Multi-Temporal Data for Land Use Change Analysis Using a Machine Learning Approach (Google Earth Engine)

Aji, A.,[1*] Husna, V. N.[1] and Purnama, S. M.[2]
[1]Department of Geography, Faculty of Social Sciences, Universitas Negeri Semarang, Indonesia
E-mail: ajiananto@gmail.unnes.ac.id[*]
[2]Surveying and Mapping Study Programme, Politeknik Sinar Mas Berau Coal, Indonesia
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v20i4.3145

Abstract
Land use and land cover change have significant impacts on climate, the environment, and natural ecosystems. This research analyzes land use change over time using Google Earth Engine (GEE) and provides recommendations for land use planning based on the results. This research aligns with priorities related to SDG issues, specifically the maintenance of degraded terrestrial ecosystems that negatively impact the livelihoods of the organisms that inhabit them. The study utilized Sentinel-2 with Multi Spectral Instrument, Level 2A time series from 2019 - 2023, which were processed using cloud computing and a classification method utilizing SMILE Random Forest. The classification model achieved an accuracy value of 95%. The calculation results indicate a total land use and land cover (LULC) area of 1007.96 hectares. The largest change in land use occurred in fields, decreasing from 10.64% to 4.96%, or 57.18 hectares, while the smallest change occurred in dryland forest, at 2.18 hectares. The total predicted LULC change area was 558.02 hectares.

Keywords: Cloud Processing, Machine Learning, Random Forest, Google Earth Engine, Land Use/Land Cover

1. Introduction

The demand for land is increasing (on the demand side), while the availability of land remains unchanged (on the supply side). To meet this demand, policies regarding land use and land cover (LULC) must change. LULC refers to the physical composition and characteristics of land elements on the Earth's surface. The LULC system is complex, consisting of natural, social, and economic spatial components. Effective management and planning of land use and land cover (LULC) can significantly mitigate extreme events, climate change, and land degradation [1] [2] and [3]. Retaining water in the soil and allowing vegetation to absorb water can prevent flooding, landslides, and erosion [4]. Changes in land use and land cover (LULC) have a significant impact on humidity and temperature conditions, as well as the composition of the atmospheric boundary layer (ABL). It is crucial to consider the effects of LULC on these factors when studying the environment. Land use has a significant impact on local weather, climate, and air quality [5]. Proper land conservation practices can prevent catastrophic events such as floods, landslides, and erosion during land cover change [6]. It is crucial to consider the effects of land use on the environment to avoid these negative impacts. Therefore, it is imperative to implement proper land conservation practices.

Remote sensing technology provides efficient and cost-effective information on land use change studies from a distance. It is recognized as the most important data source for mapping land cover and monitoring land cover change over time [7]. Urban morphology can be easily identified through remote sensing by examining trends in development direction and regional development based on land use conditions at a given time [8]. The use of multi-temporal data provides comprehensive information on urban development, which is essential for land development monitoring purposes. This research confidently employs the Google Earth Engine (GEE) platform as a powerful data processing and analysis tool.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

The GEE platform provides detection and analysis tools for agricultural land use change. The text employs a passive tone and impersonal construction, avoiding first-person perspectives unless necessary. The text is free from grammatical errors, spelling mistakes, and punctuation errors. By using temporal Sentinel-2 data and the Random Forest classification method, GEE can analyze changes in the conversion of rice fields to other sectors by mapping rice field land cover [9]. The language used in the text is clear, objective, and value-neutral, avoiding biased, emotional, figurative, or ornamental language. Additionally, GEE can monitor urban development. High-level, standard language with consistent technical terms is used, and common sentence structure is adhered to. A model for assessing urban change was created using the U-Net architecture and deep learning techniques through the open source TensorFlow library on the GEE cloud platform [10]. The LULC mapping provided by GEE produces quick and precise spatial products that are useful for monitoring, analysis, and government policy recommendations.

2. Study Area

This research focuses on Kendal District in Central Java Province, Indonesia (Figure 1). The district has undergone significant land cover change, particularly the transition from vegetation to built-up land. Land use and land cover changes (LULC) are happening in almost all areas of Kendal regency, some of which are not in balance with the land's capacity [12]. One of the consequences of this land cover change is flooding, which has been a major concern in recent years. The Kendal sub-district is highly susceptible to flooding, and this susceptibility is steadily increasing. The extent of flooding has also risen due to an extreme weather event with very high rainfall at the end of 2022 [13]. The land's ability to cope with these climatic conditions is inadequate for its function, resulting in various events such as flooding and landslides. The affected areas begin from the southern part of Kendal District. Monitoring land cover poses a challenge in managing the area [14].

