Jurnal Ilmu Pertanian Indonesia (JIPI), October 2025
(Indonesian Journal of Agricultural Sciences)
ISSN 0853-4217 EISSN 2443-3462

Vol. 30 (4): 765–772
http://journal.ipb.ac.id/index.php/JIPI
DOI: 10.18343/jipi.30.4.765

OPEN ACCESS Check for updates

Land Use Analysis Using Machine-Learning Based on Cloud Computing Platform

Syukur Toha Prasetyo, Fahmi Arief Rahman*, Sinar Suryawati, Slamet Supriyadi, Eko Setiawan

(Received December 2023/Accepted July 2025)

ABSTRACT

Land use analysis can provide a foundation for successful and efficient regional planning and environmental monitoring. The application of machine-learning on a cloud computing platform (Google Earth Engine, GEE) in land use analysis enables efficient and rapid processing of spatial data on a wide scale. It overcomes the constraints inherent in conventional approaches. The purpose of this study was to identify land use and estimate its level of accuracy using GEE and a Random Forest machine-learning method. The data utilized were the administrative boundaries of Bangkalan Regency (1:25,000) and Landsat 8 SR L2 C2 T1 satellite images from 2022. Satellite image analysis using the Random Forest algorithm on the GEE platform with the JavaScript API, including masking, cloud masking, class and sampling, training, and testing sample data. Land use study using the Random Forest algorithm yielded the following results in order of area: vegetation 65,040.39 ha (49.98%), agricultural land 31,817.16 ha (24.45%), settlements 20,578.05 ha (15.81%), open land 6,683.94 ha (5.14%), and water bodies 6,021.09 ha (4.63%). The accuracy test in GEE revealed an overall accuracy (OA) of 91.39% and a kappa score of 88.39%, or 0.88. At the same time, validation in the field gave an OA of 88.68% and a Kappa of 85.53%. The findings of this study can be applied to land use evaluation and fundamental decision-making.

Keywords: land use, random forest, geographic information system, remote sensing

INTRODUCTION

Land is a resource that can change function due to a variety of reasons, including population expansion, terrain changes, land value, and community socioeconomic conditions (Susanti *et al.* 2020). Thus, land use analysis can be useful in regional planning and environmental monitoring (Zulfajri *et al*. 2021). Proper land use analysis is critical for understanding patterns and changes to promote more effective, efficient, and sustainable planning. Land use is defined as human activities that are directly tied to land and may be easily understood through mapping land use classification (Lestari and Arsyad 2018). Land use classification can be done using either conventional methods or remote sensing. Remote sensing utilizing satellite imagery is more effective and efficient than previous methods (Firmawan and Nirmala 2021).

Remote sensing is commonly utilized in land use analysis, but it has not been widely integrated with cloud computing platform-based machine-learning algorithms like Google Earth Engine (GEE). The platform provides efficient and rapid processing of spatial data at scale, overcoming the limits of traditional methods that necessitate high-performance equipment, substantial data storage, and sophisticated analytic processes. Arifin (2017) conducted research on land use changes following the construction of the Suramadu bridge in Bangkalan District, and Rahman and Adiputra (2022) studied the conversion of agricultural land into settlements in Bangkalan Regency, both using conventional methods that require high-specification devices, satellite imagery, and manual interpretation for a long time. The GEE platform is more convenient than traditional methods of processing data from various satellites in terms of speed, flexibility, and cost, and it includes machine-learning algorithms such as Random Forest (RF) to speed up the classification process (Febriani *et al*. 2022; Suryono *et al*. 2022).

The purpose of this study was to categorize land use and determine the accuracy of classification results using GEE and a Random Forest machine-learning algorithm. The findings of this study can serve as a reference for future land use planning that is faster, low-cost, more effective, and efficient.

