MALAYSIAN JOURNAL OF SCIENCE
MJS
ISSN 1394-3065

Regular Issue

Malaysian Journal of Science 45(2): 42-54 (June 2026) https://mjs.um.edu.my

Utilization of Landsat-8 Image Classification Results with the Random Forest Algorithm to Determine Groundwater Recharge Areas (Case Study: Singgahan and Montong Districts, Tuban Regency)

Moh. Singgih Purwanto[1ab], Adi Susilo[2a*], Agus Naba[3a], Ayi Syaeful Bahri[4b], Muhammad Ari Sanggam[5c], and Noorlaila Hayati[6c]

Abstract: Groundwater is highly beneficial for society, and fulfills essential daily living needs and activities. With increasing population growth, groundwater use has increased, necessitating the creation of a map depicting the distribution of groundwater recharge areas. This map is expected to aid the government and communities to identify all the locations of groundwater distribution and ensure the sustainability of groundwater discharge in the area. The method employed involves using Landsat-8 image classification results to produce land cover maps of the study area. Land cover data processing was conducted using the open platform Google Earth Engine with a Random Forest algorithm. The classification of Landsat-8 images resulted in six land cover classes with an overall accuracy of 99.60% and a Kappa value of 0.994. Overlay and weighting were performed using parameters that determine the distribution of groundwater recharge areas. The parameters utilized included supporting data such as rainfall, soil type, and slope, which were obtained from DEMNAS data. The overlay and weighting results indicate that the potential recharge area covered 13,010.33 hectares (59.21%), the transition zone covered 8,935.27 hectares (40.66%), and the discharge area covered 28.64 hectares (0.13%).

Keywords: *Landsat-8, image classification, random forest, groundwater recharge area.*

1. Introduction

Groundwater, which is also known as subsurface water, is a natural resource that is crucial for humanity. According to Law Number 7 of 2004 on Water Resources, Chapter 1, Article 1, Paragraph 4 emphasizes that groundwater is defined as water contained in layers of soil or rocks below the surface of the land (Undang-Undang no 7, 2004). The benefits of groundwater are significant for society, as it meets essential daily living needs. With an increasing population growth, groundwater use has risen accordingly. Similar to other natural resources, groundwater is being exploited at an increasing rate worldwide. It is generally preferred as a source of water for household purposes because its availability is relatively constant and it is more naturally protected from pollution. Therefore, information regarding the distribution of water recharge areas in a region, such as Tuban Regency, is essential.

Tuban Regency is located in East Java, where land dips into bowls, then rises again into rocky limestone stretches that run sideways across the map. Based on data from the Meteorology, Climatology, and Geophysics Agency in 2021, rain falls here most months, around 1100 to 1500 millimeters each year. That same report counted roughly 90 to 120 wet days annually. Water bubbles out of the ground in places like Singgahan and Montong, spots home to natural flows such as Krawak spring. Bowl-shaped terrain mixed with porous rock makes seepage likely, feeding those outlets. Additionally, there are springs in several sub-districts within this district, such as the Krawak spring located between the Singgahan and Montong sub-districts. There has been no previous research in this area, and the presence of springs, supported by the basin topography and karst areas, justifies the need for research on the recharge areas in these two sub-districts. To determine the potential recharge areas, spatial analysis is required. One of the accurate methods for conducting spatial analysis is Remote sensing, which is a long-established and effective tool for land cover monitoring, due to its ability to quickly, widely, precisely, and easily provide information about spatial variability on the Earth's surface (Shin et al., 2022). Landsat-8 pictures step into the role perfectly for sorting ground types as they carry red, green, blue eyes, plus near and far infrared sight, each tuned to catch how surfaces bounce back light. Because certain bands react strongly to what covers the land, spotting differences becomes possible (Zulfajri et al., 2021).

There are several previous studies related to the research being conducted, including the work of Purwanto et al. (2022). Their approach pulled in Landsat-8 OLI/TIRS satellite images along with factors like rainfall, soil type, land cover, and slope. They used a points-based system to analyze land cover in the study area of Tanah Datar Regency, North Sumatra. The study concluded that the use of rainfall data, slope, land cover, and soil type can effectively indicate recharge areas. On investigating surface

Authors information:
[a]Brawijaya University, St. Veteran Malang, Lowokwaru, Malang, Indonesia. E-mail: singgih16@student.ub.ac.id[1]; adisusilo@ub.ac.id[2]; anaba@ub.ac.id[3]
[b]Geophysics Engineering Department, Institut Teknologi Sepuluh Nopember, St. Teknik Kimia, Sukolilo, Surabaya, Indonesia. E-mail: singgih@geofisika.its.ac.id[1]; syaeful_b@geofisika.its.ac.id[4]
[c]Geomatics Engineering Department, Institut Teknologi Sepuluh Nopember, St. Teknik Kimia, Sukolilo, Surabaya, Indonesia. E-mail: sanggamari@geofisika.its.ac.id[5]; noorlaila@geodesy.its.ac.id[6]
*Corresponding Author: adisusilo@ub.ac.id

Received: June, 2024
Accepted: November, 2024
Published: June, 2026

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

plants, some kinds stood out that allowed water to seep down into the dirt layers without much resistance. Soil type also has an impact due to its permeability, and slope data influences the formation of land where groundwater accumulates. However, rainfall data had a minor influence.

Furthermore, the study by Zulfajri et al. (2021) classified land cover using Landsat-8 OLI data in Pidie Regency, Aceh Province, with a Random Forest algorithm. The algorithm used 70% of the data for training and 30% for testing. The accuracy of the results was tested using the Confusion Matrix method. The study found that Landsat-8 OLI imagery is limited when cloud cover is present, affecting the quality and classification results. Therefore, careful attention is needed when selecting and determining the training data area. Instead of falling into typical traps with decision trees, the Random Forest method handled overfitting well. The results showed an accuracy of 89.53% and a kappa value of 0.91, indicating that the classification was reliable and accurate, making it useful as a reference for future research. Based on previous studies, this work differs from earlier findings as it dives into groundwater recharge zones by pairing Landsat-8 images with a Machine Learning approach, the Random Forest model, which is new in Singgahan and Montong, Tuban Regency. Additionally, the parameter data used in this study are the most recent, covering the year 2022 in the study area.

Based on the background and several previous studies, research will be conducted on the spatial analysis of groundwater recharge areas in Singgahan and Montong Districts, Tuban Regency. Satellite pictures from Landsat-8 form the base for sorting how land is covered there. Sorting these images uses Machine Learning, specifically Random Forest, to label surface types more accurately followed by a validation of the classification results against the actual physical conditions (ground truth) at the research location. Subsequently, overlay and weighting will be conducted using parameters that determine the distribution of groundwater recharge areas, including rainfall, soil type, and slope (Vasileva, 2019). It is hoped that this map of groundwater recharge area distribution will be useful for the government and local communities in identifying groundwater distribution locations, especially in Singgahan and Montong Districts, Tuban Regency.