3. Methodology

Figure 2 illustrates the comprehensive research methodology, encompassing data pre-processing, survey validation, accuracy test classification, analysis, and recommendations. The research began with data pre-processing. Sentinel-2 data was utilized during this stage, based on its availability (time series) in Google Earth Engine (GEE). Radiometric correction and cloud masking were applied to enhance data quality.

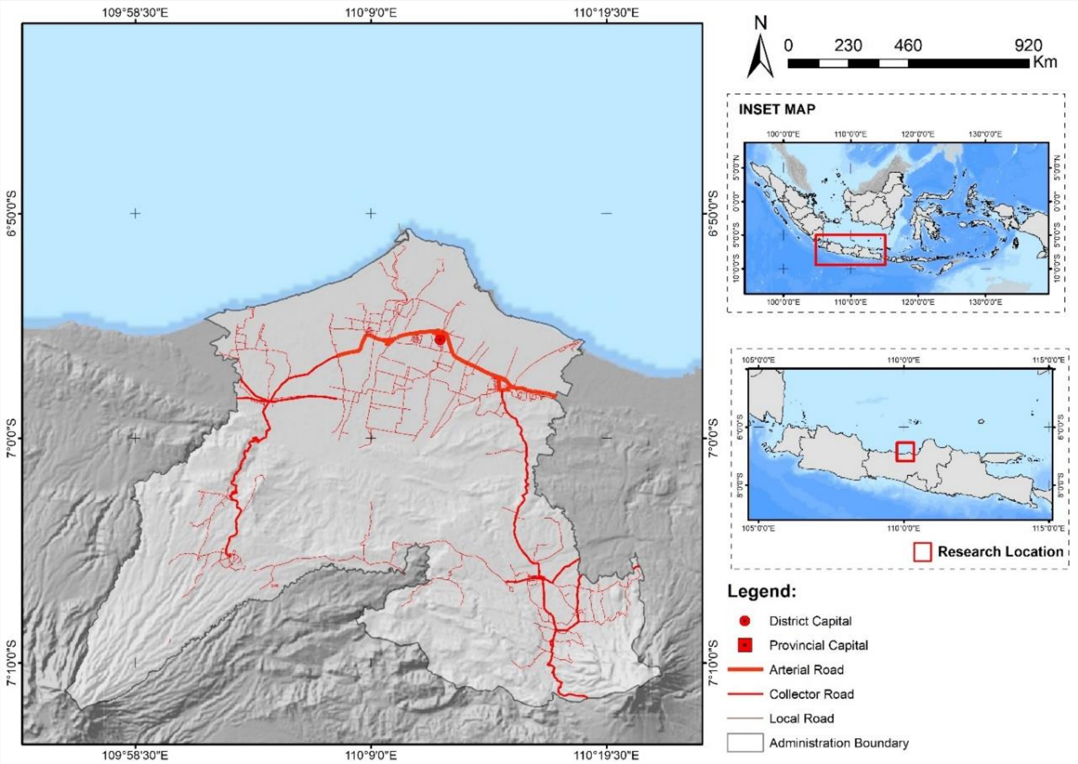


Figure 1: Kendal district, Indonesia

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

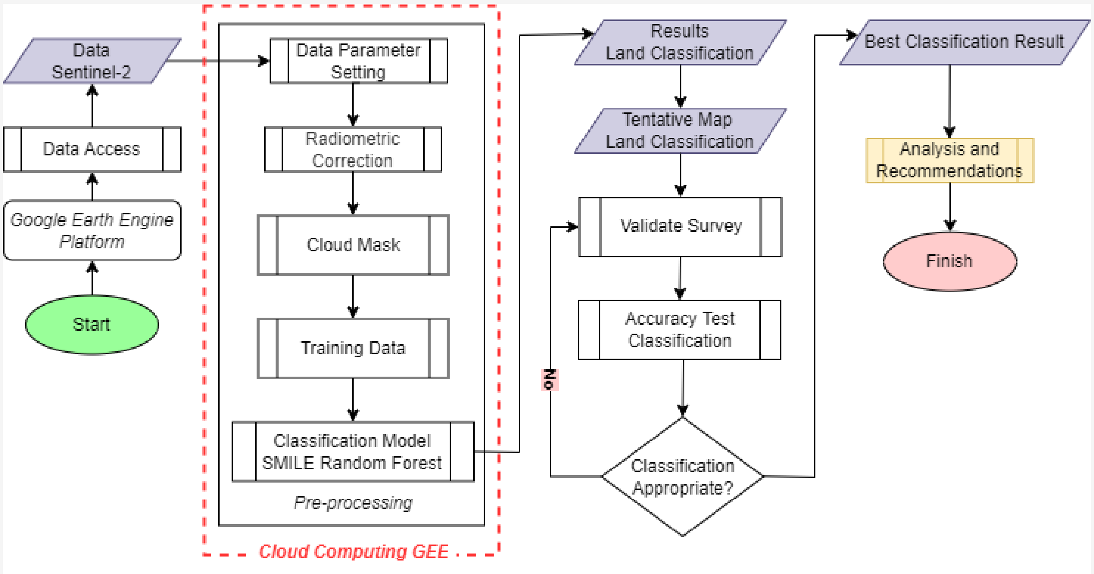


Figure 2: LULC detection methodology

```
var trainSample = image2022.sampleRegions({
  collection: train,
  scale: 50,
  properties: ['class'],
});
```

