COMPARISON OF MACHINE LEARNING ALGORITHMS FOR LAND USE AND LAND COVER ANALYSIS USING GOOGLE EARTH ENGINE (CASE STUDY: WANGGU WATERSHED)

Septianto Aldiansyah1, Randi Adrian Saputra2
1Department of Geography, Faculty of Mathematics and Natural Sciences, Universitas Indonesia
2Department of Geography Education, Faculty of Teacher Training and Education, Universitas Halu Oleo

*e-mail: septiantoaldisnayah863@gmail.com

Received: 27.10.2022; Revised: 31.12.2022; Approved: 31.12.2022

Abstract. Human population growth and land use and land cover (LULC) change have always developed side by side. Considering selection of a good Machine Learning (ML) classifier algorithm is needed considering the high estimation of LULC maps based on remote sensing. This study aims to produce a LULC classification of Landsat-8 and Sentinel-2 images by comparing the accuracy performance of three ML algorithms, namely: Classification and Regression Tree (CART), Random Forest (RF), and Support Vector Machine (SVM). Dataset comparison ratios were also explored to find the LULC classification results with the best accuracy. Sentinel-2 is better than Landsat-8 regarding Overall Accuracy (OA) and Coefficient Kappa. The comparison ratio of the training and testing datasets with a good level of accuracy is 70:30 on both images with the average OA Landsat-8 and Sentinel-2 being 92.09% and 94.21%, respectively. The RF algorithm outperforms CART and SVM in both types of satellite imagery. The mean OA of the CART, RF, and SVM classifiers was 92.03%, 94.74%, 83.54% on Landsat-8, 93.14%, 96.15%, and 93.34% on Sentinel-2, respectively.

Keywords: Google Earth Engine; Land Use Land Cover; Classification and Regression Tree; Random Forest; Support Vector Machine

1 INTRODUCTION

Future studies will be easier to understand by understanding Land Use and Land Cover (LULC) forms of various scales for various global phenomena such as drought, flooding, erosion, migration, and climate change. Continuous and accurate LULC analysis is an integral part of sustainable development activities carried out in certain areas. LULC maps become an important component for various scientific studies involving the effects of climate change on river flows and watersheds (Sridhar et al., 2019), geomorphology (Sujatha & Sridhar, 2018), groundwater management (Sridhar et al., 2018; Xiao et al., 2021; Xiao et al., 2021; Xiao et al., 2022), social knowledge for natural resource management (Sridhar et al., 2021), and monitoring agricultural land (Sridhar et al., 2017; Jamali et al., 2019; Rahman et al., 2020). A land can use the LULC map in determining suitable land for agricultural activities so that watershed management can be more sustainable (Cihlar, 2000, Renschler & Harbor, 2002).

Remote sensing is the most commonly used method of mapping land cover and tracking its changes over time (Phan et al., 2020). With an increasing population and the need to develop new areas to meet the demand for food production, energy generation, and water security, the hydrological and water resource modeling community is interested in integrating and evaluating land use changes and their effects on watershed areas (Kang et al., 2019; Setti et al., 2020).

Limitations in making low-resolution land cover maps in a wide area of course involve large amounts of data. Hefty storage capacity, processing power, and flexibility are required to implement a diversified approach (Xie et al., 2019). This problem has been addressed by the new Google Earth Engine (GEE)
technology. GEE is a form of integration of remote sensing and big data into a high-performance cloud-based platform and enables fast and easy computing of satellite imagery (Gorelick et al., 2017; Sidhu et al., 2018; Tamiminia et al., 2020; Kolli et al., 2020; Aldiansyah et al., 2021; Rahmi et al., 2022).

