Accuracy Assessment of Land Use Change Analysis Using Google Earth in Sadar Watershed Mojokerto Regency

F A Islami1,2, S D Tarigan2, E D Wahjunie2 and B D Dasanto3

1Watershed Management Study Program, Graduate School, IPB University
2Department of Soil Science and Land Resources, Faculty of Agriculture, IPB University
3Department of Geophysics and Meteorology, Faculty of Mathematics and Science, IPB University

Email: fikryaislami@gmail.com

Abstract. Mojokerto Regency is one of the regencies in East Java with a high population growth rate of around 0.96%, thus encouraging significant land use changes on built-up areas. Classifying remote sensing imageries to obtain reliable and accurate land use and land cover (LULC) information remains a challenge that depends on many factors such as complexity of the landscape, the remote sensing data selected, image processing, and classification methods. This study examined the accuracy assessment of LULC classification using Google Earth in Sadar Watershed, Mojokerto, East Java Indonesia for the years 2010, 2015, and 2020. The land use was classified into five categories; those are agriculture land (paddy field, field, and plantation), non-agriculture land (forest land, bushland, grazing land), bare land, settlement land, and water bodies. Around 85 random points were generated in ArcGIS and verified with Google Earth. The results showed that the Overall Accuracy of LULCC for 2010 was 80.2% and Kappa Coefficient was 0.74; for 2015, the Overall Accuracy was 85.3% and Kappa Coefficient was 0.8, and for 2020, the Overall Accuracy was 84.0%, and Kappa Coefficient was 0.79. All accuracy is considered as good categorized and acceptable in both overall accuracy and Kappa Coefficient.

1. Introduction
Change in land use and land cover (LULC) is the primary driver of environmental changes [1,2]. The biggest challenge is how to preserve the environment and increase economics and social benefits, so for this reason it is important to understand the patterns and trends of land use change on the regional, local and global scales.

Information on land use change is important for planning and monitoring the development of an area, so that can be utilized optimally while maintaining its sustainability and minimizing land conflicts [3]. According to[4], land use information cannot always be interpreted directly from the land cover, therefore complementary information is needed such as information from remote sensing imagery, supporting information on surrounding locations.

The accuracy produced by remote sensing is based on the image data used [5]. In this study using Sentinel 2-A imagery to identify land use in 2020 which has 13 bands that have 10 meters resolution for 4 bands: 20 meters resolution for 6 bands and 60 meters resolution for 3 bands. Sentinel 2-A images have the advantage of higher spatial resolution, shorter temporal resolution, and multispectral image
with 13 spectral bands that are better than Landsat 8 image [6], while for the year 2010 and 2015 using Landsat 5 image.

Accuracy assessment is an important step in the for processing land use change analysis. It contains the information value of the resulting data to a user. The overall accuracy of the classified image compares how each of the pixels is classified versus the definite land cover conditions obtained from their corresponding ground truth data. Producer’s accuracy measures errors of omission, which is a measure of how well real-world land cover types can be classified. User’s accuracy measures errors of commission, which represents the likelihood of a classified pixel matching the land cover type of its corresponding real-world location [20]. The objectives of this research are identifying the classification of land use change in the Sadar Watershed, Mojokerto Regency, East Java, Indonesia; and to calculate the accuracy of land use.

2. Method

2.1. Study area
Mojokerto Regency is one of the regencies in East Java that has a high population growth rate, so that it triggers a rapid change in land use. The population growth rate in Mojokerto Regency reaches 0.96% which has a big impact on land conversion [8]. The research location is in the Sadar Watershed (111°20'13" - 111°40'47" E and 7°18'35" - 7°47'30" S), Mojokerto Regency, East Java, Indonesia. The Sadar Watershed is a priority watershed in the downstream of the Brantas Watershed and must be taken seriously because of the flooding problem. The study area is presented in figure 1. The types of data used in this study is physical data included in table 1.
### Table 1. Data types and sources.

| Data                                      | Data Source                                                                 | Note                                      |
|-------------------------------------------|-----------------------------------------------------------------------------|-------------------------------------------|
| Digital Elevation Model                   | DEM from: https://earthexplorer.usgs.gov/                                   | Watershed delineation                      |
| Topographic Map                           | Topographic maps and river networks through the Geospatial Information Agency (BIG) http://tanahair.indonesia.go.id/ |                                           |
| River Network Map                         |                                                                             |                                           |
| Landsat 5 and Sentinel 2A Image           | Landsat Image from: https://glovis.usgs.gov/                                 | Land use change analysis in Sadar Watershed for the year 2010, 2015 and 2020 |
| Coordinates of rain station and random points | Ground check at Sadar Watershed                                             | Watershed delineation, accuracy test      |

#### 2.2. Data analysis method

Land use is analyzed using the supervised classification method, before the classification, the image data was cut according to the study area. The method was divided into three steps, there were composite band for Landsat and Sentinel image, supervised classification, and accuracy assessment. The most often problem of mapping using remote sensing data is the selection of the classification method to be used in image classification [20]. This method has the advantage of quantitatively evaluating variance and correlation of spectral response patterns when classifying unknown pixels. While accuracy assessment is done by spreading random points using ArcGIS software.

