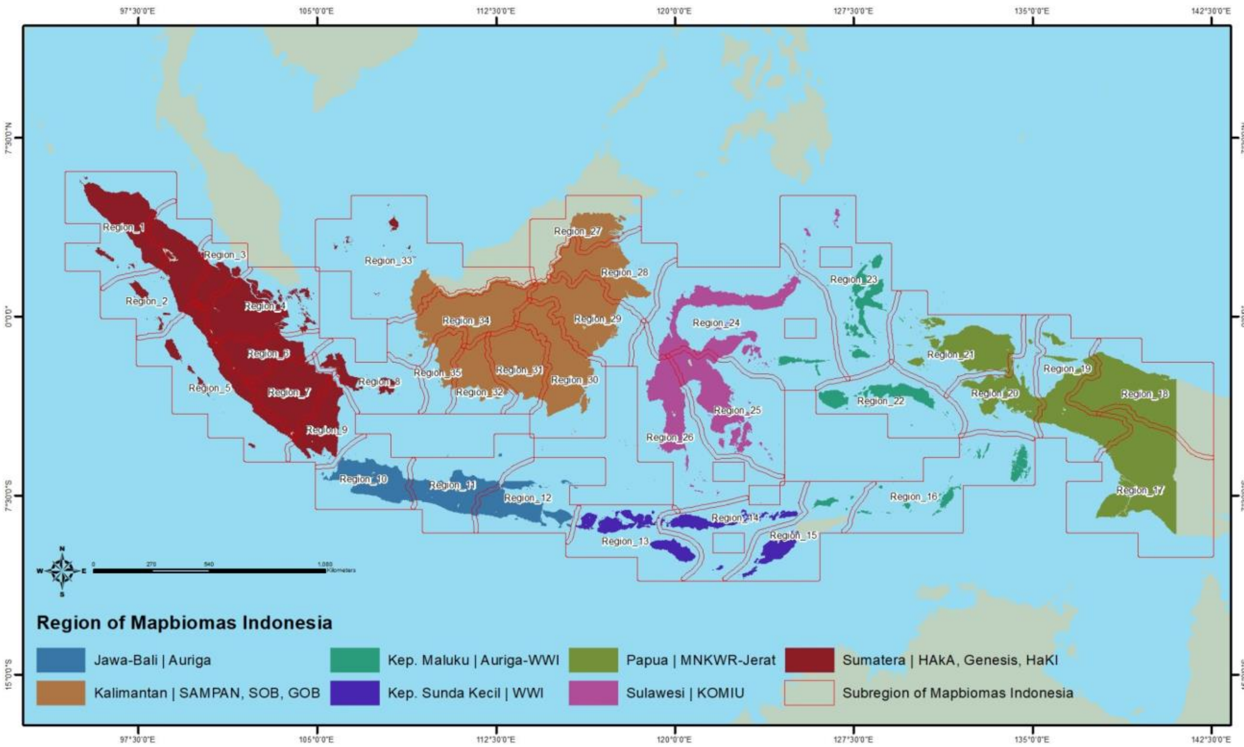


Figure 1 the Mapbiomas Indonesia regions.

### 1.3. Mapbiomas Indonesia Network

The Mapbiomas Indonesia platform has been developed by a network of civil society organizations under the coordination of Auriga Nusantara. Mapping of the various regions and cross-sectoral themes for Collection 1 was divided among the network partners. Coordinating institutions in Sumatra were HAkA (*Hutan, Alam dan Lingkungan Aceh*), Genesis, and HaKI (*Hutan Kita Institute*). Kalimantan was coordinated by Sampan, SOB (*Save Our Borneo*), and GoB (*Green Our Borneo*). Sulawesi and Papua were coordinated by KOMIU (*Kompas Peduli Hutan*), and MNUKWR and JERAT respectively. Meanwhile, Auriga Nusantara and WWI coordinated mapping of the Java-Bali, Lesser Sundas (Lombok and Nusa Tenggara), and Moluccas regions. Auriga Nusantara and WWI also provided technical coordination for mapping of the cross-sectoral themes, including forest plantations, oil palm plantations, mining, mangroves, and aquaculture.

Every step in Mapbiomas Indonesia development, for both regional and cross-sectoral themes, was coordinated and supported by the Mapbiomas Brazil network.

### 1.4. Application

Mapbiomas Indonesia is designed as a land cover and land use change monitoring platform applicable for:

a. mapping and measuring transitions in land cover and land use,
b. calculating gross and net forest loss and expansion,
c. monitoring water resources and their interactions with land cover and land use classes,
d. monitoring agricultural land and industrial crop expansion,
e. monitoring infrastructure expansion,
f. monitoring protection estates,
g. monitoring forest concessions and land-based permits,
h. regional planning.

## 2. METODOLOGY

### 2.1. General Process

Figure 2 below gives the illustration regarding the general protocols applied by Map Biomas Indonesia to produce the land-use and land-cover data for Collection 1.