2. Research Description

Research Location

The research location for the assignment is in Singgahan District and Montong District, Tuban Regency, East Java Province, Indonesia. Astronomically, Singgahan District is located at coordinates 6° 55' - 7° 1' South Latitude and 111° 44' - 111° 48' East Longitude. Singgahan District covers an area of 79.05 km² and consists of 12 villages. Meanwhile, Montong District is located at coordinates 6° 55' - 7° 1' South Latitude and 111° 48' - 111° 55' East Longitude. Montong District covers an area of 147.98 km² and consists of 13 villages. Both sub-districts in the research area contain numerous water sources, such as Krawak, Nglirip Waterfall, and Nganget Hot Springs. However, no research or exploration regarding groundwater recharge has been conducted in these areas, and that gap sparked the motivation for this study.

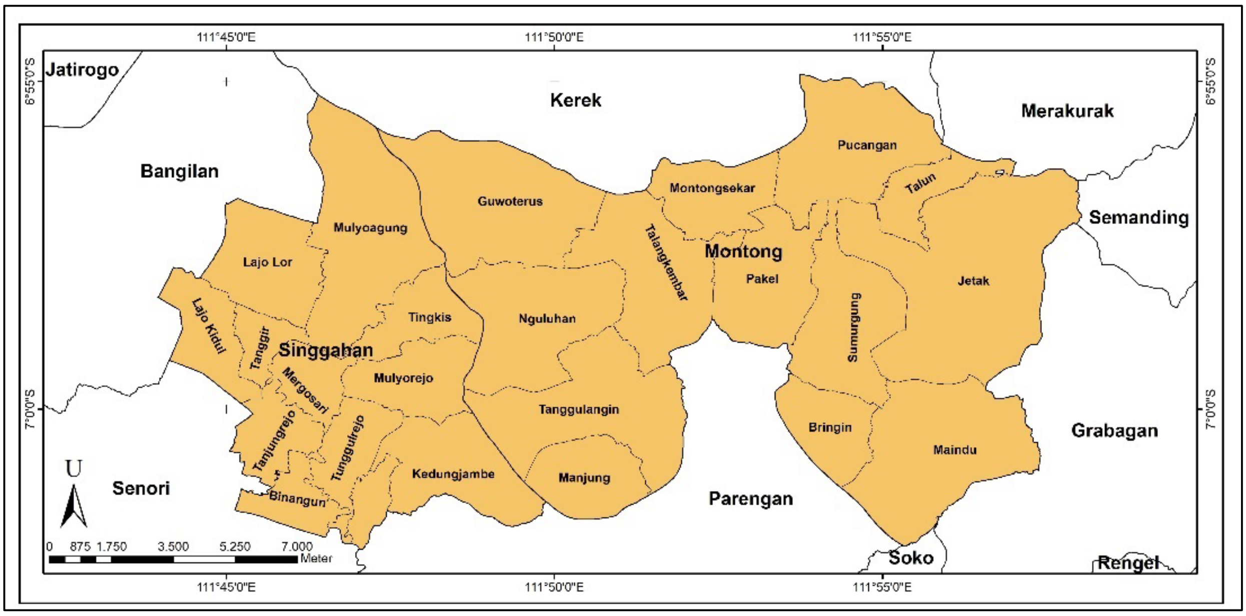


Figure 1. Research Location Map.

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Data and Equipment

In conducting this research, the data used were as follows:

1. Landsat-8 level 2 tier 1 satellite image data at 30-meter spatial resolution, obtained from the Google Earth Engine platform, covers the whole research area with the date filters specific and cloud cover less than 40%, sourced from USGS and accessed through earthexplorer.usgs.gov.
2. Vector map of the administrative boundaries of Singgahan District and Montong District at a scale of 1:25,000, obtained from Inageoportal accessed through Ina-Geoportal (indonesia.go.id).
3. Data on rainfall levels for the Tuban Regency area, 2022, sourced from BMKG.
4. Soil type data using soil type shapefile data obtained from the Ministry of Environment and Forestry.
5. Slope data using Digital Elevation Model National data with a spatial resolution of 8 meters obtained from Inageoportal.

The equipment used in this research includes:

1. A laptop and mouse.
2. Google Earth Engine for satellite image processing; geotaging software on mobile phones for validation sampling points; ArcGIS for studying patterns, showing results, and arranging layouts; and Microsoft Office 365 for number and word processing.

Stages of Processing Land Cover Maps with Landsat-8 Image Data

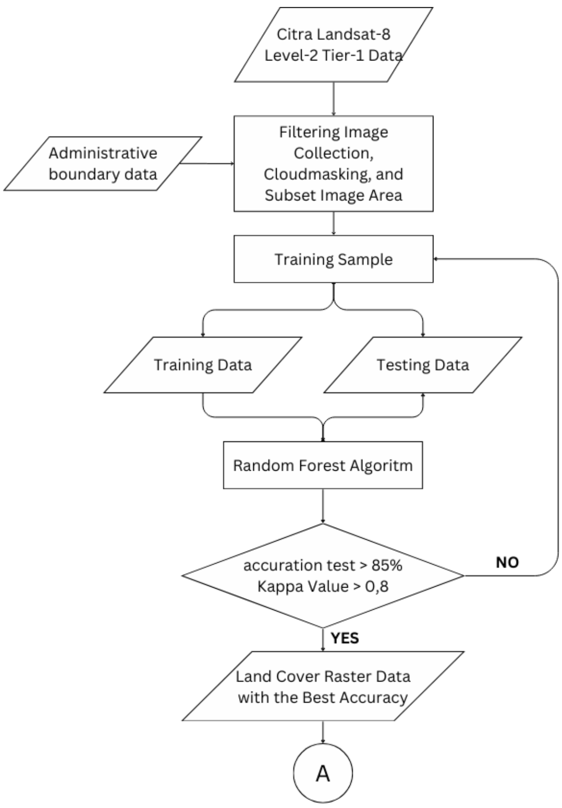


Figure 2. Flowchart of Land Cover Map Processing

The following is an explanation of the flow diagram above:

1. Create a script for accessing Landsat-8 Level 2 Tier 1 data by applying several filters, like time filters, cloud filters, and area filters. The selected image is cropped using the boundary of the study area extracted from InaGeoportal and given in shapefile format (subset data). Thus, an image without any cloud cover and adjusted to the study location area is obtained.
2. Create training samples by taking point samples in each land cover class, consisting of six land cover classes: water bodies, rice fields, settlements, moors, plantations, and dense forests.
3. Write a script to divide the training sample data into training data and testing data, with a ratio of 0.7 for training data to 0.3 for testing data.
4. Perform classification by creating a Random Forest algorithm script. In the classification process, band stacking is carried out by adding the Modified Normalized Difference Vegetation Index (MNDVI), Modified Normalized Difference Water Index (MNDWI), and Simple Ratio (SR) indices. These three indices are used to differentiate several land covers that will be classified. The MNDVI index is used to differentiate forest areas from grassland and agriculture (Wang & Fu, 2023). The MNDWI index is used to increase the contrast between water and built-up land, where water features will have increased values and built-up land will have decreased values (Ramji, 2022). The Simple Ratio (SR) is an index that has a good correlation with foliage density and is sensitive to three main external factors: sun and display geometry, soil background, and atmospheric effects (Melillos & Hadjimitsis, 2020).
5. Conduct an accuracy test after the classification results are available, displaying the accuracy value with Overall Accuracy (OA) and Kappa values. The classification accuracy is considered adequate if the OA value is >85% and the Kappa value is >0.80. If the classification accuracy is less than 85%, reclassification will be performed by improving the training area (Wiggers et al., 2020).
6. The result from processing the Landsat-8 satellite images is a raster map showing the land cover classification for the Singgahan District and Montong District in Tuban Regency.