Figure 3: Determination of sample training class

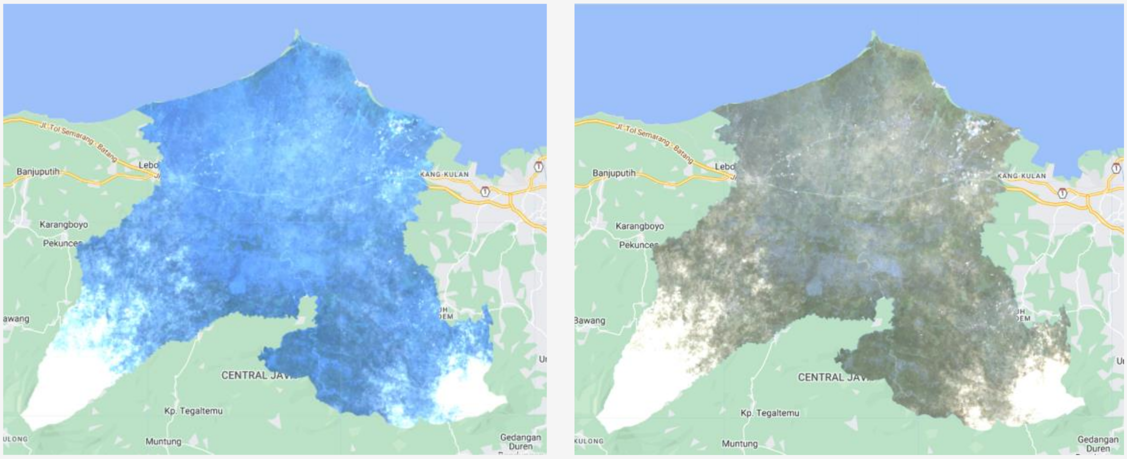
All Sentinel-2 data processing activities were conducted using cloud computing and cloud processing via the Google Earth Engine (GEE) platform. The random forest method was then employed to classify the Sentinel-2 data. The accuracy of the classification results was tested through field validation activities. Field observations documented the existing ground conditions during this phase. The image interpretation results were compared with the actual conditions in the field to measure the accuracy of the classification. The analysis and recommendation stage clearly explains the results of land change detection through a model built using a cloud processing approach. The model and system provide accurate and reliable spatial information in the form of land change in Kendal District. The results of land change detection can effectively assist stakeholders in making informed recommendations and decisions regarding land use, taking into account the carrying capacity, capacity and needs of the community.

The Random Forest method is a powerful machine learning technique that is commonly used for data classification, regression, and clustering. It employs ensemble learning, where a multitude of decision trees are constructed and merged to generate highly accurate predictions. The Random Forest (RF) method was confidently chosen, tested, and utilized for classification. This method has consistently demonstrated its success in classifying multitemporal satellite data, as evidenced by its performance in [15]. In fact, in numerous studies, RF has outperformed other classification algorithms such as CART, SVM, kNN, and MLC [16]. This nonparametric classification method is based on machine learning and is under control. It involves creating decision trees that evaluate each individual pixel to determine its class. For more information, please refer to the provided source [17].

In 2019 and 2023, the LULC class variable classified various objects, including Dryland Forest, Mixed Crops, Settlements, Industrial Buildings, Irrigated Rice Fields, Fields, Mixed Plantations, Water, Shrubs, and Plantations, using random sample selection. The classification object samples are trained using a scale of 50 demonstrated in Figure 3. Each feature in the training dataset represents the value of pixels with a size of approximately 50 x 50 metres in the image. The Random Forest method is used for machine learning-based classification, based on the training results.

```
var classifier = ee.Classifier.smileRandomForest(10).train({
  features: training,
  classProperty: 'class',
  inputProperties: before.bandNames()})
```

Figure 4: Training sample classification



(a) (b)

Figure 5: Radiometric correction results (a) Uncorrected image (b) Corrected image

Figure 4 depicts the SMILE (Statistical Machine Intelligence and Learning Engine) random forest method selects 10 representative decision trees to perform LULC classification. Each decision tree is built randomly and uses different subsets of training data to generate variation. The combined prediction results from each tree produce a highly accurate final result.

