Celebes Agricultural 3(2): 82 - 99 February (2023)

Research Article

High Heterogeneity LULC Classification in Ujung Kulon National Park, Indonesia: A Study Testing 11 Indices, Random Forest, Sentinel-2 MSI, and GEE-based Cloud Computing

Klasifikasi LULC Berheterogenitas Tinggi di Taman Nasional Ujung Kulon, Indonesia: Sebuah Studi dalam Menguji 11 Indeks, Random Forest, Sentinel-2 MSI, dan GEE berbasis Cloud Computing

Rahmat Asy'Ari[1,2,*], Aulia Ranti[1,2], Azelia Dwi Rahmawati[1,2], Moh Zulfajrin[3], Lina Lathifah Nurazizah[4], Made Chandra Aruna Putra[2,5], Zayyaan Nabiila Khairunnisa[2,6], Faradilla Anggit Prameswari[7], Rahmat Pramulya[8,9], Neviaty P. Zamani[6,10], Yudi Setiawan[9,11], Ajat Sudrajat[12], Anggodo[12]

[1] Department of Forest Management, IPB University

[2] IPB Sustainable Science Research Students Association (IPB SSRS Association), IPB University

[3] Department of Soil Science and Land Resource, IPB University

[4] Department of Agronomy and Horticulture, IPB University

[5] Department of Marine Science and Technology, IPB University

[6] Department of Civil and Environmental Engineering, IPB University

[7] Department of Geography, Faculty of Mathematics and Natural Science (FMIPA), University of Indonesia

[8] Faculty of Agriculture, Teuku Umar University

[9] Center for Environmental Research, IPB University

[10] Center for Transdisciplinary And Sustainability Sciences, IPB University

[11] Department of Forest Resources Conservation and Ecotourism, IPB University

[12] Ujung Kulon National Park, Ministry of Environment and Forestry, Indonesia

*email: rahmatasyariasyari@apps.ipb.ac.id

Received: October 2022
Accepted: February 2023
Published: February 2023

Abstract: The Ujung Kulon National Park (UKNT) is one of the national parks on the island of Java and has an essential role in saving endemic species in Indonesia. As a form of national park conservation effort, the completeness of LULC spatial data is a primary database that is indispensable in determining national park management policies. Therefore, this research was conducted to map the LULC (Land Use - Land Cover) in the forest landscape with high heterogeneity in UKNT. Sentinel-2 MSI (Multispectral Instrument) image data were classified using the Random Forest (RF) classification algorithm and tested using 11 index algorithms. The classification process takes place on a cloud computing-based geospatial platform, Google Earth Engine (GEE). This test resulted in 10 LULC classes; water had the broadest percentage of 45.44%. Meanwhile, the primary forest has an area of 21,868.41 or about 19.53% of the total area of the national park. However, there are some discrepancies in the spatial information generated by this classification process, so it is considered necessary to evaluate the combination of indexes, training data, and classification algorithms to limit the classification area. Therefore, this study is expected to be considered for further research related to LULC in high-heterogeneity landscapes.

Keywords: LULC, Indices, Random Forest, Sentinel-2, Ujung Kulon NP

Abstrak: Taman Nasional Ujung Kulon (UKNT) merupakan salah satu taman nasional yang berada di Pulau Jawa dan memiliki peran penting dalam penyelamatan spesies endemik di Indonesia. Sebagai bentuk upaya konservasi taman nasional, kelengkapan data spasial LULC merupakan basis data dasar yang sangat diperlukan dalam menentukan kebijakan pengelolaan taman nasional. Oleh karena itu, penelitian ini dilakukan untuk memetakan LULC (Land Use - Land Cover) pada lanskap hutan yang berhetoroginitas tinggi di UKNT. Data citra Sentinel-2 MSI (Multispectral Instrument) diklasifikasikan menggunakan algoritma klasifikasi Random Forest (RF) serta diuji menggunakan 11 algoritma indeks. Proses klasifikasi berlangsung pada platform geospasial berbasis cloud computing yaitu Google Earth Engine (GEE). Pengujian ini menghasilkan 10 kelas LULC dan tubuh air merupakan kelas dengan persentase terluas yaitu 45,44%. Sedangkan hutan primer memiliki luas sebesar 21.868,41 atau sekitar 19,53% dari total luas kawasan taman nasional. Akan tetapi, terdapat beberapa ketidaksesuaian informasi spasial yang dihasilkan proses klasifikasi ini sehingga dinilai perlu adanya evaluasi pada kombinasi indeks, data training, algoritma klasifikasi hingga pembatasan area klasifikasi. Penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi penelitian selanjutnya terkait LULC pada lanskap

Citation:
Asy'Ari R, Ranti A, Rahmawati AD, Zulfajrin M, Nurazizah LL, Putra MCA, Khairunnisa ZN, Prameswari FA, Pramulya R, Zamani NP, Setiawan Y, Sudrajat A, Anggodo. 2023. High heterogeneity LULC classification in Ujung Kulon National Park, Indonesia: A study testing 11 indices, random forest, sentinel-2 MSI, and GEE-based cloud computing. *Celebes Agricultural*. 3(2): 82-99. doi: 10.52045/jca.v3i2.381

p-ISSN: 2723-7974
e-ISSN: 2723-7966
doi: 10.52045/jca.v3i2.381

Website:
https://ojs.untika.ac.id/index.php/faperta

heterogenitas tinggi.

Kata kunci: LULC, Indices, Random Forest, Sentinel-2, Ujung Kulon NP

INTRODUCTION

Ujung Kulon National Park (UKNT) is one of the national parks on the island of Java and is designated as a cultural heritage site by UNESCO. The conservation of the Javan rhino (*Rhinoceros sondaicus*) is one of the reasons for the designation of the Ujung Kulon area as a national park. The Javan rhinoceros is one of the large endemic mammals threatened with extinction and is highly sensitive to habitat disturbance (Brook et al., 2014; Sandground, 1933). This makes habitat protection a priority to ensure these animals' survival and avoid human activities that cause conflict (Santiapillai and Mackinnon 1990). In addition, this national park stores a variety of species, namely 35 species of mammals, 5 species of primates, 240 species of birds, 65 species of reptiles (Milto and Lukin 2020), 22 species of amphibians, 72 species of insects, 142 species of fish and 33 types of coral reefs.

Uncontrolled human activities, as economic demands, often impact the environment. For example, land conversion and forest encroachment are problems that often occur in forest landscapes in Indonesia (Harada et al., 2015; Austin et al., 2019; Dauvergne, 1994). This action causes the fragmentation of forest areas and harms the sustainability of the surrounding ecosystem (Armenteras et al., 2006) and includes the protection of endemic species in it (Simamora et al., 2021; Yanuar & Chivers, 2009). This is a world problem, especially in countries with tropical forests (Allen and Barnes, 1985; Dauvergne, 1994). Nijman (2013) explains that forest fragmentation and deforestation have long occurred in forest areas on the island of Java. For example, UKNT has various dynamics of land use change, which was previously a forest area into an agricultural area. In the UKNT area, diverse introduced land uses such as agriculture, plantations, community settlements, and others exist. Economic factors are one of the causes of this action, so people depend on forest areas. This will undoubtedly have a fatal impact on efforts to protect endangered animals and the biodiversity contained therein.

Until now, the presentation of spatial-based land use information has used the latest technology to provide convenience in the identification process and post-field analysis (Chatelain et al., 1996). This considers higher costs, time, and human resources if it does not involve remote sensing technology. However, over time, remote sensing technology in the form of software is considered less practical in terms of geospatial data analysis. This is because the software requires considerable storage resources and specifications for high-quality computer equipment. Therefore, this concept is considered inefficient if applied globally, involving various types of data such as satellite imagery. Several years earlier, the largest company Google launched the Google Earth Engine (GEE) platform, which provides a wide range of geospatial analysis capabilities on a global scale (Gorelick et al. 2017). In addition, this platform can provide access to satellite imagery such as Landsat, Sentinel, ALOS, SRTM, ASTER, and others (Kumar & Mutanga, 2018; Qu et al., 2021).