METHODS

Research Site

The study focused on land use analysis using GEE, and the findings were validated by fieldwork in Bangkalan Regency, one of the districts on Madura Island's western tip. The district's coordinates are 1120 40'−1130 08' E and 60 51'−7 0 11' LS, with an area of

Study Program of Agrotechnology, Faculty of Agriculture, University of Trunojoyo Madura, Bangkalan 69162
* Corresponding Author:
Email: fahmi.rahman@trunojoyo.ac.id

1,260.15 km² and elevation of 2−100 m above sea level (Central Statistics Agency 2023).

Tools and Materials

This investigation was conducted using Lenovo laptops (8 GB of RAM), digital cameras, Avenza Maps, stationery, and Google Maps. Data was processed and analyzed using the GEE platform, followed by QGIS and Microsoft Office. The administrative map of Bangkalan Regency, which comes from the Indonesian Terrain Map (RBI) on a scale of 1:25,000 from https://tanahair.indonesia.go.id; and Landsat 8 Surface Reflectance Level 2 Collection 2 Tier 1 satellite imagery from January 1 to December 31, 2022, with a spatial resolution of 30 m and a temporal resolution of 16 d from the GEE platform.

Data Engineering and Analysis

- **Google Earth Engine**

GEE is a cloud computing platform that processes data. The data processing procedure began with the upload of Bangkalan Regency's administrative map to GEE assets. The map was presented on the map panel by making a Landsat 8 Surface Reflectance Level 2 Collection 2 Tier 1 satellite request with JavaScript API programming.

- **Masking and cloud masking**

Masking is a filter that mixes satellite photos from various time periods (Novianti 2021). At this point, it mixes Landsat 8 satellite imagery with January 1–December 31, 2023. Cloud masking is a technique for reducing cloud cover and obtaining cloud-free satellite images (Rizaldi *et al.* 2023). Through cloud masking, the study region of Bangkalan Regency was achieved with minimal or no cloud cover.

Class creation, sampling, and classification process

Class creation determines land use classes by modifying existing formulae. This study looked at five types of land: water bodies, agricultural land, settlements, vegetation, and open ground. Each land class was sampled by marking it with points and polygons. The collected samples passed through the process training area before being divided into 70% training samples and 30% testing samples. The next step was to classify land use using supervised machine learning and the random forest technique. This technique worked by generating several decision trees that forecast the sample class based on the most mergers of each tree (Rizaldi *et al*. 2023).

Test accuracy

Correctness tests were used to assess the correctness of land use classification results (Marlina 2022). The GEE platform included a script for doing overall accuracy and Kappa analysis; thus, the results were available instantly after running the script. Mathematically, the OA and kappa values can be derived using the confusion matrix generated by GEE's land use analysis results. The kappa analysis can be calculated using the equations below:

$$\text{Kappa (K)} = \frac{N\ \sum_{i=1}^{r} \text{Xii} - \sum_{i=1}^{r} (\text{Xi+x+1})}{N^2 - \sum_{i=1}^{r} (\text{Xi+x+1})}$$

where:

N = number of validated land closure points
*X*i+ = number of points of validation results in the type of land closure i
X+i = number of points of interpretation results in the type of land closure i
*X*ii = number of types of land cover i–i interpretation result (diagonal row)
r = number of land use types

The values in the Kappa analysis range from 0 to 1, with the interpretation of the Kappa values presented in Table 1.

Field Check Validation Test

The validation test was carried out by comparing the categorization findings and directly seeing the field conditions. The validation process was carried out at 53 places throughout the Bangkalan Regency. The validation criteria included classified places such as water bodies, agricultural land, settlements, vegetation, and open land. The validation test results were also used to compare the GEE classification results to the validation results in terms of OA and Kappa values. In addition to validation, the researcher visualized the processed data by comparing satellite sightings, categorization results using Random Forests, and field sightings.

Map Overlay

The map overlay concluded this set of investigations. At this point, spatial data analysis was

Table 1 Kappa values (Landis and Koch 1997)

K value	Interpretation
< 0.00	Very Poor
0.00–0.20	Poor
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81 1.00	Very Good

performed, including land use analysis using the Bangkalan map visuals and colors of categorization results in QGIS 3.32 2023 software. The application did data modeling and map layout based on the findings of land categorization, resulting in a map of land cover and its attributes. In addition, the area of each land use class could be calculated using QGIS.