The GEE platform provides independent satellite imagery of various spatial resolutions. The platform is built in JavaScript and Python in handling coding (Shelestov et al., 2017; Mateo-García et al., 2018; Pimple et al., 2018) using MapReduce architecture for parallel processing which is a technique for breaking large amounts of data into small parts and processing them across multiple devices. LULC classification results using non-parametric Machine Learning (ML) methods such as Classification and Regression Trees (CART), Support Vector Machine (SVM), and Random Forest (RF) have very accurate accuracy (Bar et al., 2020; Liu et al., 2020; Tassi & Vizzar, 2020). GEE was chosen because of its broad capabilities in various LULC-based research fields. Midekisa et al. (2017) showed that GEE was able to produce multi-temporal LULC maps on the African continent. Kolli et al. (2020) demonstrated the ability of GEE in analyzing the rate of land use change around Lake Kolleru, India by utilizing the RF algorithm and obtaining an Overall Accuracy (OA) of 95.9 with a Kappa Coefficient (k) of 0.94. Rahman et al. (2020) also analyzed the performance of the RF and SVM algorithms for the urban area of Bhola and the rural area of Dhaka in Bangladesh with an accuracy of 96.9% and 98.3%, respectively. The application of GEE has also reached agricultural areas for crop mapping (Shelestov et al., 2017), comparative analysis of algorithms, and multi-temporal data sets over a wider area (Aguilar et al., 2018). Despite its advantages, in some cases, CART often experiences overfitting when the model fits the training data but fails to generalize the invisible test data (Lawrence & Wright, 2001). RF tends to require more time in executing data sets to obtain a low Redistribution Error Rate (Prajwala, 2015). Whereas SVM is limited by the number of data sets (Rudrapal & Subhedar, 2015) and class targets. In general, GEE often focuses on climate change, LULC change analysis, and monitoring of water resources with time series analysis (Wang et al., 2018; Workie & Debella, 2018; Jamei et al., 2019). However, GEE has computational limitations in terms of time, storage and memory. Tamiminia et al. (2020) describe these limitations in terms of large computing and memory capabilities; considering the time constraints involved, it is better to use a batch system.

The rapid increase in demand for LULC data and wide area coverage. An understanding of the methods and performance of machine learning that has been widely used, such as GEE is required. The purpose of this study is to classify LULC from the Landsat-8 and Sentinel-2 multispectral satellites, compare the ML algorithm on the GEE platform and compare training data and testing data for the Landsat-8 and Sentinel-2 multispectral satellites to produce a classification with the highest accuracy.

2 MATERIALS AND METHODOLOGY

2.1 Study Area

The Wanggu watershed is a priority watershed in Southeast Sulawesi Province because of its strategic location, namely, in the upstream part there are forest areas, ex-transmigration of Java, West Java, and Bali as well as spontaneous transmigration (transmigration carried out by local communities themselves).

This area spans an area of 453.44 km² based on the UTM 51S projection. This watershed has a strategic function and role because it flows through Kendari City, downstream, and is the center of government, downstream, and is the center of government, education and economy, industrial area, and clean water source.
2.2 Data
Earth Observation Data (EOD) includes satellite imagery from popular platforms such as Landsat and Sentinel stored on GEE’s cloud-based platform. Landsat and Sentinel data can be accessed via USGS (The United States Geological Survey) at GEE. This study uses Landsat-8 surface reflectance Tier 1 data which is atmospherically corrected using LASRC (Landsat-8 Surface Reflectance Code) and Sentinel-2 level 1C data. The use of this data is due to constraints such as cloud cover. Each data is selected with a cloud cover criterion of <10% for each year and the images are combined into a single image. Six Bands (Bands 2-7) of Landsat-8 with 30 m resolution and Nine Bands (Bands 2-8, and 11-12) of Sentinel-2 with 10 m resolution were used in this study (Table 2-1). The LULC class is divided into five main classes: forest, water body, built area, open land, and vegetation. Agricultural and plantation areas are considered vegetation, while rivers and ponds are considered bodies of water.