This cloud computing-based platform uses Google's storage resources (Gorelick et al., 2017), so it can process geospatial data with a reasonably large scope (Fattore et al., 2021). In addition, this platform can also involve various types of classification algorithms such as random forest (RF), support vector machine (SVM), and others, as well as index algorithms such as NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), NDBI (Normalized Difference Built-Up Index), and more. Tamiminia et al. (2020) explained that the development of GEE had created a lot of enthusiasm for researchers to study geospatial data science. This makes the GEE platform widely applied in various fields, such as agriculture, forestry, and others (Mutanga & Kumar 2019). For example, the Aprilianti et al. (2021) study involved the GEE platform in identifying the index characteristics of plantation vegetation in Bogor Regency, Indonesia, while Floreano & de Moraes (2021) applied the GEE platform in mapping LULC in Rondônia State, Brazil.

The application of cloud computing-based mapping technology in national park management agencies is minimal. In addition, the exploration and application of the GEE platform in recent years have not yet been maximized (Tamiminia et al., 2020). Detailed-scale mapping needs to be done as area control and data availability to ensure the sustainability of the conservation program (Qu et al., 2021; Galicia and García-Romero, 2007). This is because the community borders the UKNT forest area to the north and east, vulnerable to land conversion. Therefore, this basic information research is vital to map the highly heterogeneous LULC (land use - land cover) in the UKNT area using cloud computing technology, as well as several classification algorithms and indexes. Information and recommendations from the results of this study are expected to be used as consideration for the national park agency, local government, and central government in determining forest area conservation policies. In addition, the results of this study can be used as supporting information for further research.

MATERIALS AND METHODS

Research Location

This research takes a case study in Ujung Kulon National Park/UKNT, Pandeglang Regency, Province of Banten (Figure 1). The study duration lasted 2 months, namely December 2021 - January 2022. Data collection took place on every land use within the UKNT area.

Data Sources

This study uses the primary data from Sentinel-2 MultiSpectral Instrument (S2-MSI) images that can map land cover on earth. The S2-MSI image is a medium-resolution image of 10 - 60 meters. This image has 13 total bands: aerosol band, visible band, red edge, NIR (Near Infrared), and SWIR (Shortwave Infrared). The use of bands is adjusted to the needs of the index algorithm formula used in the study. In addition, the field data used is GCP (ground checkpoint; Table 1) for each mapped land use. The GCP data is used as a reference for classifying LULC.

Table 1. Ground checkpoint

No	Land use	GCP	
		Training data	Validation data
1.	Water body	47	20
2.	Primary forest	142	30
3.	Mixed forest	117	30
4.	Bamboo forest	31	10
5.	Mangroves forest	104	30
6.	Bare land	47	25
7.	Built-up area	70	15

8.	Crop area	69	30
9.	Swamp forest	25	20
10.	Schrub	45	19
	Total	470	231
	Grand Total	701	

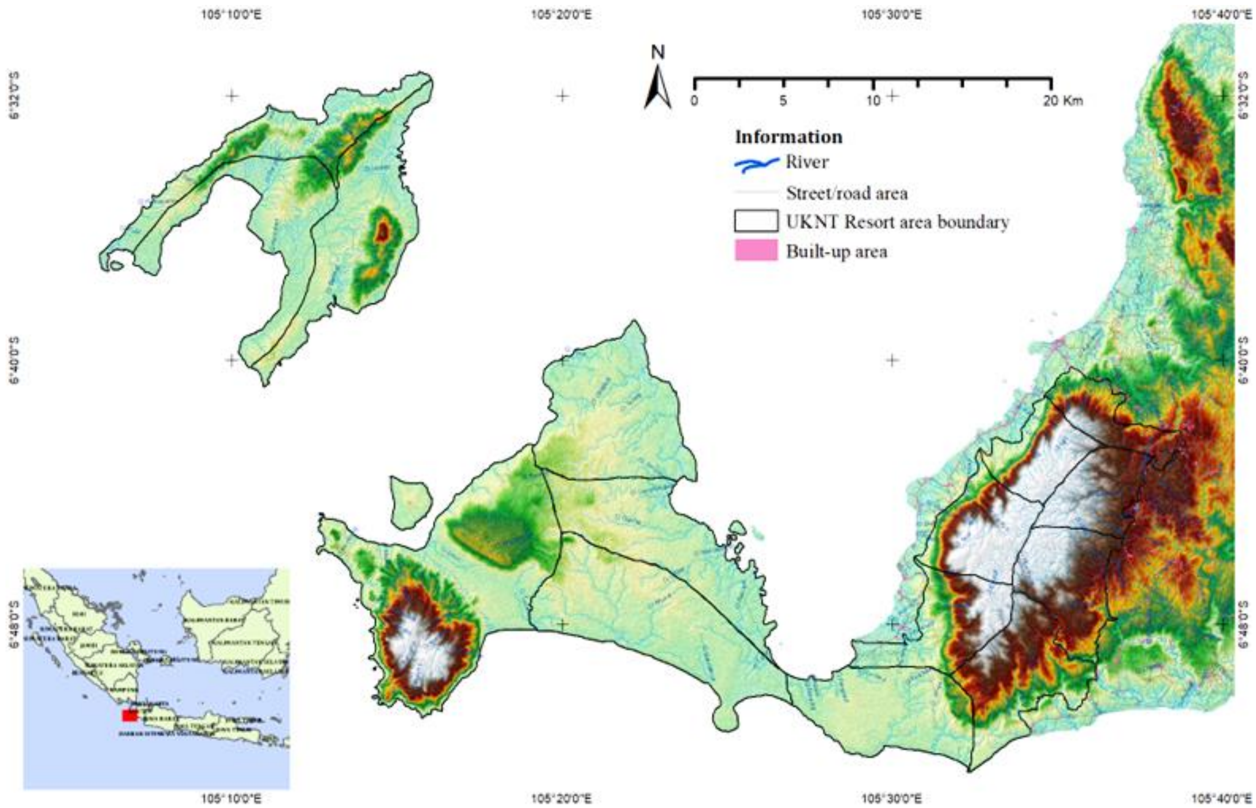

Figure 1. The map of the research site

Land Use - Land Cover Classification in UKNT

The GEE platform was involved in this study to conduct land use classification analysis. The efficiency of spatial data processing offered is one reason the authors are involved in this platform. This platform's integration and data processing capabilities, for example, can run the RF (random forest) classification algorithm and index algorithm. All analysis steps are written in a GEE worksheet with the Javascript programming language and start from calling the image data source, defining training and validation data (Table 1), cloud masking and filtering (using Sentinel-2 quality assurance band/QA60, maximum cloud cover 5%), date filtering (January 2020 to December 2022), involvement of index bands (11 indexes; Table 2), and exporting the raster outcomes. Involvement of the index algorithm in this classification process to help translate the characteristic wavelengths reflected by the earth's surface. The wavelength information is obtained from the values of each band in the satellite sensor.