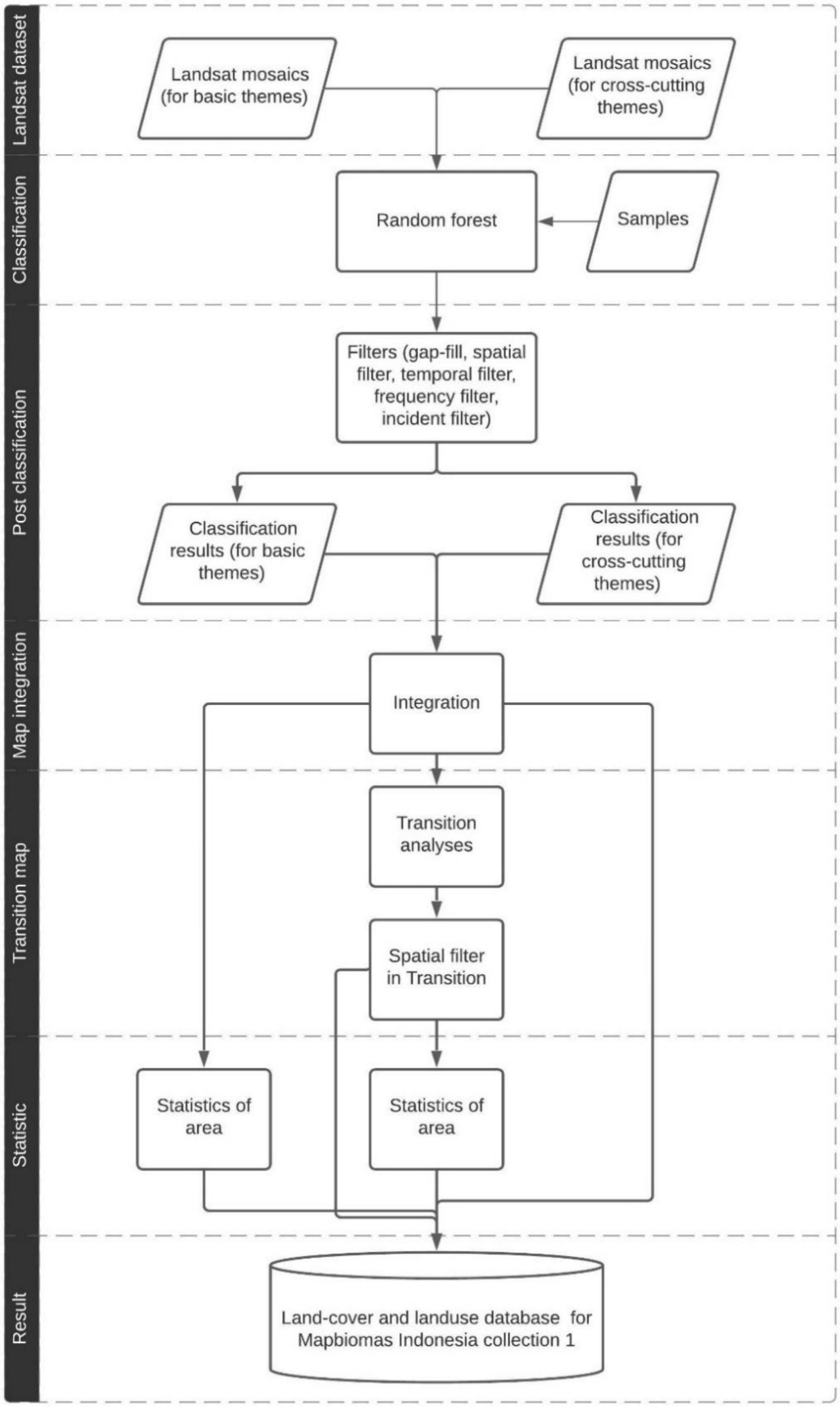



Figure 2 the general methodologies and steps implemented in the Google Earth Engine for producing the Mapbiomas Indonesia collection 1 data.

## 2.2. Google Earth Engine (GEE)

Google defines its Google Earth Engine as "a platform for petabyte-scale scientific analysis and visualization of geospatial datasets, both for public benefit and for business and government users". The Mapbiomas imagery processing sequence is based on Google technology, including image processing in cloud computing infrastructure, programming with JavaScript and Python through the Google Earth Engine, and data storage using Google Cloud Storage.

## 2.3. Landsat

Since the first launch in July 23rd 1972, Landsat satellites have continuously acquired space-based images of the Earth's land surface, providing data that serve as valuable resources for land use/land change research [1]. Landsat satellites have become the world's longest time series data provider. Moreover, low cost (free data of Landsat), having 16 days of temporal resolution, computational capabilities and analysis ready data (ARD) have stimulated the uptake of Landsat data and a growth in Landsat time series research and applications [2].

The availability of satellite observations influences surface monitoring capabilities. Higher temporal resolution observations enable more reliable change detection and surface monitoring [3]. Hence, Mapbiomas Indonesia built the analysis on Landsat satellites data which came from three generations of Landsat satellites to cover 20 years period (2000-2019) for Collection 1 as denoted by Table 1.

The Landsat imagery collections were accessed through the Google Earth Engine sourced from the National Aeronautics and Space Administration (NASA) and the United States Geologic Service (USGS). Landsat quality used by Mapbiomas Indonesia collection 1 was Level-1 Collection 1 Tier 1 top of the atmosphere reflectance (TOA).

Table 1 timeline and history of the Landsat missions [1].

| No | Satellite | Launched | Decommissioned | Sensors | Note |
|---|---|---|---|---|---|
| 1 | Landsat 1 | July 23, 1972 | January 6, 1978 | Return Beam Vidicon (RBV) and Multi Spectral Scanner (MSS) | - |
| 2 | Landsat 2 | January 22, 1975 | February 25, 1982 | Return Beam Vidicon (RBV) and Multi Spectral Scanner (MSS) | - |
| 3 | Landsat 3 | March 5, 1978 | March 31, 1983 | Return Beam Vidicon (RBV) and Multi Spectral Scanner (MSS) | - |
| 4 | Landsat 4 | July 16, 1982 | December 14, 1993 | Thematic Mapper (TM) and Multi Spectral Scanner (MSS) | - |
| 5 | Landsat 5 | March 1, 1984 | June 5, 2013 | Thematic Mapper (TM) and Multi Spectral Scanner (MSS) | Utilized by Mapbiomas Indonesia |
| 6 | Landsat 6 | October 5, 1993 | October 5, 1993 | Enhanced Thematic Mapper (ETM) | Failed to reach orbit |
| 7 | Landsat 7 | April 15, 1999 | Still active | Enhanced Thematic Mapper Plus (ETM+) | Utilized by Mapbiomas Indonesia |
| 8 | Landsat 8 | February 11, 2013 | Still active | Operational Land Imager (OLI) and the Thermal InfraRed Sensor (TIRS) | Utilized by Mapbiomas Indonesia |

## 2.4. Spectral Parameter

It is well known that monitoring land-cover change requires high-resolution imagery ( ~ 30-m resolution or better) in order to accurately quantify areas and rates of change [4]. Landsat images used to be categorized as high-resolutions if it was back in the 70s – 80s. Although Landsat images has become low resolutions in today's standard, Landsat satellites were embedded with multispectral sensors which could be sensitive to recognize land surface object according to the relative spectral response (RSR).

