

Machine Learning-Based Rice Field Mapping in Kulon Progo using a Fusion of Multispectral and SAR Imageries

Yusri Khoirurrizqi ^{1,*}, Rohmad Sasongko ¹, Nur Laila Eka Utami ¹, Amanda Irbah ¹, Sanjiwana Arjasakusuma ^{1,2}

¹ Faculty of Geography, Universitas Gadjah Mada, Sekip Utara, Bulaksumur, Yogyakarta 55281

² Pusat Studi Lingkungan Hidup, Universitas Gadjah Mada, Kompleks Gedung PSLH-EFSD UGM, Jl. Kuningan, Caturtunggal, Depok, Sleman, Yogyakarta 55281

* Correspondence: y.khoirurrizqi@mail.ugm.ac.id

Citation:

Khoirurrizqi, Y., Sasongko, R. Utami, N.L.E., Irbah, A., & Arjasakusuma, S. (2023). Machine Learning-Based Rice Field Mapping in Kulon Progo using a Fusion of Multispectral and SAR Imageries. *Forum Geografi*, 37(2), 134-148.

Article history:

Received: 21 November 2022

Accepted: 13 October 2023

Published: 27 December 2023

Abstract

The land-conversion of rice fields can reduce rice production and negatively impact food security. Consequently, monitoring is essential to prevent the loss of productive agricultural land. This study uses a combination of Sentinel-2 MSI, Sentinel-1 SAR, along with SRTM (elevation and slope data) to monitor rice fields land-conversion. NDVI, NDBI and NDWI indices are transformed from the annual median composite Sentinel-2 MSI images used to identify different rice fields with another object. A monthly median composite of SAR images from Sentinel-1 data are used to identify cropping patterns of rice fields in the inundation phase. The classification is performed by using the Random Forest machine learning algorithm in the Google Earth Engine (GEE) platform. Random Forest classification is run using 1000 trees, with a 70:30 ratio of training and testing data from sample features extracted by visual interpretation of high-resolution Google Earth imagery. In this study, Random Forest classification is effective in computing a high amount of multi-temporal and multi-sensory data to map rice-field land conversion with an accuracy rate of 96.16% (2021) and 95.95% (2017) for mapping paddy fields. From the multitemporal rice field maps in 2017–2021, a conversion of 826.66 hectares of rice-fields to non-rice fields was identified. Based on the spatial distribution, the conversion from rice-field to non-rice field is higher at the area near the roads, built area and Yogyakarta International Airport. Therefore, it is important to assess and ensure that National Strategic Projects are managed with due regard to environmental impacts and food security.

Keywords: rice field, land conversion, remote sensing, multi-censor, machine learning.

1. Introduction

The conversion of rice fields into other land uses in Indonesia can reduce rice production and negatively impact food security (Vinet *et al.*, 2020). The possible conversion of agricultural land in the Kulon Progo area is estimated to be 300 hectares due to airport megaprojects, industrial estates and settlements (Stathopoulos *et al.*, 2023). To prevent the loss of productive agricultural land to other uses, it is essential to have accurate information on the dynamics of rice fields (Bahagia *et al.*, 2020).

The conversion of rice fields can be monitored using multi-temporal remote sensing data. Monitoring using time series remote sensing data in the same area can illustrate the dynamics of agricultural land change more accurately, for example the study mapping the intensity of crop yields by (Parelius, 2023) and the detection of changes in agricultural land (Hossain *et al.*, 2023). Using optical remote sensing data in the tropics presents several challenges due to the unique characteristics of this region. Remote sensing data in the tropics is regularly affected by extensive cloud cover, making it challenging to obtain cloud-free optical remote sensing data (Jia *et al.*, 2014). Clouds can limit the availability of cloud-free imagery, affecting the frequency and consistency of data collection. The tropical regions are located close to the Equator, resulting in a relatively high sun angle throughout the year. The intense illumination can generate strong reflectance and glint effects, making it challenging to capture consistent and accurate optical measurements. Additionally, the presence of tall vegetation can cast long shadows, which can complicate the interpretation of remotely sensed data. Therefore, Synthetic Aperture Radar (SAR) remote sensing data can be an alternative to provide cloud-free data (Blickensdörfer *et al.*, 2022). SAR data also can acquire data both during the day and at night; hence, it can facilitate round-the-clock monitoring and enhance the temporal coverage of remote sensing observations. Likewise, SAR sensors mounted on satellite platforms typically have a wide swath width, allowing for large-scale coverage of areas in a single acquisition (Kilbride *et al.*, 2023).

Image Synthetic Aperture Radar (SAR) is a remote sensing technology that adopts the working principle of radar which uses microwave waves (0.75–130 cm) to map the Earth's surface (Yusrina *et al.*, 2019). SAR imagery is a remote sensing system that employs microwaves with the advantage of being able to produce images that are not affected by weather conditions. It has the ability to can penetrate clouds and can work at night (Yusrina *et al.*, 2019). SAR images are used in various tasks, such as segmentation, classification, target recognition, denoising, change



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

detection, etc. The change detection task relating to SAR images is to determine the change information of two SAR images captured in the same area but at different times. Change detection is applied in disaster monitoring, the supervision of a country's resources and changed target detection (He *et al.*, 2023). The radar signal exhibits a smaller scattering intensity on water objects, making it more sensitive to detecting water content and it is suitable for mapping rice fields that have an inundation phase (Arjasakusuma *et al.*, 2020). Including SAR data in a classification model can increase the accuracy of crop mapping due to the increased availability of cloud cover independent data. Moreover, the physical and structural properties of SAR signals above the crop canopy will complement the optical information from multispectral sensors (Blickensdörfer *et al.*, 2022). The combination of data derived from multisensory and multitemporal imagery will better explain the dynamics of changes in productive land. Nonetheless, the integration of multi-sensor remote sensing data between radar and optical requires a method that is capable of processing data with different value distributions (Hadibasyir *et al.*, 2020).

Machine learning is an effective method for processing data with complexities (different value distributions) and large spatial dimensions, for instance multi-sensor remote sensing data (Maxwell *et al.*, 2018). Machine learning is defined as a field of study that presents computers with the ability to learn without being explicitly programmed. Machine learning (ML) is used to teach machines how to handle data more efficiently. Occasionally, after viewing the data, we cannot interpret the extract information obtained from the data. In that case, we apply machine learning. Machine learning relies on different algorithms to solve data problems (Mahesh, 2020) and has been used in agriculture for several years. The crop prediction process was completed by consolidating the preceding data and the present data belonging to a particular month to verify the accuracy of climatic data. Therefore, remote sensing data processing using machine learning could be of immense value as a food crop prediction method (Triscowati *et al.*, 2020).