Table 2. List of involved indexes

No	Method	Formula	Reference
1	Modified Normalized Difference Vegetation Index (MNDVI)	MNDVI = (Red Edge 2 - Red Edge 1) / (Red Edge 2 + Red Edge 1)	Xu 2006
2	Enhanced Vegetation Index (EVI)	EVI = G ((NIR - Red) / (NIR + C1 x Red - C2 x Blue + L))	Huete et al., 2002
3	Soil Adjusted Vegetation Index (SAVI)	SAVI = 1.5 (NIR - Red)/ (NIR + Red + 0.5)	Rouse jr. et al. 1973
4	Atmospherically Resistant Vegetation Index (ARVI)	ARVI = (NIR - (Red - (Blue - Red))) / NIR + (Red - (Blue - Red)))	Kauffman and Tanre 1992
5	Specific Leaf Area Vegetation Index (SLAVI)	SLAVI = NIR / (Red + SWIR)	Lymburner et al., 2000
6	Index- Based Built- up Index (IBI)	IBI = ((NIR)/NIR + Red)) + ((Green)/Green + SWIR1))	Xu 2008
7	Green Normalized Difference Vegetation Index (GNDVI)	GNDVI = (NIR - Green) / (NIR + Green)	Gitelson et al. 1996
8	Normalized Difference Built-up Index (NDBI)	NDBI = (SWIR - NIR) / (SWIR + NIR)	Zha et al. 2003
9	Augmented Normalized difference water index (ANDWI)	ANDWI = (Blue + Green + Red - NIR - SWIR1 - SWIR2) / (Blue + Green + Red + NIR + SWIR1 + SWIR2	Rad et al. 2021
10	Modified Normalized Difference Water Index (MNDWI)	MNDWI = (Green - SWIR1) / (Green - SWIR1)	Xu 2006
11	Land Surface Water Index (LSWI)	LSWI = (NIR - SWIR) / (NIR + SWIR)	Xiao et al., 2002

Accuracy Measurement

Classification of land use types sometimes produces spatial information that is not following actual conditions (Story and Congalton 1986; Congalton 1991). This requires measurement of accuracy as a form of assessment by comparing the classification results with the validation data. This accuracy measurement will conclude the importance of the involvement of the index algorithm combination in the LULC classification process. In this study, the measurement of accuracy involves the equations of kappa statistics, user accuracy, producer accuracy, and overall accuracy (OA). A total of 231 data will be involved in calculating the accuracy, which is separate from the previous training data (Table 1; Congalton and Green 2019). This is to anticipate the bias in the results of the RF classification tested using the same data (Pal 2005).

$$Kappa\ Statistic\ = \frac{N\ \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} X_{i+} X_{+1}}{N\ \sum_{i=1}^{r} X_{i+} X_{+i}}$$

$$User's\ Accuracy = \frac{X_{ii}}{X_{+i}} \times 100\%$$

$$Producer's\ Accuracy = \frac{X_{ii}}{X_i} \times 100\%$$
$$Overall\ Accuracy = \frac{\sum_{i=1}^{r} X_{ii}}{N} \times 100\%$$

Table 3. Interpretation of kappa values

Kappa values	Class interpretation
<0.00	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect

RESULTS

The Condition of the Study Location

The national park congress in Bali in 1982 was the forerunner to establishing this national park. During the congress, Indonesia introduced UKNT as one of five national parks established to support the conservation of endangered species. These animals are threatened with extinction and included in the conservation policy of 25 priority animals from the Indonesian government, such as the Javan rhino (*Rhinoceros sondaicus*; Santiapillai and MacKinnon 1990), Javan banteng (*Bos javanicus*), Javan leopard (*Panthera pardus melas*), Javan gibbon (*Hylobates moloch*; Smith et al. 2017), muncak (*Muntiacus muntjak*), Javan eagle (*Nisaetus bartelsi*), and surili (*Presbytis comata*; Heriyanto and Iskandar 2004). Apart from the 25 priority animals, there are many more protected and sheltered animals in the UKNT forest landscape. Therefore, the UKNT forest area can be said to be the complete national park for protected animals compared to other national parks on the island of Java.

Efforts to preserve Indonesia's biodiversity in natural forests continue to be carried out to protect the genetic resources contained therein. However, this process is not free from problems directly related to the surrounding community's economy. Land conversion, forest encroachment, wildlife theft, and uncontrolled use of natural tourism areas are problems that often occur in this national park. In addition to socio-economic factors, this national park also has a threat, namely the natural tsunami disaster from Mount Anak Krakatau. In 2018, for example, an eruption of Mount Anak Krakatau resulted in a tsunami that hit the northwest part of this national park. This certainly has an impact on forest cover in the tsunami-affected area.

Land Use Classification in UKNT

The GEE classification results in nine land use classes: water body, primary forest, mixed forest, bamboo forest, mangrove forest, bare land, built-up area, crop area, and swamp forest (Figures 2 and 3). The classification results reveal that the water body is the LULC class with the largest area of about 45.44%, and the built-up area is the lowest LULC class with an area of 144.45 ha or about 0.13% (Table 4). Within the UKNT area, the water body is located northwest to the south of this national park. This area includes parts of the river to marine conservation areas. In addition, the second largest LULC class is mixed forest area which covers about 19.53% or approximately 25,616.28 ha of the total area of UKNT.

It is known from a field survey conducted at one of the UKNT resorts that mixed forest consists of large trees mixed with overlapping plants such as honje and tepus (Figure 2E). These clumps of plants cover the entire forest floor very tightly in this area, making it difficult for humans to reach them. According to Purwaningsih and Atikah (2018), the gaps formed by several fallen trees naturally allow seeds to germinate. In addition, this mixed forest is a food corridor for UKNT's main animal, the Javan rhino. This makes the LULC class very important in supporting the habitat of the endangered Javan rhino.

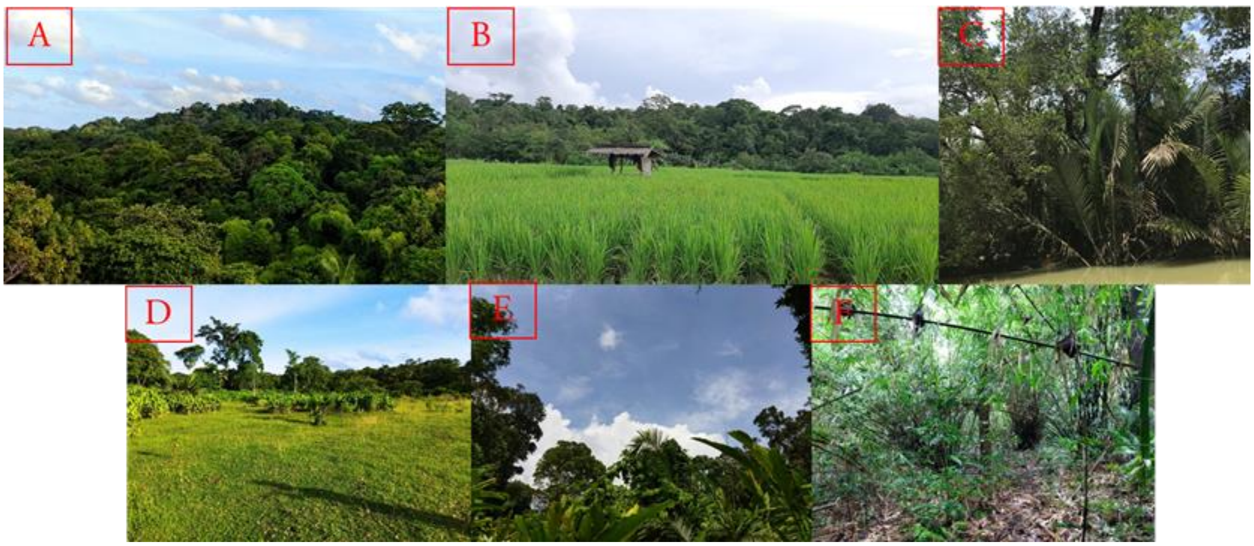

Figure 2. Primary forest (a), crop area (b), mangrove forest (c), bare land (d), mixed forest (e), bamboo forest (f).

