Land use land cover change analysis using minimum distance classification (MCD) and spectral angle mapper (SAM) method

Erina Sukmawati*[1], Ardiansyah[2], Afik Hardanto[2]

[1]Post Graduate Programme, Universiti Malaysia Terengganu, Kuala Nerus 21030, Terengganu, Malaysia

[2]Department of Agricultural Engineering, Jenderal Soedirman University, Purwokerto 53123, Indonesia

*Corresponding Author e-mail: p5918@pps.umt.edu.my

Received: 06-06-2024 | Revised: 20-06-2024 | Accepted: 28-06-2024

Abstract

Increasing population growth followed by massive land use change. These conditions can lead to disturbances in the hydrological function of a watershed. Several spatial approaches are used to determine the extent of land-use change, including Minimum Distance Classification (MDC) and Spectral Angle Mapper (SAM) algorithms. This study aims to gain insights into the performance of the two algorithms in classifying land use change. The study site was conducted in the Putih Sub-watershed, which is part of the upstream Serayu watershed. Land cover classification is divided into three classes (i.e. vegetation/forest, agriculture, and settlement. Spatial data from Landsat 7 and 8 imagery over 10 years (2009-2019) is applied as material for the analysis. Based on the results of the MDC classification, the area of agricultural land increased between 2009 (63%) and 2014 (70%), but decreased in 2019 (66%). While in SAM the same pattern was observed, namely an increase in utilized agricultural area from 2009 (63%) to 2014 (75%) before declining by 73% in 2019. The SAM method (RMSE=75.14) is more accurate than the MDC (RMSE=105.13), although both methods share the same pattern. It is possible that the increase in agricultural land and settlements effect to the severe erosion in the Serayu watershed, which requires the attention of policy makers and related stakeholders.

Keywords: hydrological function, land use change, machine learning, watershed, water conservation.

1. INTRODUCTION

Land use change has occurred significantly over the past two decades, especially dominating in the developing regions. Potapov et al. (2022) reported that the largest areas of forest reduction occurred in Africa, Asia and North America. Meanwhile, the smallest area of land conversion is on the Australian Continent. One of the main factors leading to forest conversion is the expansion of human settlements, with the highest percentages in Africa (68%) and Asia (73%). A decrease in forest area can affect environmental functions, such as: biophysical and biodiversity. Changes in hydrological functions were reported in South America due to forest conversion (D'Almeida et al., 2007). Meanwhile, the impact of deforestation is also recorded to change soil characteristics and reduce its function in the tropics (Veldkamp et al., 2020). The disturbance of invertebrate biodiversity in buffer zones also occurs due to the reduction of forest area in riparian areas and tropical biodiversity (Silva-Araújo et al., 2020; Yesuf et al., 2019).

Forests are important areas in watersheds. Forests function primarily as water catchment areas. Forest destruction in a watershed will cause the watershed to become a critical area. Based on the condition of the watershed, the government has determined that 108 watersheds in Indonesia are in critical condition and are prioritized to be immediately addressed and restored (Aryani et al., 2020; Fatahilah, 2013). One of these critical watersheds is the Serayu watershed, which administratively runs from Wonosobo Regency southward to Cilacap Regency. This critical condition is possible due to high land use change, especially in the upstream area (Marhendi & Pramono, 2019). Therefore, information on land use change in the upstream area can be used as an instrument in mitigating the hydrological dynamics that occur in the Serayu watershed.

The Serayu watershed covers an area of approximately 358,514 ha, so the process of monitoring land use change takes a long time. Land use change needs a lot of time and effort. The development of geographic information systems currently makes it easier in determining these spatial changes (Talukdar et al., 2020). One of them is the use of satellite imagery and machine learning in determining and classifying land use. Therefore, this study tested two supervised method algorithms in classifying land use, especially in the upstream area of the Serayu watershed. The objective of this study is to obtain information on land use change in the Putih Subwatershed which is part of the upper Serayu Watershed, and to analyze the performance of the land use change determination using two different algorithms.