Table 2-1: Landsat-8 and Sentinel-2 band information

<table>
<thead>
<tr>
<th>Imagery Data Layer</th>
<th>Source</th>
<th>Band Used</th>
<th>Central Wavelength (µm)</th>
<th>Band Width (µm)</th>
<th>Spatial Resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8</td>
<td>Google Earth</td>
<td>Blue (Band 2)</td>
<td>0.482</td>
<td>0.060</td>
<td>30</td>
</tr>
<tr>
<td>Operational Land Imager surface reflectance Tier 1</td>
<td>Google (GEE)</td>
<td>Green (Band 3)</td>
<td>0.561</td>
<td>0.067</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red (Band 4)</td>
<td>0.655</td>
<td>0.038</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Near-Infra-Red (Band 5)</td>
<td>0.865</td>
<td>0.028</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short-Wave Infra-Red 1 (Band 6)</td>
<td>1.609</td>
<td>0.085</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short-Wave Infra-Red 2 (Band 7)</td>
<td>2.200</td>
<td>0.186</td>
<td>30</td>
</tr>
<tr>
<td>Sentinel-2 MSI: MultiSpectral Instrument, Level-1C</td>
<td>Google (GEE)</td>
<td>Blue (Band 2)</td>
<td>0.496</td>
<td>0.066</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green (Band 3)</td>
<td>0.560</td>
<td>0.036</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red (Band 4)</td>
<td>0.664</td>
<td>0.031</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red-Edge 1 (Band 5)</td>
<td>0.704</td>
<td>0.015</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red-Edge 2 (Band 6)</td>
<td>0.740</td>
<td>0.015</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red-Edge 3 (Band 7)</td>
<td>0.782</td>
<td>0.020</td>
<td>20</td>
</tr>
</tbody>
</table>
The Orthocorrected image with the least amount of initial cover serves as the main input for classification. After importing satellite data into GEE, the next step is to remove cloud shadows and cloud cover either because it is cloudy or without data using cloud masking (Mateo-García et al., 2018). This technique is suggested and can be optimally performed by GEE (Zurqani et al., 2018; Aldiansyah et al., 2021).

The next satellite image in Composites every year becomes one image using a median filter. The median value is assigned to each pixel in all images and results in one image for the entire image collection. Each ML algorithm uses 50-75 training data. Each training data is placed on the appropriate land cover type according to the interpretation of image data that has gone through the median filter composite. The generated training data is then split randomly into training data and testing data using predetermined ratios for each LULC map. The LULC training data uses a total of 899 and 802 feature points, respectively on Landsat-8 and Sentinel-2 images.

<table>
<thead>
<tr>
<th>Band Combination</th>
<th>Eigenvalue</th>
<th>Standard Deviation</th>
<th>Eigen Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-Infra-Red (Band 8)</td>
<td>0.835</td>
<td>0.106</td>
<td>10</td>
</tr>
<tr>
<td>Short-Wave Infra-Red 1 (Band 11)</td>
<td>1.610</td>
<td>0.091</td>
<td>20</td>
</tr>
<tr>
<td>Short-Wave Infra-Red 2 (Band 12)</td>
<td>2.202</td>
<td>0.175</td>
<td>20</td>
</tr>
</tbody>
</table>

### 2.3 Algorithms

#### Classification and Regression Tree (CART)

CART is a binary decision regression tree developed by Breiman et al. (2017) and used for simple decision-making in logical if-then scenarios. CART works recursively by separating nodes until they reach a terminal node, based on a predetermined threshold. The "classifier.smileCart" technique in the GEE library is used in this study to perform CART classification.

#### Random Forest Classifier (RF)

RF is one of the most commonly used classifiers in constructing an ensemble classification (Brieman, 2001) by combining many CART trees. Several decision trees are generated by RF randomly from training data sets and variables. The optimal number of
calculated trees ranges from 100-500, and the optimal number of variables calculated is the square of the set of variables (Belgiu & Drăguț, 2016). The "classifier.RandomForest" technique in the GEE library is used in this study to classify RF.