Table 4. Comparison of LULC area of classification and improvement

No.	Land use	Area	
		Hectares (ha)	Percentage (%)
1	Water body	49,971.27	44.63
2	Primary forest	21,868.41	19.53
3	Mixed forest	25,616.28	22.88
4	Bamboo forest	6,039.32	5.39
5	Mangroves forest	1,897.88	1.69
6	Bare land	595.66	0.53
7	Built-up area	144.45	0.13
8	Crop area	1,938.02	1.73
9	Swamp forest	871.88	0.78
10	Shrub	3,029.40	2.71
	Total	111,972.57	100.00

The forest area in UKNT is the last remaining habitat for the Javan rhino on the island of Java. In addition to mixed forest, GEE has identified four other forest classes, namely primary forest (19.53%), bamboo forest (5.39%), mangrove forest (1.69%), and swamp forest (0.78%) (Table 4). In some forest classes, this also has a vital role in habitat conservation efforts in UKNT. It is known that the Javan rhino wallowing activity was also recorded in this type of forest. Primary forest is a forest that is quite widely spread in the UKNT area. This type of forest consists of large trees, and the forest floor is overgrown with rare invasive plants. According to Hariyadi et al. (2012) and Febriana et al. (2019), the langkap plant (*Arenga obtusifolia*) distributed in the UKNT forest is one of the problems that threaten the habitat of the Javan rhino. In Santosa et al. (2013), the land overgrown by rare plants was converted into rhizome plantation land as a feeding corridor for herbivores in this national park.

UKNP is essential in the surrounding community, which is a buffer zone. Agriculture, plantations, and tourism are one of the uses of the area by the surrounding community. The RF algorithm detected the presence of agricultural regions within the UKNT area with an area of about 1,938.02 ha (Table 4). This agricultural expansion should not occur in conservation forest areas, given the importance of maintaining the ecosystem function of the national park. This is in accordance with the research of Wandani et al. (2022), which explains that 1,556.82 ha of community-owned farms are located in eight resorts in UKNT. This is due to the increasing number of residents in the villages surrounding the national park, triggering continued dependence on UKNT resources. A similar condition occurs in Meru Betiri National Park. Due to income uncertainty, forests were illegally destructed (Harada et al. 2015). Dauvergne (1994) also describes the World Bank's investigation of deforestation, whose factors are primarily land clearing for agriculture, commercial logging, firewood collection, and animal husbandry. Around 2012 – 2017, Dwiyahreni et al. (2021A) reported that the intense access and community activities contributed to a 10.5% forest cover declination in the UKNT-protected area (Dwiyahreni et al. 2021B).

LULC Classification based on Resort Division

According to the Regulation of the Directorate General of KSDAE-KLHK, UKNT, in the area's management, is divided into several resorts. This national park has 14 resorts as a processing unit with the smallest area. The LULC classification covers all resorts in UKNT and provides an overview of the land cover dynamics at each resort. Li et al. (2017) explained that spatial information regarding changes in LULC is one indicator that can assess the effect of human interaction with the environment. Therefore, any differences in the composition of LULC in each resort are indirectly influenced by the landscape and the dynamics or accessibility of the surrounding population (Dauvergne 1994; Carr 2009; Carr et al. 2005).

Each resort has six resorts, namely Legon Bajo, Peucang Island, Legon Pakis, Cibunar, Karang Ranjang, and Handeleum Island, which are dominated by bodies of water (Figure 4). Meanwhile, the other six resorts, namely Kalejetan Resort, Ketapang, Ranca Pinang, Cibadak, Padali, and Taman Jaya, are dominated by primary forest, and the remaining two resorts are mixed forest (Figure 4). The difference in the dominance of the LULC class between these resorts occurs because five of the 14 resorts do not have water areas. These five resorts are located in the mountainous region of Mount Honje and are directly adjacent to community-owned land. These five resorts also get a strong influence from community activities around the national park's boundaries. In addition, two other resorts influence human activities in the vicinity. This influence can be seen from the expansion of agricultural land that occurred at Legon Pakis Resort (260.62 ha), Tamanjaya Resort (168.59 ha), Ketapang Resort (182.77 ha), Kopi

Resort (77.97 ha), Padali Resort (115.84 ha), Resort Cibadak (205.34 ha), Resort Ranca Pinang (143.25 ha). Therefore, the expansion of community-owned agriculture within the UKNT resort area should not occur. Given the essential function of the area to be protected from forest destruction activities. In addition, several big cats make this area a habitat, so it is considered very prone to conflict with farmers on agricultural land in the UKNT forest area (Gunawan et al. 2017).

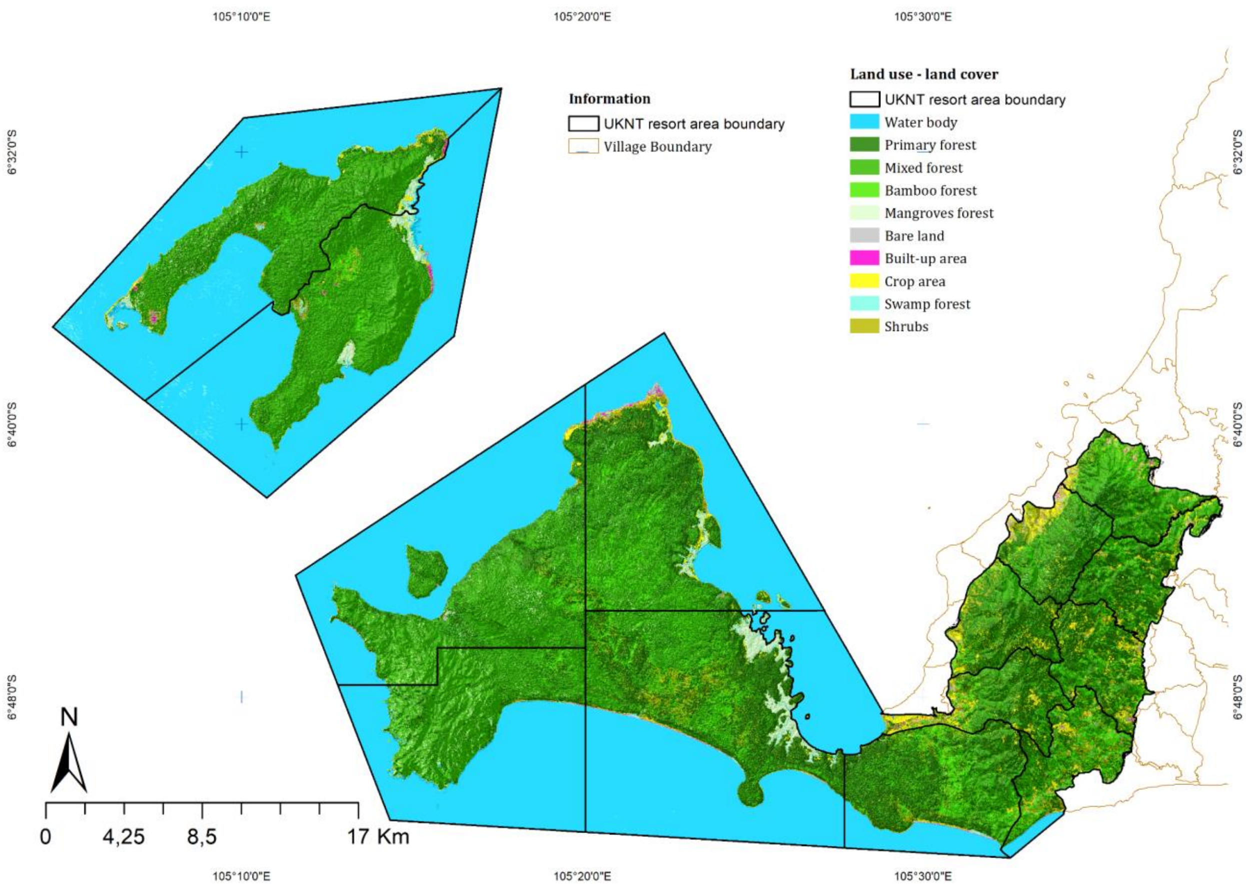

Figure 3. LULC classification results

The other LULC classes have distribution in each resort, although they have a slightly smaller composition. Open land, for example, only has about 0.53% of the total area and is more widely distributed in Resorts Ranca Pinang (84.13 ha), Kopi (83.32 ha), Ketapang (73.67 ha). This open land class is characterized by the presence of an artificial grazing area created by UKNT to assist in monitoring herds of banteng (*Bos javanicus*). This can be seen in Figure 2, part D, which depicts one of the pastures in the P. Handeleum Resort. In addition, open land is associated with shrubs scattered in Resort H. Handeuleum and is a former area affected by the tsunami in 2018. It was previously known that this area was prone to tsunami activity that damaged stretches of coastal forest (Widiyanto and Hsiao, 2020; Widiyanto et al., 2020). Borrero et al. (2020) explained that the impact of tsunami activity as high as 10 meters hit the coastal forest as far as 400 meters into the interior of the UKNT coastal forest area.