The sensors aboard each of the Landsat satellites were designed to acquire data in different ranges of frequencies along the electromagnetic spectrum (figure 1). According to the USGS (https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qt-news_science_products=0#qt-news_science_products), each specific band of Landsat satellites can be useful to detect a specific object as shown by table 2 and table 3.

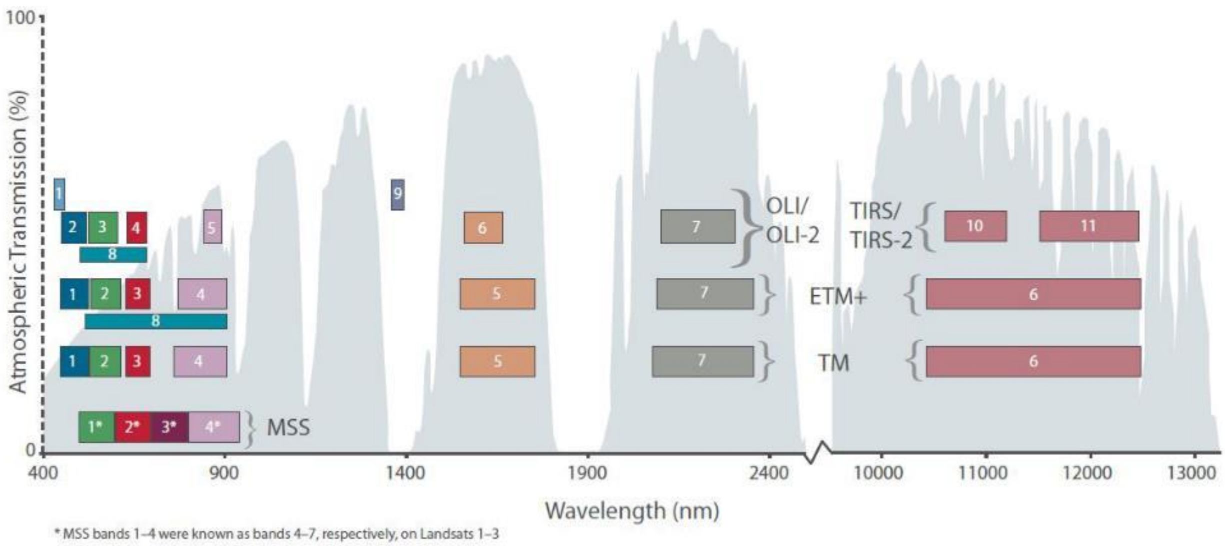



Figure 3 the bandpass wavelength for the sensors on all Landsat satellites.

Table 2 Landsat 8 Operational Land Image (OLI) and Thermal Infrared Sensor (TIRS).

| No | Band | Wavelength (micrometers) | Useful for mapping | Spatial resolution (meters) |
|---|---|---|---|---|
| 1 | Band 1 - coastal aerosol | 0.43-0.45 | Coastal and aerosol studies | 30 |
| 2 | Band 2 - blue | 0.45-0.51 | Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation | 30 |
| 3 | Band 3 - green | 0.53-0.59 | Emphasizes peak vegetation, which is useful for assessing plant vigor | 30 |
| 4 | Band 4 - red | 0.64-0.67 | Discriminates vegetation slopes | 30 |
| 5 | Band 5 - Near Infrared (NIR) | 0.85-0.88 | Emphasizes biomass content and shorelines | 30 |
| 6 | Band 6 - Short-wave Infrared (SWIR) 1 | 1.57-1.65 | Discriminates moisture content of soil and vegetation; penetrates thin clouds | 30 |
| 7 | Band 7 - Short-wave Infrared (SWIR) 2 | 2.11-2.29 | Improved moisture content of soil and vegetation; penetrates thin clouds | 30 |
| 8 | Band 8 - Panchromatic | 0.50-0.68 | 15 meter resolution, sharper image definition | 15 |

| | | | | |
|---|---|---|---|---|
| 9 | Band 9 - Cirrus | 1.36-1.38 | Improved detection of cirrus cloud contamination | 30 |
| 10 | Band 10 - TIRS 1 | 10.60-11.19 | 100 meter resolution, thermal mapping and estimated soil moisture | 100 |
| 11 | Band 11 - TIRS 2 | 11.50-12.51 | 100 meter resolution, improved thermal mapping and estimated soil moisture | 100 |

Table 3 Landsat 4-5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+).

| No | Band | Wavelength (micrometers) | Useful for mapping | Spatial resolution (meters) |
|---|---|---|---|---|
| 1 | Band 1 - blue | 0.45-0.52 | Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation | 30 |
| 2 | Band 2 - green | 0.52-0.60 | Emphasizes peak vegetation, which is useful for assessing plant vigor | 30 |
| 3 | Band 3 - red | 0.63-0.69 | Discriminates vegetation slopes | 30 |
| 4 | Band 4 - Near Infrared | 0.77-0.90 | Emphasizes biomass content and shorelines | 30 |
| 5 | Band 5 - Short-wave Infrared | 1.55-1.75 | Discriminates moisture content of soil and vegetation; penetrates thin clouds | 30 |
| 6 | Band 6 - Thermal Infrared | 10.40-12.50 | Thermal mapping and estimated soil moisture | L5: 120 (30), L7: 60 (30) |
| 7 | Band 7 - Short-wave Infrared | 2.09-2.35 | Hydrothermally altered rocks associated with mineral deposits | 30 |
| 8 | Band 8 - Panchromatic (Landsat 7 only) | 0.52-0.90 | 15 meter resolution, sharper image definition | 15 |

To enhance the ability to distinguish the land surface objects, Mapbiomas Indonesia not only used the original bands of Landsat but also the spectral indices generated from Landsat as the parameters that are used by the algorithms to do classification analyses.