4. Result and Discussion

4.1 Radiometric Correction

Radiometric correction is necessary for remotely sensed imagery as it contains reflections from the Earth's surface that are impacted by Rayleigh and aerosol diffusion. This correction effectively removes the diffuse atmospheric contribution, resulting in improved visual quality of the image and increased effectiveness of the reflectance threshold. TOA radiometric correction is the preferred method for converting radiance values to reflectance, ensuring accurate and reliable results demonstrated in Figure 5. The atmospheric correction significantly enhances the visual clarity and detail of the image [18]. It amplifies the greenness of vegetation, such as forests and farmland, and brightens open land and settlements. This correction is particularly valuable for visual analysis, as it enables precise identification of each object.

4.2 Cloud Mask

Automatic tools for classifying cloud shadows are commonly based on geometric identification methods that involve thresholding a single spectral band, reflectance difference or ratio, or derived index [19]. The purpose of cloud removal is to provide comprehensive information about the Earth's surface and to minimize the impact of information bias on objects obscured by clouds demonstrated in Figure 6. The multitemporal cloud removal procedure was confidently employed by performing cloud removal on selected images over a one-year period using Scene Classification Layer (SCL) data. The procedure confidently utilised commonly used spectral bands, namely B2 (blue), B3 (green), B4 (red), B8 (NIR) and B11 (SWIR).

4.3 Land Use/Land Cover Classification

The Kendal District's land use/land cover (LULC) classification was confidently obtained using time series data from 2019 to 2023. The LULC classification includes a diverse range of categories such as Dryland Forest, Mixed Crops, Settlement, Industrial Buildings, Irrigated Rice Fields, Fields, Mixed Plantation, Water, Shrubs, and Plantation. The classification objects were evaluated Figure based on rigorous parameters such as Number of Trees, Variables per Division, and Max Nodes [20]. A novel script has been developed in the GEE environment to achieve the objective. The script quickly evaluates multiple input parameter combinations and selects the combination with the highest accuracy using the overall accuracy and kappa index. The classification model has an accuracy value of 95%, as demonstrated in Table 1.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

[1] [2]

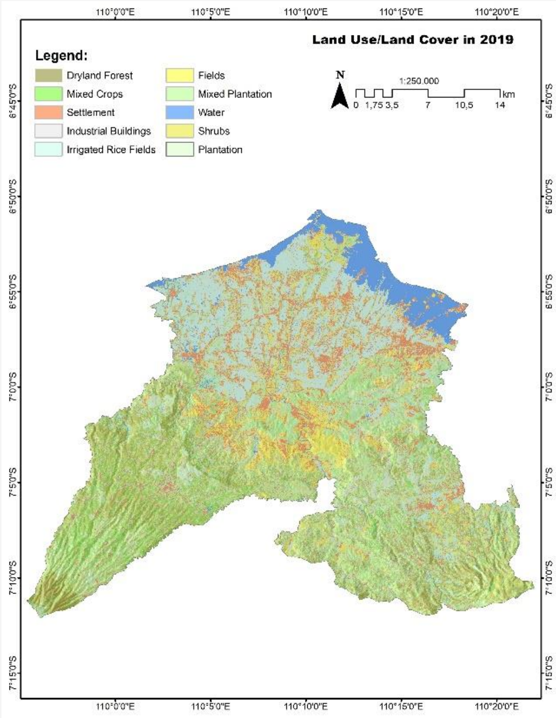
Figure 6: Comparison of cloud removal results (a) Cloudy image, (b) Cloud free image