Meanwhile, another study revealed that UKNT, within 60 km of the tsunami epicentre, destroyed coastal forests along the coast of Panaitan Island and as far as 200 meters inland (Muhari et al., 2019). In addition, it is essential to note that the impact of the damage caused by this tsunami is threatening the habitat and existence of endangered animals in the world, namely *R. sondaicus* (Setiawan et al. 2018). The field survey conducted by the author shows that this area is an expanse of heath forest and overgrown with shrubs and shows natural forest succession. This is different from the built-up area, which is the smallest LULC class and has an area of about 0.13% of the total area. The area spread throughout the resort ranging from 3 - 29 ha. This LULC class is known to have a distribution in the Legon Pakis Resort area of around 7.32 ha. It is connected to access roads belonging to rural communities along the west coast of Ujung Jaya Village, Pandeglang Regency.

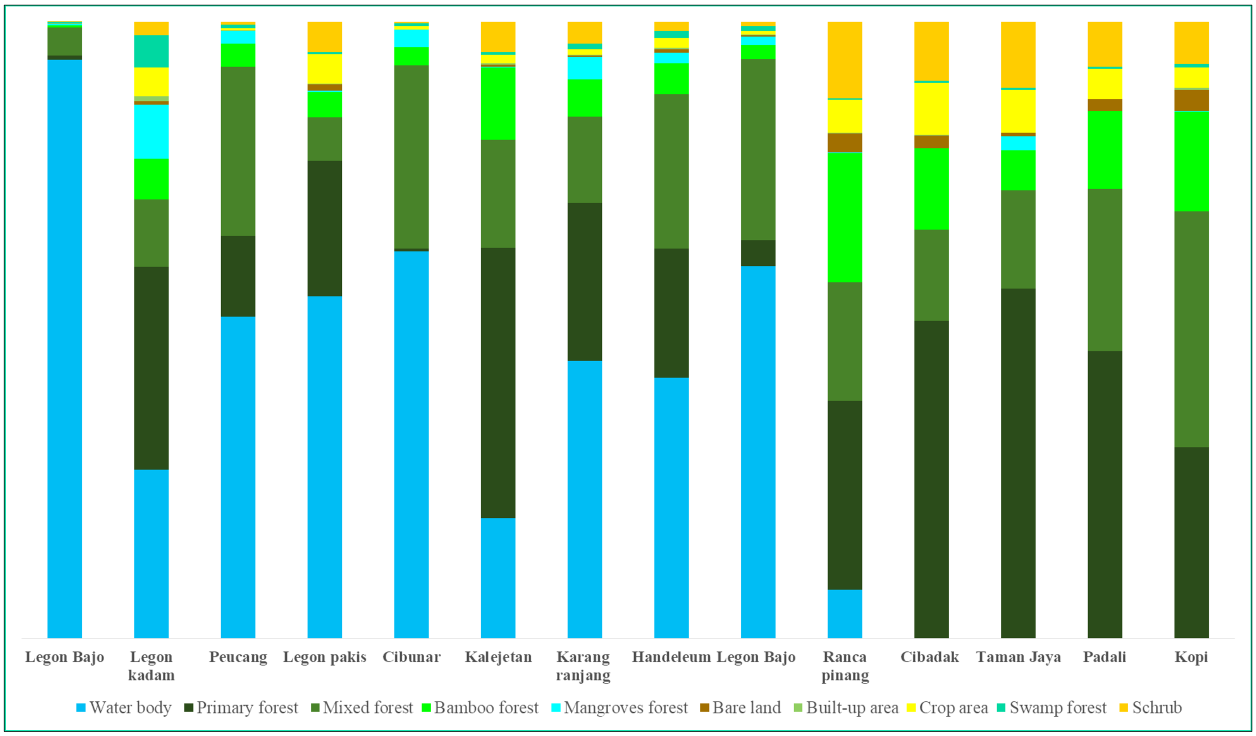

Figure 4. LULC based on UKNT's resorts

Spatial information on the distribution of LULC at each resort is beneficial for managing national parks in strategic conservation plans at the resort level (Galicia and García-Romero 2007). The roles and functions of knowing the LULC information include information on the distribution of vacant land for banteng grazing land, primary forest integrity as a habitat for herbivores and big cats, mixed forest expanses for feed corridors, mangrove forests and swamps for conservation of aquatic animals, built-up land and agriculture for mitigation. Wildlife conflicts, and so on. In addition, the resort-level LULC class information provides an overview of the changes in LULC in this national park. By extracting spatial information, it can also provide an overview of deforestation activities in UKNT in each particular period and changes due to being affected by other natural disasters. In its development, UKNT has experienced various natural disasters, such as forest fires which are thought to be caused by the existence of a peat dome on Panaitan Island and the impact of the tsunami waves along the coast of Handeleum Island.

Classification Accuracy

The LULC classification generally provides spatial information that requires further accuracy testing. This is used to test the accuracy of the classification results to provide spatial information that can be used reliably by users. This accuracy was tested using the Overall Accuracy (OA) and Kappa Statistics (KS) formulas and obtained values of 80.08% and 0.78, respectively. A total of 231 field validation data were tested in this accuracy assessment to get the Substantial class (Table 3). Meanwhile, according to Scepan (1999), the best accuracy value can be seen from the OA value, which reaches 85%, so the OA value generated from the study does not reach the limit value. This is due to the discrepancy between the classification results and the validation data, which is 46 data. This proves that further studies are needed regarding this method to produce the best accuracy value (including exceeding 85%) in classifying LULC with high heterogeneity. In addition, an evaluation of index involvement needs to be carried out, considering the different capabilities of each index and assessing its contribution to the classification process through variable importance analysis (Figure 5).

Constraints and Limitations: Index Assessment and Classification Process

Based on the classification results, seven resorts should not have agricultural land but are classified as having agricultural land. This is a misclassification that occurred, so this spatial information needs improvement. This is defined because agriculture is easy to identify in the area division in UKNT, while the other LULC classes cannot be identified directly. Almost all LULCs have distribution in all resorts. However, some classes of LULC should not have distribution in any of the resorts. For example, agriculture and mangrove forests are LULC classes that are not spread throughout the resort.

Meanwhile, swamp forests and built-up land require further investigation by verifying the classification results and comparing them in the field. This is due to the construction of offices or supporting operational facilities for conservation and tourism activities at every resort in UKNT. Therefore, the built-up land class is found in all resorts in UKNT. In the case of mangroves, this study found 1,897.88 ha, whereas, in previous studies, the area of mangroves in this national park was 547.34 ha (Asy'Ari et al., 2022). This is considered necessary for an evaluation of the LULC classification process. In the same case, the classification process should be carried out at each resort by considering the representation of the LULC class to reduce the influence of the LULC class, which should not be in its area.