**CAI (Cellulose Absorption Index)** is used to discriminating plant litter from soil. It is defined as a spectral variable which describes the depth of the lignocellulose absorption feature in the shortwave infrared region [5].

$\mathbf{CAI = 0.5\ (R_{2.0} + R_{2.2}) - R_{2.1}}$

where $R_{2.0}$, $R_{2.1}$, and $R_{2.2}$ are reflectance factors in bands at 2.00–2.05, 2.08–2.13, and 2.19–2.24 µm, respectively.

**EVI (Enhanced Vegetation Index)** was developed as a standard satellite vegetation product for the Terra and Aqua Moderate Resolution Imaging Spectroradiometers (MODIS). EVI provides improved sensitivity in high biomass regions while minimizing soil and atmosphere influences. **EVI2** evaluates a 2-band EVI without a blue band [6].

**EVI2 = 2.5 * ((NIR – R) / (NIR + 2.4 * R +1)**

where NIR is the near infrared band and R is the red band.

**GCVI (Green Chlorophyll Vegetation Index)** is a vegetation index to quantify greenness and to understand vegetation density [7].

**GCVI = (NIR / G) – 1**

where NIR is the near infrared band and G is the green band.

**NDFI (Normalized Differencing Fraction Index)** is a new spectral index for enhanced detection of forest canopy damage caused by selective logging and forest fires [8, 9, 10, 11].

$$\mathbf{NDFI = (GV_{shade} - (NPV + Soil)) / (GV_{shade} + NPV + Soil)}$$

where GV is green vegetation [12] NPV is the non-photosynthetic vegetation [13]. Soil is the soil fraction derived from SMA (spectral mixture analysis) [9, 14]. $GV_{shade}$ is the shade-normalized GV fraction given by,

$$\mathbf{GV_{shade} = GV / (100\text{-}Shade)}$$

**NDVI (Normalized Difference Vegetation Index)** is used to quantify vegetation greenness and is useful in understanding vegetation density and assessing changes in plant health. NDVI is calculated as a ratio between the red (R) and near infrared (NIR) values in traditional fashion [15, 16].

$$\mathbf{NDVI = (NIR - R) / (NIR + R)}$$

**NDWI (Normalized Difference Water Index)** is an index proposed for remote sensing of vegetation liquid water from space. NDWI is used to monitor changes in water content of leaves [17].

$$\mathbf{NDWI = (NIR - SWIR1) / (NIR + SWIR1)}$$

**PRI (Photochemical Reflectance Index)** is a method to remotely assess photosynthetic efficiency using narrow-band reflectance [18, 19].

$$\mathbf{PRI = [R_{531} - R_{570}]/[R_{531} + R_{570}]}$$

where R indicates reflectance and numbers indicate wavelength nanometers at the center of the bands).

**SAVI (Soil Adjusted Vegetation Index)** is used to correct Normalized Difference Vegetation Index (NDVI) for the influence of soil brightness in areas where vegetative cover is low. Landsat Surface Reflectance-derived SAVI is calculated as a ratio between the R and NIR values with a soil brightness correction factor (L) defined as 0.5 to accommodate most land cover types [20].

$$\mathbf{SAVI = ((NIR - R) / (NIR + R + L)) * (1 + L)}$$

**SRTM (Shuttle Radar Topography Mission)** 1 Arc-Second Global is elevation data offer worldwide coverage of void filled data at a resolution of 1 arc-second (30 meters) and provide open distribution of this high-resolution global data set (USGS, https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects).

### 2.5. Statistic Reducer

Spectral data of Landsat are stored in something termed as the digital number (DN) which is exactly similar with most other raster formats. The DN ranges are between 0-255 if the data quality was 8 bits (for instance: Landsat 5 & Landsat 7) and 0-4095 if the data quality was 12 bits (for instance: Landsat 8). Basically, the DN is numbers, and the statistic reducers were applied to capture several alternative variations in the spectral data that can enrich the spectral parameters used by the classifier.

Table 4 list of spectral parameters contained in the Landsat Mosaics of Mapbiomas Indonesia.

| | Name | Statistical reducer | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median | Median_dry | Median_wet | Amplitude | Standard deviation | Minimum | Maximum |
| **Band** | Blue | | | | | | | |
| | Green | | | | | | | |
| | Red | | | | | | | |
| | NIR | | | | | | | |
| | SWIR1 | | | | | | | |
| | SWIR2 | | | | | | | |
| | Temp | | | | | | | |
| **Index** | NDVI | | | | | | | |
| | CAI | | | | | | | |
| | EVI2 | | | | | | | |
| | GCVI | | | | | | | |
| | NDWI | | | | | | | |
| | PRI | | | | | | | |
| | SAVI | | | | | | | |
| | Hallcover | | | | | | | |
| **Fraction** | GV | | | | | | | |
| | NPV | | | | | | | |
| | Soil | | | | | | | |
| | Cloud | | | | | | | |
| | Shade | | | | | | | |
| **MEM Index** | GVS | | | | | | | |
| | NDFI | | | | | | | |
| | SEFI | | | | | | | |
| | WEFI | | | | | | | |
| | FNS | | | | | | | |

Table 5 list of spectral parameters considered by Mapbiomas Indonesia for classifying basic themes.