Stages of Random Forest Processing

The land cover data processing in this research is done by applying the Random Forest algorithm. The outcome of the classifier, Random Forest, is chosen because it is a highly efficient and powerful classification method. Furthermore, it is capable of handling large training datasets as well as estimating missing data. The steps of utilizing Random Forest in this work are as follows:

1) The initial steps involve pre-processing as outlined below:
   a. Cloud masking: This accomplishes removing pixels covered by clouds or their shadows, through functions like cloudmask and maskL8sr. At this point, bitwise masking is done on the Cloud Shadow bit (bit 3) and Clouds (bit 5), both set to 0, meaning cloud-free pixels.

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Only pixels where bits 3 and 5 are both 0 (clear condition) are kept.

b. Advanced masking: This step utilizes the QA_PIXEL band to define and identify the pixel quality condition information. The first five bits (0-4) are inspected, where bit 0 represents fill, bit 1 is dilated cloud, bit 2 is cirrus, bit 3 is cloud, and bit 4 is cloud shadow. The results are compared with 0, and clear pixels are retained. Saturation masking is also used to detect saturated pixels.

c. Radiometric correction: Applied to optical and thermal bands to convert digital values into representative physical values. For the optical bands of Landsat-8, a scale factor correction is applied by multiplying each pixel by 0.0000275 and subtracting 0.2. This converts Digital Number (DN) values into surface reflectance. For the thermal bands, pixels are multiplied by 0.00341802, and 149.0 is added to convert DN into surface temperature in Kelvin.

2) Adding vegetation and water indices.

a. NDVI (Normalized Difference Vegetation Index): Used to measure vegetation density in a region by utilizing the ratio between the Near-infrared (NIR) band 5 and red bands using band 4. The NDVI value ranges from -1 to 1, where values closer to 1 indicate dense vegetation, while values closer to -1 indicate areas without vegetation, such as water or bare land. The equation is:

$$NVDI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

b. MNDWI (Modified Normalized Difference Water Index): Used to detect surface water presence, such as lakes, rivers, or water bodies. This index compares the GREEN band and shortwave infrared (SWIR). In Landsat-8 images, GREEN corresponds to band 3, and SWIR corresponds to band 6. Positive MNDWI values indicate water presence, while negative values indicate dry land or vegetation:

$$MNDWI\ 0 = \frac{GREEN - SWIR}{GREEN + SWIR} \tag{2}$$

c. MNDVI (Modified Normalized Difference Vegetation Index): Measures vegetation with more smoothness, reducing the impact of reflected light variation. It works by averaging the visible bands, RED, GREEN, and BLUE, compared to NIR:

$$MNVDI = \frac{NIR - (\frac{RED + GREEN + BLUE}{3})}{NIR + (\frac{RED + GREEN + BLUE}{3})} \tag{3}$$

d. SR (Simple Ratio): A simpler vegetation index is calculated by the ratio of reflected NIR and RED light. This index signals the presence of vegetation where higher values characterize more green or denser vegetation.

$$SR = \frac{NIR}{RED} \tag{4}$$

3) Satellite image acquisition: Landsat-8 Level 2 satellite data from April 1, 2022, to October 31, 2023, have been acquired. Only scenes with a cloud cover of less than 40% that mainly cover the chosen Region of Interest (ROI) are kept after filtering the data.

4) Model development: A model is constructed based on the Bagging, or Bootstrap Aggregating, concept. The first step is to train a model with a satellite image sample data to predict the land cover classes using spectral indices and bands. The model is designed to produce 100 decision trees to stabilize and improve the accuracy of the prediction. 70% of the data is utilized for training purposes, and the remaining 30% is kept for testing. During this step, the decision trees are formed by the use of random data subsets. Each tree identifies different features to divide the data, for instance, classifying the areas as forest if NDVI > x. The classification output of each tree is aggregated through a voting method to lower bias.

5) Training and feature association: The model was trained with the given training data, in which the input features consist of spectral bands and vegetation indices. Then the model links these features to different land cover classes, such as forest, water bodies, etc.

6) Testing and accuracy evaluation: Model accuracy is evaluated using the testing data, and various accuracy matrices are computed, such as:

a. Overall Accuracy: The ratio of correct predictions to the total data.

b. User Accuracy: The percentage of correct predictions for each class.

c. Producer Accuracy: The percentage of true data correctly classified by the model.

d. Kappa Coefficient: Measures model accuracy compared to random classification.

7) Application of the model: The trained model is applied to the entire study area, predicting land cover classes for each pixel based on the spectral features of the image. The results are displayed as land cover classes with different colours.

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Stages of Processing Parameters for Determining Groundwater Recharge Area

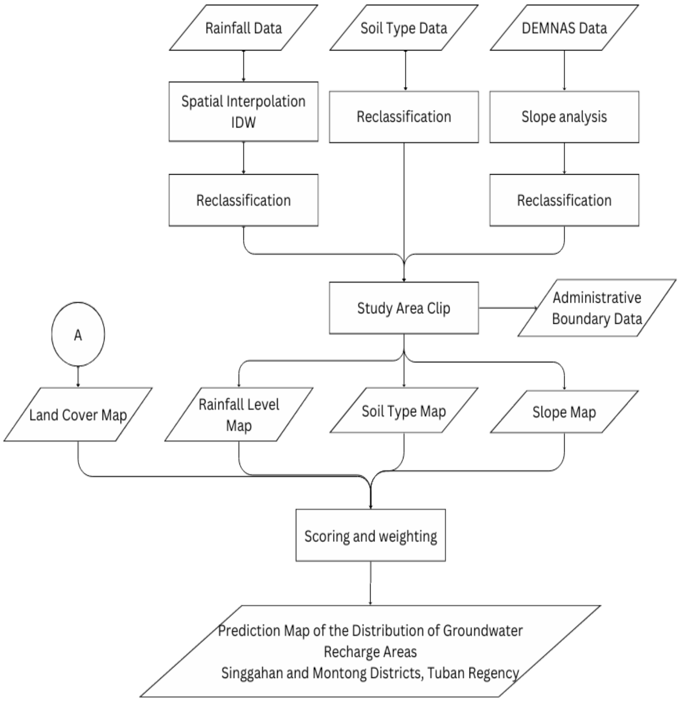


Figure 3. Parameter Data Flow Diagram

The following is an explanation of the following flow diagram:

1) Rainfall Data
Rainfall data from several rain gauge stations in Tuban Regency are obtained from the official BMKG website. The station points undergo an Inverse Distance Weighted (IDW) interpolation process to estimate values for the surrounding areas. Reclassification is then conducted into three classes: <2000 mm/year, 2000-3000 mm/year, and >3000 mm/year.