Support Vector Machine (SVM)
SVM is one of the Supervised Machine Learning algorithms used to solve regression and classification problems. The SVM classifier works by creating an ideal hyperplane in the training phase to group multiple classes with inaccurate pixels. The extreme points/vectors are selected to build the hyperplane. This extreme point is called the support vector. Parameters Cost, Gamma, and Kernel function (Adelabu et al., 2015) is used in this study. The cost value used is 10. A higher cost value indicates less incorrect data in the classification. The Gamma value used is 0.5, while the type of Kerney used is Radial or RBF. The "classifier.libsvm" technique in the GEE library is used in this study to classify SVM.

2.4 Accuracy Assessment
Assessment of accuracy on the dataset using training data grouping and data testing. A search was conducted on the comparison of datasets to find a good level of accuracy in classifying images. The ratio of accuracy tested is 50:50, 60:40, 70:30, and 80:20. The confusion matrix is used to validate and evaluate the accuracy of image classification. Overall Accuracy (OA) and Coefficient Kappa (k) were also used to test the accuracy of the classification results. The OA and k are calculated form the following equations:

\[ OA = \left( \frac{P_c}{P_n} \right) \times 100 \]  
(2.1)

where \( P_c \) is the number of pixels classified correctly and \( P_n \) is the total number of pixels.

\[ k = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} x_{+i})} \]  
(2.2)

where \( r \) is the number of rows and columns in the error matrix, \( x_{ii} \) is the number of observation in row \( i \) and column \( i \), \( x_{i+} \) is the marginal total of row \( i \), \( x_{+i} \) is the marginal total of column \( i \), and \( N \) is the total number of observation.

User Accuracy (UA) and Producer Accuracy (PA) are also calculated for each LULC class in the Confusion matrix. Validation with UA is determined by the ratio of the correctly categorized pixels in that class to the total number of pixels classified. Similarly, PA is determined by the ratio of properly categorized pixels to the total number of pixels in the reference data in each class. The classifier that has the best performance will be selected for further image classification for Spatio-temporal change analysis.

3 RESULTS AND DISCUSSION
3.1 LULC Classification Using GEE
This study examines the performance of the ML technique in classifying LULC using Landsat-8 surface reflectance Tier 1 data with a resolution of 30 m and Sentinel-2 data with a resolution of 10 m. Figure 3-2 and Figure 3-3 demonstrate ML algorithms such as CART, RF, and SVM in classifying LULC maps for 2016, 2018, and 2020 on the GEE platform. Orthorectified images with minimal cloud cover and pixels that were damaged due to cloudy conditions were removed from all available images using the Cloud Masking algorithm. Gaps in cloudy images are filled using temporal aggregation methods such as median, mean, and minimum/maximum. This study uses the median value for the collection of Landsat-8 and Sentinel-2 images for 1 year.

The results of the study show that for a more accurate classification it is better to use data from MSI Sentinel-2 when compared to Landsat-8. Explanatory factors that make Sentinel-2's performance better are the presence of additional image channels, especially Red-Edge 1 band, and more detailed spatial resolution. Down-sampling the original spatial resolution of Sentinel-2 data had an approximately equal decreasing effect on the classification performance leaving out the Red-Edge
bands. With almost the same band and by reducing the spatial resolution to 20 m, classification using Sentinel-2 performs better than Landsat-8. This can be an alternative if using MSI Sentinel-2 for mapping with critical processing power or limited by storage capacity. Whereas Landsat-8 works well because it relies more on the ability of Band Blue (B2) to classify forests. Similar studies have also reported the power of the Red-Edge 1 band on the classification of plant and tree species (Immitzer et al., 2016) and the prediction of biophysical variables (Korhonen et al., 2017).