The combination of index algorithms in this analysis provides an overview of one of the causes of misclassification. We have included several indices considered to have a significant contribution opportunity, evident in Figure 5. 11 indices were involved in classifying 9 LULC classes and were assessed to be highly heterogeneous. Figure 5 has proven the contribution of these 11 indices in assisting the RF classification algorithm in classifying 9 LULC classes in UKNT. Starting from the ANDWI index is the highest index, with a contribution of 14.10%, and the LSWI index has the lowest contribution value of 5.06%. In addition to the index, insufficient training data is also considered to influence the classification process. This is a challenge for research on LULC classification in large-scale primary forest areas due to difficult access to LULC classes that will be used as training data.

Every surface of the earth, specifically the LULC class, has a threshold value when viewed from the point of view of the characteristics of the index algorithm (Aprilianti et al. 2021). There are several indices involved that have the same characteristics between more than one LULC class. This occurs in the LULC class between shrubs and agriculture and between built-up and open land. In some resort locations, a LULC class should not exist. For example, the agricultural LULC class at Resort Karang Ranjang, P. Handeleum, P. Peucang, Cibunar, Legon Kadam, and Legon Bajo. The six resorts should not have the distribution of agricultural land because community access is difficult and far from community villages. This also occurs in the LULC class of built-up land identified along the north to the west coast of Resort P. Handeleum. The entire class spread in several heath forest locations on the former tsunami's coast. This causes the classification algorithm to experience errors in interpreting the earth's surface and provides inappropriate spatial information. This is caused by the high heterogeneity in the forest landscape, so the classification algorithm experiences confusion in the decision-making process (Ghimire et al., 2010).

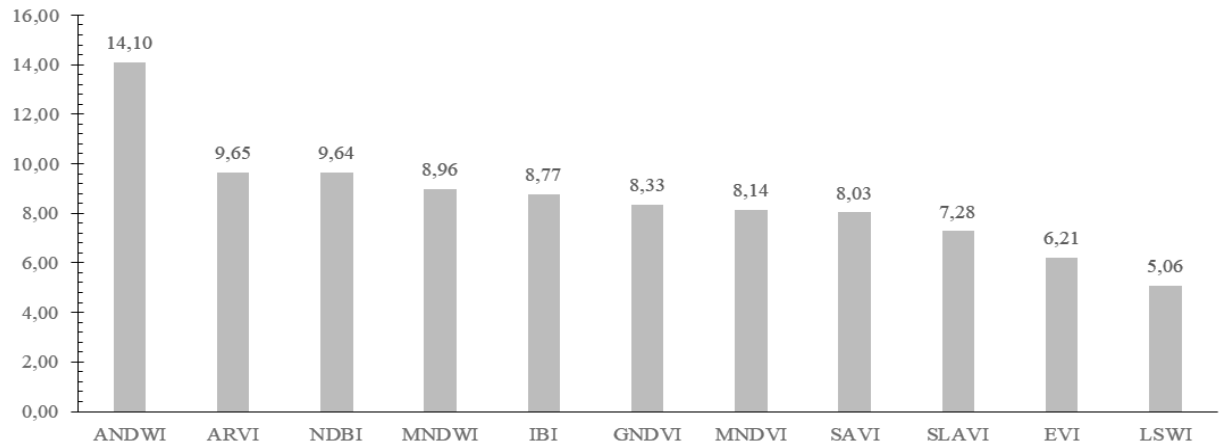

Figure 5. Important variables in the classification process using 11 indices

The heterogeneity of the forest landscape in UKNT affects the indexes' involvement in the classification process. This is identified at the time of selecting which indices to include. The level of landscape heterogeneity influences the classification process (Qu et al. 2021); hence, researchers must be careful in classifying. Ghimire et al. (2010) also revealed that this is a challenge in the LULC classification process due to the high intra-class variability and heterogeneous landscape artifacts. The index's ability to assess the earth's surface or distinguish each earth's surface based on its pixel value is one of the most critical reasons before starting the classification process. Therefore, this error requires further investigation into the cause of this misclassification to prove confidence in the resulting accuracy. This was also carried out to test the ability of the RF classification algorithm to classify in high heterogeneity landscapes, which previously had better capabilities than the support vector machine algorithm (SVM; Pal 2005; Belgiu and Dragut, 2016).

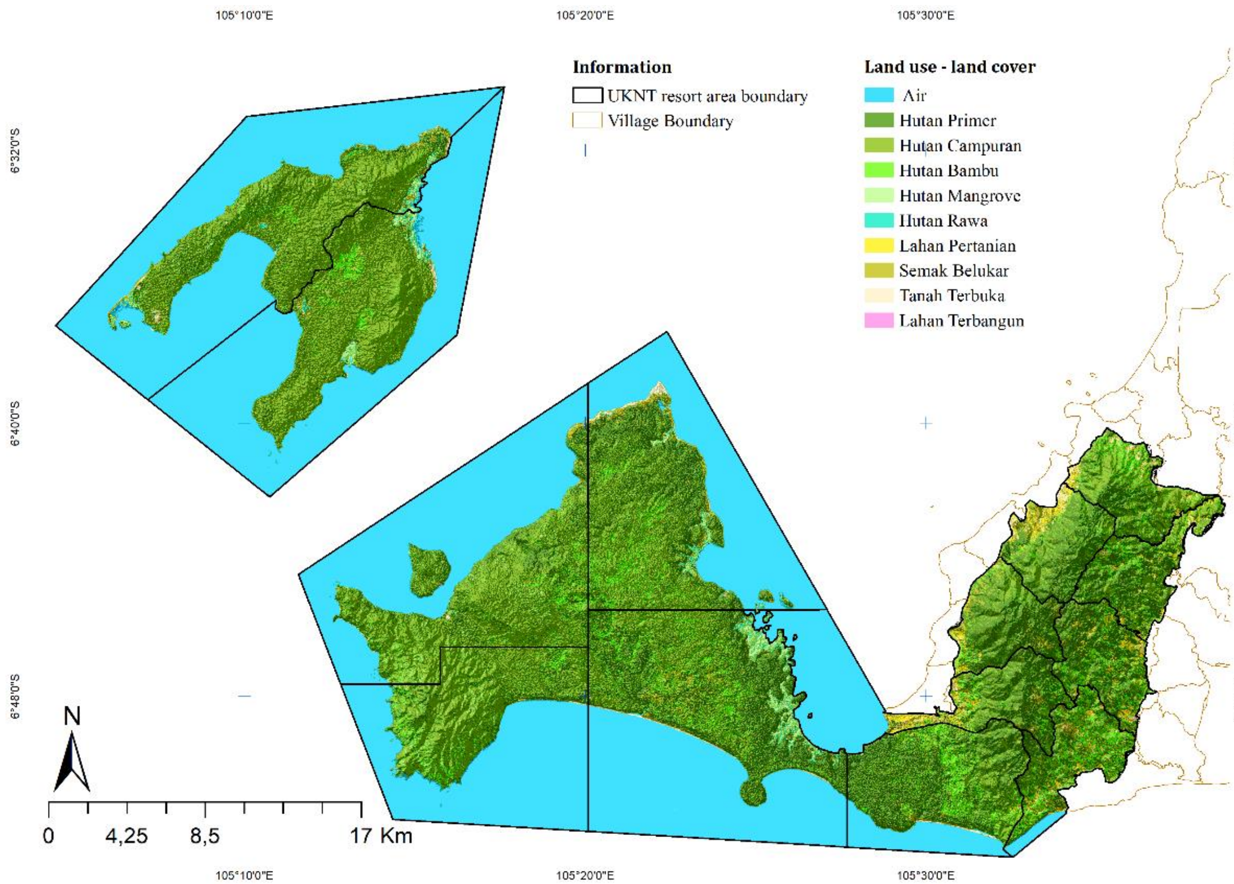

Figure 6. Spatial distribution of improved LULC

Table 5. LULC class area

LULC Class	Area (ha)	Percentage (%)
Water body	494,64	44,74
Primary forest	218,04	19,72
Mixed forest	257,61	23,30
Bamboo forest	57,68	5,22
Mangroves forest	18,34	1,66
Bare land	7,13	0,64
Built-up area	0,11	0,01
Crop area	7,22	0,65
Swamp forest	4,43	0,40
Schrub	40,42	3,66
Total	1105,61	100,00

To ensure the suitability of the classification results, the improvement of the RF classification results is carried out in the ArcMap software. Improvements to the LULC class take precedence over the LULC class, which should not have distribution across multiple locations. There are four classes of LULC whose distribution is improved: swamp forest, water bodies, agricultural land and built-up land. The results of the spatial and extensive LULC improvements are presented in Figure 6 and Table 5. Although there is no visible change spatially, the change in land area value is very important in the LULC classification. Spatial information from the results of this improvement can then be used in further research to make policy decisions.