Currently, the combination of remote sensing and machine learning has only been applied to identify rice cropping patterns using optical sensor data (Rafif *et al.*, 2021). However, the identification changes in rice fields land-conversion has not been conducted. In India, with the support of machine learning, a smart, reliable yet low-cost novel solution has been developed which can automate several procedures from the watering of crops to predicting the health of crops. With the assistance of machine learning, this system is also capable of predicting the conditions and the type of crop that is appropriate for a particular type of soil and environment (Stathopoulos *et al.*, 2023).

In this study, we implemented the supervised random forest method to conduct multi-temporal and multi-sensor analysis with the aim of classifying rice field and non-rice field objects. When rice field classification analysis is carried out on all the pixels in the images, the amount of data is large and unstructured (Wu *et al.*, 2023). Concerning this type of data, machine learning is recommended, as it will be complicated to devise models manually. Random forest is a machine learning method that is not sensitive to multicollinearity (Zollini *et al.*, 2023). By using random forest, temporal features engineering can be obtained as far as possible, from which the best features are selected using the variable importance plotting function (Triscowati *et al.*, 2020). Random forest was chosen for rice field classification because it can handle large amounts of data, is not sensitive to multicollinearity and therefore enables the modelling process to include as many variables as possible, and as it is fast and not sensitive to overfitting. Furthermore, it can select any explanatory variables that affect the predictions, using the varimplot function (Xu *et al.*, 2023).

In this study, we focus on classifying rice field and non-rice field objects in Kulon Progo Regency, Yogyakarta Special Region, Indonesia, as a case study. To achieve this aim, we used the supervised random forest method to analyse multi-temporal and multi-sensor data and classify rice field and non-rice field objects in Kulon Progo Regency, Yogyakarta Special Region, Indonesia (Fu *et al.*, 2023). This study aims to identify rice field agricultural land based on rice planting patterns in Kulon Progo from 2017 to 2021. This analysis is performed using machine learning by way of utilising a combination of multi-temporal radar and optical satellite imagery.

2. Research Methods

2.1. Study Area

The study occurred in Kulon Progo Regency, Yogyakarta Special Region, Indonesia, at location 7°38'42" to 7°59'3" N and 110°1'37" to 110°16'26" E (Figure 1). The data source employed to produce the map of the study area is retrieved from Sentinel-2 and Google Satellite Imagery. The study area is dominated by lowland coastal plains in the south and denudational mountains in the north. This region was dominated by urban, cropland and rice field land cover. Kulon Progo

supplies rice for the Yogyakarta Special Region, with rice productivity of approximately 120,000 tons per year in 2020. Kulon Progo was chosen as the study area because a massive land conversion is taking place owing to the Yogyakarta International Airport (YIA) megaproject. The new construction project began at the end of 2017. Based on this, this study will also conduct a distance sensitivity analysis of the presence of new airports, toll roads and built-up areas to ascertain the effect on rice production in Kulon Progo. The distribution of airports, roads and built-up areas in Kulon Progo Regency is shown in Figure 2.

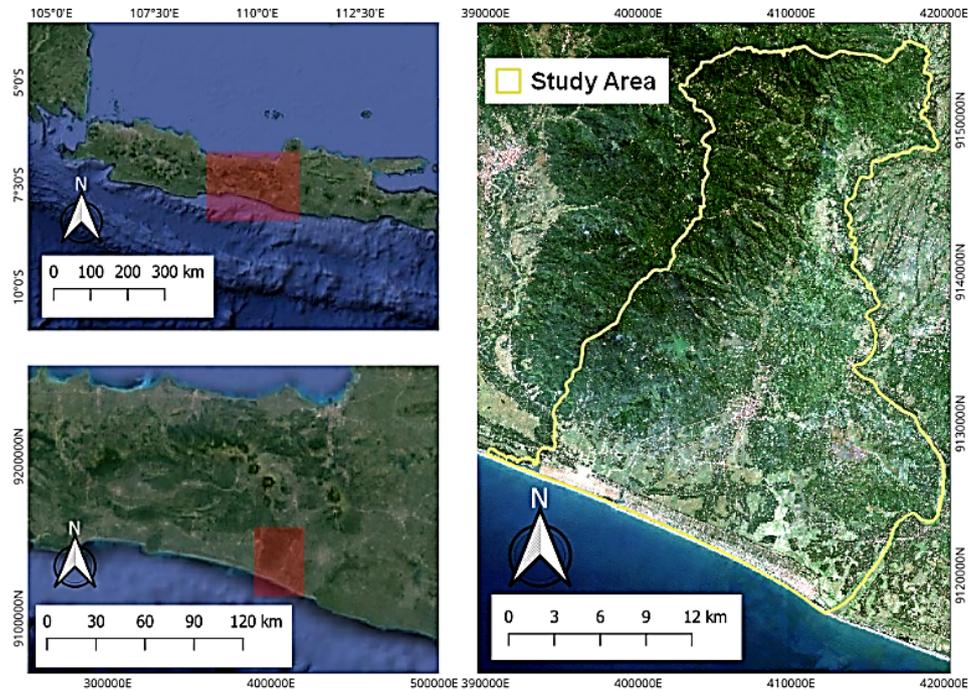


Figure 1. The Study Area Map (source: Sentinel-2 and Google Satellite Imagery).

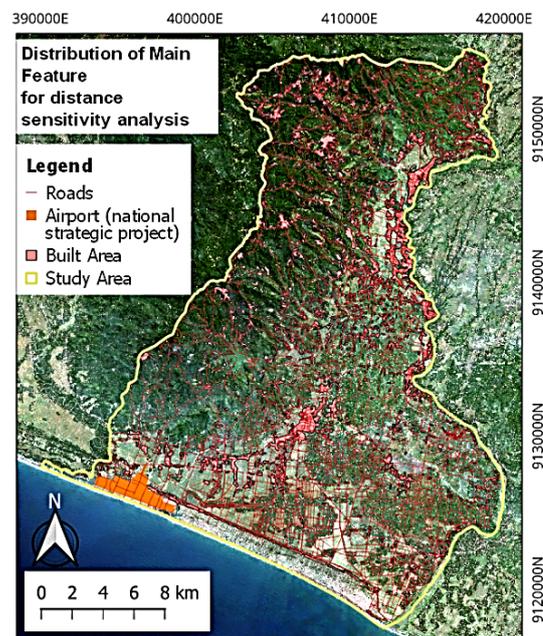


Figure 2. The Distribution of the Main Feature for Distance Sensitivity Analysis Map.