2. METHODS

Study area

This study was conducted in 2021 by analyzing spatial data from a decade previously divided into three data periods: 2009, 2014, and 2019. Spatial data was taken from three villages, namely Serang, Kejajar, and Jengkol, which are administratively located in Wonosobo Regency. Serang, Kejajar and Jengkol villages (administratively located in Wonosobo Regency) are hidrologically located in Putih Sub Watershed, while data processing was carried out in February 2021 at the Unsoed Agricultural Technology Laboratory.

Datasets

The tool used was spatial data processing software (QGIS 3.10.00 Coruna). Materials included image data downloaded from the United States Geological Survey (USGS; https://earthexplorer.usgs.gov/). Google Image was used as a calibration alternative for ground checks that were not possible due to Covid-19 conditions.

Data Analysis

This study was conducted with several stages of study which are basically divided into four parts, which are: spatial data collection, processing of digital mapping, classification process, and calibration (Figure 1). The processing of the digitized maps includes acquisition, georeferencing, digitization, and atmospheric correction. For the classification process, two supervised methods are used, namely Minimum Distance Classification (MDC) and Spectral Angle Mapper (SAM). The principle of the supervised method is that the process of determining the category of the preferred class is based on selecting a training region that represents each category. Statistics resulting from the training data for each category are then used as a basis for classification. The result of the supervised process is when the classes selected by the analyst are spectrally well separated and the selected training regions are truly representative of the entire dataset. Therefore, spectrally separating the land cover categories is the most important part of supervised methods.

The principle of the MDC algorithm is to calculate the spectral similarity between the unrecognized pixels and the class of the training area identifier, the more similar the closest spectral distance is (Sun & Ongsomwang, 2020). The MDC algorithm does not considers class variations, so the classification process can be faster than other algorithms (e.g. Maximum likelihood). Meanwhile, the SAM algorithm is used when the reflectance data has been calibrated and is relatively insensitive to exposure. Pixels that are farther than the specified maximum angle limit in radians are not classified. The SAM algorithm assumes reflectance data, but if light data is used, the error is generally insignificant as the original data is still close to zero (Shafri et al., 2007).

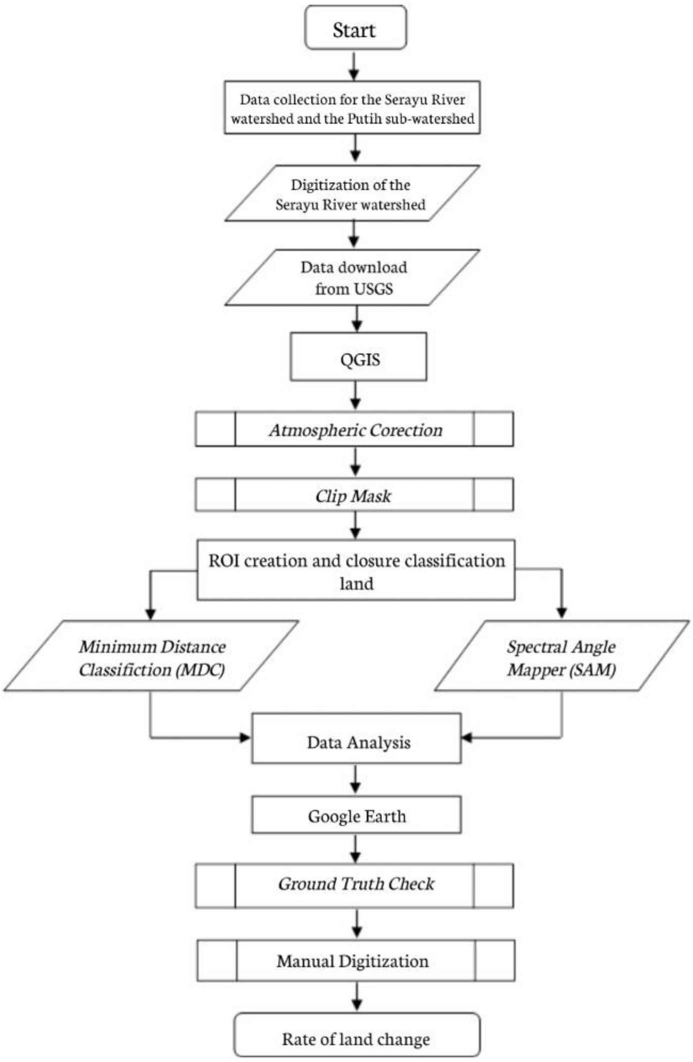

Figure 1. Flowchart of the method

3. RESULTS

The classification process is implemented by classifying based on the reference color in the training area or ROI (region of interest) generating process. The ROI area is created as a reference pixel for the classification class. The combination of bands on the RGB map is used to facilitate class selection so that the boundaries are visible (Khan & Das, 2021; Table 1).