2) Soil Type Data
Soil type data was taken from the Ministry of Environment and Forestry, which provided, in the form of a shapefile, data depicting Indonesian Soil types. This data was then further processed by cropping using the area of the research location to generate a map of soil types of Singgahan and Montong Districts, Tuban Regency.

3) DEMNAS data
Slope data is acquired from DEMNAS data with a spatial resolution of 8 meters. A slope analysis process is conducted to determine the slope values, followed by a reclassification process into five classes: <8% (flat), 8-15% (ramps), 15-25% (wavy), 25-40% (slightly steep), and >40% (steep).

4) Scoring and Weighting: Each zone is marked out based on the administrative boundaries of the study locale and subsequently changed from raster to vector data for the overlay operation. The 'Weight' column represents the extent of each parameter's contribution to the determination of the Groundwater Recharge Area, whereas the 'Value' column in the table represents the score given to each classification according to the parameters set. The score indicates the degree of probability that the area is a recharge zone. This research employs a scoring procedure to ascertain recharge areas, wherein every parameter is given a score and weighted differently according to its significance. The different layers are then combined, and the area with the highest value is considered the recharge area. There are four maps as parameters, and they are from the Indonesian Ministry of Forestry's 2009 regulations, with some changes. The classification of parameters of groundwater recharge areas, together with the scores and weights employed in this research, is consistent with these regulations as stated in Table 1 (Menkeu RI, 2009). However, the parameters for rainfall and soil type have been changed to be in line with the research area.

Table 1. Scoring and Weighting of Parameters Determining Groundwater Recharge Areas.

Parameter	Weight	Classification	Value
Land Cover	30%	Residential areas, rice fields	1
		Field, moor	2
		Shrubs, meadows	3
		Production forests, plantations	4
		Dense forest	5
Rainfall	15%	<2000	1
		2000 - 3000	2
		>3000	3
Soil Type	35%	Grummosol	1
		Litosol, Mediterran	2
		Latosol, Kambisol, Mollisol, Gleisol	3
		Aluvial	4
		Regosol, Rezina	5
Slope	20%	>40%	1
		25 – 40%	2

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Parameter	Weight	Classification	Value
		15 – 25%	3
		8 – 15%	4
		<8%	5

The final weighting is the accumulated result of the weight percentage multiplied by the value of each parameter, which then produces a final value based on equation 5 below.

$$S_T = B_1\,(S_1) + B_2\,(S_2) + \cdots . + B_n\,(S_n) \qquad (5)$$

S_T = Total score
$B_1, B_2, \ldots, B_n$ = Weight of each parameter
$S_1, S_2, \ldots, S_n$ = Score for each parameter
n = Number of parameters

If all data processing has been completed, the result will be a map showing the distribution of groundwater recharge areas in Singgahan and Montong Districts, Tuban Regency.

3. Result and Analysis

Land Cover

Land cover data processing based on Landsat-8 imagery is conducted using the Google Earth Engine (GEE) platform employing the Random Forest algorithm (Arikan, 2023). Subsequently, an accuracy test is performed using an error matrix (confusion matrix) to determine the accuracy level resulting from the Landsat-8 image classification process using the Random Forest algorithm. Figure 4 below shows the land cover map of Singgahan and Montong sub-districts resulting from the classification of Landsat-8 images by the Random Forest (RF) algorithm. Meanwhile, the area and percentage distribution of land cover in the two sub-districts are presented in Table 2.

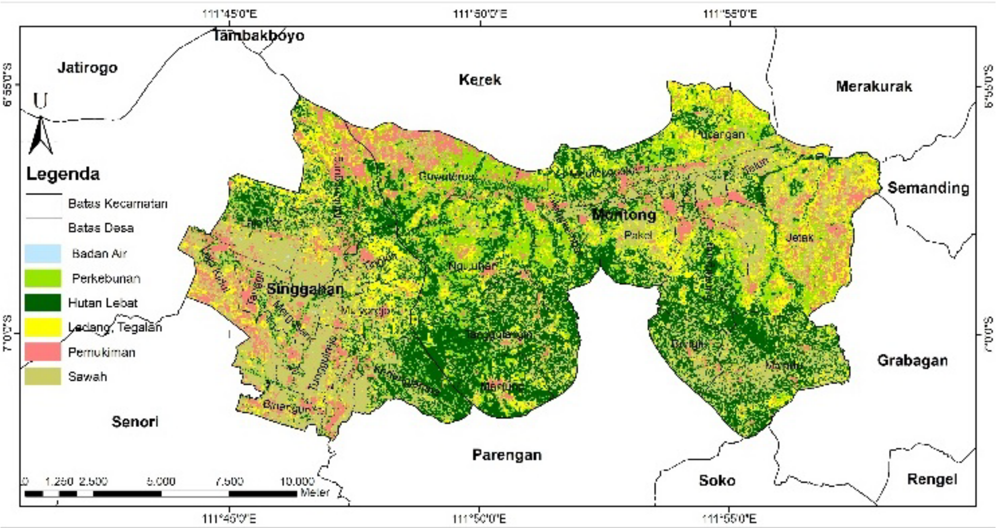


Figure 4. Land Cover Map Using the RF Algorithm for the Singgahan and Montong District Areas

Table 2 shows the area, percentage of land cover types, and the classification of values for each land cover type. A scoring range of 0 to 5 is used, depending on how well each land cover type can function as a recharge area or an area with high permeability. In this study, a score of 0 is assigned to the land cover parameter representing water bodies. This parameter is given a score of 0 because water bodies are areas where water cannot infiltrate the soil. Hence, a score of 0 in the context of this parameter is indicative of non-permeable soil in that area. Dense forest land cover is given a score of 5 because a dense forest is an excellent area for soil absorption. In this case, a plant root factor and plant root together with the soil in the forest area are loose so that the surface water can be absorbed well and retained underground. So, it is concluded that the dense forest area has great potential to become a recharge area. This entire classification is based on the Regulation of the Minister of Forestry of the Republic of Indonesia Number: P.32/MENHUT-II/2009. It was found that the land cover in Montong and Singgahan sub-districts is predominantly comprised of rice fields and dense forests, with percentages of 28.05% and 28.22%, respectively. The area covered by paddy fields is 6,159.37 hectares, while the area of dense forests is 6,196.65 hectares.