2.2. Data acquisition and pre-processing

This study uses Harmonised Sentinel-2 MSI Top-of-Atmosphere (Figure 3), Sentinel-1 GRD C-Band, in conjunction with NASA SRTM Digital Elevation provided (Figure 4) by Google Earth Engine Datasets. After 2022-01-25, Sentinel-2 scenes with PROCESSING_BASELINE '04.00' or above have their DN (value) range shifted by 1000. The HARMONISED collection modifies data in newer scenes to align with the range in the older scenes. Subsequently, Sentinel-2 data is

pre-filtered to acquire a cloud-free image. Cloud-free images are obtained by attaching the ee.Filter.it tool to filter the images used where the property value applied is 'CLOUDY_PIXEL_PERCENTAGE' less than 20%. In particular, only images with a cloudy pixel percentage of less than 20% are retained. Then, if within the specified timeframe there are no images with cloud conditions of less than 20%, masking will be conducted to produce images with cloud conditions of less than 20%. The part of the image that is masked is then median composite to patch the part of the image that is masked because of the cloud filter. The one-year time span for the median composition was chosen because of the high cloud cover; thus, eliminating it requires multi-temporal imagery over a prolonged period. Sentinel-2 is then calculated for their indices which includes the Enhanced Vegetation Index (EVI), Normalised Difference Vegetation Index (NDVI), Normalised Difference Water Index (NDWI) and Land Surface Water Index (LSWI). Those indices are utilised to differentiate rice fields with other objects, for example built area, high vegetation and bare soil (Tong *et al.*, 2023).

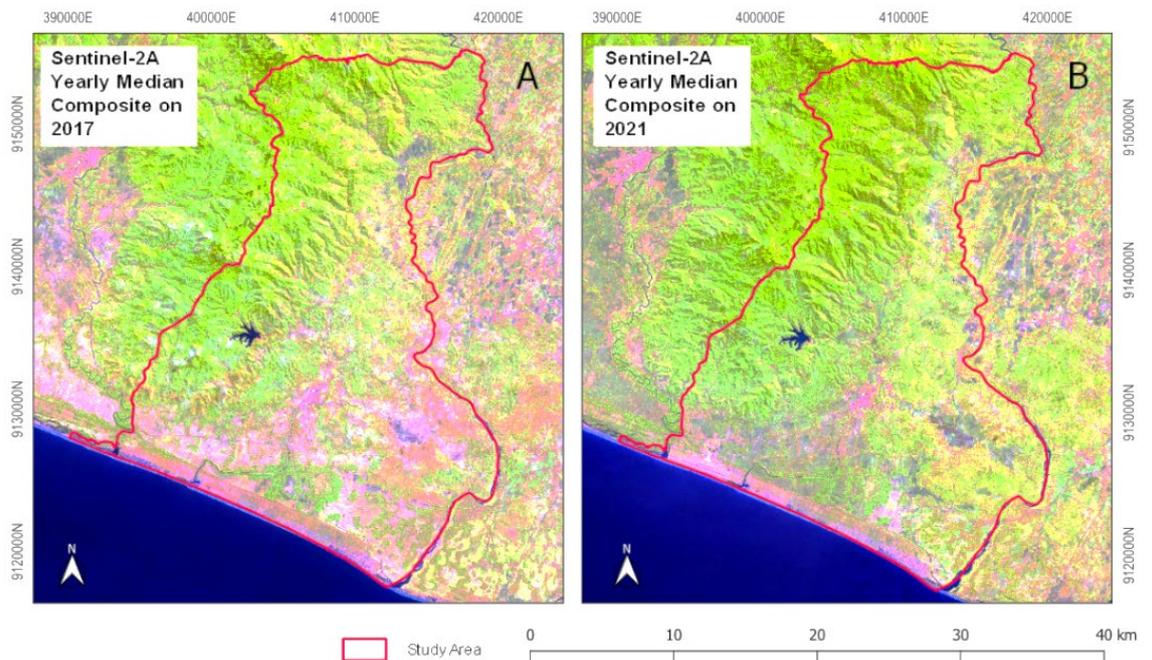


Figure 3. Sentinel-2 Annual Median Composite Image On 2017 (A) and 2021 (B) with RGB colour composite SWIR1, NIR, Blue (11, 8, 3).

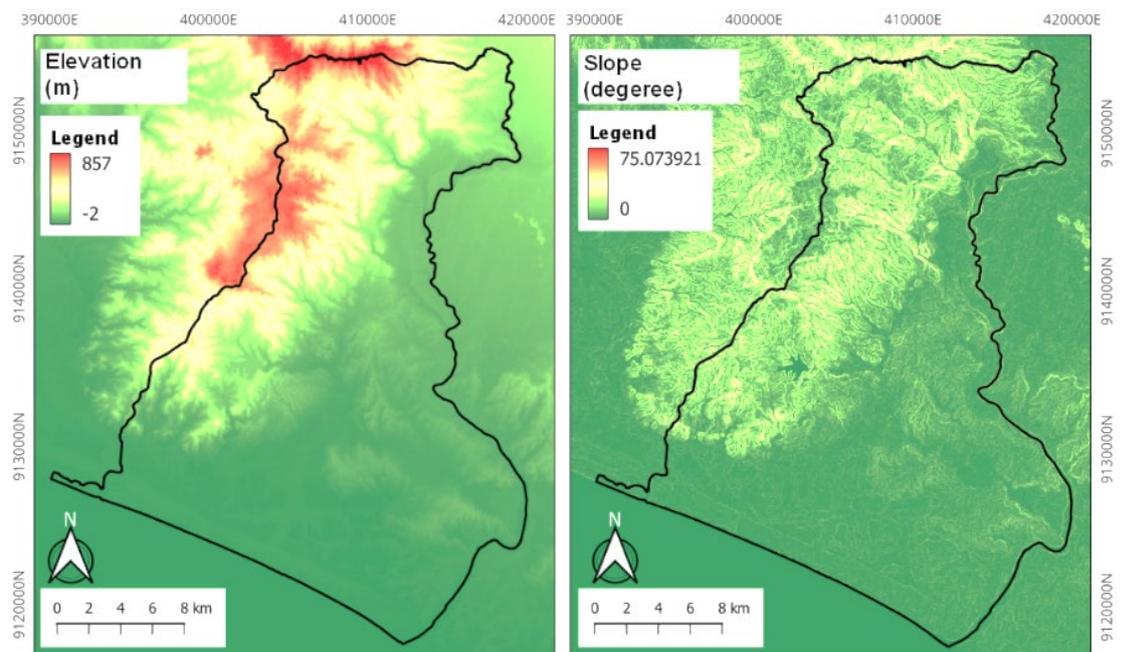


Figure 4. Elevation and slope data of study area obtained from SRTM DEM On 2000.