Table 1. Band combinations of Landsat 7 and 8 images for several land covers

Land Cover Type	RGB	
	Landsat 7	Landsat 8
Agriculture	5 4 1	6 5 2
Settlement	7 5 3	7 6 4
Vegetation/Forest	5 4 3	4 5 3

The combination of bands that were used in this study is a combination of infrared, one of them is on Landsat 7 with a combination of RGB 5-4-3 in the image that has been selected and processed by atmospheric correction (Figure 2), the combination of infrared bands shows the red color, which is vegetation, then for further grouping based on the brightness of the color, for dark red is dense vegetation, while for light red is agriculture, while the others are settlements. The RGB combination is used to generate an ROI which is a guidelines for the pixels to be classified, after that proceed to the formation of macroclusters and microclusters in the classification of the band set, for classification according to the algorithm being used.

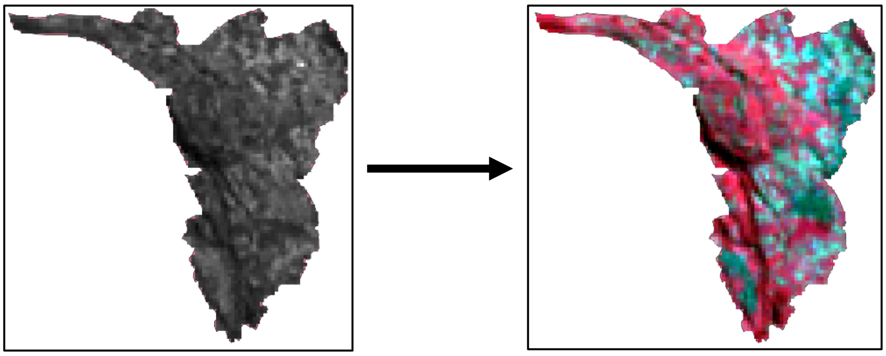

Figure 2. Combination of infrared bands (e.g. vegetation 5-4-3) on Landsat 7 Satellite Image (b) from the selected image and atmospheric correction (a).

Land use change using the MDC method in all the sites indicates that agricultural land is more dominant than plantation and forest areas (Figure 3). Jengkol village has the highest agricultural land area (441.7-432.5 ha), followed by Kejajar village (265.5-276.8 ha) and Serang village (132.5-216.7 ha). The pattern of agricultural land change in the three villages has been relatively constant over the past 10 years (Table 2). The increase in settlement area in the three villages was followed by a decrease in forest and plantation area, although the decrease in forest and plantation area was higher than the increase in settlement area (Figure 4). This indicates that changes in forest and plantation land occurred due to two factors at one point, that is the expansion of settlement areas and agricultural areas. Population growth is assumed to be the cause of increased household needs and agricultural activities (Li et al., 2021; Tendaupenyu et al., 2017).

Table 2. Land use change over 10 years from three villages in the Putih Subwatershed using the MDC method.

Land Use	Serang Land area (ha)			Kejajar Land area (ha)			Jengkol Land area (ha)		
	2009	2014	2019	2009	2014	2019	2009	2014	2019
Settlement	27.3	52.9	73.3	51.0	103.9	153.6	17.4	39.7	49.9
Agriculture	132.5	216.6	216.7	265.5	276.8	236.3	441.7	442.2	432.5
Vegetation/Forest	140.4	30.6	10.2	171.7	107.6	98.3	94.7	71.9	71.5
Total land area	300.2	300.2	300.2	488.3	488.3	488.3	553.8	553.8	553.8