Table 2. Area and Land Cover Score Using the RF Algorithm in the Montong and Singgahan Districts

Land Cover	Area (Ha)	Percentage	Value
Water body	25,42	0,12%	0
Rice field	6159,37	28,05%	1
Settlement	2297,91	10,46%	2
Farm, Moor	3550,72	16,17%	3

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Plantation	3729,84	16,98%	4
Dense Forest	6196,65	28,22%	5

Accuracy testing of the land cover results was conducted by splitting the data into two sets: 70% of the data was used for training, and 30% was used for testing. Calculations were then performed using a confusion matrix, which was constructed to estimate the values for producer's accuracy, user's accuracy, and overall accuracy (Wiggers et al., 2020). By using the training data and Google Earth Engine software, the following confusion matrix results were obtained:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 188 & 0 & 0 & 0 & 0 \\ 0 & 0 & 111 & 0 & 0 & 0 \\ 0 & 1 & 1 & 159 & 1 & 2 \\ 0 & 0 & 0 & 0 & 257 & 4 \\ 0 & 0 & 0 & 0 & 0 & 531 \end{bmatrix}$$

Figure 5. Confusion Matrix for Land Cover Classification Using Random Forest Algorithm

Based on the confusion matrix results, calculations for overall accuracy and the Kappa value were performed to assess the accuracy of the land cover classification. The classification accuracy standards used are as follows: (1) Adequate if the OA value is 85-89%; (2) Moderate if the OA value is 90-94%; (3) High if the OA value is >95% (Zulfajri et al., 2021). If the algorithm achieves an overall accuracy (OA) value >85% and a Kappa value >0.80, it is considered to have high accuracy, and the data processing can proceed. The calculations yielded an Overall Accuracy (OA) value of 99.60% and a Kappa value of 0.994. The results of the classification accuracy test are presented in Table 3.

Table 3. Image Classification Accuracy Test Results with RF

Land Cover	User's Accuracy	Producer's Accuracy
Water body	1,000	1,000
Rice field	0,994	1,000
Settlement	1,000	1,000
Farm, Moor	0,994	1,000
Plantation	1,000	0,989
Dense Forest	0,994	0,996
Overall Accuracy	0,996	
Kappa	0,994	

Therefore, the land cover map resulting from image classification using the RF algorithm was selected as a parameter in determining groundwater recharge areas. The land cover parameter holds a weight of 30% with five classification classes as outlined in Table 2.

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Rainfall

Rainfall parameters were compiled based on the average rainfall data for 2021-2022 collected from 15 rainfed stations around the research area, sourced from BMKG. The rainfall parameters were classified into three classes: <2000 mm/year, 2000-3000 mm/year, and >3000 mm/year, with a weight of 15% to determine groundwater recharge areas.

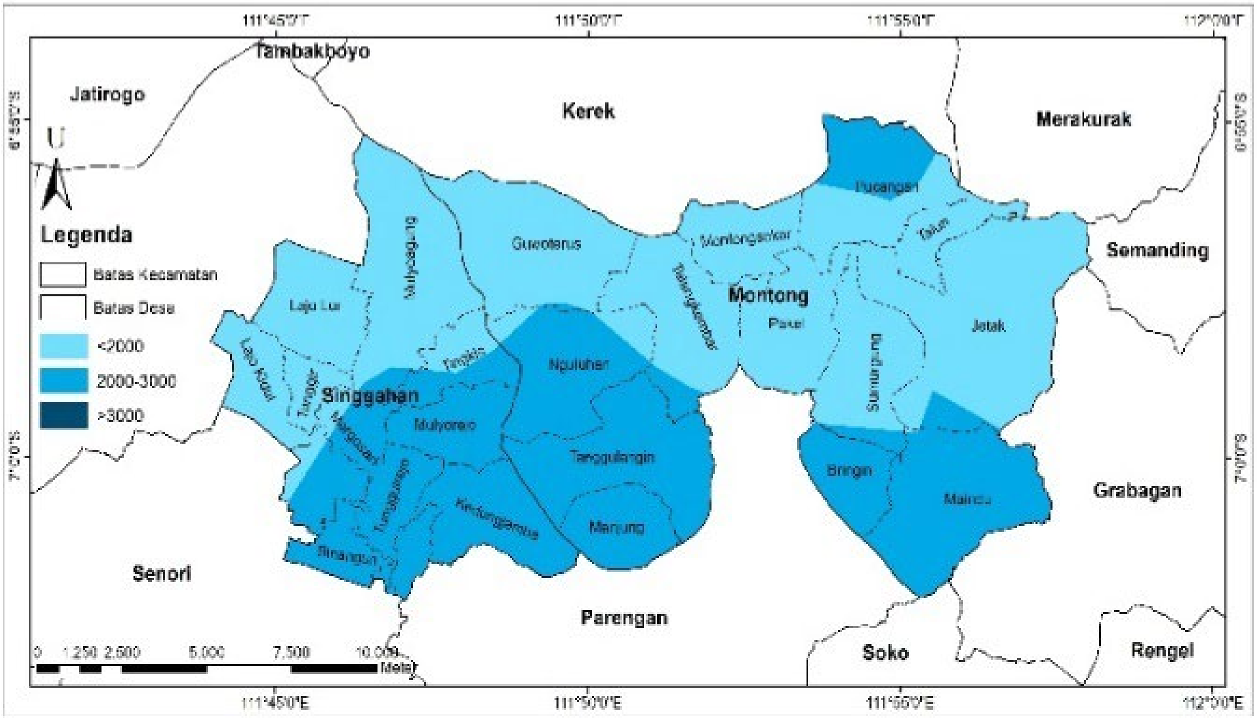


Figure 6. Land Cover Map Using the RF Algorithm for the Singgahan and Montong District Area

Based on Figure 6, it is evident that Singgahan and Montong sub-districts only have two bulk level classes spread across the two sub-districts. This is because the rainfall spread over the research area does not exceed 3000 mm/year. The area and percentage of rainfall data are presented in Table 4.

Table 4. Area and Rainfall Score for Singgahan and Montong Districts

Rainfall (mm/year)	Area (Ha)	Percentage	Value
<2000	12374,64	56,34%	1
2000-3000	9588,66	43,66%	2
>3000	0	0%	3

Soil Type

Parameters of soil type were sourced from a vector map of soil types, which was then overlain with the administrative boundaries of the research area, namely Singgahan and Montong Districts. Soil type is a decisive factor in locating recharge areas because it depends on the grain size of the soil. Bigger grain sizes widen the spaces between particles, thereby increasing porosity. Additionally, higher organic content in the soil can affect vegetation density in the area. Soil Research Institute in Bogor has classified various soil types in Indonesia, with one notable classification introduced by Dudal and Soepraptohardjo in 1957 (Fiantis et al., 2017), including the following soil groups: (1) Mediterranean soils, which are compact in texture and have low organic content, leading to reduced water absorption capacity; (2) Litosols, derived from weathered rock, with thin, rocky soil layers that are less capable of retaining water; (3) Gleysols, which possess a loose structure and high organic content, resulting in good porosity and making them effective absorbers; (4) Cambisols, which have a relatively fine texture and are moderately capable of functioning as areas for water absorption and storage.

The soil type map of Singgahan and Montong Districts is displayed in Figure 6 below.