Sentinel-1 data is used to create monthly composite medians and smoothing filtering is undertaken to remove noise generated from radar images (Becker *et al.*, 2023). Smoothing filtering is performed using the FocalMean tool in Google Earth Engine. There are several parameters used in the tool to perform smooth filtering on the Sentinel-1 image (Tripathi *et al.*, 2023). These parameters include the radius, kernel type and units used. The radius utilised for smoothing filtering is 20 metres, the type of kernel used is a circle, whilst the unit applied is metres. Sentinel-1 data is divided into two polarisations, namely VV (vertical-vertical) and VH (vertical-horizontal). VV and VH data were each undertaken to produce monthly composite medians.

Additionally, Digital Elevation Model data is also utilised to obtain the elevation and slope. The data collected is then used as input to classify paddy and non-paddy fields (Onojeghuo *et al.*, 2023). This stage is then carried out on data for 2017 and 2021. The DEM data employed is NASA SRTM Digital Elevation with a spatial resolution of 30m. SRTM data is then developed into elevation data by selecting the SRTM dataset to be simply elevation data (Tzepkenlis *et al.*, 2023). Subsequently, the elevation data is used to build a slope model using the ee.Terrain.slope tool.

2.4. Accuracy Test

The accuracy test involves using a confusion matrix to create producer accuracy, along with user accuracy, overall accuracy and the Kappa value to assess the accuracy of the information on the map (Movchan *et al.*, 2023). The confusion matrix is developed using the errorMatrix tool in Google Earth Engine. ErrorMatrix is developed by calculating a 2D error matrix to compare two data, where one data comprises the actual values (sample dataset) and the other data contains the predicted values (classification results). The x-axis includes the actual data or correct sample datasets, while the y-axis contains the predicted values (classification results). The following is equation 1, 2 and 3 used to calculate the overall accuracy value and Kappa value:

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \tag{1}$$

$$\text{Expected Accuracy} = \frac{(TP + FP) \times (TP + FN) \times (FN + TN) \times (FP + TN)}{(TP + TN + FN + FP)^2} \tag{2}$$

$$\text{Kappa} = \frac{\text{Overall Accuracy} - \text{Expected Accuracy}}{1 - \text{Expected Accuracy}} \tag{3}$$

Where TP is the number of rice pixels that are actually classified as rice pixels, TN is the number of non-rice pixels that are basically classified as non-rice pixels, FP is the number of non-rice pixels that are incorrectly classified as rice and FN is the number of rice pixels classified as non-rice pixels.

The testing data concerning model performance measurement is obtained from 30% of the randomly selected features from visual interpretation of Google Map high resolution satellite view. The optimal percentage for data training and data testing in machine learning depends on various factors, such as the size of the dataset, the complexity of the model, besides many other factors. Therefore, there is no single optimal percentage pertaining to all types of problems (Pham *et al.*, 2023). However, a common ratio that is repeatedly exploited for data training and data testing is 70:30 or 80:20. This ratio has strong empirical evidence in machine learning literature and is frequently used as the default ratio for various types of problems. Several studies have verified that the 70:30 ratio performs well as regards classification problems with small to medium-sized datasets. The final classification result is validated using rice-field reference data from the Director General of Agricultural Infrastructure and Facilities, Ministry of Agriculture of the Republic of Indonesia. The reference data applied is first confirmed by means of visual interpretation to ensure that it matches the field data.

2.5. Post Classification and Data Analysis

Pixel-based classification created the ‘salt and pepper’ effect in the classification result due to distribution of an isolated pixel. Consequently, sieve tools are used after classification to fix this issue (Brown *et al.*, 2023). The salt and pepper effect results in the image containing numerous spots owing to the pixel-based classification that is undertaken (Tzepkenlis *et al.*, 2023). Therefore, it is necessary to perform a filter to reduce the effects of salt and pepper. One filter that can be used to reduce the effects of salt and pepper is the sieving tool found in the QuantumGIS software (Vinet *et al.*, 2020). The sieve tool works by removing raster polygons that are smaller than a predefined threshold and replacing them with the values of the largest neighbouring polygons. This is beneficial in minimising the effects of salt and pepper.

After minimising the effects of salt and pepper, data analysis was then performed using post-classification comparative change detection. This analysis is carried out per pixel through a classification scheme that uses two identical images at different times (Chen *et al.*, 2023). This analysis is used to observe changes in the use of paddy fields that occur from 2017 to 2021. There are four classes resulting from this analysis including no-change (rice-field), gain rice-field, loss rice field, as well as no-change (non-rice-field). No-change (rice-field) denotes that paddy fields have not changed, gain rice-field means that there is a conversion from non-paddy fields to paddy fields, loss rice fields signifies that there is a conversion from paddy fields to non-paddy fields, whereas no-change (non-rice-field) represents non-paddy fields that have not undergone conversion. This process is completed using the Semi-Automatic Classification Plugin in the QGIS software. The purpose of this analysis is to comprehend the magnitude of changes from paddy fields to non-paddy fields and the distribution of these changes. A flowchart that describes the stages of the work in this study is presented in Figure 5.