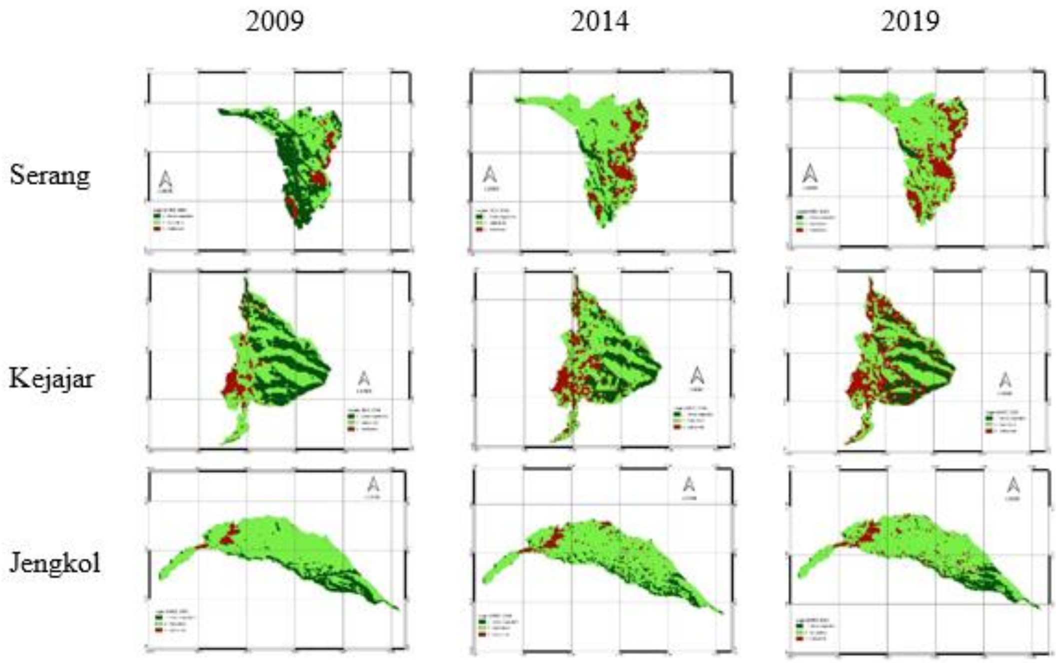

Figure 3. MDC method classification results of land use change in Serang, Kejajar, and Jengkol villages over the last 10 years.

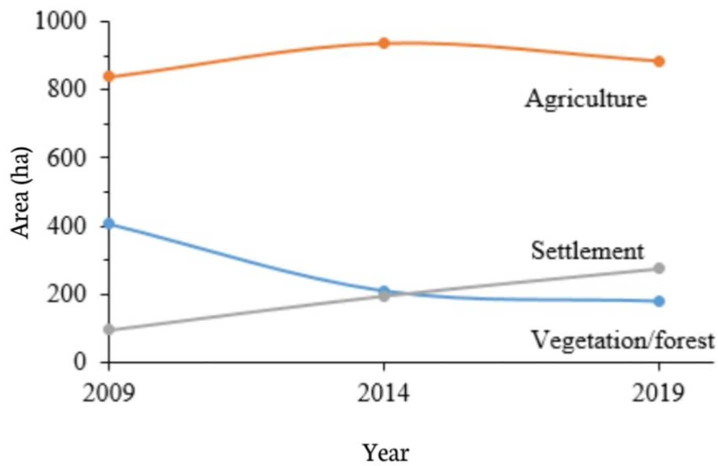

Figure 4. Pattern of change in land use by MDC method for 10 years.

The same pattern of change is also shown in the SAM method, but the trend of increasing residential area does not occur significantly compared to the MDC method (Figure 5). The difference in the results of the two methods, especially on the residential area, could be due to the different approaches regarding the coverage of the accessed color area and based on the angle point of the pixels (Attri et al., 2015; Singh et al., 2022). The similarity of the area data is seen in the three land uses, where the dominant area is agriculture, which is highest in Jengkol Village (443.2-450.0 ha) followed by Kejajar Village (267.9-352.6 ha) and Serang Village (133.3-216.3 ha). Regarding the pattern of change, the decline in forest and plantation area was more dominantly caused by agricultural activities than by the expansion of settled areas (Figure 6).