Research Flow Chart

This study followed a systematic procedure (Figure 1) for creating a land cover map of Bangkalan Regency with Landsat 8 satellite imagery and the GEE platform. The technique included cloud masking, class formation, and sample collecting, followed by Random Forest classification. A field check validated accuracy, which was then verified using the Kappa coefficient. If the classification accuracy exceeds 80% (Kappa > 0.8), the result advances to the overlay step and culminates in the creation of a land coverage map. This method provides accurate, data-driven land cover classification with geographical validation.

RESULTS AND DISCUSSION

Land Use Conditions in 2022

Land use analysis from Landsat 8 satellite data in 2022 using a Random Forest method yielded the following percentage area results: vegetation (49.98%), agricultural land (24.45%), settlements (15.81%), open land (5.14%), and water bodies (4.63%). Vegetation can occur in main forests, secondary forests, plantations, and land uses other than those included in this study. The district's vegetation was evenly dispersed. Except for the western, northern, and a portion of the south and southeast, Bangkalan's land usage consisted of water bodies. The Bangkalan area was evenly distributed in terms of agricultural land, settlements, and open ground, except for the outer (water body) and the middle, which was also a mountainous area (Figure 2).

According to Table 2, the land use in Bangkalan Regency in 2022 was dominated by vegetation and agricultural land. According to Nuraida *et al*. (2022), vegetation is a collection of many tree varieties that coexist and interact in a given location. The vegetation

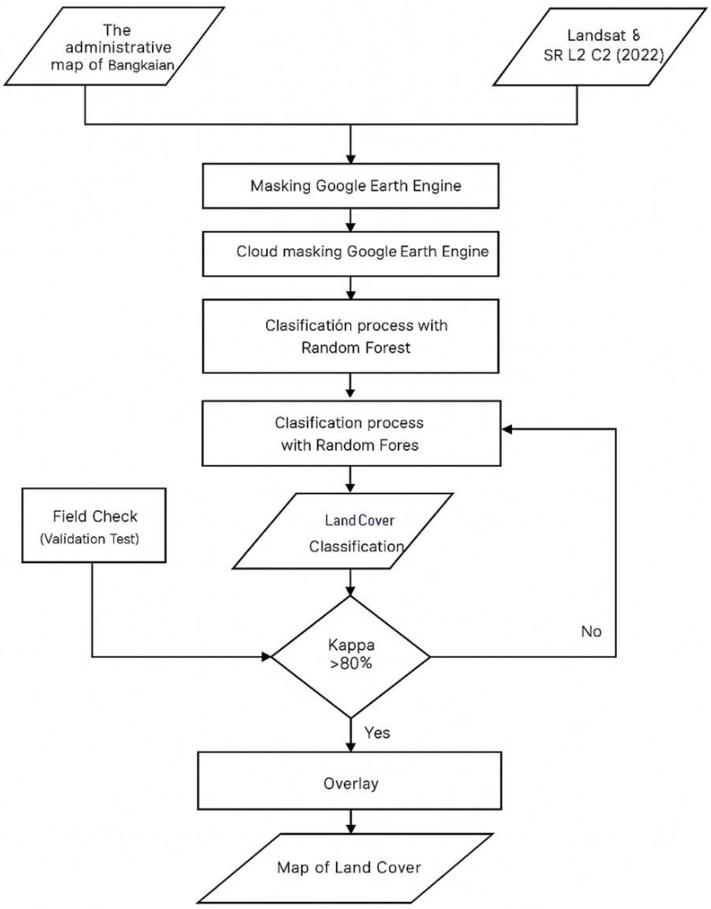


Figure 1 Research flow chart.