Figure 3-2 and Figure 3-3 show that the classification algorithm generalizes most of the vegetation as open land in 2016 for CART. In the SVM algorithm, most of the open land and water bodies were incorrectly classified as vegetation areas in 2016 and 2018. This is because the type of land cover has a reflection that is identical to the vegetation area. Nonetheless, SVM is quite good at classifying built-up areas and vegetation in 2020.

In Figure 3-1 it can be seen that forest land cover classes and vegetated areas tend to change from time to time, followed by an increase in the built-up area in both imagery types. The forest area decreased to 7.99 km² and 45.23 km² respectively, while the built-up expanded to 35.88 km² in 2020. Vegetation land cover tends to increase to 40.62 km² and 23.6 km² respectively. The open land experienced decline of 12.48 km² and 15.07 km² respectively, most of which were closed from forest to become built-up. The same thing happened to the land cover of water bodies which continued to shrink until 3.5 km² stopped becoming land.

Human activities have been shown to be responsible for driving these changes resulting in landscape changes that negatively impact ecosystem services and human well-being. (Lambin, 1997; Shiferaw et al., 2019). Main drivers of global environmental change, but not limited to, speed of urbanization (Sreenivasulu & Bhaskar, 2010), increase in population (Hassan & Nazem, 2016), other socio-economic development activities resulting in deforestation (Torbick, 2006), loss of biodiversity (Sala et al., 2000), change to arable agricultural land (Cihlar, 2000, Renschler & Harbor, 2002), and air resources.

### 3.2 Comparison of Classification Performance

The performance of the most commonly used classifier in evaluating the accuracy and effectiveness of all classifiers through Overall Accuracy (OA). This accuracy represents the number of pixel data correctly classified by the classifier algorithm into percentage form. In addition, the Confusion matrix, and UA and PA are also used to measure the class performance of each classifier. The model with the best capability is selected based on OA and k. The performances of the CART, RF, and SVM classifiers are compared in Table 2-2 and Table 2-3 in terms of OA and k with different forms of dataset division.

Exploration of training tests and data testing was carried out on all algorithms in 2016, 2018, and 2020. The comparison of training and testing datasets with a good level of accuracy was 70:30. The average OA on Landsat-8 and Sentinel-2 was 92.09%, and 94.21%, respectively, followed by a ratio of 60:40, 80:20, and 50:50 on Landsat-8, and a ratio of 80:20, 50:50, and 60:40 on Sentinel-2. Another study also showed that the ratio of dataset comparison in the best RF algorithm was 70:30 (Saha et al., 2021; Adelabu et al., 2015).

From Table 2-2 and Table 2-3, it is also seen that the classifier using RF outperforms the classifier with CART and SVM in both types of satellite imagery. In OA, the mean CART, RF, and SVM classifiers were 92.03%, 94.74%, and 83.54% on Landsat-8. Meanwhile, the average OA of the CART, RF, and SVM classifiers was 93.14%, 96.15%, and 93.34% on Sentinel-2. When compared with CART and SVM, the RF classifier shows the highest producer accuracy and user accuracy, as proven by other studies (Shetty, 2019).
Figure 3-1: LULC changes using RF classifier of Sentinel-2 and Landsat-8 for the years 2016, 2018, and 2020.

Table 2-2: Kappa coefficient and overall accuracy from Landsat-8 for various ratios for comparison of training data and testing data