CONCLUSIONS

This study succeeded in mapping LULC in Ujung Kulon National Park (UKNT), which possessed a high level of heterogeneity. The compactness of 11 indexes through the random forest algorithm produces 10 LULC classes and shows that water bodies have the most significant percentage of the area and the least amount of built-up land. In this case, the high level of heterogeneity disrupts the LULC classification process so that the accuracy values only reach 80.08% (OA) and 0.78 (KS). This condition creates a classification error that requires classification correction before the spatial information dissemination of the LULC data is carried out. Therefore, the combination of the index, training data, classification algorithm (RF), and determination of the classification area is considered to need further evaluation to solve the confusion in research on this topic. It is hoped that this research will become a reference material for further research and consideration for UKNT management.

REFERENCES

Allen JC & Barnest DF. 1985. The causes of deforestation in developing countries. *The Annals of the American Association of Geographers*, 75(2): 163-184. https://doi.org/10.1111/j.1467-8306.1985.tb00079.x

Aprilianti HS, Ari RA, Ranti A & Aslam MF. 2021. Identification and classification of cloud computing-based vegetation index values on several lands used in Bogor Regency, Indonesia. *IOP Conference Series: Earth and Environmental Science*, 918: 012011. https://doi.org/10.1088/1755-1315/918/1/012011

Armenteras D, Rudas G, Rodriguez N, Sua S, & Romero M. 2006. Patterns and causes of deforestation in the Colombian Amazon. *Ecological Indicators*, 6(2):353–368. https://doi.org/10.1016/j.ecolind.2005.03.014

Austin KG, Schwantes A, Gu Y, & Kasibhatla PS. 2019. What causes deforestation in Indonesia? *Environmental Research Letters*, 14(2). https://doi.org/10.1088/1748-9326/aaf6db

Asy'ari R, Rahmawati AD, Sa'diyya N, Gustawan AH, Setiawan Y, Zamani NP, Pramulya R. 2022. Mapping mangrove forest distribution on Banten, Jakarta, and West Java Ecotone Zone from Sentinel-2-derived indices using cloud computing based Random Forest. *The Journal of Natural Resources and Environment Management (JPSL)*, 12(1): 97-111. https://doi.org/10.29244/jpsl.12.1.97-111

Belgiu M & Drăguţ L. 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114: 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Brook SM, Dudley N, Mahood SP, Polet G, Williams AC, Duckworth JW, van Ngoc T, & Long B. 2014. Lessons learned from the loss of a flagship: The extinction of the Javan rhinoceros *Rhinoceros sondaicus annamiticus* from Vietnam. *Biological Conservation*, 174:21–29. https://doi.org/10.1016/j.biocon.2014.03.014

Borrero JC, Solihuddin T, Fritz HM, Lynett PJ, Prasetya GS, Skanavis V, Husrin S, Kongko W, Istiyanto DC, Daulat A, & Purbani D. 2020. Field survey and numerical modelling of the December 22, 2018 Anak Krakatau tsunami. *Pure and Applied Geophysics*, 177(6): 2457-2475. https://doi.org/10.1007/s00024-020-02515-y

Carr D. 2009. Population and deforestation: why rural migration matters. *Progress in Human Geography*, 33(3): 355-378. https://doi.org/10.1177/0309132508096031

Carr DL, Suter L, & Barbieri A. 2005. Population dynamics and tropical deforestation: state of the debate and conceptual challenges. *Population and Environment*, 27(1): 89-113. https://doi.org/10.1007%2Fs11111-005-0014-x

Chatelain C, Gautier L, & Spichiger R. 1996. A recent history of forest fragmentation in southwestern Ivory Coast. *Biodiversity & Conservation*, 5(1): 37–53. https://doi.org/10.1007/BF00056291

Congalton R & Green K. 2019. *Assessing the Accuracy of Remotely Sensed Data Principles and Practices*, *Third Edition*. United States: CRC Press.

Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1): 35-46. https://doi.org/10.1016/0034-4257(91)90048-B

Dauvergne P. 1993. The Politics of Deforestation in Indonesia. *Pacific Affairs*, 66(4):497. https://doi.org/10.2307/2760676

Dwiyahreni AA, Fuad HA, Muhtar S, Soesilo TE, Margules C, & Supriatna J. 2021. Changes in the human footprint in and around Indonesia's terrestrial national parks between 2012 and 2017. *Scientific reports*, 11(1): 1-14. https://doi.org/10.1038/s41598-021-83586-2

Dwiyahreni AA, Fuad HA, Soesilo TE, Margules C, & Supriatna J. 2021. Forest cover changes in Indonesia's terrestrial national parks between 2012 and 2017. *Biodiversitas*, 22(3): 1235-1242. https://doi.org/10.13057/biodiv/d220320

Fattore C, Abate N, Faridani F, Masini N, & Lasaponara R. 2021. Google earth engine as multi-sensor open-source tool for supporting the preservation of archaeological areas: The case study of flood and fire mapping in Metaponto, Italy. *Sensors*, 21(5):1–27. https://doi.org/10.3390/s21051791

Febriana I, Kusmana C, & Rahmat UM. 2020. Komposisi jenis tumbuhan dan analisis sebaran langkap (*Arenga obtusifolia* Mart.) di Taman Nasional Ujung Kulon. *Journal of Natural Resources and Environmental Management*, 10(1): 52-65. https://doi.org/10.29244/jpsl.10.1.52-65

Floreano IX & de Moraes LAF. 2021. Land use/land cover (LULC) analysis (2009–2019) with Google Earth Engine and 2030 prediction using Markov-CA in the Rondônia State, Brazil. *Environmental Monitoring and Assessment*, 193(4): 1-17. https://doi.org/10.1007/s10661-021-09016-y

Galicia L & García-Romero A. 2007. Land use and land cover change in highland temperate forests in the Izta-Popo National Park, Central Mexico. *Mountain Research and Development*, 27(1): 48-57. https://doi.org/10.1659/0276-4741(2007)27[48:LUALCC]2.0.CO;2

Ghimire B, Rogan J, & Miller J. 2010. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters*, 1(1): 45-54. https://doi.org/10.1080/01431160903252327

Gitelson AA, Kaufman YJ, Merzlyak MN. 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote sensing of Environment,* 58(3): 289-298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, & Moore R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gunawan H, Iskandar S, Sihombing VS, & Wienanto R. 2017. Conflict between humans and leopards (*Panthera pardus melas* Cuvier, 1809) in Western Java, Indonesia. *Biodiversitas Journal of Biological Diversity*, 18(2): 652-658. https://doi.org/10.13057/biodiv/d180229

Harada K, Prabowo D, Aliadi A, Ichihara J, Ma HO. 2015. How can social safeguards of REDD+ function effectively conserve forests and improve local livelihoods? A case from Meru Betiri National Park, East Java, Indonesia. *Land*, 4(1):119–139. https://doi.org/10.3390/land4010119

Hariyadi AR, Priambudi A, Setiawan R, Purnama H, & Yayus A. 2012. Optimizing the habitat of the Javan rhinoceros (*Rhinoceros sondaicus*) in Ujung Kulon National Park by reducing the invasive palm *Arenga obtusifolia*. *Pachyderm*, 52: 49-54.