| | Name | Statistical reducer | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median | Median_dry | Median_wet | Amplitude | Standard deviation | Minimum | Maximum |
| **Band** | Blue | | | | | | | |
| | Green | | | | | | | |
| | Red | | | | | | | |
| | NIR | | | | | | | |
| | SWIR1 | | | | | | | |
| | SWIR2 | | | | | | | |
| **Index** | GCVI | | | | | | | |
| | EVI2 | | | | | | | |
| | NDFI | | | | | | | |
| | NDVI | | | | | | | |
| | NDWI | | | | | | | |
| | SAVI | | | | | | | |
| | SEFI | | | | | | | |
| | textG | | | | | | | |
| **Fraction** | GV | | | | | | | |
| | NPV | | | | | | | |
| | Soil | | | | | | | |
| | Shade | | | | | | | |
| | Gvshade | | | | | | | |

### 2.6. Landsat Mosaic

The terminology of mosaicking in GIS describes a process to merge several layers of data into a single dataset. Mapbiomas Indonesia mosaicked all spectral parameters into a single dataset termed as the Landsat mosaic that was then used by the classifier as the main component to learn the characteristics of the samples to run the classification process. The intended Landsat mosaic was also collecting all available Landsat in a certain grid within a certain period, then the best pixels were selected to be combined into a single Landsat grid which contained the best pixels.

Mapbiomas Indonesia uses mapping units based on the International Map of the World on the millionth scale (IMW) at a scale of 1:250,000 to create Landsat mosaics. Each grid covers an area of $1^{o}$ latitude x $1.5^{o}$ longitude with a total of 286 grids to cover the whole of Indonesia. The Landsat mosaic used by Mapbiomas Indonesia is annual (per year) and cloud-free. The acquisition period set by Mapbiomas Indonesia is generally in between April-September every year, coinciding with Indonesia's dry season, so the cloud noise is less. However, the acquisition period can be adjusted or extended to cover a wider range of Landsat data. Mapbiomas Indonesia also applies a quality assessment (QA) band and median reducer to create a cloud-free annual Landsat mosaic.

### 2.7. Algorithm

Mapbiomas Indonesia applied a supervised classification approach to produce land-cover and land-use data. Supervised classification is a technique with the concept of segmenting the spectral domain into regions that can be associated with the ground cover classes of interest to a particular application [21]. Supervised classification requires previously classified reference samples in order to train the classifier and subsequently classify unknown data [22].

Random forest [23] was a machine learning algorithm chosen by the Mapbiomas Indonesia as the classifier. The random forest classifier contained a number of decision trees on the given parameters (in this case spectral parameters) that was used to learn the characteristics or patterns of the given samples. Random forest does not only rely on one decision tree, but also considers the decision from other trees based on the majority votes to predict the analyzed data in a classification process.

### 2.8. Classification System

Land cover refers to expanses of objects covering the Earth's surface, whereas land use relates to human interaction with land or types of activities taking place on the Earth's surface. A classification system is an abstract representation with the names, codes and definitions of the classes, the well-defined diagnostic criteria (classifiers) used to distinguish different types of land cover, and the relationship among land cover classes [24]. A classification system is a concept to interpret or to define various classes of land-use and land-cover (LULC) in the real condition and to simplify them into only several classes that can well represent those actual classes.

Mapbiomas Indonesia applied classification concepts published by the UN Food and Agriculture Organization (FAO), and chose a hierarchical structure of classification system to accommodate different levels of information, starting with structured broad-level classes, which allow further systematic subdivision into more detailed sub-classes [25]. In level 1, the classification system was divided into six LULC classes: (1) Forest, (2) non-Forest Natural Formation, (3) Farming, (4) non-Vegetation Area, (5) Water, and (6) non-Observed. Each class in level 1 consisted of two classes in level 2 except the non-Forest Natural Formation and non-Observed classes. For level 3, the Natural Forest class (in level 2) was separated into the Mangrove and Forest Formation sub-classes.

Table 6 the Mapbiomas Indonesia collection 1 classification system.

| No. | ID | Classification | Natural/Anthropic | Land cover/Land use | Regional/ Cross sectoral |
|---|---|---|---|---|---|
| | 1 | **1. Forest** | **Natural** | **Land cover/Land use** | **Regional** |
| | 2 | 1.1 Natural Forest | Natural | Land cover | Regional |
| 1 | 3 | 1.1.1 Forest Formation | Natural | Land cover | Regional |
| 2 | 5 | 1.1.2 Mangrove | Natural | Land cover | Cross sectoral |
| 3 | 9 | 1.2 Plantation Forest | Anthropic | Land use | Cross sectoral |
| 4 | 10 | **2. Non-Forest Natural Formation** | **Natural** | **Land cover** | **Regional** |
| | 13 | 2.1 Other Non-Forest Natural Formation | Natural | Land cover | Regional |
| | 14 | **3. Farming** | Anthropic | Land use | Regional |
| 5 | 35 | 3.1 Palm Oil | Anthropic | Land use | Cross sectoral |
| 6 | 21 | 3.2 Agriculture | Anthropic | Land use | Regional |
| | 22 | **4. Non-Vegetated Area** | **Natural/Anthropic** | **Land cover/Land use** | **Regional** |
| 7 | 30 | 4.1 Mining | Anthropic | Land use | Cross sectoral |
| 8 | 25 | 4.2 Other Non-Vegetated Area | Natural/Anthropic | Land cover/Land use | Regional |
| | 26 | **5. Water** | **Natural/Anthropic** | **Land cover/Land use** | **Regional** |
| 9 | 31 | 5.1 Aquaculture | Anthropic | Land use | Cross sectoral |
| 10 | 33 | 5.2 River, Lake and Ocean | Natural | Land cover | Regional |
| 11 | **27** | **6. Non-Observed** | **None** | **None** | **None** |