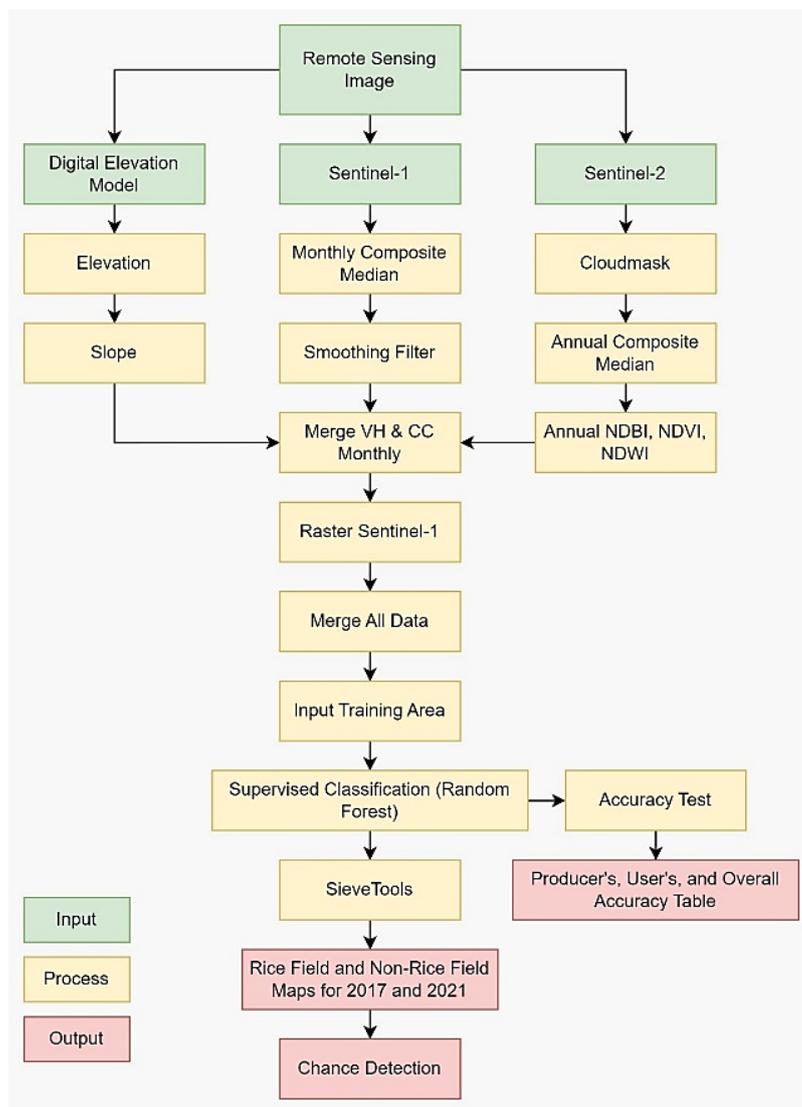


Figure 5. Flow Diagram of Research Method.

3. Results and Discussion

3.1. Indices Transformation

Normalised Difference Vegetation Index (NDVI), Normalised Difference Build-Up Index (NDBI) and Normalised Difference Water Index (NDWI) were used in this study to obtain temporal data used to identify irrigated areas as rice fields. NDVI is exploited to improve the appearance of rice fields in imagery by way of a comparison of vegetation greenness levels as an indicator of biomass in the form of rice in paddy fields. Nevertheless, the NDBI will emphasise the appearance of a building on remote sensing imagery enabling it to be distinguished between other use covers and building land use which can facilitate the sampling of mapping paddy fields (Giofandi, 2020). Similarly, NDWI is utilised to increase the appearance of water enabling rice

fields that are in the inundation phase to be detected (Hidayati *et al.*, 2018). Figure 6 illustrates the result of the NDVI, NDBI and NDWI image transformation on both 2017 and 2021.

Figure 6 illustrates that quite considerable changes in the use of rice fields have occurred in the area around the airport because of the shift in land use from previously being rice fields to an airport. Moreover, on account of the construction of access infrastructure to the airport in the form of roads in the area around the airport, several rice fields have been lost. The high and low pixel values of an image can be distinguished by their hue, the high pixel values have a bright hue while low pixel values have a darker hue. In the results of the NDBI image, the brightness level in the airport area is higher in 2021 than in 2017. As a result of changes in land use in the airport area, it can also be understood from the NDWI results that, in respect of the airport area, the soil moisture level has decreased. In the image, it is noticeable that several areas that were previously paddy fields had a high level of brightness in 2017, subsequently becoming darker in 2021. Regarding the results of this image transformation, it is apparent that the distribution of paddy field use is in gently sloping areas, south and southwest of the Menoreh Hills. No paddy fields were located in the Menoreh Hills area.

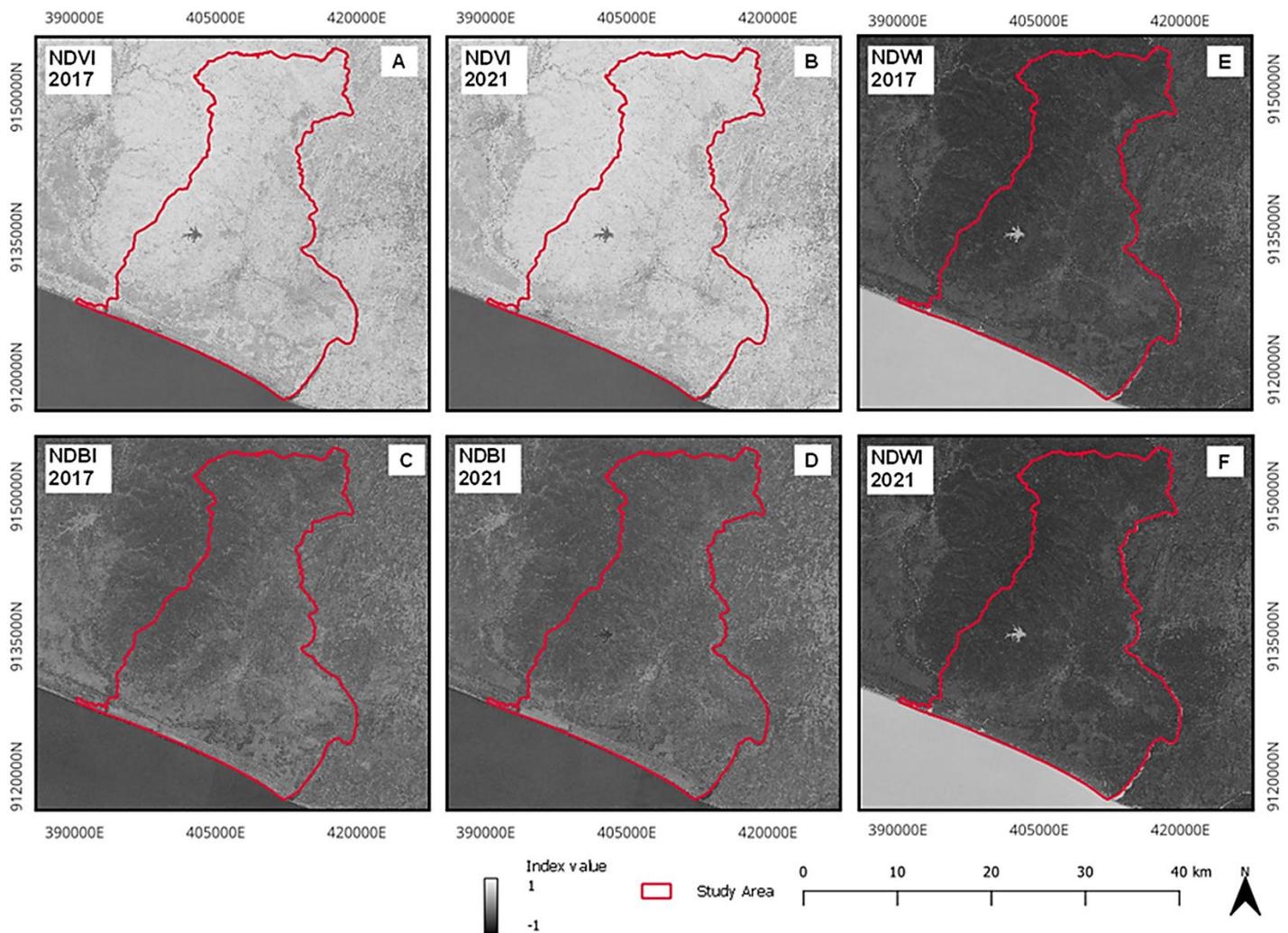


Figure 6. Image Result of The Transformation Indices of (A) NDVI 2017, (B) NDVI 2022, (C) NDBI 2017, (D) NDBI 2021, (E) NDWI 2017, (F) NDWI 2021