The most frequently used method to obtain information from remote sensing data is multispectral classification. Multispectral classification is a digital automation technique that has been widely used to map land use and land cover [5]. Multispectral image classification is performed using two classification methods, supervised classification, and unsupervised classification. According to [7], the advantages of supervised classification are having control over informational classes based on training samples, and control over classification accuracy. While the drawbacks of this method are that the sample used is not necessarily representative and the presence of an unidentified spectral class.

#### 2.3. Land use classification

Land use classification is a classification providing information on land cover and the types of human activity involved in land use [21]. Land use classification method contain of two, there are supervised and unsupervised classification [22]. In this study used supervised classification. Supervised classification can be defined as the process of samples of known distinctiveness to classify pixels of unknown identity. This method generates a map with each pixel assigned to a class based on its multispectral composition.

#### 2.4. Accuracy Assessment

The field data is used for the accuracy assessment process for the year 2020, the classification results of the Sentinel-2A satellite image. A total of 85 random points were used for the validation and accuracy of the classification results. Determination of the random points is carried out to identify area that represents each desired land cover class and build a numerical description of the spectral properties of each land cover. Random points are selected based on field data and analyzed into statistical information on land use types.

The users of LULC maps need to know the level of accuracy of the map so that it can be used more efficiently and correctly [9,10]. According to [11], the interpretation accuracy of land use and land cover change is not allowed below 80%. The most widely promoted classification accuracy is in the form of error matrix which can be used to derive a series of descriptive and analytical statistics [12,9,13-15].
The step is a very effective way to show accuracy in that the accuracies of each category are plainly described along with both the errors of commission errors and errors of omission errors present in the classification [13]. Overall accuracy, producer’s accuracy, user’s accuracy, and Kappa statistics are generally reported, and these terms have been explained in detail in many studies [17,9-10,13-16].

User’s accuracy = \frac{\text{Number of Correctly Classified Pixels in each Category}}{\text{Total Number of Reference Pixels in that Category (The Row Total)}} \times 100 \quad (1)

Producer accuracy = \frac{\text{Number of Correctly Classified Pixels in each Category}}{\text{Total Number of Reference Pixels in that Category (The Column Total)}} \times 100 \quad (2)

Total (overall) accuracy = \frac{\text{Total Number of Correctly Classified Pixels (Diagonal)}}{\text{Total Number of Reference Pixels}} \times 100 \quad (3)

Kappa Coefficient (T) = \frac{(TS \times TCS) - \sum (\text{Column Total} \times \text{Row Total})}{TS^2 - \sum (\text{Column Total} - \text{Row Total})} \quad (4)

**Table 2.** Rating criteria of Kappa statistics.

| No | Kappa statistics | Strength of agreement |
|----|------------------|-----------------------|
| 1  | <0               | Poor                  |
| 2  | 0.00 – 0.20      | Slight                |
| 3  | 0.21 – 0.40      | Fair                  |
| 4  | 0.41 – 0.60      | Moderate              |
| 5  | 0.61 – 0.80      | Substantial           |
| 6  | 0.81 – 1.00      | Almost perfect        |

3. Results and Discussions

3.1. Land use classification

Land use change is an interaction between humans and their environment, where the main factor or environmental focus is land, while human attitudes and policy responses to land will determine the steps of their activities, so that they will leave marks on the land in the form of land use [3].

The main factor of land use change is the increase in population. Land use in rural and urban areas shows different complexity. The difference in complexity is supported by landscape objects, cultural landscapes, ecosystems, production systems, etc.

Based on the Regional Spatial Plan (RTRW) of Mojokerto Regency in 2012-2032, the Sadar Watershed has included in the hydrometeorological natural disasters prone area, in this case the flood disaster. The Sadar Watershed itself includes Mojoanyar District, Bangsal District, Mojosari District, Pungging District and Ngoro District which are areas with a high population growth rate. According to data from the Central Statistics Agency for Mojokerto Regency [8], the population growth rate in 2020 reached 0.96%, so in this case it is necessary to conduct a study related to changes in land use due to population growth to minimize the impact of floods. The land use of Mojokerto Regency for the year 2010, 2015, and 2020 will be explained in table 3 below.
Table 3. Land use change in Sadar Watershed.