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

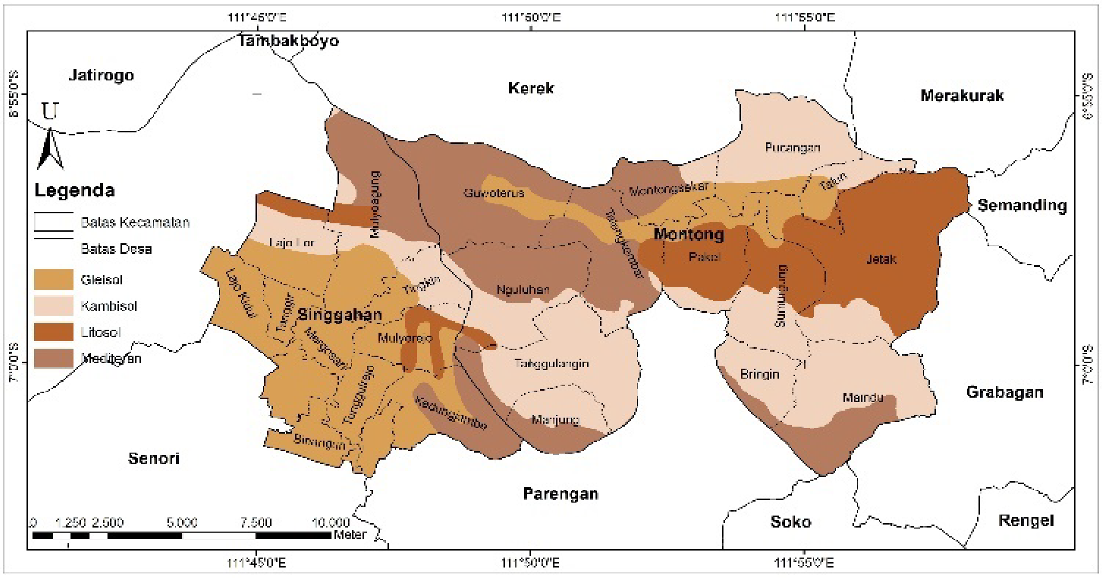


Figure 7. District Soil Type Map for Singgahan and Montong District Area

Figure 7 shows that Singgahan and Montong Districts consist of four types of soil spread across several villages, with the area and value of each type of soil outlined in Table 5 below.

Table 5. Area and Rainfall Score for Singgahan and Montong Districts

Soil Type	Area (Ha)	Percentage	Value
Litosol	3780,66	17,21%	2
Mediterranean	5561,97	25,32%	2
Cambisole	6952,93	31,66%	3
Gleisol	5667,74	25,81%	3

Soil type carries the maximum influence as a factor in identifying groundwater recharge zones. It accounts for 35% of the total impact. This is mainly because various soil types differ in their inherent ability to allow water to infiltrate.

Slope

The slope parameters are derived from slope analysis of DEMNAS data. Subsequently, the slope levels were reclassified into five classes: <8% (flat), 8-15% (ramps), 15-25% (wavy), 25-40% (slightly steep), and >40% (steep). The land slope map in the Singgahan and Montong Districts can be observed in Figure 7.

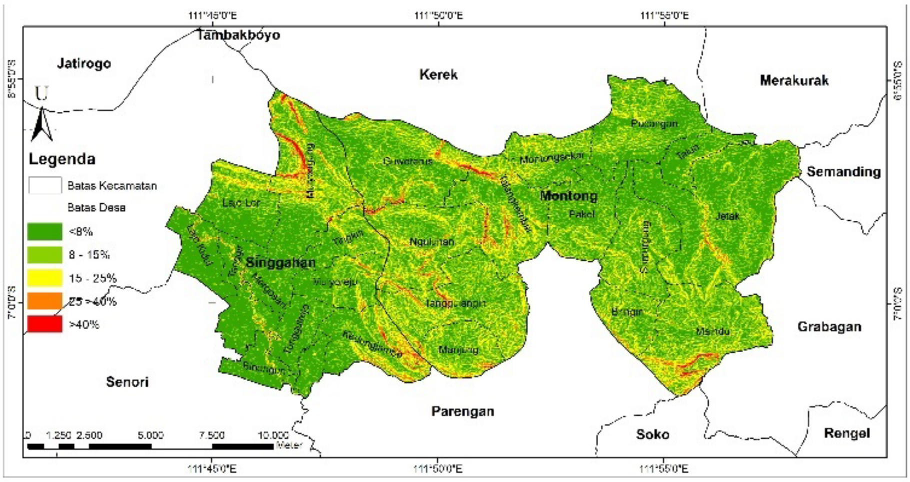


Figure 8. Slope Map of the Singgahan and Montong District Areas

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

From Figure 8, it is evident that most of the slopes in the Singgahan and Montong Districts are characterized by flat (<8%) and gentle slopes (8-15%). The slope parameter holds a weight of 20%. The area and percentage of slope in the Singgahan and Montong Districts are presented in Table 6 below.

Table 6. Area and Slope Score for Singgahan and Montong Districts

Slope	Area (Ha)	Percentage	Value
>40%	102,98	0,47%	1
25-40%	610,33	2,78%	2
15-25%	3039,51	13,84%	3
8-15%	7697,16	35,05%	4
<8%	10511,46	47,86%	5

Map of the Distribution of Groundwater Recharge Areas

The map depicting the distribution of groundwater recharge areas was generated through the overlay of the four determining parameters: land cover, rainfall, soil type, and slope. The overlay was made according to the scoring and weighting values as shown in Table 1. Subsequently, class determination was carried out according to research by Purwanto et al. (2022), where groundwater recharge areas were divided into three classes: Recharge Area, Transition Zone, and Discharge Area, as outlined in Table 7.

Table 7. Groundwater Recharge Area Classification Interval Value for Singgahan and Montong Districts

Classification of Groundwater Recharge Areas	Interval
Recharge Area	2,63 – 3,85
Transition Zone	1,42 – 2,63
Discharge Area	0,20 – 1,42

Following the process of determining the value intervals for each class, a map illustrating the groundwater recharge area was generated, as depicted in Figure 8. In the map, the recharge area is represented in green, the transition zone in yellow, and the discharge area in red.

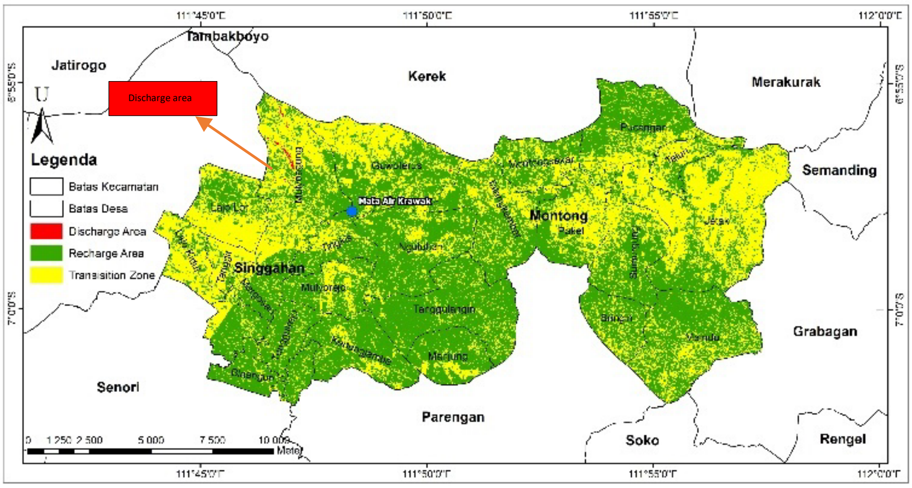


Figure 9. Regional Groundwater Recharge Area Distribution Map, Singgahan and Montong Districts

The area and percentage distribution of groundwater recharge areas are presented in Table 8. The largest percentage being 59.21%, covering an area of 13010.33 ha for the recharge area, followed by 40.66% covering 8935.27 ha for the transition zone, and 0.13% covering 28.64 ha for the discharge area (Northwest area).