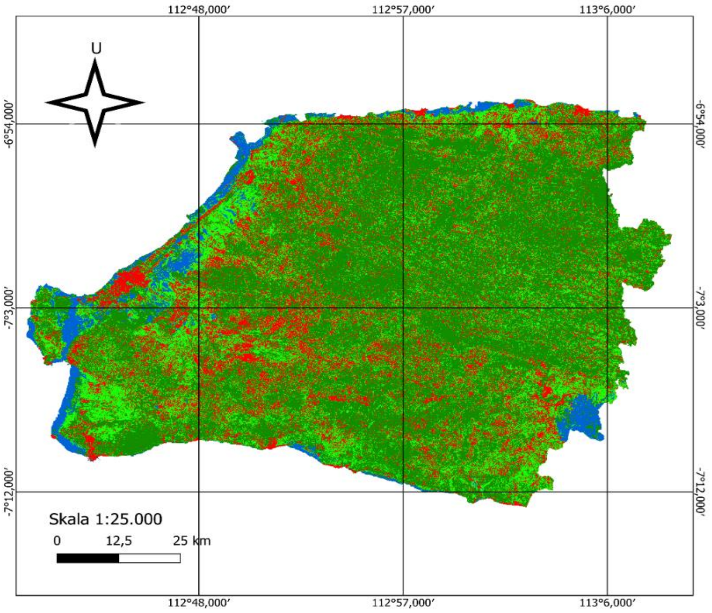


Figure 1 Land use map of Bangkalan Regency in 2022.

Table 2 Land use in Bangkalan Regency in 2022

Types of land use	Area (ha)	Percentage (%)
Water bodies	6,021.09	4.63
Farmland	31,817.16	24.45
Settlement	20,578.05	15.81
Vegetation	65,040.39	49.98
Open land	6,683.94	5.14
Total	130,141	100

section, which results from land use classification, is divided into forests and plantations. The vegetated area was 65,040.39 ha, or 49.98% of the total area of the Regency. Agriculture was the second major land use, accounting for 31,817.16 ha (24.45%). Climate, soil, water, and vegetation all play a role in determining agricultural land use. There were two types of agricultural land: wetland farming land and dry land. The classification of agricultural area in Bangkalan Regency in 2022 included wetland agriculture, such as rice fields, marshes, and mangrove forests. Dryland agriculture included fields, gardens, and rain-fed rice fields.

Residential areas classified using the Random Forest algorithm land use have a total area of 20,578.05 ha (15.81%), which included homes in towns and villages, industrial estates, built land, and public facilities. In general, settlements are regions where humans live. These locations can be found in towns, villages, or industrial zones (Tosiani *et al*. 2020). Meanwhile, open land and water bodies covered 6,683.94 ha (5.14%) and 6,021.09 ha (4.63%), respectively. The classification included open ground at Jaddih Hills, a former mining site, as well as open regions throughout Bangkalan Regency. Open land appeared without vegetation, whether naturally or because of human activities (Tosiani *et al*. 2020). Meanwhile, the map's water bodies represent wetlands such as rivers, marshes, and ponds. A body of water is the appearance of water or waters in a certain location, such as rivers, oceans, ponds, reservoirs, and lakes (Tosiani *et al*. 2020).

Accuracy and Validation

The accuracy of land use analysis results in Bangkalan Regency was tested using a Random Forest algorithm and a confusion matrix in the form of Overall Accuracy (OA) and Kappa. Both were used to assess the effectiveness of the random forest method classification model against a basic metric table. OA was calculated by adding the percentage of correct classifications. Kappa (K) was a measure of agreement between actual observations and forecasts of future land use classification. These statistics are useful for evaluating model performance, measuring classification agreement, and comparing alternative land-use models.

Based on the accuracy test results from GEE, the score has an overall accuracy of 91.39% and a Kappa value of 88.39%, or 0.88 (Table 3). The Kappa score suggests that Bangkalan Regency's land use classification results are 88.39% accurate. The kappa index ranges between 0.81 to 1.00, indicating an excellent predicate (Susanti *et al*. 2020). This denotes a classification model using a Random Forest algorithm capable of avoiding 88% of random classification errors, implying that the results of its interpretation are adequate for analysis (Zulfajri *et al*. 2021). The findings of land use classification using the Random Forest classification algorithm were evaluated to determine visual truth in the field. This is accomplished through a field survey or by visiting the observation site personally. Fifty-three observation sites (Figure 3) were utilized to compare the visual appearance to the classification findings in the GEE.