Table 3. Land use change over 10 years from three villages in the Putih Subwatershed using the SAM method.

Land Use	Serang Land area (ha)			Kejajar Land area (ha)			Jengkol Land area (ha)		
	2009	2014	2019	2009	2014	2019	2009	2014	2019
Settlement	27.1	47.0	70.9	53.1	56.9	90.1	17.5	22.4	34.6
Agriculture	133.3	216.3	214.8	267.9	352.6	316.2	443.2	441.5	450.0
Vegetation/Forest	139.8	36.9	14.4	167.2	78.8	82.0	93.2	89.8	69.2
Total land area	300.2	300.2	300.2	488.3	488.3	488.3	553.8	553.8	553.8

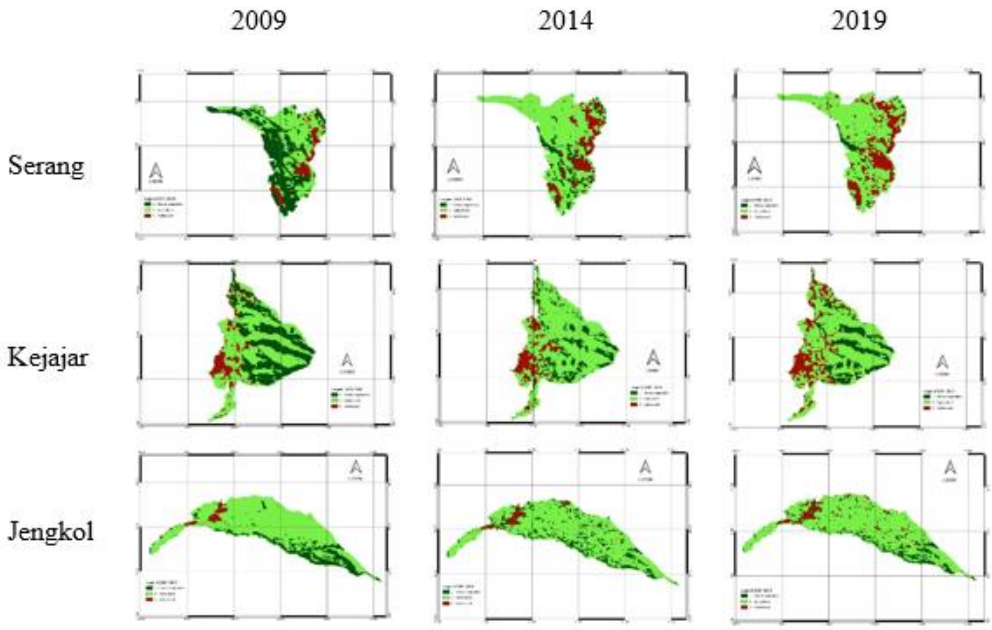

Figure 5. SAM method classification results on land use change in Serang, Kejajar, and Jengkol villages over the last 10 years.

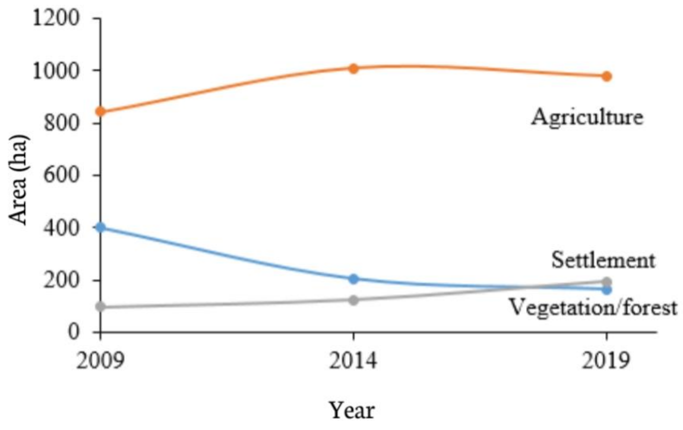

Figure 6. The pattern of land use change using the MDC method for 10 years

When compared, the two methods have different performance. Using manual classification of satellite image data as observation data, it can be seen that the land use change classification method using the SAM method provides a lower Root Mean Square Error (RMSE) value than the MDC method, which is 75.14 (SAM) and 105.13 (MDC). However, both methods have less precise estimation values when compared to the observation data (Figure 7).