Table 8. Groundwater Recharge Area Classification, Area, and Percentage for Singgahan and Montong Districts

Classification	Area(Ha)	Percentage
Recharge Area	13010,33	59,21%
Transition Zone	8935,27	40,66%
Discharge Area	28,64	0,13%

Based on the results of the map processing of groundwater recharge area distribution, it is clear that the research location, especially Singgahan and Montong Districts, has great potential to be used as groundwater recharge areas. This is due to the fact that the groundwater recharge parameters have a high weighting. Land cover parameters correlation figures show that high potential zones are generally located in areas of dense vegetation such as forests, plantations, and fields. This is because plants help the infiltration process so that raindrops can flow properly. Regarding the soil type, the infiltration rate is basically conditioned by the permeability of the soil. Soils containing a high percentage of clay have low permeability, which is the general rule. Hence, water will flow better in soils with less clay and more coarse textures like Cambisols and Gleysols. The slope parameter is the primary determinant of groundwater recharge areas since the slope can influence infiltration of water, where gentle slopes will allow more water to be absorbed and retained. Most of the slopes in the Singgahan District are flat, thus having the highest groundwater recharge potential. Besides the parameters mentioned above, the physical form of the research area is also

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

considered, which mainly comprises karst formations with high rock porosity or porous regions, leading to high permeability and more effective infiltration compared to other land types. Thus, the relatively low average rainfall over the past two decades does not significantly affect the recharge potential. Validation beyond overlaying the four parameters is necessary. Currently, there are no published maps of groundwater recharge areas from relevant ministries or local government authorities, making the final validation of the created recharge area map challenging.

The research results, when compared to the Groundwater Basin map of Indonesia (MyPatriot, 2017) in Figure 10, reveal that the Singgahan and Montong sub-districts are located within the groundwater basin area (marked by red dots in the orange region). Although the groundwater basin data are presented on a regional scale, these data can serve as supporting evidence, which indicates that the research findings are sufficiently accurate. The accuracy was supported by the fact that this study identifies Montong and Singgahan as areas with potential to become recharge zones with more specific results. This delineates the level of potential, including recharge areas (green), transition zones (yellow), and discharge areas (red) (Figure 9).

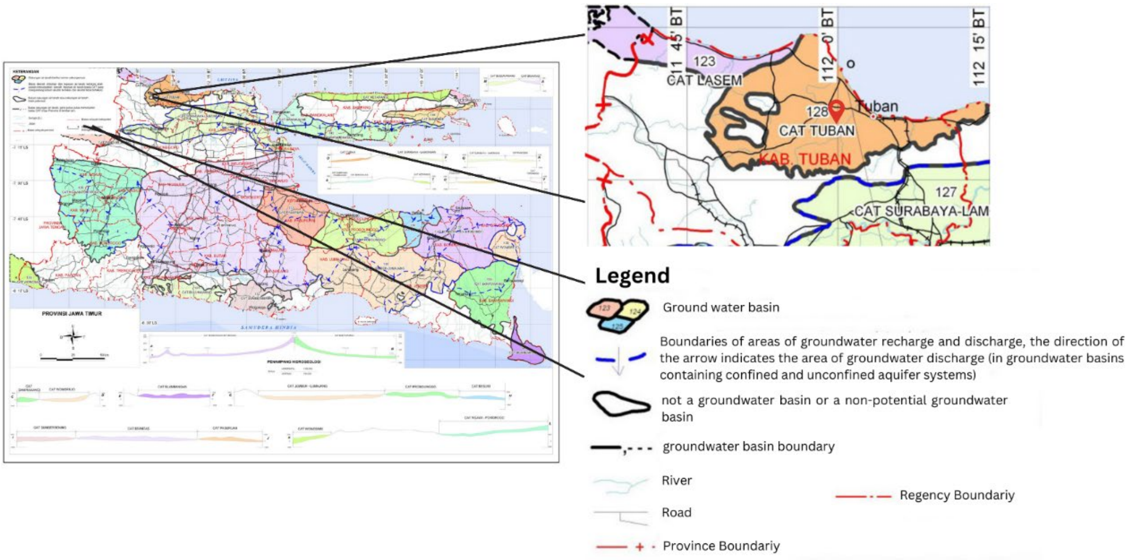
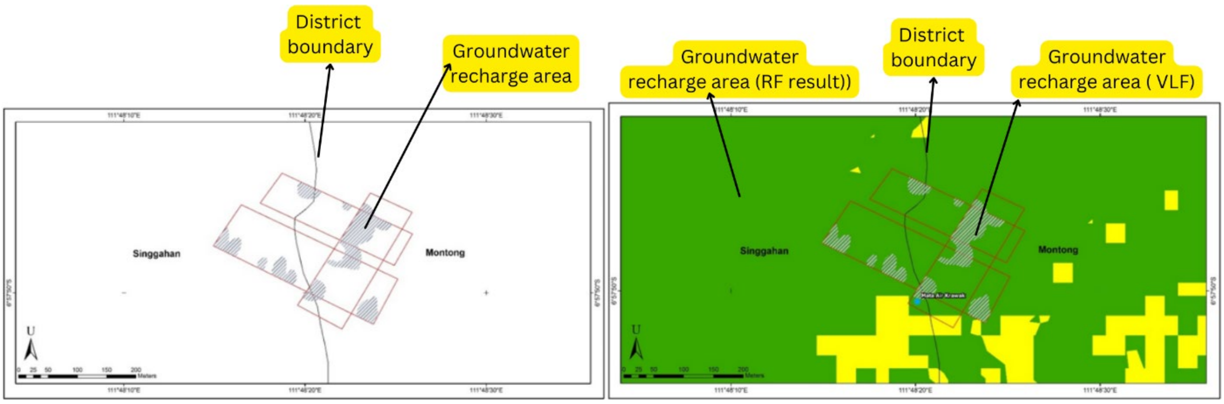


Figure 10. Map of Groundwater Basin in Indonesia

The validation was conducted by comparing the satellite image processing results with data from research using the VLF geophysical method (Purwanto et al., 2024). Figure 11a shows the distribution map of groundwater recharge areas in the Krawak spring region, Singgahan sub-district, Tuban Regency. The satellite image and VLF results were combined (Figure 11b). The results indicate that the green areas represent groundwater recharge zones, while the hatched areas (gray) denote recharge zones based on the VLF geophysical method. The results of this study are validated; however, the limited scope of the previous research resulted in the comparison data being less optimal, which is a limitation of this study. Under these circumstances, it is expected that future research will cover a wider area, allowing for a more accurate and comprehensive mapping of recharge areas in the Singgahan and Montong sub-districts. This will provide a valuable reference for well placement and drought mitigation efforts in the future.