### 2.9. Classification

Image classification is the process of categorizing and labelling groups of pixels or vectors within an image based on specific rules. The categorization rules were devised using one or more spectral or textural characteristics [26]. Mapbiomas Indonesia determined to classify the Landsat imageries into 10 classes of LULC, consisting of:

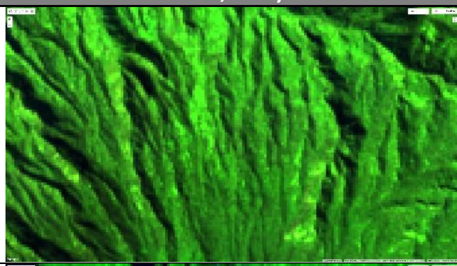
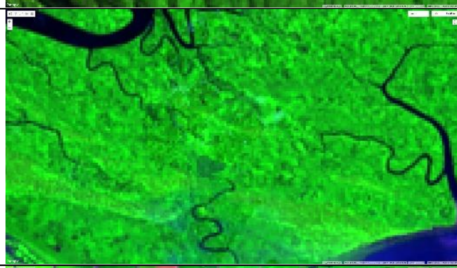
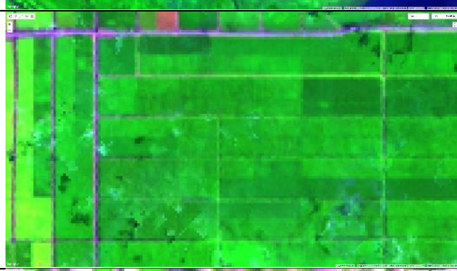
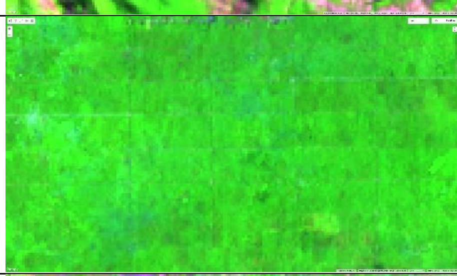
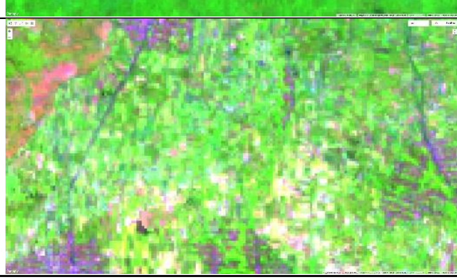
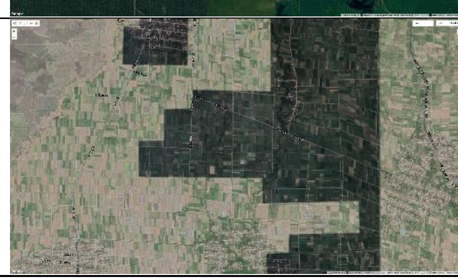
a. five basic classes (forest formation, other non-forest natural formation, agriculture, non-vegetated area, and river/lake/ocean), and
b. five cross-cutting classes (mangrove, plantation forest, oil palm, mining, and aquaculture).

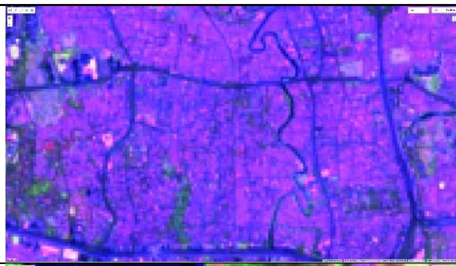
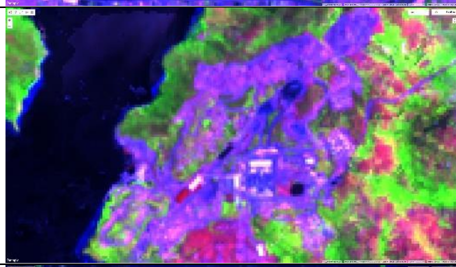
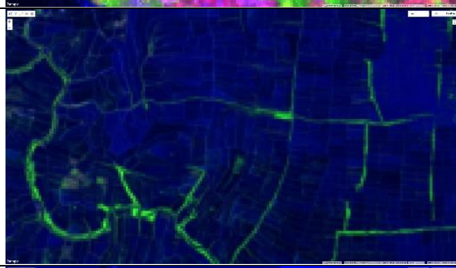
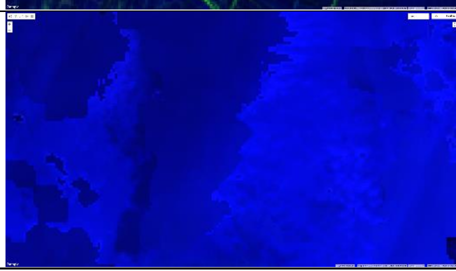
Five cross-cutting classes were analyzed separately (as described in appendices), while the five basic classes were analyzed using general methods. Once Landsat mosaics and parameters (for the Random Forest Classifier) had been completely processed, samples representing five basic classes were collected by considering the key interpretations (table 7) . The samples were taken at locations that did not experience land-cover changes during the period 2000-2019 by considering the distributions, sizes, and proportions of the samples, and these samples were referred to as 'stable samples'.

The characteristics of the stable samples were analyzed by the Random Forest classifier which was considering the predefined spectral parameters set previously to classify the entire study area. The number of trees in the Random Forest Classifier varies from 50 to 100. The classification process was run in the Google Earth Engine platform, the analyses were conducted for each defined region in Indonesia and for each year in the defined time series (2000-2019).

The classifications resulting from the process utilizing the stable samples were re-checked and evaluated. Improvements were conducted by adding more samples for each class which were termed as the complimentary samples. Once the stable and complimentary samples had been defined, the classification process was sent back to the Random Forest Classifier to obtain the final classification result.