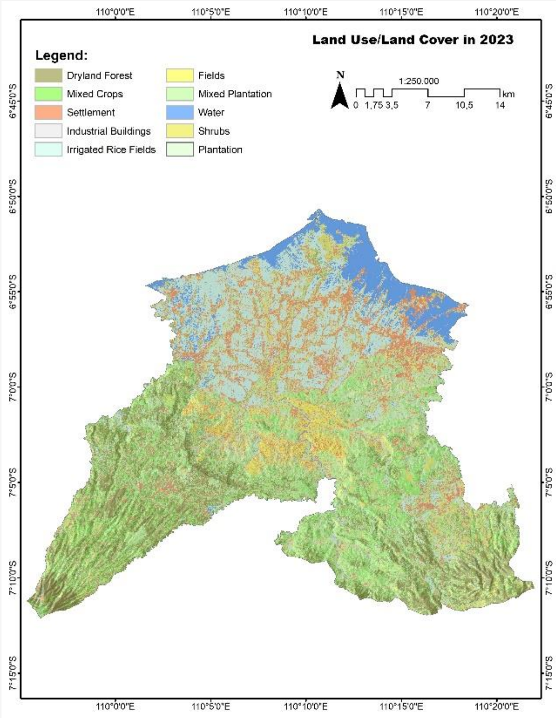
Table 1: Classification accuracy

		Reference Data										Total Rows
		Dryland Forest	Mixed Farms	Settlements	Industrial Buildings	Irrigated Rice Fields	Fields	Mixed Plantation	Waters	Shrubs	Plantation	
Classification	Dryland Forest	107	0	0	0	0	0	0	0	0	0	107
	Mixed Farms	8	30	0	0	0	0	2	0	0	1	41
	Settlements	0	0	40	0	0	0	0	0	0	1	41
	Industrial Buildings	0	0	0	5	0	0	0	0	0	0	5
	Irrigated Rice Fields	0	0	0	0	26	0	0	0	0	0	26
	Fields	0	0	0	1	0	14	1	0	0	0	16
	Mixed Plantation	0	1	0	0	0	0	44	0	0	0	45
	Waters	0	0	2	0	1	0	0	258	0	0	261
	Shrubs	0	0	0	5	0	0	2	0	17	0	24
	Plantation	1	0	0	0	0	0	1	0	0	19	21
	Column Total	116	31	42	11	27	14	50	258	17	21	587
											Accuracy	95.40 %

Figure 7 illustrates the 2019 and 2023 LULC maps, which depict the land use and land cover with 10 categories, including Dryland Forest, Mixed crops, Settlements, Industrial buildings, Irrigated rice fields, Fields, Mixed plantations, Water, Shrubs, and Plantations. Kendal District underwent significant changes in LULC over four years, including dryland forest, mixed farms, settlements, industrial buildings, water bodies, and shrubs. The land area for each LULC object in 2019 and 2023 is depicted in Figure 8, with a total LULC area of 1007.96 hectares. Irrigated Rice Fields dominated the landscape in 2019, covering 216.04 hectares, while industrial buildings occupied the smallest area, covering only 1.75 hectares. In 2023, Irrigated Rice Fields will continue to be the LULC object with the largest area, covering 208.32 hectares. The Industrial Building will occupy the smallest area, covering only 1.80 hectares.

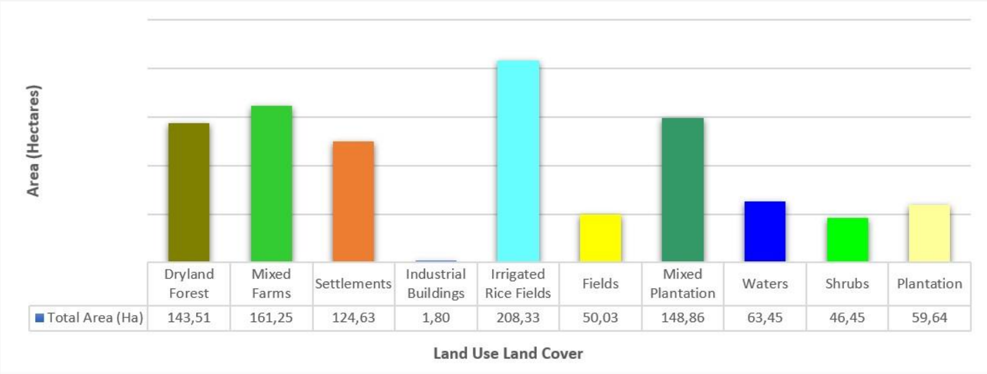
Figure 9 demonstrates that irrigated rice fields, mixed plantation, and dryland forest have the highest percentage in both 2019 and 2023 LULC. These findings are in line with the statistical data of Kendal Dalam Angka in 2023, which highlights the agricultural sector as one of the major contributors to the Gross Regional Domestic Product (GRDP) of Kendal Regency, accounting for 18.56% [21]. The government confidently supports this condition with the existence of Kendal Regency Regional Regulation Number 11 of 2020. The regulation aims to protect sustainable food agricultural land in Kendal Regency by amending Kendal Regency Regional Regulation Number 13 of 2013.



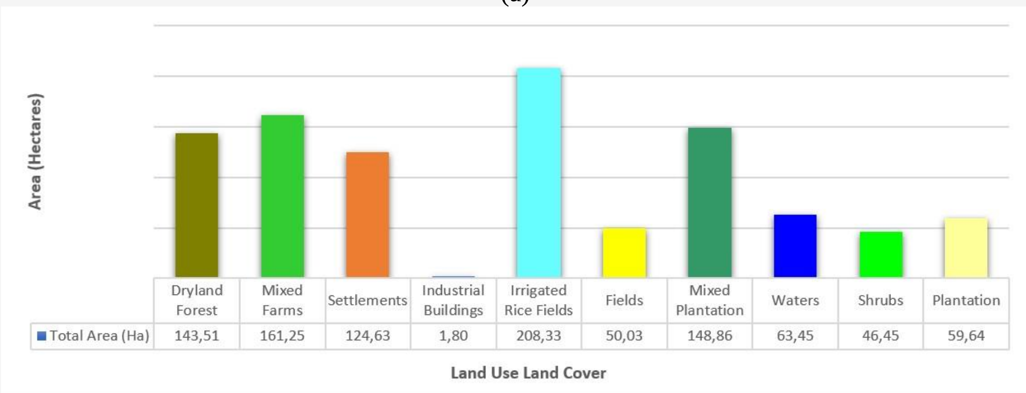


(a) (b)