Heriyanto NM, & Iskandar S. 2004. Status populasi dan habitat sural (*Presbytis comata* desmarest) di kompleks hutan Kalajeten-Karangranjang, Taman Nasional Ujung Kulon. *Jurnal Penelitian Hutan dan Konservasi Alam*, 1(1): 89-98. https://doi.org/10.20886/jphka.2004.1.1.89-98

Huete A, Didan K, Miura T, Rodriguez EP, Gao X & Ferreira LG. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment,* 83(1-2): 195-213. https://doi.org/10.1016/S0034-4257(02)00096-2

Kaufman YJ & Tanre D. 1992. Atmospherically resistant vegetation index (ARVI) for EOSMODIS. IEEE Transactions on Geoscience and Remote sensing, 30(2): 261-270. https://doi.org/10.1109/36.134076

Kumar L & Mutanga O. 2018. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sensing*, 10(10). https://doi.org/10.3390/rs10101509

Li W, Dong R, Fu H, Wang J, Yu L, & Gong P. 2020. Integrating Google Earth imagery with Landsat data to improve 30-m resolution land cover mapping. *Remote Sensing of Environment*, 237: 111563. https://doi.org/10.1016/j.rse.2019.111563

Lymburner L, Beggs PJ & Jacobson CR. 2000. Estimation of canopy-average surface- specific leaf area using Landsat TM data. *Photogrammetric Engineering and Remote Sensing*, 66(2): 183-192.

Milto KD & Lukin YA. 2020. A revised herpetofauna of Ujung Kulon National Park, West Java, Indonesia. *Russian Journal of Herpetology*, 27(6):353–368. https://doi.org/10.30906/1026-2296-2020-27-6-353-368

Muhari A, Heidarzadeh M, Susmoro H, Nugroho HD, Kriswati E, Wijanarto AB, Imamura F, & Arikawa T. 2019. The December 2018 Anak Krakatau Volcano tsunami as inferred from post-tsunami field surveys and spectral analysis. *Pure and Applied Geophysics*, 176(12): 5219-5233. https://doi.org/10.1007/s00024-019-02358-2

Mutanga O & Kumar L. 2019. Google earth engine applications. *Remote Sensing*, 11(5). https://doi.org/10.3390/rs11050591

Nijman V. 2013. One hundred years of solitude: Effects of long-term forest fragmentation on the primate community of Java, Indonesia. In Marsh LK & Collins CA (Eds.) *Primates in Fragments: Complexity and Resilience*, Springer New York. hlm 33–45. https://doi.org/10.1007/978-1-4614-8839-2_3

Rad AM, Kreitler J, Sadegh M. 2021. Augmented Normalized Difference Water Index for improved surface water monitoring. *Environmental Modelling & Software*, 140: 105030. https://doi.org/10.1016/j.envsoft.2021.105030

Rouse jr JW, Haas RH, Schell JA & Deering DW. 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA special publication. 351(1974): 309

Sandground JH. 1933. Two new helminths from *Rhinoceros sondaicus*. *The Journal of Parasitology*, 19(3): 192-204. https://doi.org/10.2307/3271925

Santiapillai C & MacKinnon K. 1990. *Conservation and Management of Rhinos in Indonesia*. WWF - Asia Programme.

Santosa Y, Mamat Rahmat U, Prasetyo LB, & Kartono AP. 2013. Javan rhino (*Rhinoceros sondaicus* desmarest 1822) utilization distribution and habitat selection in Ujung Kulon National Park. *Jurnal Manajemen Hutan Tropika*, 19(1):31–38. https://doi.org/10.7226/jtfm.19.1.31

Scepan J. 1999. Thematic validation of high-resolution global land-cover data sets. *Photogrammetric Engineering and Remote Sensing*, 65: 1051-1060.

Setiawan R, Gerber BD, Rahmat UM, Daryan D, Firdaus AY, Haryono M, Khairani KO, Kurniawan Y, Long B, Lyet A, & Muhiban M. 2018. Preventing global extinction of the Javan rhino: tsunami risk and future conservation direction. *Conservation Letters*, 11(1): 12366. https://doi.org/10.1111/conl.12366

Simamora TI, Purbowo SD, & Laumonier Y. 2021. Looking for indicator bird species in the context of forest fragmentation and isolation in West Kalimantan, Indonesia. *Global Ecology and Conservation*, 27. https://doi.org/10.1016/j.gecco.2021.e01610

Smith JH, King T, Campbell C, Cheyne SM, & Nijman V. 2018. Modelling population viability of three independent Javan gibbon (*Hylobates moloch*) populations on Java, Indonesia. *Folia Primatologica*, 88(6): 507-522. https://doi.org/10.1159/000484559

Story M & Congalton RG. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3): 397-399.

Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, & Brisco B. 2020. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164:152–170. https://doi.org/10.1016/j.isprsjprs.2020.04.001

Pal M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1): 217-222. https://doi.org/10.1080/01431160412331269698

Purwaningsih P & Atikah TD. 2018. Diversitas floristik dan struktur vegetasi di hutan Gunung Payung, Taman Nasional Ujung Kulon. *Berita Biologi*, 17(3): 335-349. http://dx.doi.org/10.14203/beritabiologi.v17i3.3434

Qu L, Chen Z, Li M, Zhi J, & Wang H. 2021. Accuracy improvements to pixel-based and object-based LULC classification with auxiliary datasets from google earth engine. *Remote Sensing*, 13(3). https://doi.org/10.3390/rs13030453

Wandani RA, Asy'Ari R, Setiawan Y, & Anggodo A. 2022. Deteksi ekspansi padi pada lanskap hutan di Taman Nasional Ujung Kulon, Indonesia menggunakan algoritma RF dan Sentinel-2 multispectral instrumen. *National Multidisciplinary Sciences*, 1(2):235–245. doi:10.32528/nms.v1i2.64. https://doi.org/10.32528/nms.v1i2.64

Widiyanto W & Hsiao SC. 2020. Potential tsunami hazard in Ujung Kulon National Park. *IOP Conference Series: Materials Science and Engineering*, 982(1): 012034. https://doi.org/10.1088/1757-899X/982/1/012034

Widiyanto W, Hsiao SC, Chen WB, Santoso PB, Imananta RT, & Lian WC. 2020. Run-up, inundation, and sediment characteristics of the 22 December 2018 Sunda Strait tsunami, Indonesia. *Natural Hazards and Earth System Sciences*, 20(4): 933-946. https://doi.org/10.5194/nhess-20-933-2020

Xiao X, Boles S, Frolking S, Salas W, Moore Iii B, Li C, He L & Zhao R. 2002. Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using vegetation sensor data. *International Journal of Remote Sensing*, 23(15): 3009-3022. https://doi.org/10.1080/01431160110107734

Xu H. 2006. Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing,* 27(14): 3025-3033. https://doi.org/10.1080/01431160600589179

Xu H. 2008. A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing*, 29(14): 4269-4276. https://doi.org/10.1080/01431160802039957

Yanuar A & Chivers DJ. 2010. Impact of forest fragmentation on ranging and home range of siamang (*Symphalangus syndactylus*) and agile gibbons (*Hylobates agilis*). In Gursky S & Supriatna J (Eds.) *Indonesian Primates*, Springer New York. 97–119. https://doi.org/10.1007/978-1-4419-1560-3_7

Zha Y, Gao J & Ni S. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing,* 24(3): 583-594. https://doi.org/10.1080/01431160304987