The transformation of the NDBI image in 2017 should reveal a lower building brightness index level than in 2021, notwithstanding that 2021 has a higher built-up area than 2017. This is most likely due to buildings with very high reflectance levels, specifically airports causing the visualisation of other buildings that have lower reflection levels. Hence, the hue becomes darker while the visual appearance of the airport becomes exceedingly bright. However, in 2017, the buildings were not large enough to provide high spectral reflections; therefore, the distribution of building reflections was more even with a moderately high level of brightness. Granting the visualisation of the 2017 NDBI transformed image is undoubtedly brighter than in 2021 concerning the built-up area, the 2021 image still exhibits a higher maximum pixel value.

3.2. Monthly Median Composite of SAR Images

Radar VH and VV polarisation can be used to detect wetlands, where wetlands are identical to the presence of paddy fields in the inundation phase (Gupta *et al.*, 2022). The results of the VV and VH polarisations show the distribution of wet paddy fields with the aim of employing to detect sampling locations with the aim of classifying paddy fields. Figure 7 illustrates the radar polarisation image of VV and VH with RGB colour composites composed from median composite months 10, 11 and 12.

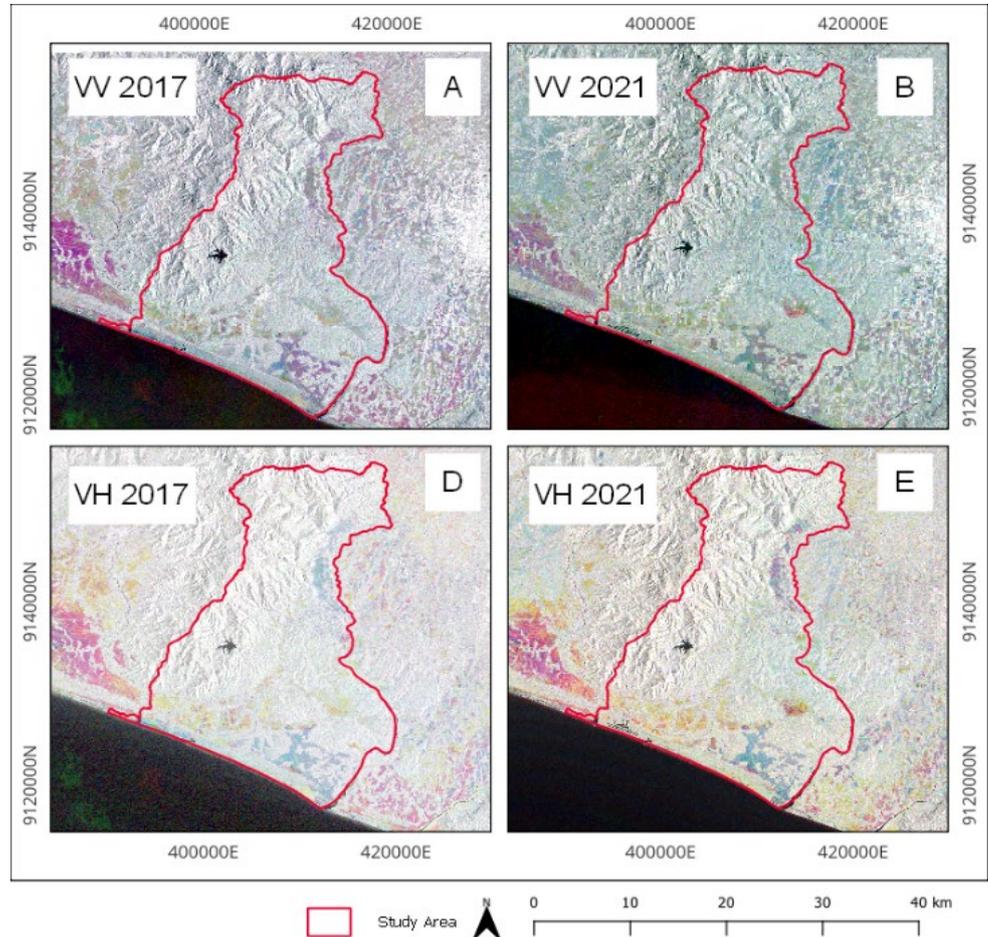


Figure 7. Radar Polarisation Images of VV On 2017 (A), VV On 2021 (B), VH On 2017 (D) and VH On 2021 (E).

Those particular months are chosen to visualise the data because they have a rainy season that, it is assumed will inundate the rice field enabling the radar sensor to detect it. The selection of sample training is carried out using visual interpretation. Visual interpretation is undertaken based on existing interpretation keys (Movchan *et al.*, 2023). Key interpretations include hue/colour, size, shape, texture, pattern, shadow, site and association. Based on this, a sample training location was obtained to identify rice and non-rice field objects (Figure 9). Training samples used to identify rice and non-rice fields were applied in 2017 and 2021. Both years made use of the same training sample, although in 2021, the airport object was added as a non-rice field object that previously did not exist in 2017.

In relation to the results of the VV and VH polarised radar images, it can be observed that the distribution of paddy fields is significant in the southern region of Kulon Progo Regency. From both these images it can be noted that the changes in land use that occurred in the Temon District and its surroundings, precisely in the area where the airport was built, triggered changes in the distribution of the wetlands. Examining the previous image, it appears to have a striking colour which indicates that the wetlands are almost uniform with the dry lands. This appearance can be clearly observed, particularly in the VV polarised radar images.