<table>
<thead>
<tr>
<th>Year</th>
<th>Classifier</th>
<th>50:50 OA (%)</th>
<th>k</th>
<th>60:40 OA (%)</th>
<th>k</th>
<th>70:30 OA (%)</th>
<th>k</th>
<th>80:20 OA (%)</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>CART</td>
<td>83.33</td>
<td>0.79</td>
<td>89.58</td>
<td>0.86</td>
<td>92.31</td>
<td>0.90</td>
<td>89.36</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>89.17</td>
<td>0.86</td>
<td>93.75</td>
<td>0.91</td>
<td>93.85</td>
<td>0.92</td>
<td>93.62</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>89.17</td>
<td>0.85</td>
<td>87.5</td>
<td>0.84</td>
<td>87.88</td>
<td>0.84</td>
<td>85.11</td>
<td>0.81</td>
</tr>
<tr>
<td>2018</td>
<td>CART</td>
<td>89.47</td>
<td>0.86</td>
<td>88.17</td>
<td>0.84</td>
<td>90.48</td>
<td>0.88</td>
<td>89.13</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>89.47</td>
<td>0.86</td>
<td>93.56</td>
<td>0.92</td>
<td>93.75</td>
<td>0.92</td>
<td>93.49</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>88.6</td>
<td>0.85</td>
<td>93.55</td>
<td>0.92</td>
<td>93.55</td>
<td>0.92</td>
<td>93.48</td>
<td>0.91</td>
</tr>
<tr>
<td>2020</td>
<td>CART</td>
<td>90.91</td>
<td>0.88</td>
<td>93.59</td>
<td>0.92</td>
<td>93.22</td>
<td>0.91</td>
<td>89.74</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>89.9</td>
<td>0.87</td>
<td>93.59</td>
<td>0.92</td>
<td>96.61</td>
<td>0.95</td>
<td>89.74</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>90.91</td>
<td>0.87</td>
<td>91.03</td>
<td>0.88</td>
<td>87.18</td>
<td>0.79</td>
<td>92.86</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Table 2-3: Kappa coefficient and overall accuracy from Sentinel-2 for various ratios for comparison of training data and testing data

<table>
<thead>
<tr>
<th>Year</th>
<th>Classifier</th>
<th>50:50 OA (%)</th>
<th>k</th>
<th>60:40 OA (%)</th>
<th>k</th>
<th>70:30 OA (%)</th>
<th>k</th>
<th>80:20 OA (%)</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>CART</td>
<td>83.83</td>
<td>0.73</td>
<td>84.21</td>
<td>0.8</td>
<td>92.54</td>
<td>0.90</td>
<td>90.63</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>90.29</td>
<td>0.79</td>
<td>89.48</td>
<td>0.86</td>
<td>95.59</td>
<td>0.94</td>
<td>93.19</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>81.81</td>
<td>0.81</td>
<td>89.83</td>
<td>0.79</td>
<td>93.18</td>
<td>0.85</td>
<td>90.48</td>
<td>0.88</td>
</tr>
<tr>
<td>2018</td>
<td>CART</td>
<td>93.87</td>
<td>0.88</td>
<td>82.24</td>
<td>0.77</td>
<td>93.23</td>
<td>0.91</td>
<td>93.33</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>97.95</td>
<td>0.93</td>
<td>85.98</td>
<td>0.82</td>
<td>96.05</td>
<td>0.96</td>
<td>96.08</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>96.47</td>
<td>0.92</td>
<td>91.66</td>
<td>0.85</td>
<td>94.2</td>
<td>0.92</td>
<td>92.04</td>
<td>0.89</td>
</tr>
<tr>
<td>2020</td>
<td>CART</td>
<td>87.73</td>
<td>0.77</td>
<td>83.75</td>
<td>0.78</td>
<td>93.65</td>
<td>87.5</td>
<td>92.41</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>92.62</td>
<td>0.91</td>
<td>90</td>
<td>0.87</td>
<td>96.82</td>
<td>0.95</td>
<td>92.45</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>90.32</td>
<td>0.87</td>
<td>89.41</td>
<td>0.86</td>
<td>92.64</td>
<td>0.85</td>
<td>91.67</td>
<td>0.89</td>
</tr>
</tbody>
</table>