Figure 7: LULC classification results (a) LULC 2019 (b) LULC 2023



(a)



(b)

Figure 8: Area of land change (a) LULC 2019 (b) LULC 2023

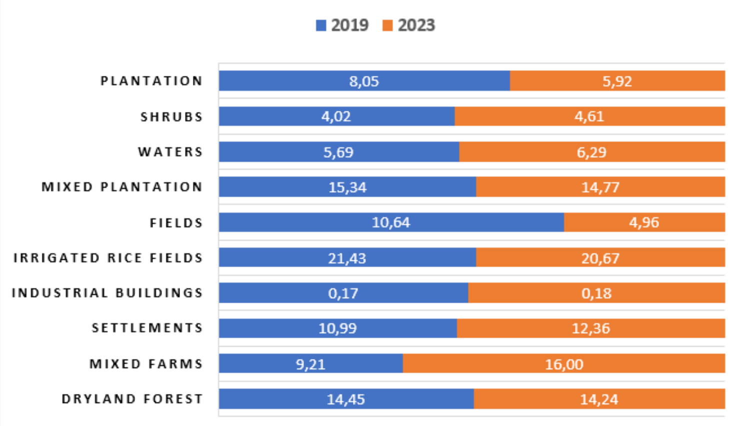


Figure 9: The percentage of land change

The Kendal District Regulation No. 11/2020 states that the area of sustainable paddy fields is 226.66 hectares, consisting of 218.39 hectares of paddy fields and 827 hectares of drylands spread over 19 sub-districts [22]. The LULC model results confidently predict a reduction in irrigated paddy area from 21.43% to 20.67% or 216.04 hectares in 2019 to 208.32 hectares in 2023. The LULC change detection clearly shows a decrease in several land types, including dryland forest, irrigated paddy, fields, mixed gardens and plantations. The largest change in land use is in fields from 10.64% to 4.96% or 57.18 hectares, while the smallest change is in dryland forest at 2.18 hectares. The shift in land use pattern from irrigated rice fields to ordinary fields was caused by the prolonged dry season. Changes in land use and land cover in Kendal District are very dynamic, especially in irrigated rice fields. The conversion of agricultural land to industrial use increased significantly between 2014 and 2018 [11]. This map confidently displays the changes in land use and land cover (LULC) by comparing two sets of data: before and after. The 'before' data is sourced from the 2019 classification results, while the 'after' data is from the 2023 classification. The LULC objects are then remapped based on their values, which range from 0-9 or 10 LULC classification objects. The LULC change map distinguishes between changed objects (shown in red) and fixed objects (shown in grey) by calculating the difference in each LULC object value.

Land change is present in almost all areas of Kendal District, with the highest accumulation of change occurring in the southern region. The LULC classification results indicate that the land change in the southern area is primarily due to Mixed Crops and Mixed Plantation, which are highly susceptible to human activities. In Indonesia, mixed plantations are typically associated with rural settlements or homesteads and are traditionally cultivated by local people.

The distribution of land use and land cover (LULC) change areas is clearly shown in Figure 10. The red colour highlights the areas that have undergone changes, while the grey colour indicates the areas that remain consistent with no changes. The LULC change detection results reveal that the affected area totalled 558.02 hectares, representing a significant level of dynamism over a period of 4 years. The LULC change detection results reveal that the affected area totalled 558.02 hectares, representing a significant level of dynamism over a period of 4 years. The LULC change detection results reveal that the affected area totalled 558.02 hectares, representing a significant level of dynamism over a period of 4 years. This accounts for 54.6% of the total area. The analysis unequivocally demonstrates that Kendal District is undergoing significant changes in land use, which can be attributed to various factors, particularly human activities and local land use policies. Human activities, such as urbanisation, intensive agriculture, and infrastructure development, have been identified as the primary drivers of these changes.

The high rate of change in land use underscores the urgent need for implementing sustainable land use policies. Continuous monitoring of these changes provides a basis for designing policies that focus on sustainable land use, taking into account land capacity and carrying capacity. With a clear understanding of the distribution of changes, policymakers can design effective strategies to manage land, minimize negative impacts, and maintain the balance of the environment and natural resources.