| Class | Land Use       | 2010   | 2015   | 2020   |
|-------|----------------|--------|--------|--------|
|       | Area (km²)     | %      | Area (km²) | %      | Area (km²) | %      |
| 1     | Agriculture    | 151.6  | 39.3   | 148.9  | 38.6   | 81.5  | 21.1   |
| 2     | Non-Agriculture| 76.31  | 19.8   | 71.9   | 18.6   | 72.2  | 18.7   |
| 3     | Bare Land      | 53.27  | 13.8   | 66.8   | 17.3   | 89.2  | 23.1   |
| 4     | Settlement Land| 44.46  | 11.5   | 48.3   | 12.5   | 106.3 | 27.5   |
| 5     | Water Bodies   | 60.55  | 15.7   | 50.3   | 13.0   | 37.3  | 9.7    |

Based on table 3, the land use conditions in the Sadar Watershed of Mojokerto Regency in 2010 were mostly dominated by agricultural land consisting of paddy field, field, and plantation, covering an area of 151.6 km² or 39.3%, while non-agricultural land use covering an area of 234.6 km² or 60.74%. In 2015, the area of agricultural decreased by 148.9 km² or about 38.55%, while the use of non-agricultural land was 237.4 km² or 61.45%. In 2020, the area of agricultural land decreased drastically by 81.5 km² or around 21.1%, while the use of non-agricultural land was 304.86 km² or 78.91%. It can be concluded that there is a land conversion from agricultural land to non-agricultural land such as bare land and built-up land. Changes in land use between 2010-2020 will be explained in figure 2, figure 3, and figure 4 below.

3.2. Accuracy assessment of land use change analysis
The accuracy assessment is used to determine the accuracy result of the land use classification [18]. The most used method is the contingency method to determine the level of accuracy, or also called the contingency matrix (confusion matrix). According to [19], the contingency matrix contain some information, there are producer's accuracy, overall accuracy, and kappa accuracy. Producer's accuracy and user's accuracy are estimators of overall accuracy. Producer's accuracy is the accuracy seen from the side of the map producer, while user's accuracy is the accuracy seen from the user's side of the map. The accuracy test of the classification results will be described in table 4 below.

![Figure 2. Land use change of Sadar Watershed for the year 2010-2020 in km².](image-url)
Figure 3. Changes in land use for the year 2010-2020.

Table 4. Accuracy assessment calculation of land use change analysis.

| LU Class     | Agri | Non-Agr | Bare land | Built-up land | Water bodies | Total User |
|--------------|------|---------|-----------|---------------|--------------|------------|
| Agriculture  | 26   | 5       | 0         | 0             | 0            | 31         |
| Non-agriculture | 3   | 13      | 0         | 0             | 0            | 16         |
| Bare land    | 2    | 0       | 7         | 1             | 0            | 10         |
| Built-up land | 0   | 3       | 3         | 10            | 0            | 16         |
| Water bodies | 0    | 0       | 0         | 0             | 13           | 13         |
| Total Producer | 31  | 21      | 10        | 11            | 13           | 86         |

Users accuracy = \( \frac{\text{Number of Correctly Classified Pixels in each Category}}{\text{Total Number of Reference Pixels in that Category (The Row Total)}} \times 100 \) (agriculture)

= \( \frac{26}{31} \times 100 = 83.9\% \)

Producer accuracy = \( \frac{\text{Number of Correctly Classified Pixels in each Category}}{\text{Total Number of Reference Pixels in that Category (The Column Total)}} \times 100 \)

= \( \frac{26}{31} \times 100 = 83.9\% \)

Overall accuracy = \( \frac{\text{Total Number of Correctly Classified Pixels (Diagonal)}}{\text{Total Number of Reference Pixels}} \times 100 \)

= \( \frac{26+13+7+10+13}{86} \times 100 = 80.2\% \)

Corrected Sample = \( 26+13+7+10+13 = 69 \)

Kappa Coefficient = \( \frac{(69x86-(31x31)+(16x21)+(10x10)+(16x11)+(13x13))}{(86^2-(31x31)+(16x21)+(10x10)+(16x11)+(13x13))} = 0.74 \) (Substantial)
Figure 5. Correctly and incorrectly classified data.

Table 5. Recap of accuracy assessment.

| Year | Overall Accuracy (%) | Kappa Coefficient |
|------|----------------------|-------------------|
| 2010 | 80.23                | 0.74              |
| 2015 | 85.33                | 0.8               |
| 2020 | 84.00                | 0.79              |

Based on overall accuracy and kappa coefficient, Overall Accuracy for the year 2010 is 80.23% and kappa coefficient is 0.74; in 2015, Overall Accuracy is 85.33% and kappa coefficient is 0.8; and for 2020, Overall Accuracy is 84% and kappa coefficient is 0.79, can be concluded that the land use classification has a great correlation.