(a) (b)

Figure 11. (a) Map of the Distribution of Groundwater Recharge Areas Using VLF; (b) Overlay map of VLF results and satellite imagery (Modified).

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

4. Conclusion

Based on the results obtained, it was observed that Landsat-8 image classification for land cover maps can be determined using several algorithms. The Random Forest algorithm was employed and it is found that it exhibits the highest Overall Accuracy (OA) value of 99.75% and a Kappa value of 0.99. Subsequently, to delineate the groundwater recharge area, four parameters are considered, including land cover, rainfall, soil type, and slope. Scoring and weighting are conducted, resulting in a recharge area of 13010.33 ha, a transition zone of 8935.27 ha, and a discharge area of 28.64 ha. From the processing results, it can be concluded that the research location, namely Singgahan and Montong Districts, harbors a significant potential to serve as a source of groundwater recharge area. The findings of this research can be beneficial for the government and local communities in preserving and upholding the primary function of groundwater recharge areas. Based on the overall results, the author has several recommendations. First, the study should be followed up by incorporating additional parameters, such as the type of rock in the research area, since rock type is closely related to the water infiltration process. Moreover, the threshold values for the classification of potential groundwater recharge areas need to be derived from an official standard or reference.

5. Acknowledgements

Thankful to God Almighty, whose mercy and grace were abundantly granted through the duration of the research to be completed without any hindrance. A heartfelt thanks to parents whose unending prayers and support have been the source of strength in this journey. Besides that, the author highly appreciates Prof. Dr. Ir. Adi Susilo, M.Si., Ph.D., Prof. Dr. Eng. Agus Naba, S.Sc., M.T., and Dr. Ayi Syaeful Bahri, S.Si., M.T., as Promoter and Co-Promoters, for their consistent guidance, support, and pearls of wisdom throughout the research work. Moreover, the author would like to extend gratitude to Mr. Ari Sanggam and Mrs. Dr. Ing. Noorlaila Hayati, S.T., M.T., for their help and contributions towards the research to a great extent.

6. References

Adugna T., Xu W., & Fan, J. (2022). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sensing*, *14*(3), 1–22. https://doi.org/10.3390/rs14030574

Arikan D. (2023). Land cover generation using random forest algorithm on Google Earth Engine platform. Proceedings of the International Turkic World Congress On Science and Engineering, Bishkek Kyrgyztan.

Badan Meteorologi, Klimatologi, dan Geofisika. (2021). Badan Meteorologi, Klimatologi, dan Geofisika. Retrieved from https://www.bmkg.go.id/

Forestier H. (2007). Ribuan Gunung, Ribuan Alat Batu: Prasejarah Song Keplek, Gunung Sewu, Jawa Timur (No. 7). Kepustakaan Populer Gramedia.

Lachaud A., Adam M., & Mišković I. (2023). Comparative Study of Random Forest and Support Vector Machine Algorithms in Mineral Prospectivity Mapping with Limited Training Data. Minerals, 13(8). https://doi.org/10.3390/min13081073

Melillos G., & Hadjimitsis D. (2020). Using simple ratio (SR) vegetation index to detect deep man-made infrastructures in Cyprus. https://doi.org/10.1117/12.2557893

Mypatriot.id. (2017). Peta cekungan air tanah Indonesia tahun 2017. Retrieved from https://mypatriot.id/patgtl/datapubliclist

Peraturan Menteri Kehutanan RI. (2009). P.32/MENHUT-II/2009 tentang Tata Cara Penyusunan Teknik Rehabilitasi Hutan dan Lahan Daerah Aliran Sungai (RTkRHLDAS). Jakarta: Menteri Kehutanan RI.

Purwanto M. S., Haq, Z. M., Fachyesi, A. S., Dewi, S., & Amalina, S. (2022). Penentuan recharge area pada Kabupaten Tanah Datar menggunakan citra Landsat-8 dan sistem informasi geografis (SIG). Journal Geosaintek, 8(3), 242–248.

Purwanto M. S., Susilo, A., Bahri, A. S., Naba, A., Sari, U. I., & Almais, A. T. W. (2024). Mapping underground river flows in karst areas with the VLF-EM method (Case study of the Krawak Region, Singgahan Tuban). IOP Conference Series: Earth and Environmental Science, 1307(1), Article 012005. https://doi.org/10.1088/1755-1315/1307/1/012005

Ramji Satyaji Rao Y. (2022). Mapping and Change Detection of Water Bodies in the Godavari Delta using Geospatial Technology. *ADBU-Journal of Engineering Technology*, *11*(1), 1–6. www.earthexplorer.usgs.gov

Shin N., Saitoh T. M., Takeuchi Y., Miura T., Aiba M., Kurokawa H., Onoda Y., Ichii K., Nasahara K. N., Suzuki R., Nakashizuka T., & Muraoka H. (2022). Review: Monitoring of land cover changes and plant phenology by remote-sensing in East Asia. *Ecological Research*, *38*(1), 111–133. https://doi.org/10.1111/1440-1703.12371

Fiantis D. (2017). Morfologi dan Klasifikasi Tanah. Padang. Lembaga Pengembangan Teknologi Informasi dan Komunikasi (LPTIK) Universitas Andalas.

Undang-undang Nomor 7 Tahun 2004 tentang Sumber Daya Air. Jakarta, Indonesia: Pemerintah Pusat.

Vasileva T. (2019). An assessment of potential groundwater recharge zones in Bulgaria. *Geologica Balcanica*, *48*(1), 43–61. https://doi.org/10.52321/geolbalc.48.1.43

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)

Wang S., & Fu G. (2023). Modelling soil moisture using climate data and normalized difference vegetation index based on nine algorithms in alpine grasslands. *Frontiers in Environmental Science*, *11*(February), 1–12. https://doi.org/10.3389/fenvs.2023.1130448

Wiggers M. J., Nuarsa I. W., & Putra I. N. (2020). Monitoring perubahan penggunaan lahan pesisir di Kecamatan Batu Layar, Kabupaten Lombok Barat pada tahun 2002 dan 2019. Journal of Marine Research and Technology, 3(2), 68-74.

Zulfajri, Danoedoro P., & Murti S. H. (2021). Klasifikasi tutupan lahan data Landsat-8 Oli menggunakan metode Random Forest. Journal Penginderaan Jauh Indonesia, 03(01), 1–7. http://journal.mapin.or.id/index.php/jpji/issue/archive

DOI:https//doi.org/10.22452/mjs.vol45no2.5
Malaysian Journal of Science 45(2): 42-54 (June 2026)