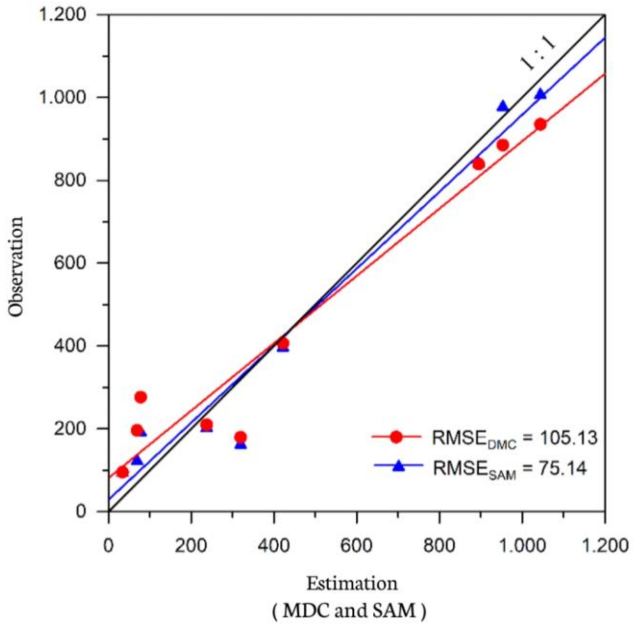

Figure 7. Predicted value of the two approaches to the observed value of land use change classification.

4. DISCUSSION

A comparison of the use of the two algorithms, which are: 1: (i) The MDC method has the advantages of making classification simpler, pixels are grouped at the closest color, and avoiding misinterpretation because it uses the shortest distance in pixel grouping. While the disadvantages of the method, that are: the classification of pixels depends on color so that bias often occurs, if the standard deviation or threshold of pixel distance has been determined then some pixels may not be classified, and if the distance is more than the average distance of each category then the pixels cannot be classified; (ii) The SAM method has several advantages compared to MDC, those are: classification is not influenced by color but by spectral color, not influenced by solar lighting factors because the angle between vectors does not depend on the length of the vector, an effortless and fast method for mapping the similarity of the image spectrum with the reference spectrum, does not require the assumption of statistical distribution of input data in performing classification. While the weakness of the SAM methods are: only uses the direction of the spectrum and the pixels are grouped into the most relevant class.

5. CONCLUSION

The use of algorithms with supervised classification methods have advantages and disadvantages. The classification process of land use change using the supervised algorithm can be performed faster than the manual method, however, the selection of the method must consider the availability of data, data quality, and the size of the study area. Although they have different levels of accuracy, both methods provide the same information related to land use change in the Putih Subwatershed, which is a decrease in the forest and plantation area caused by an increase in the area of settlements and agriculture. This is a concern for policy makers and stakeholders in order to maintain the function of the upper watershed as a water catchment area. Further observations need to be made concerning the optimal function of the upper watershed area as an environmental and socioeconomic buffer area for the community.

ACKNOWLEDGMENT

This study activity is part of research on water resources management funded by BLU Unsoed (No. 3716/UN23.14/PN.01.00/2018).

REFERENCES

Aryani, N., Ariyanti, D. O., & Ramadhan, M. (2020). Pengaturan Ideal tentang Pengelolaan Daerah Aliran Sungai di Indonesia (Studi di Sungai Serang Kabupaten Kulon Progo). Jurnal Hukum Ius Quia Iustum, 27(3). https://doi.org/10.20885/iustum.vol27.iss3.art8

Attri, P., Chaudhry, S., & Sharma, S. (2015). Remote Sensing & GIS based Approaches for LULC Change Detection – A Review. International Journal of Current Engineering and Technology, 5(5).