3.3. Classification Result

The final classification result pertaining to rice fields in Kulon Progo is shown in Figure 8. The classification is undertaken by using sample data that is shown in Figure 9. The sample data is

separated into two classes: rice field and non-rice field. The non-rice field class consists of water bodies, vegetation, built area, bare soil and ponds. The samples are taken in polygons and points, the points samples are used to obtain small building areas like housing. The result demonstrates that distribution of rice fields predominantly occurs in low land areas. In the southern area, the paddy field patterns found tend to be clustered in large patches, while in the central and northern areas they tend to be clustered in small patches. This classification result underwent a reclassification process to remove misclassification of paddy fields in the coastal area. Subsequently, the result is confirmed using a confusion matrix and the summary of accuracy shown in Table 1 and Table 2. In 2021, the overall accuracy value was slightly higher (96.1630%) than in 2017 (95.9461%). By reason of the relatively small difference in accuracy value between these two classification periods, a change detection analysis could be performed correctly. Nevertheless, the result of the change detection can also be impacted by this different accuracy value. Table 1 explains that the lowest accuracy occurs in the producer's accuracy in the rice field class. This situation shows that the classification error is primarily caused by overclassifying rice field (non-rice field area is categorised as rice field).

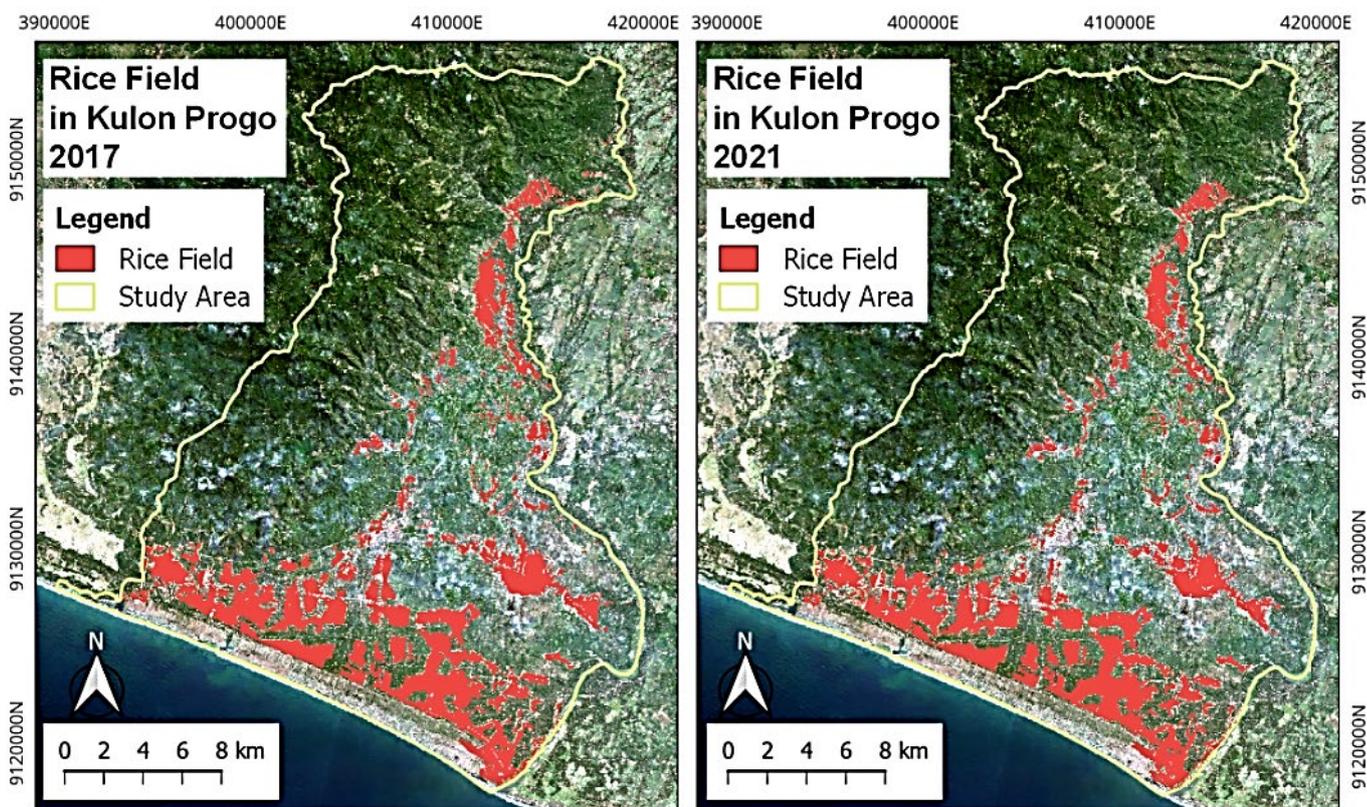


Figure 8. Rice Field Classification Map for 2017 and 2021.

Table 1. User's and Producer's Accuracy of Classification Result.

	2017		2021	
	Rice Field	Non-Rice Field	Rice Field	Non-Rice Field
User's accuracy (%)	91.5206	96.6543	94.1028	96.4776
Producer's accuracy (%)	81.4037	98.6155	80.3162	99.0751

Table 2. Overall Accuracy and Kappa Value of Classification Result.

	2017	2021
Overall accuracy (%)	95.9461	96.1630
Kappa statistic	0.8380	0.8444

The mapping of agricultural land in Kulon Progo was previously conducted using visual interpretation based on physical land parameters, such as land use and landform. In their research, agricultural rice fields were distinguished based on their water availability, for instance irrigated or

rain-fed rice fields. However, this study has limitations given that it was not possible to differentiate between different types of rice fields using digital classification. The classification of rice field types was addressed by means of an ecological landform method, which could not be achieved using digital classification based on Random Forest classification in this study. Nevertheless, the digital classification based on machine learning in this research has the advantage of providing faster information, particularly for multi-temporal analysis and change detection. Despite the limitation in distinguishing types of paddy field, the digital classification approach offers valuable insights and advances the process of analysing land cover changes over time.