The field validation yielded an overall accuracy score of 88.68% and a Kappa of 0.85, representing 85.53%. The Kappa number falls between 0.81 to 1.00, indicating that the Kappa analysis produced excellent results. These data also show that the classification and interpretation results are appropriate because they passed the recommended level of accuracy, which is

Table 3 Process results of confusion matrix using Google Earth Engine

Land class	Water bodies	Farmland	Settlement	Vegetation	Open land	Total
Water bodies	128	2	0	3	0	133
Farmland	4	15	3	4	2	28
Settlement	0	3	136	3	5	147
Vegetation	1	4	1	232	1	239
Open land	0	6	3	2	18	29
Total	133	30	143	244	26	576
OA			91.84			
Kappa			88.39			

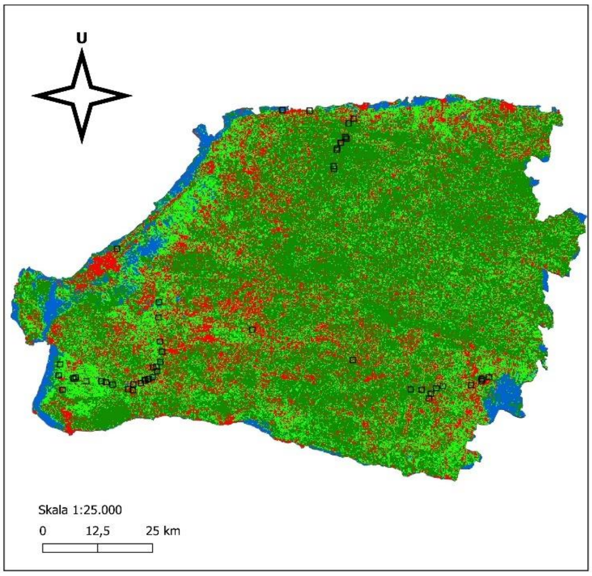


Figure 2 Validation point distribution.

Table 4 Process results of confusion matrix through field validation tests

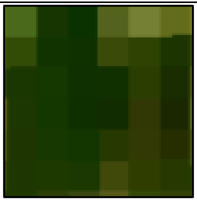
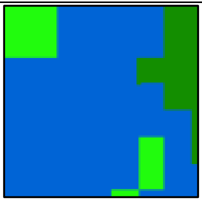
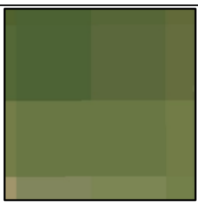
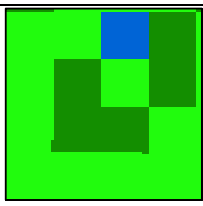
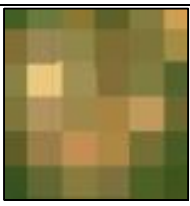
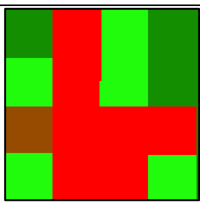
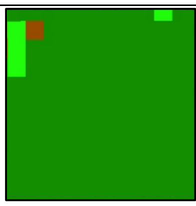
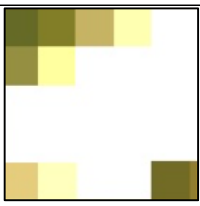
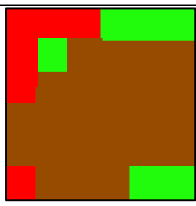
Land class	Water bodies	Farmland	Settlement	Vegetation	Open land	Total
Water bodies	5	0	0	0	0	5
Farmland	1	15	0	0	0	16
Settlement	0	1	9	0	1	11
Vegetation	1	0	1	10	1	13
Open land	0	0	0	0	8	8
Total	7	16	10	10	10	53
OA			88.68			
Kappa			85.53			

85%. (Asra *et al*. 2020) The overall accuracy and Kappa values derived from processing using GEE and field check validation have the same predicate, which is very good. Both were separated by differences in scores of 3.16 for overall accuracy and 2.86 for Kappa analysis (Table 4).