D'Almeida, C., Vörösmarty, C. J., Hurtt, G. C., Marengo, J. A., Dingman, S. L., & Keim, B. D. (2007). The effects of deforestation on the hydrological cycle in Amazonia: A review on scale and resolution. International Journal of Climatology, 27(5), 633–647. https://doi.org/10.1002/joc.1475

Fatahilah, M. (2013). Kajian Keterpaduan Pengelolaan Daerah Aliran Sungai (DAS) Garang Provinsi Jawa Tengah. Jurnal Geografi, 10(2).

Khan, F., & Das, B. (2021). Land Use Classification and Land Cover Assessment Using Accuracy Matrix for Dhamtari District, Chhattisgarh, India. Suranaree J. Sci. Technol., 29(3).

Li, C., Yang, M., Li, Z., & Wang, B. (2021). How Will Rwandan Land Use/Land Cover Change under High Population Pressure and Changing Climate? Applied Sciences, 11(12), 5376. https://doi.org/10.3390/app11125376

Marhendi, T., & Pramono, A. A. (2019). Prediksi Peningkatan Debit Puncak Sungai Serayu Berdasarkan Perubahan Penggunaan LahanPrediksi Peningkatan Debit Puncak Sungai Serayu Berdasarkan Perubahan Penggunaan Lahan. JRST (Jurnal Riset Sains dan Teknologi), 3(1), 19. https://doi.org/10.30595/jrst.v3i1.3048

Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., Zalles, V., Li, X., Khan, A., Stolle, F., Harris, N., Song, X.-P., Baggett, A., Kommareddy, I., & Kommareddy, A. (2022). The Global 2000-2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results. Frontiers in Remote Sensing, 3, 856903. https://doi.org/10.3389/frsen.2022.856903

Shafri, H. Z. M., Suhaili, A., & Mansor, S. (2007). The Performance of Maximum Likelihood, Spectral Angle Mapper, Neural Network and Decision Tree Classifiers in Hyperspectral Image Analysis. Journal of Computer Science, 3(6), 419–423. https://doi.org/10.3844/jcssp.2007.419.423

Silva-Araújo, M., Silva-Junior, E. F., Neres-Lima, V., Feijó-Lima, R., Tromboni, F., Lourenço-Amorim, C., Thomas, S. A., Moulton, T. P., & Zandonà, E. (2020). Effects of riparian deforestation on benthic invertebrate community and leaf processing in Atlantic forest streams. Perspectives in Ecology and Conservation, 18(4), 277–282. https://doi.org/10.1016/j.pecon.2020.09.004

Singh, K. T., Singh, N. M., & Devi, T. T. (2022). A Remote Sensing, GIS Based Study on LULC Change Detection by Different Methods of Classifiers on Landsat Data. In A. K. Dikshit, B. Narasimhan, B. Kumar, & A. K. Patel (Eds.), Innovative Trends in Hydrological and Environmental Systems (pp. 107–117). Springer Nature. https://doi.org/10.1007/978-981-19-0304-5_9

Sun, J., & Ongsomwang, S. (2020). Multitemporal Land Use and Land Cover Classification from Time-Series Landsat Datasets Using Harmonic Analysis with a Minimum Spectral Distance Algorithm. ISPRS International Journal of Geo-Information, 9(2), 67. https://doi.org/10.3390/ijgi9020067

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. Remote Sensing, 12(7), 1135. https://doi.org/10.3390/rs12071135

Tendaupenyu, P., Magadza, C. H. D., & Murwira, A. (2017). Changes in landuse/landcover patterns and human population growth in the Lake Chivero catchment, Zimbabwe. Geocarto International, 32(7), 797–811. https://doi.org/10.1080/10106049.2016.1178815

Veldkamp, E., Schmidt, M., Powers, J. S., & Corre, M. D. (2020). Deforestation and reforestation impacts on soils in the tropics. Nature Reviews Earth & Environment, 1(11), 590–605. https://doi.org/10.1038/s43017-020-0091-5

Yesuf, G., Brown, K. A., & Walford, N. (2019). Assessing regional-scale variability in deforestation and forest degradation rates in a tropical biodiversity hotspot. Remote Sensing in Ecology and Conservation, 5(4), 346–359. https://doi.org/10.1002/rse2.110