JURNAL TEKNOINFO
Volume 16, Nomor 2, Juli 2022, Page 233-237
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

THE APPLICATION OF MACHINE LEARNING USING GOOGLE EARTH ENGINE FOR REMOTE SENSING ANALYSIS

Muhammad Iqbal Habibie

Agency for the Assessment and Application of Technology (BPPT)
Jl. M.H Thamrin No. 8 Jakarta Pusat DKI Jakarta, Indonesia
Email: muhammad.iqbal@bppt.go.id

Abstract

The spatial dimensions and temporal resolutions of the change detection analyses have been limited by traditional methodologies (i.e., desktop computing, open source). For decades, Remote Sensing (RS) have been collected large amounts of data, which are difficult to manage and analyzed using standard software packages and desktop computing resources. For this, Google developed the Google Earth Engine (GEE) cloud computing to successfully meet the issues of large data analysis. GEE is a cloud-based computing as a planetary-scale geospatial platform for Earth science data and analysis, allows these spatiotemporal constraints to be lifted and handle massive amounts of geodata over wide areas and to monitor the environment over long periods of time. We summarize the GEE data catalog's big geospatial data such as Climate and weather for surface temperature, climate, atmospheric and weather. It also contains Imagery like Landsat, Sentinel, MODIS and High-resolution Imagery and Geophysical information contains of terrain, land cover, cropland, and other geophysical data. Furthermore, supervised machine and unsupervised machine algorithms were used for several applications for Land Use Land Cover (LULC), hydrology, urban planning, natural disasters, and climate assessments. The research describes the utilization to resolve the big data using machine learning algorithm.

Keyword: Remote Sensing, GEE, cloud-based, machine learning, geospatial data

1. INTRODUCING

The spatial dimensions and temporal resolutions of the change detection analyses have been limited by traditional methodologies (i.e., desktop computing, open source). For decades, Remote Sensing (RS) have been collected large amounts of data, which are difficult to manage and analyzed using standard software packages and desktop computing resources. For this, Google developed the Google Earth Engine (GEE) cloud computing to successfully meet the issues of large data analysis. GEE is a cloud-based computing as a planetary-scale geospatial platform for Earth science data and analysis, allows these spatiotemporal constraints to be lifted and handle massive amounts of geodata over wide areas and to monitor the environment over long periods of time [1]. GEE is presently available to registered users via two web-based platforms: Code Editor and maps. The GEE Code Editor allows users to do analysis and modification using programming (JavaScript or Python), but the maps show the current satellite images, precipitation, Data Elevation Model (DEM), etc. This platform enables users to develop and run several algorithms and rapid calculations that allow for easy global scale analysis. Users can produce systematic data products utilizing interactive apps after developing algorithms in Google Earth Engine. Overlay analysis, matrix calculations, map calculations, image processing, time-series analysis, raster & vector conversion, statistics, and other complicated geographic data analysis are all possible. GEE system that does not require application development, web programming, or HTML. Users enable accurate and timely mapping and monitoring of vegetation from year to year.

We summarize the GEE data catalog's big geospatial data such as Climate and weather for surface temperature, climate, atmospheric and weather. It also contains Imagery like Landsat, Sentinel, MODIS and High-resolution Imagery and Geophysical information contains of terrain, land cover, cropland, and other geophysical data. Furthermore, supervised machine and unsupervised machine algorithms were used for several applications for Land Use Land Cover (LULC), hydrology, urban planning, natural disasters, and climate assessments. Previous research studies in Indonesia [2][3] suggested some implications of new research subjects [4], Analysis of Drought Potential Distribution[5][6], Potential Fishing Zones Estimation[7], and alternative technologies utilizing remote sensing [8].

In this study, we analyzed the area of interest to evaluate the land based on management parameters that limit the accessibility and can be assessed using a Google Earth Engine. The research describes the utilization to resolve the big data using machine learning algorithm.