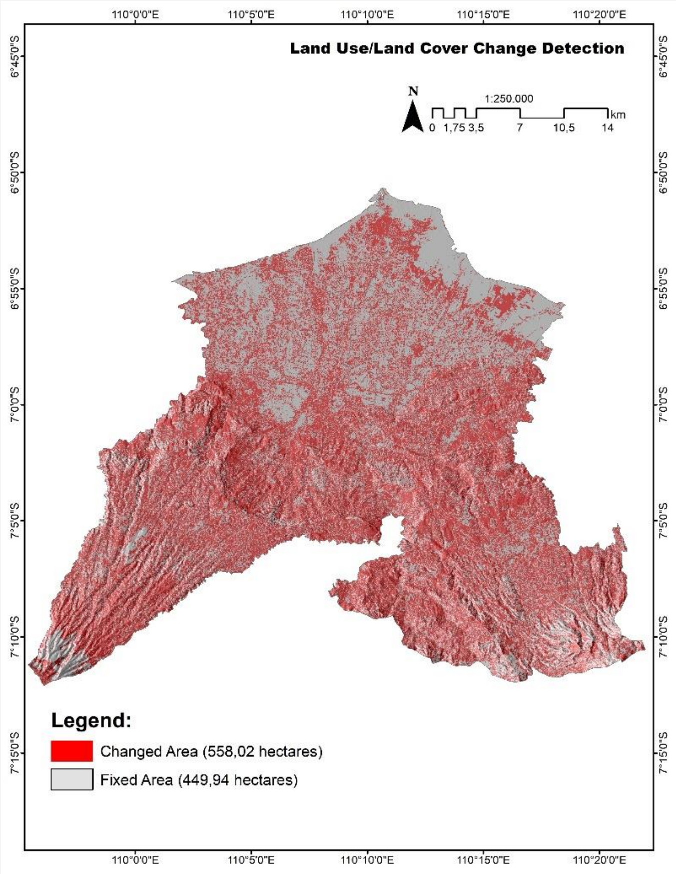


Figure 10: LULC change map

5. Conclusion

Based on the mapping of LULC (Land Use and Land Cover) changes from 2019 to 2023 using 10 classes of land cover classification according to the Indonesian national standard, the results of data processing using cloud processing with machine learning-based random forest classification method have shown significant ability in quickly and accurately detecting temporal land changes. High accuracy, measured by a classification accuracy of 0.9 and a kappa accuracy of 0.9 from 2019 to 2023, confirms the reliability of the method used. The analysis of the classification results shows that the highest changes occur in the mixed plant and settlement classes. This indicates significant dynamics in land use during the given period, which should be the main focus of future environmental planning and management. To improve classification accuracy, future work should integrate methods up to the level of deep learning. This approach is expected to provide higher accuracy and classification precision, strengthening the usefulness of mapping results for environmental planning and natural resource management in the future.

References

[1] Cihlar, J., (2000). Land Cover Mapping of Large Areas from Satellites: Status and Research Priorities. *International Journal of Remote Sensing*, Vol. 21(6–7), 1093–1114. https://doi.org/10.1080/014311600210092.

[2] Yan, W. Y., Shaker, A. and El-Ashmawy, N., (2015). Urban Land Cover Classification Using Airborne LiDAR Data: A Review. *Remote Sensing of Environment*, Vol. 158, 295-310. https://doi.org/10.1016/j.rse.2014.11.001.

[3] Aldileemi, H., Zhran, M., & El-Mewafi, M. (2023). Geospatial Monitoring and Prediction of land Use/Land Cover (LULC) Dynamics Based on the CA-Markov Simulation Model in Ajdabiya, Libya. *International Journal of Geoinformatics*, Vol. 19(12), 15–29. https://doi.org/10.52939/ijg.v19i12.2973

[4] Stürck, J., Schulp, C. J. E. and Verburg, P. H., (2015). Spatio-Temporal Dynamics of Regulating Ecosystem Services in Europe – The Role of Past and Future Land Use Change. *Applied Geography*, Vol. 63, 121–35. https://doi.org/10.1016/j.apgeog.2015.06.009.

[5] Sun, W., Liu, Z., Zhang, Y., Xu, W., Lv, X., Liu, Y., Lyu, H., Li, X., Xiao, J. and Ma, F., (2020). Study on Land-use Changes and Their Impacts on Air Pollution in Chengdu. *Atmosphere*, Vol. 11(1). https://doi.org/10.3390/atmos11010042.

[6] Pollner, J., Kryspin-Watson, J. and Nieuwejaar, S., (2013). *Disaster Risk Management and Climate Change Adaptation in Europe and Central Asia Europe and Central Asia*. Region and Sustainable Development Department. 1-66. https://www.preventionweb.net/files/15518_gfdrrdrmandccaeca1.pdf.

[7] Noi Phan, T., Kuch, V. and Lehnert, L. W., (2020). Land Cover Classification Using Google Earth Engine and Random Forest Classifier-the Role of Image Composition. *Remote Sensing*, Vol. 12(15). https://doi.org/10.3390/rs12152411.