The accuracy of UA and PA in each class is presented in Figure 3-4 and Figure 3-5. It can be seen that compared to other classes, vegetation and water body classes have good performance, with UA and PA above 80%. In contrast to RF, SVM has difficulty identifying water bodies efficiently. This is because the appearance in the image has very few pixels so it is not enough to train classification accurately and tends to produce poor performance when compared to other land uses. In UA and PA, the SVM classifier mostly performs better than CART on the vegetation class.
It is very difficult to distinguish between vegetation land classes (agriculture/plantation) and open land in the 30 m resolution Landsat-8 imagery because of the mixed pixels. However, this is not a problem with Sentinel-2 imagery. Multiple land use classifications are combined at the same time and allow superior classification for small areas and more land cover types. In this study, the Sentinel-2 image shows a better and clearer image appearance. When compared between Landsat-8 and Sentinel-2, Sentinel-2 produces the highest accuracy results due to the more detailed spatial resolution and the number of band combinations that are applied more than Landsat-8.

Random Forests generally incorporate many soft linear boundaries on the decision surface. In SVM and CART, misclassification occurs between several classes, and SVM works best when training data is inputted less or less frequently. In this study, when the number of training data inputs is the same as the RF and CART algorithms, the SVM class targets will overlap because each data point exceeds the number of training data specimens that can be processed so that SVM must be run with separate training samples to get optimal results. In line with what was explored by Cervantes et al. (2007), Rudrapal & Subhedar, (2015), and Shetty (2019), that the number of class targets that are too large will also worsen SVM performance, so that in this research the number of suitable classes explored by SVM is 5 classes. Whereas CART and RF classifier do not show significant changes in accuracy if the number of training samples is reduced (Deng & Wu, 2013). This may be a better option if there are less data available (Shetty, 2019).

![Figure 3-2: LULC Maps from Landsat-8 Imagery with CART, RF, and SVM algorithms for 2014, 2018, and 2021](image-url)
Each algorithm has advantages and disadvantages. RF is more resistant and less affected by the presence of parameters, whereas SVM is more sensitive to hyperparameters (Chang et al., 2019). In this study, RF did outperform all classifier algorithms in terms of UA and PA, regardless of training data size, followed by CART, then SVM. However, the SVM and CART classifiers performed better for water bodies and forest land cover. Vegetated land cover was partially misclassified as forest by the SVM and CART classifiers. Some studies claim that SVM outperforms CART (Shao & Lunetta, 2012) or vice versa (Congalton & Green, 2019). In such study, as suggested by Congalton & Green (2019) it is best to choose lesser complex and faster algorithm of the two for classification. The better performance of CART over SVM could be favoured due to the quality of training samples used. The use of multispectral images in GEE can simplify the process of classifying large study areas. This can make performance during image pre-processing more efficient with these implemented methods and algorithms.

Figure 3-3: LULC Maps from Sentinel-2 Imagery with CART, RF, and SVM algorithms for 2014, 2018, and 2021
Figure 3-4: Accuracy of the user for each land class using CART, RF, and SVM classifiers: (a) Landsat-8, (b) Sentinel-2
4 CONCLUSION

The type of classifier affects the accuracy of the classification of LULC data on satellite images. The comparison between training data and testing data is good at 70:30 for both types of satellite imagery. Overall Accuracy and Coefficient Kappa RF classifier outperforms CART and SVM on both types of satellite imagery. Sentinel-2 imagery performs better classification than Landsat-8 because Sentinel imagery has red-edge bands, which makes it possible to classify vegetation better than Landsat. The high-resolution pixel size also makes Sentinel-2 outperform Landsat-8 in terms of accuracy.

ACKNOWLEDGEMENTS

We would like to thank the University of Indonesia for facilitating this research.

AUTHOR CONTRIBUTIONS

Author contributions are as follows: Septianto Aldiansyah: imagery processing, conceptualization, methodology, results analysis, map layouting. Randi Andrian Saputra: imagery processing, results analysis, prepare draft manuscript.

REFERENCES


Figure 3-5: Accuracy of the producer for each land class using CART, RF, and SVM classifiers: (a) Landsat-8, (b) Sentinel-2.


Sreenivasulu, V., & Bhaskar, P. U. (2010). Change detection in land use and land cover using remote...


210