JURNAL TEKNOINFO
Volume 16, Nomor 2, Juli 2022, Page 233-237
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

2. RESEARCH METHODS

Muara Enim Regency is one of the areas in South Sumatra Province. Geographically, it is located between 4º – 6º south latitude and 104º – 106º east longitude. The administrative area of Muara Enim Regency is divided into 20 sub-districts consisting of 326 villages/kelurahan, namely 310 villages and 16 sub-districts. The capital is in Muara Enim District. The topography of Muara Enim Regency is quite diverse, ranging from lowlands to highlands. Most of the sub-districts are in low-lying areas with an altitude of less than 100 meters above sea level (asl) which includes 20 (twenty) sub-districts, with an area of 7,058.41 km² (77.22 percent) of the area of Muara Enim Regency. Five other sub-districts are located at an altitude of more than 10 meters above sea level (masl), namely Lawang Kidul District (100-50 m asl), Tanjung Agung District (500-800 mdpl), Semende Darat Tengah District (100 m asl), Semende Darat Laut Subdistrict (500-1000 m asl) and Semende Darat Ulu Subdistrict (>100 m asl).

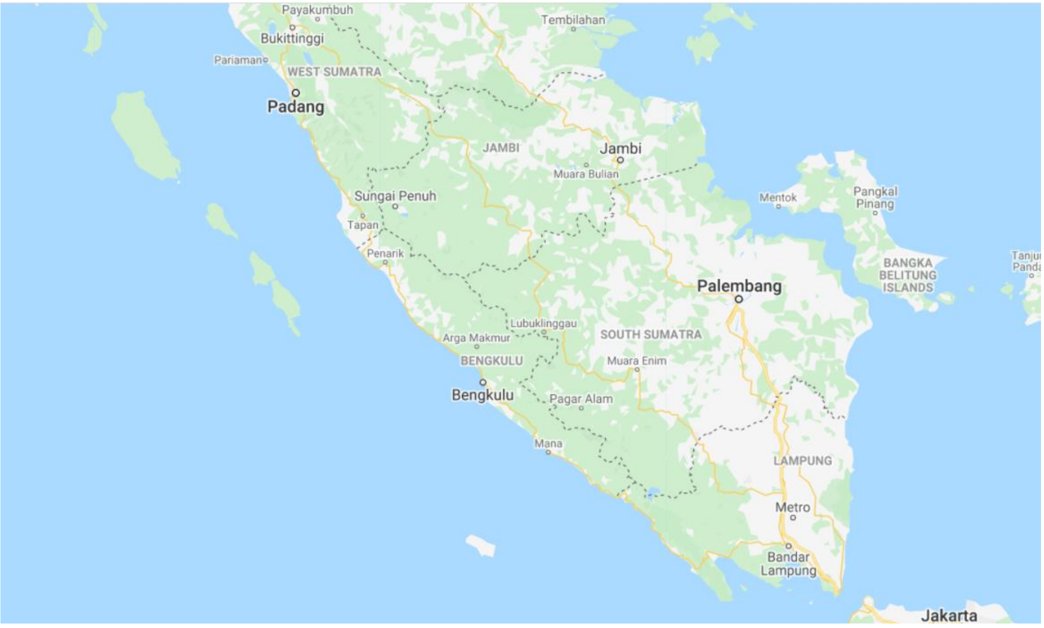

Figure 1. Muara Enim Regency

Land use in Muara Enim is divided into 2 (two) major groups, namely Protected Areas and Cultivation Areas. Protected area is an area designated with the main function of protecting environmental sustainability which includes natural resources, artificial resources, and historical and cultural values of the nation in the interest of sustainable development. This area is basically an area that based on carrying capacity analysis has limitations to be developed because of the limiting factors that serve as criteria (slope, soil type, rainfall, altitude; as well as volcanic hazard zone, soil movement vulnerability zone, and water conservation zone). very high potential).

2.1. Landsat 5 TM

The datasets were used Tier 1 calibrated top-of-atmosphere (TOA) reflectance from Landsat 5 TM Collection 1. The image metadata is used to obtain the calibration coefficients. Landsat 5 Collection 1 datasets were available from January 1984 to May 2012.