Validation of land use classes was also accomplished by comparing satellite sightings and classification findings to GEE and field observations. The satellite sighting was represented by a depiction of bands 4, 3, and 2 in natural hues. This band combination was used to get classification results using a Random Forest technique, which identifies pixels of the same color. The classification findings were displayed in the form of colors representing the land use classes in Bangkalan Regency. Meanwhile, the appearance in the field takes the form of images obtained while validating and monitoring directly. Table 5 shows a comparison between satellite sightings, categorization results, and field sightings.

The results of land use analysis employing remote sensing methods combined with machine learning and cloud computing platforms can aid with environmental monitoring, natural resource management, and regional planning in a timely, cost-effective, and efficient manner. The findings of this study could also

Table 5 Land use class visualization

Land Class	Visualization		
	Landsat 8	Classification Results	Field
Water bodies	 Has a dark green color	 Represented by dark blue	 Representing rivers, ponds, and reservoirs
Farmland	 Has a pastel green color	 Represented by light green	 Guardian agricultural areas such as rice fields
Settlement	 Has a light brown to dark brown color	 Represented by red	 Representing urban, residential, and industrial areas
Vegetation	 Has a dark green color similar to a body of water but more concentrated	 Represented by dark green	 Representing vegetation areas such as forests and moorlands
Open land	 Has a white to cream color	 Represented by the color brown	 Represents an area of open land without vegetation or settlements on it

help sustainable development that preserves natural functions while increasing economic value. This platform could also provide accurate and current information for a variety of applications, including

precision agriculture, protected area monitoring, and degraded land rehabilitation. Overall, land use identification using remote sensing is critical to promoting sustainable natural resource management and development. The information gained can also be used to make current and future land-use decisions.

CONCLUSION

It is possible to utilize machine-learning random forest methods for land use analysis in Bangkalan Regency in 2022 using the Google Earth Engine platform, and the interpretation is suitable for analysis. The Kappa accuracy test for random forest yields 0.88 and 0.85 from validation 0.85, indicating a flawless prediction (nearly perfect). In 2022, vegetation and agricultural land account for 65,040.39 ha (49.98%) and 31,817.16 ha (24.45%), respectively, followed by settlements of 20,578.05 ha (15.81%), open land of 6,683.94 ha (5.14%), and water bodies of 6,021.09 ha (4.63%). The findings of this study could be used to evaluate land use and to guide current and future land use decision-making.

REFERENCES

Arif N, Wahyuni FS. 2016. Penggunaan metode machine learning untuk pengenalan pola tutupan lahan pada citra satelit. In: *Seminar Nasional Teknologi Informasi dan Multimed*. STMIK AMIKOM Yogyakarta, 6–7 Februari 2016.

Arifin S. 2017. Kajian perubahan penggunaan tanah pasca pembangunan jembatan Suramadu di Kecamatan Bangkalan, Kabupaten Bangkalan, Provinsi Jawa Timur. Yogyakarta (ID). Sekolah Tinggi Pertahanan Nasional.

Asra R, Mappiasse MF, Nurnawati AA. 2020. Penerapan model CA–Markov untuk prediksi perubahan penggunaan lahan di Sub–DAS Bila tahun 2036. *Agrovital: Jurnal Ilmu Pertanian*. 5(1): 1–10. https://doi.org/10.35329/agrovital.v5i1.630

[BPS] Badan Pusat Statistika. 2023. *Kabupaten Bangkalan Dalam Angka* 2023. Jakarta (ID): Badan Pusat Statistik.