Table 7 key interpretations in collecting samples.

| No | Landsat composite (false colour: SWIR1; NIR; Red) | High resolution imagery | Sample for |
|---|---|---|---|
| 1 |  |  | Forest |
| 2 |  |  | Mangrove |
| 3 |  |  | Plantation forest |
| 4 |  |  | Other non-forest natural formation |
| 5 |  |  | Palm oil |
| 6 |  |  | Agriculture |

| | | | |
|---|---|---|---|
| 7 |  |  | Other non-vegetated area |
| 8 |  |  | Mining |
| 9 |  |  | Aquaculture |
| 10 |  |  | River, Lake and Ocean |

### 2.10. Post Classification

Post-classification processes are effective to stabilize the results and to reduce the biases raised in the classification process. Hence, Mapbiomas Indonesia ran a circle process of post-classification consisting of a gap-fill filter, spatial filter, temporal filter, frequency filter, and incident filter.

#### 2.10.1. Gap-Fill Filter

Landsat mosaicking caused a removal effect of data on areas containing cloud or shadow. The classifier, which was the Random Forest, did not classify the 'no-data' and these 'no-data' remained when the classification was completed. The gap-fill filter was applied to fill 'no-data' with data by considering the preceding data.

#### 2.10.2. Spatial Filter

The spatial filter was used to prevent changes in classification values in groups of pixels. The filter is made based on "connectPixelCount" where the function positions the connected pixel components with the same pixel values. This filter requires at least five connected pixels as a minimum connection value.

#### 2.10.3. Temporal Filter

The temporal filter was used to identify unwanted transitions occurring over three to five years. The filter would examine and change the central position of non-sequential pixels for reclassification according to prior and subsequent classes.

#### 2.10.4. Frequency Filter

The frequency filter considered the occurrence of a particular class across the time series. All class occurrences with a persistence of less than 10% would be filtered and classified as non-class.

#### 2.10.5. Incident Filter

The incident filter was used to stabilize pixel values that changed too frequently over the 20 years. This usually occurred at boundaries between classes. Pixel values that had changed more than eight times would be replaced by stable pixel values in the time series.

### 2.11. Integration

The basic theme products that had undergone filtering processes for each year from 2000-2019 were then integrated with cross-sectoral themes by applying a set of specific hierarchical prevalence rules.

### 2.12. Transition

Transition was the dynamic of LULC, where a certain class changed to be other classes in certain periods. Mapbiomas Indonesia analyzes the transition of LULC based on time periods as: (a) per year, (b) per 5 years, (c) per 10 years, (d) all observed years.

### 2.13. Statistics

The statistics of the mapped LULC classes were calculated based on several spatial units, such as administrative boundaries, forest estates, watersheds, forest and peatland moratorium areas, concession areas included in zonal statistics.

### 2.14. Validation

Validation strategies have been initially based on comparative analyses with existing reference maps for specific regions and years. Over time, validation will be enhanced further through accuracy analyses based on statistical techniques using independent sample points with visual interpretation for the whole of Indonesia and for the whole time series.

## ANNEX

### Annex 1 – Coastal Zone

Mapbiomas Indonesia represented the coastal zone by classifying the mangrove and aquaculture classes as cross-cut themes. The scope was defined as areas having elevation 0-35 meters above sea level measured continuously from the coastal line referred to the elevation data of the SRTM Digital Elevation 30m.

The classifications for the coastal zone were begun by preparing the annual cloud-free Landsat mosaics to cover the period of 2000-2019. Several spectral parameters (table 6) were mosaicked, then sent to the classifier Random Forest.

Supervised classification was conducted by directing the collected samples on the objects represented by the mangrove and aquaculture classes onto the Landsat images. The classification process was run in the Google Earth Engine platform where the spectral characteristics captured in the samples were analyzed by the classifier and then classified all study areas.

Once the classification process was completed, the results were corrected by running a chain of filters in the post-classification process. A gap-Fill filter, temporal filter and frequency filter were implemented before the results of the coastal zone were integrated with the other classes.

Table 8 list of spectral parameters considered by Mapbiomas Indonesia for classifying the coastal zone.

| | Name | Statistical reducer | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median | Median_dry | Median_wet | Amplitude | Standard deviation | Minimum | Maximum |
| Band | Green | | | | | | | |
| | Red | | | | | | | |
| | NIR | | | | | | | |
| | SWIR1 | | | | | | | |
| | SWIR2 | | | | | | | |
| Index | EVI2 | | | | | | | |
| | NDSI | | | | | | | |
| | NDVI | | | | | | | |
| | NDWI | | | | | | | |
| | MNDWI | | | | | | | |
| Slope | | | | | | | | |

**Annex 2 – Industrial Plantation**

The industrial plantation classified by Mapbiomas Indonesia refers to two large-scale land-based plantation industries: forest plantation (Hutan Tanaman Industri/HTI) and oil palm (smallholder plantations are not included).

Similar to others, the classification process was started by preparing the Landsat mosaics. The classifier was also the Random Forest Classifier, but the spectral parameters used were different as listed in Table 9 below. The samples were guided to collect on represented areas of the oil palm and forest plantation only. The collected samples trained the classifier that then classified the industrial plantation classes throughout the study area.

As the final step before the result was sent to the integration process, the temporal and spatial filters were applied to correct the classification results of the industrial plantation classes.

Table 9 spectral parameters considered by Mapbiomas Indonesia for classifying the industrial plantation.

| | Name | Statistical reducer | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median | Median_dry | Median_wet | Amplitude | Standard deviation | Minimum | Maximum |
| **Band** | Blue | | | | | | | |
| | Green | | | | | | | |
| | Red | | | | | | | |
| | NIR | | | | | | | |
| | SWIR1 | | | | | | | |
| | SWIR2 | | | | | | | |
| | TIR1 | | | | | | | |
| | TIR2 | | | | | | | |
| **Index** | CAI | | | | | | | |
| | EVI2 | | | | | | | |
| | NDVI | | | | | | | |
| | NDWI | | | | | | | |
| | LAI | | | | | | | |

## Annex 3 – Industrial Mining

The methodology and process for industrial mining areas were similar to other cross-cut themes except the spectral parameters used for the classifier. It was begun by mosaicking the spectral parameters, collecting appropriate samples representing mining activities, and analyzing them in the Google Earth Engine platform.