Table 1. Landsat 5 TM Collection 1 Tier Toa Reflectance

Name	Description	Resolution
B1	Blue	0.45 - 0.52 µm
B2	Green	0.52 - 0.60 µm
B3	Red	0.63 - 0.69 µm
B4	Near Infrared	0.76 - 0.90 µm

JURNAL TEKNOINFO
Volume 16, Nomor 2, Juli 2022, Page 233-237
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

B5	Shortwave Infrared 1	1.55 - 1.75 µm
B6	Thermal Infrared 1	10.40 - 12.50 µm
B7	Shortwave Infrared 2	2.08 - 2.35 µm
BQA	Band QA Bitmask	

Based on the table 1, the landsat 5 TM were used

2.2 MODIS MCD12Q1 Land Cover Type
The MODIS MCD12Q1 V6 package contains worldwide land cover types derived from six distinct categorization methods at annual intervals (2001-2016). It is calculated using MODIS Terra and Aqua reflectance data that have been supervised classified. The supervised classifications are then refined further with further post-processing that incorporates past knowledge and supplementary data.

Table 1. Dataset and SOurce

Name	Description
LC_Type1	Land Cover Type 1: Annual International Geosphere-Biosphere Programme (IGBP) classification
LC_Type2	Land Cover Type 2: Annual University of Maryland (UMD) classification
LC_Type3	Land Cover Type 3: Annual Leaf Area Index (LAI) classification
LC_Type4	Land Cover Type 4: Annual BIOME-Biogeochemical Cycles (BGC) classification
LC_Type5	Land Cover Type 5: Annual Plant Functional Types classification

2.3. Machine learning
We may train a collection of Decision Tree classifiers, each on a different random subset of the train set, to enhance our method. To create a prediction, we simply collect the forecasts of all individual trees and then forecast the class with the most votes. Random Forest is the name of this method.

3. RESULT AND DISCUSSIONS

Accordingly, we classified the landcover type from MODIS using machine learning using the random forest based on 7 bands as the input parameters. Then the classifiers were train using the confusion matrix were the validation data use 5000 sample. The results were shows the landsat 5 layer and land cover classifier layer.

3.1. Landsat 5 TM
The landsat 5 dataset were processed during January 2011 until End of December 2011 and processed with the least cloud using Google Earth Engine.

JURNAL TEKNOINFO
Volume 16, Nomor 2, Juli 2022, Page 233-237
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

Figure 1. Landsat 5 TM

3.2 Classifier MCD12Q1 Land Cover Type

From the results, the landcover type from MODIS using machine learning using the random forest based on 7 bands as the input parameters.