[8] Dalilah, A., Malinda, A. R., Oktapiyansyah, R., Monicha, W. and Purnama, F., (2021). Monitoring Perubahan Penggunaan Lahan Menggunakan Citra SPOT 6 dan SPOT 7 di Kota Semarang. *Jurnal Ilmiah Ilmu Sosial*, Vol. 7(1). https://doi.org/10.23887/jiis.v7i1.23016.

[9] Mulyaqin, T., Kardiyono, K., Hidayah, I., Ramadhani, F. and Yusron, M., (2022). Deteksi Alih Fungsi Lahan Padi Sawah Menggunakan Sentinel-2 dan Google Earth Engine di Kota Serang, Provinsi Banten. *Jurnal Ilmu Pertanian Indonesia*, Vol. 27(2), 226-236. https://doi.org/10.18343/jipi.27.2.226.

[10] Zarro, C., Cerra, D., Auer, S., Ullo, S. L. and Reinartz, P., (2022). Urban Sprawl and COVID-19 Impact Analysis by Integrating Deep Learning with Google Earth Engine. *Remote Sensing*, Vol. 14(9). https://doi.org/10.3390/rs14092038.

[11] Adiyaksa, F. and Djojomartono, P. N., (2020). Evaluation of Land Use Change of Agricultural Land into Industrial Area in Kendal Regency in Period. *Journal of Geospatial Information Science and Engineering*, Vol. 3(1). https://doi.org/10.22146/jgise.55519.

[12] Febriana Saraswati, G., Suprayogi, A., Janu Amarrohman, F., (2017). Analisis Perubahan Tutupan Lahan das Blorong Terhadap Peningkatan Debit Maksimum Sungai Blorong Kendal. *Jurnal Geodesi Undip*, Vol. 6. 1-9. https://ejournal3.undip.ac.id/index.php/geodesi/article/view/16260/15686.

[13] Badan Meteorologi, Klimatologi, Dan Geofisika (BMKG), (2022). *Prakiraan Musim Hujan 2022/2023 di Indonesia*. 1-62. https://www.bmkg.go.id/berita/?p=prakiraan-musim-hujan-tahun-2022-2023-di-indonesia&lang=ID&s=detil#:~:text=Prakiraan%20Musim%20Hujan%202022%2F2023,417%20ZOM%20dari%20699%20ZOM.

[14] Organisation for Economic Co-operation and Development (OECD). (2017). *The Governance of Land Use*. https://www.oecd.org/gov/the-governance-of-land-use-in-oecd-countries-9789264268609-en.htm.

[15] Jin, Y., Liu, X., Chen, Y. and Liang, X., (2018). Land-Cover Mapping Using Random Forest Classification and Incorporating NDVI Time-Series and Texture: A Case Study of Central Shandong. *International Journal of Remote Sensing*, Vol. 39(23), 8703–8723. https://doi.org/10.1080/01431161.2018.1490976.

[16] Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sánchez, J. P., (2012). An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 67, 93-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

[17] Jamil, A. (2020). Image Classification using Machine Learning Algorithms Based on the Support Vector Machine and Random Forest. *International Journal of Geoinformatics*. Vol. 16(2), 15-22. https://journals.sfu.ca/ijg/index.php/journal/article/view/1809

[18] Ulfa, K., Hendayani, Oktavia, M. I., Pradono, K. A., Fibriawati, L., Muchsin, F., Candra, D. S., Damanik, K. W. V. D., (2020). Evaluation of Atmospheric Correction Algorithms for Sentinel-2 Over Paddy Field Area. *IOP Conference Series: Earth and Environmental Science*, Vol. 500(1), http://dx.doi.org/10.1088/1755-1315/500/1/012081.

[19] Magno, R., Rocchi, L., Dainelli, R., Matese, A., Di Gennaro, S. F., Chen C. F., Son, N. T. and Toscano, P., (2021). AgroShadow: A New Sentinel-2 Cloud Shadow Detection Tool for Precision Agriculture. *Remote Sensing*. Vol. 13(6), https://doi.org/10.3390/rs13061219.

[20] Svoboda, J., Štych, P., Laštovička, J., Paluba, D. and Kobliuk, N., (2022). Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data-A Case Study of Czechia. *Remote Sensing*, Vol. 14(5), https://doi.org/10.3390/rs14051189.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

[21] BPS, (2023). *Kabupaten Kendal Dalam Angka 2023*. Badan Pusat Statistik Kabupaten Kendal (Statistics of Kendal Regency), Indonesia. https://kendalkab.bps.go.id/publication/2023/02/28/64c90bd583090a412f244c4f/kabupaten-kendal-dalam-angka-2023.html.

[22] Peraturan Daerah Kabupaten Kendal Nomor 11 Tahun, 2020. (2020). *Perlindungan Lahan Pertanian Pangan Berkelanjutan di Kabupaten Kendal*. https://peraturan.bpk.go.id/Details/233278/perda-kab-kendal-no-11-tahun-2020.