The classification process was accomplished by applying a gap-Fill filter, temporal filter dan frequency filter as the post-classification process. Once the post classification was completed, the mining class was integrated into the other Mapbiomas Indonesia classes.

Table 10 spectral parameters considered by Mapbiomas Indonesia for classifying industrial mining areas.

| | Name | Statistical reducer | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median | Median_dry | Median_wet | Amplitude | Standard deviation | Minimum | Maximum |
| **Band** | Green | | | | | | | |
| | Red | | | | | | | |
| | NIR | | | | | | | |
| | SWIR1 | | | | | | | |
| | SWIR2 | | | | | | | |
| **Index** | NDSI | | | | | | | |
| | NDVI | | | | | | | |

## REFERENCE


[1] “Landsat—Earth observation satellites,” Reston, VA, 2016.

[2] M. A. Wulder *et al.*, “Current status of Landsat program, science, and applications,” *Remote Sens. Environ.*, vol. 225, pp. 127–147, 2019.

[3] A. V. Egorov, D. P. Roy, H. K. Zhang, Z. Li, L. Yan, and H. Huang, “Landsat 4, 5 and 7 (1982 to 2017) Analysis Ready Data (ARD) observation coverage over the conterminous United States and implications for terrestrial monitoring,” *Remote Sens.*, vol. 11, no. 4, 2019.

[4] J. R. G. TOWNSHEND and C. O. JUSTICE, “Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations,” *Int. J. Remote Sens.*, vol. 9, no. 2, pp. 187–236, Feb. 1988.

[5] P. L. Nagler, Y. Inoue, E. P. Glenn, A. L. Russ, and C. S. T. Daughtry, “Cellulose absorption index (CAI) to quantify mixed soil-plant litter scenes,” *Remote Sens. Environ.*, vol. 87, no. 2–3, pp. 310–325, 2003.

[6] Z. Jiang, A. R. Huete, K. Didan, and T. Miura, “Development of a two-band enhanced vegetation index

without a blue band," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3833–3845, 2008.

[7] M. Burke and D. B. Lobell, "Satellite-based assessment of yield variation and its determinants in smallholder African systems," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 9, pp. 2189–2194, 2017.

[8] C. Souza Jr., "Monitoring of Forest Degradation: A Review of Methods in the Amazon Basin," *Glob. For. Monit. from Earth Obs.*, pp. 171–194, 2018.

[9] C. Souza and P. Barreto, "An alternative approach for detecting and monitoring selectively logged forests in the Amazon," *Int. J. Remote Sens.*, vol. 21, no. 1, pp. 173–179, Jan. 2000.

[10] C. M. Souza, D. A. Roberts, and M. A. Cochrane, "Combining spectral and spatial information to map canopy damage from selective logging and forest fires," *Remote Sens. Environ.*, vol. 98, no. 2–3, pp. 329–343, 2005.

[11] C. M. Souza *et al.*, "Ten-year landsat classification of deforestation and forest degradation in the brazilian amazon," *Remote Sens.*, vol. 5, no. 11, pp. 5493–5513, 2013.

[12] G. P. Asner, M. Keller, R. Pereira, and J. C. Zweede, "Remote sensing of selective logging in Amazonia: Assessing limitations based on detailed field observations, Landsat ETM+, and textural analysis," *Remote Sens. Environ.*, vol. 80, no. 3, pp. 483–496, 2002.

[13] D. A. Roberts, M. O. Smith, and J. B. Adams, "Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data," *Remote Sens. Environ.*, vol. 44, no. 2, pp. 255–269, 1993.

[14] C. de Wasseige and P. Defourny, "Remote sensing of selective logging impact for tropical forest management," *For. Ecol. Manage.*, vol. 188, no. 1, pp. 161–173, 2004.

[15] E. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sens. Environ.*, vol. 185, pp. 46–56, 2016.

[16] J. G. Masek *et al.*, "A Landsat surface reflectance dataset for North America, 1990-2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 68–72, 2006.

[17] B. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.

[18] J. A. Gamon *et al.*, "NDVI canopy structure photosynthesis," *Ecol. Appl.*, vol. 5, no. 1, pp. 28–41, 1995.

[19] M. F. Garbulsky, J. Peñuelas, J. Gamon, Y. Inoue, and I. Filella, "The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies. A review and meta-analysis," *Remote Sens. Environ.*, vol. 115, no. 2, pp. 281–297, 2011.

[20] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, 1988.

[21] J. A. Richards, *Remote sensing digital image analysis: An introduction*, vol. 9783642300. 2013.

[22] X. Ceamanos and S. Valero, "Processing Hyperspectral Images," *Opt. Remote Sens. L. Surf. Tech. Methods*, pp. 163–200, 2016.

[23] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[24] H. Yang, S. Li, J. Chen, X. Zhang, and S. Xu, "The Standardization and harmonization of land cover classification systems towards harmonized datasets: A review," *ISPRS Int. J. Geo-Information*, vol. 6, no. 5, pp. 1–16, 2017.

[25] A. Di Gregorio and L. J. M. Jansen, "Land cover classification system (LCCS): classification concepts and user manual for software version 1.0." FAO, 2000.

[26] M. Shinozuka and B. Mansouri, "4 - Synthetic aperture radar and remote sensing technologies for structural health monitoring of civil infrastructure systems," in *Woodhead Publishing Series in Civil and Structural Engineering*, V. M. Karbhari and F. B. T.-S. H. M. of C. I. S. Ansari, Eds. Woodhead Publishing, 2009, pp. 113–151.