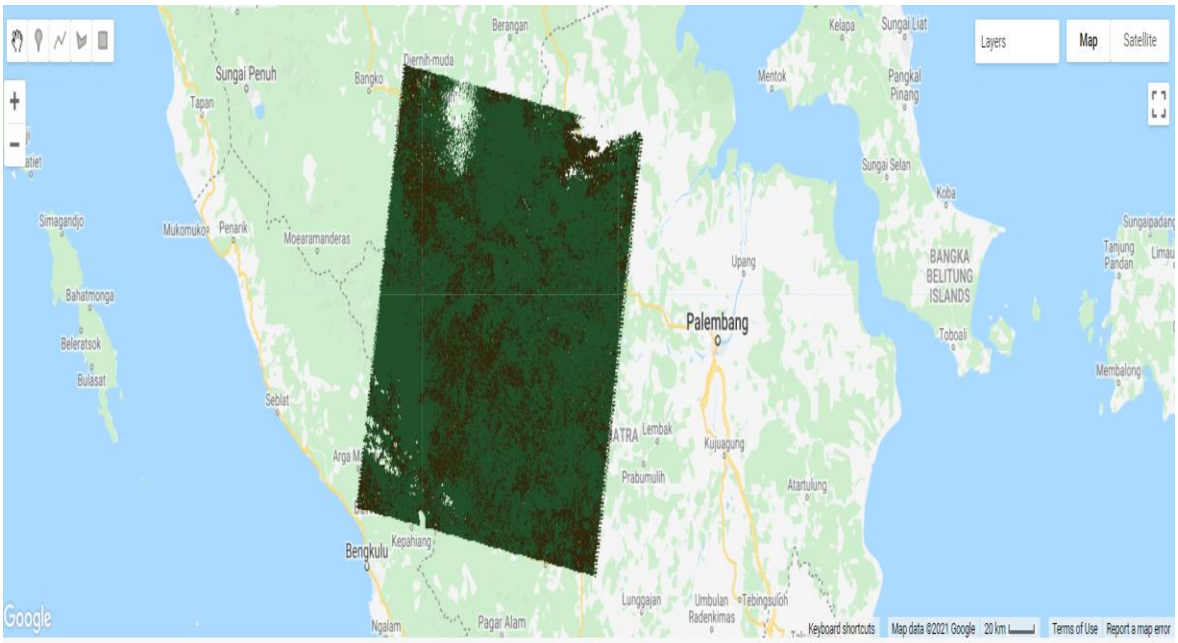

Figure 2. MCD12Q1 Land Cover Type

3.3. Machine learning

The classifiers were train using the confusion matrix were the validation data use 5000 sample. Then we calculate the overall accuracy from the confusion matrix. It shows that the training overall accuracy is 0.893 and validation overall accuracy 0.694. Research have been done using this algorithm of random forest has better accuracy [9]. This study has improvement from the previous study [10] which has overall accuracy 0.7.

4. CONCLUSION

This study demonstrates the application of machine learning using google earth engine for remote sensing analysis. The study was analyzed in Muara Enim regency which is mostly obtained as a forest area. As we can see from the results from the land cover type, mostly the land was covered by the trees. This study can be improved to improve the machine learning application using other methods and can be validated using other satellite images and other parameters like rainfall data [11].

REFERENCES

[1] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017, doi: 10.1016/j.rse.2017.06.031.

[2] N. Nurda, R. Noguchi, and T. Ahamed, *Forest productivity and carbon stock analysis from vegetation phenological indices using satellite remote sensing in Indonesia*, vol. 4, no. 3. Springer Singapore, 2020.

[3] N. Nurda, R. Noguchi, and T. Ahamed, "Change detection and land suitability analysis for extension of potential forest areas in Indonesia using satellite remote sensing and GIS," *Forests*, vol. 11, no. 4, pp. 1–22, 2020, doi: 10.3390/F11040398.

[4] M. I. Habibie, T. Ahamed, and R. Noguchi, "Deep Learning Algorithms to determine Drought prone Areas Using Remote Sensing and GIS," *2020 IEEE Asia-Pacific Conf. Geosci. Electron. Remote Sens. Technol.*, doi: 10.1109/AGERS51788.2020.9452752.

[5] F. D. Julianto and Ediyanto, "Analisis Sebaran Potensi Kekeringan Dengan Cloud Computing Platform di Kabupaten Grobogan," *J. Ilm. Tek. Geomatika IMAGI*, vol. 1, 2021.

[6] M. I. Habibie, R. Noguchi, S. Matsushita, and T. Ahamed, "Development of micro-level classifiers from land suitability analysis for drought-prone areas in Indonesia," *Remote Sens. Appl. Soc. Environ.*, vol. 20, no. June, p. 100421, 2020, doi: 10.1016/j.rsase.2020.100421.

[7] L. Vinet and A. Zhedanov, "Potential Fishing Zones Estimation Based on Approach of Area Matching Between Thermal Front and Mesotrophic Area Estimasi," *Ilmu dan Teknol. Kelaut. Trop.*, vol. 44, no. 8, pp. 951–952, 2011.

[8] M. I. Habibie, R. Noguchi, M. Shusuke, and T. Ahamed, *Land suitability analysis for maize production in Indonesia using satellite remote sensing and GIS-based multicriteria decision support system*, vol. 5. Springer Netherlands, 2019.

[9] X. Pan, Z. Wang, Y. Gao, X. Dang, and Y. Han, "Detailed and automated classification of land use/land cover using machine learning algorithms in Google Earth Engine," *Geocarto Int.*, vol. 0, no. 0, pp. 1–18, 2021, doi: 10.1080/10106049.2021.1917005.

[10] R. C. Sharma, K. Hara, and H. Hirayama, "Improvement of Countrywide Vegetation Mapping over Japan and Comparison to Existing Maps," *Adv. Remote Sens.*, vol. 07, no. 03, pp. 163–170, 2018, doi: 10.4236/ars.2018.73011.

[11] M. Enenkel *et al.*, "What rainfall does not tell us-enhancing financial instruments with satellite-derived soil moisture and evaporative stress," *Remote Sens.*, vol. 10, no. 11, 2018, doi: 10.3390/rs10111819.