4. Conclusion
The land use classification accuracy assessment in the Sadar Watershed, Mojokerto Regency, East Java has a total accuracy value of 80.23% in 2010 and a kappa coefficient of 0.74; while in 2015, the value was 85.33% and kappa coefficient was 0.8; in 2020, the total accuracy value is 84% and the kappa coefficient is 0.79. Based on these results, the accuracy and kappa coefficient values have good criteria and can be used for further analysis.

Acknowledgment
We acknowledge the support received from the Brantas Water Resources Office of Indonesia, who always support all the data that Fikry A needed. Fikry A also thanked to all the advisers for all the kindness, patience and suggestions.

References
[1] Lambin E, Turner B, Geist H, Agbola S, Angelsen A, Bruce J, Coomes O, Dirzo R, Fischer G, Folke C, George P, Homewood K, Imbernon J, Leemans R, Li X, Moran E, Mortimore M, Ramakrishnan P, Richards J, Skanes H, Steffen W, Stone G, Svedin U, Veldkamp T, Vogel C and Xu J 2001 The causes of land use and land cover change: Moving beyond the myths Glob. Environ. Change 11 261–269
[2] Goldewijk K and Ramankutty N 2004 Land cover change over the last three centuries due to human activities: the availability of new global data sets Geo. J. 61 335–344
[3] Harvini W 2017 Uji akurasi klasifikasi pengunaan lahan dengan menggunakan metode defuzzifikasi maximum likelihood berbasis Citra ALOS AVNIR-2 Bhumi 31
[4] Lillesand T, Kiefer R and Chipman J 2008 Remote Sensing and Image Interpretation: 6th Edition (Amazon: John Wiley and Son)
[5] Septiani R, Citra P and Nugroho A 2019 Perbandingan metode supervised classification dan unsupervised classification terhadap penutup lahan di Kabupaten Buleleng J. Geo 16 2-4
[6] Kawamura A, Suprayogi A and Wijaya A 2017 Analisis kesehatan hutan mangrove berdasarkan metode klasifikasi NDVI pada Citra Sentinel-2 J. Geodesi 6(1) 277-284
[7] Samaniego L, Bardossy A and Schulz K 2007 Supervised Classification of Remotely Sensed Imagery using a Modified k-NN Technique (California: IEEE Transactions on Geoscience and Remote Sensing) pp 4-7
[8] Fitriana D, Iriyanti Y, Hartono T, Aryani E, Daniati S, Suksmantri D and Maulida N 2021 Kabupaten Mojokerto dalam angka 2021 (Mojokerto: BPS Statistics of Mojokerto Regency)
[9] Smits P, Dellepaine S and Schowengerdt R 1999 Quality assessment of image classification algorithms for land cover mapping: a review and a proposal for a cost-based approach Int. J. Remote Sens. 20 1461–1486
[10] Plourde L and Congalton R 2003 Sampling method and sample placement: How do they affect the accuracy of remotely sensed maps? Photogram. Eng. Remote Sens. 69 289–297
[11] Anderson J, Hardy E, Roach J and Witmer R 1976 A Land Use and Land Cover Classification System for Use with Remote Sensor Data (Washington: Government Printing Office)
[12] Foody G 2002 Status of land cover classification accuracy assessment Remote Sens. Environ. 80 185–201
[13] Congalton R 1991 A review of assessing the accuracy of classification of remotely sensed data Remote Sens. Environ. 37 35–46
[14] Jensen J 2005 Thematic map accuracy assessment in Introductory Digital Image Processing—A Remote Sensing Perspective 3rd ed (USA: Prentice Hall) pp 495–515
[15] Liu C, Frazier P and Kumar L 2007 Comparative assessment of the measures of thematic classification accuracy Remote Sens. Environ. 107 606–616
[16] Lunetta R, Congalton R, Fenstermaker L, Jensen J and Tinney L 1991 Remote sensing and geographic information system data integration: error sources and research issues Photogram Eng. Remote Sens. 57 677–687
[17] Liu X, Skidmore A and Oosten H 2002 Integration of classification methods for improvement of land cover map accuracy ISPRS J. Photogram Remote Sens. 56 257–268
[18] Paul S 2013 Analysis of Land Use and Land Cover Change in Kiskatinaw River Watershed: A Remote Sensing, GIS & Modeling Approach Thesis (Columbia: University of Northern British Columbia)
[19] Rwangga S and Ndambuki J 2017 Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS International Journal of Geoscience 8 611-612
[20] Sari I, Christopher J, Glenn J, and Liubov Volkova J 2021 Assessing accuracy of land cover change maps derived from automated digital processing and visual interpretation in tropical forests in Indonesia International Journal of Remote Sensing 13 1446
[21] Samadder S and Shivesh K 2018 A comparison of different land-use classification techniques for accurate monitoring of degraded coal-mining areas Environmental Earth Sciences 77 713
[22] T Ganesh Kumar, Rajalashmi and D Murugar 2013 Supervised methods for land use classification International Journal of Research in Information Technology 1 64-73