Febriani N, Yunidar S, Hidayat RA, Amor G, Indrayani P. 2022. Klasifikasi citra satelit dengan metode Random Forest untuk observasi dinamika lansekap ekosistem Kabupaten Sujunjung. *El–jughrafiyah*. 2(2): 75–81. https://doi.org/10.24014/jej.v2i2.18730

Firmawan MA, Widiatmaka, Nirmala K. 2021. Identifikasi dinamika spasial penggunaan dan tutupan lahan di Kabupaten Indramayu. *Jurnal Ilmu Tanah dan Lingkungan*. 23(2): 78–84. https://doi.org/10.29244/jitl.23.2.78-84

Hossaimah H, Subari S. 2017. Percepatan alih fungsi (konversi) lahan pertanian ke non pertanian di Kecamatan Galis, Kabupaten Pamekasan. *Agrisaintifika Jurnal Ilmu–Ilmu Pertanian*. 1(2): 97–108. https://doi.org/10.32585/ags.v1i2.45

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics*. 33(1): 159–174. https://doi.org/10.2307/2529310

Lestari SC, Arsyad M. 2018. Studi penggunaan lahan berbasis data citra satelit dengan metode sistem informasi geografis (GIS). *Jurnal Sains dan Pendidikan Fisika*. 14(1): 81–88.

Marlina D. 2022. Klasifikasi tutupan lahan pada citra Sentinel–2 Kabupaten Kuningan dengan NDVI dan Algoritme Random Forest. *STRING (Satuan Tulisan Riset dan Inovasi Teknol*ogi). 7(1): 41–49. https://doi.org/10.30998/string.v7i1.12948

Novianti TC. 2021. Klasifikasi landsat 8 OLI untuk tutupan lahan di Kota Palembang menggunakan *Google Earth Engine*. *Jurnal Swarnabhumi.* 6(1): 75–85.

Nuraida D, Rosyida SZA, Widyawati NA, Sari WK, Fanani IMR. 2022. Analisis vegetasi tumbuhan herba di Kawasan Hutan Krawak. *Jurnal Biologi dan Pembelajarannya*. 9(2): 96–104. https://doi.org/10.29407/jbp.v9i2.18417

Rahman FA, Adiputra A. 2022. Analisis konversi lahan pertanian menjadi pemukiman di Kabupaten Bangkalan Tahun 2009–2019. In: *Prosiding Seminar Nasional Pendidikan Geografi FKIP UHAMKA*. Institut Pendidikan FKIP Uhamka, Jakarta Timur, 9th April 2022.

Rahmawati H. 2021. Relevansi pergeseran alih guna lahan pertanian menjadi pemukiman di Desa Simogirang, Kecamatan Prambon, Sidoarjo. *Jurnal Ekonomi Bisnis dan Akunt*ansi. 1(3): 74–86. https://doi.org/10.55606/jebaku.v1i3.171

Rizaldi A, Darmawan A. Kaskoyo H. Setiawan A. 2023. Pemanfaatan *Google Earth Engine* untuk pemantauan lahan agroforestri dalam skema perhutanan sosial. *Majalah Geografi Indonesia*. 37(1): 12–21 https://doi.org/10.22146/mgi.73923

Suryono H, Marsuhandi AH, Pramana S. 2022. Klasifikasi tutupan lahan berdasarkan Random Forest Algorithm menggunakan cloud computing platform. *Jurnal Aplikasi Statistika dan Komputasi Statatistik*. 14(3): 1–12. https://doi.org/10.34123/jurnalasks.v14i1.383

Susanti Y, Syafrudin, Helmi M. 2020. Analysis of land use change in Upper Serayu Watersheds using remote sensing and Geographic Information Systems. *Bioedukasi: Jurnal Pendidikan Biologi*. 13(1): 23–30.

Tosiani A, Mohammad AR, Sularso MGN, Lugiana M, Novita N, Lestrai NS. 2020. *Standar Operasional Prosedur (SOP): Penghitungan Akurasi dan Uncertainty Perubahan Penutupan Lahan*. Bogor (ID): IPB Pr.

Zulfajri, Danoedoro P, Murti SH. 2021. Klasifikasi tutupan lahan Data Landsat–8 OLI menggunakan metode Random Forest. *Jurnal Penginderaan Jauh Indonesia*. 3(1): 1–7. https://doi.org/10.12962/jpji.v3i1.266