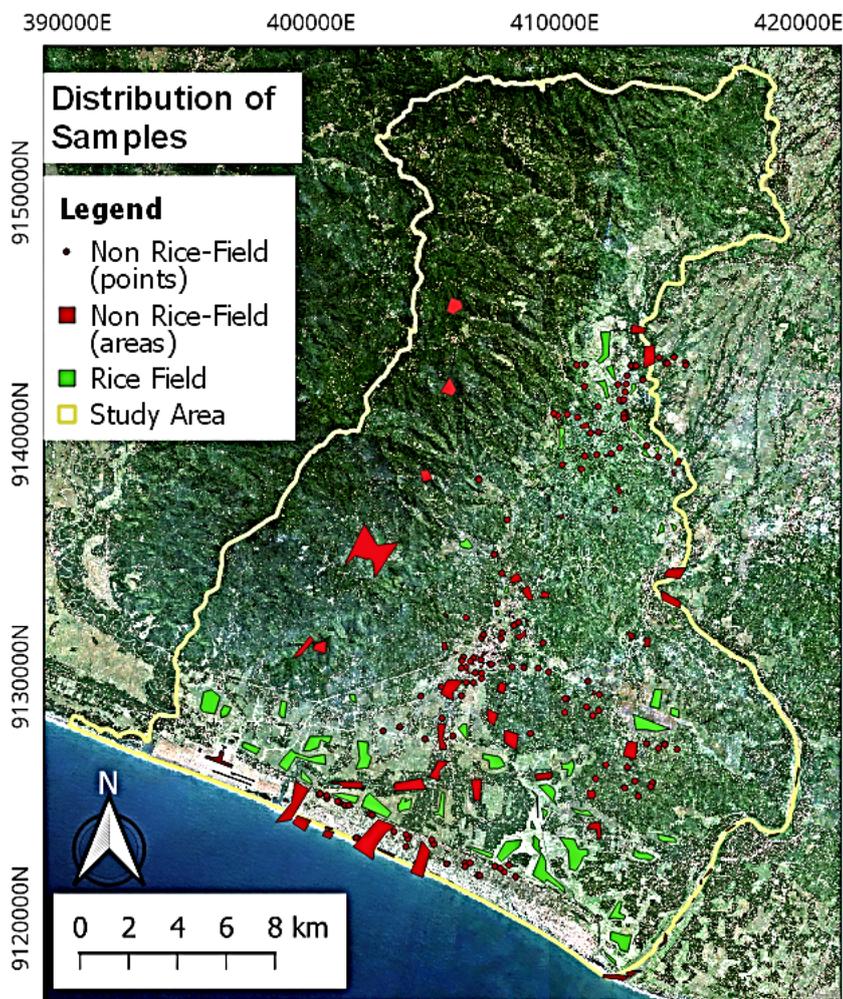


Figure 9. Distribution of Samples.

3.4. Variables Importance

Figure 10 shows 25 important variables that influenced the result of the classification of paddy fields in 2017 using the random forest method. The figure reveals that the variable that exerts the most influence is the elevation variable followed by VH4 and NDWI. The elevation variables have the most influence on the result of the classification owing to the diverse topographical conditions in Kulon Progo Regency, which are influenced by the presence of the Menoreh Hills in the north-western region of Kulon Progo. The Menoreh Hills are dominated by limestone material because of the forming process which is derived from the coral removal process. Soils with fuzzy material tend to have low soil fertility rates and low water reserves, thereby diminishing the possibility that rice plants will be able to be grown. In addition, the steep slope is also an additional factor that is affecting the low rice field in the Menoreh Mountains. The steep slope formed because of the denudational processes is a marker of massive erosion which is also an obstacle to the growth of agricultural land in the Menoreh Mountains. Based on this, it can be concluded that the Menoreh Mountains, which do are unsuitable for the development of rice agricultural land, create a significant difference in the conditions of the plain where the existence of rice agricultural land is extensive. Thus, there is a noticeable difference between the high and low elevations.

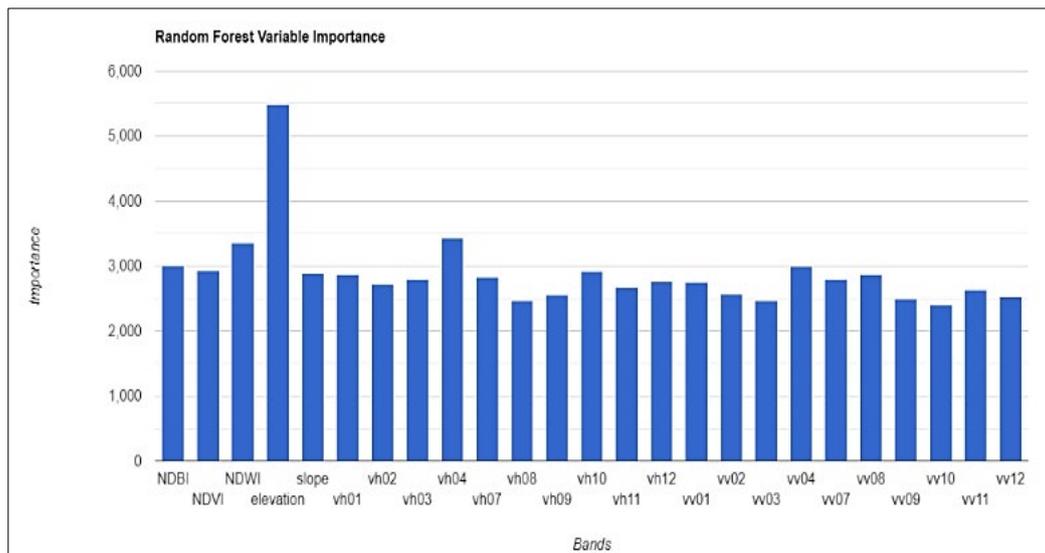


Figure 10. Random Forest Variable Importance of Rice Field Classification In 2017.

A different variable that has a considerable influence on classification results is the reflection of VH polarisation in April. April is the first month that marks the start of the dry season. Based on the Integrated Planting Calendar, the dry season in 2018 occurred from April to September. The Integrated Planting Calendar is a guideline or tool that provides spatial and tabular information related to seasonal predictions, planting, planting patterns, potential planting areas, areas prone to flooding and drought, together with attacks by plant disturbing organisms. It also provides recommendations for rice and crop varieties and needs, dosage recommendations and fertilizer needs, as well as Alsintan recommendations based on predictions pertaining to variability and climate change (Agricultural Research and Development Agency DIY, 2018). The reflection value of the rice field is influenced by the interaction of radar pulses with rice fields that have high soil thickness. This high soil moisture will reduce the reflection value of rice fields. This is also related to the results of the Normalised Difference Water Index which is an index that uses a combination of SWIR wavelengths reflecting the state of water content in vegetation and NIR which reflects the internal structure and dry matter content of vegetation. Accordingly, the combination of the two can improve accuracy in depicting the water content in vegetation. Therefore, in April it has a major influence on the results of its classification because in that month, which is the rainy season, the water content in rice plants is high, indicating that the bounce value is low and the NDWI value is high (Wang *et al.*, 2023). This low reflection value and high NDWI can be used as a marker for paddy fields allowing them to be distinguished by the appearance or cover of other lands on the surface of the Earth.

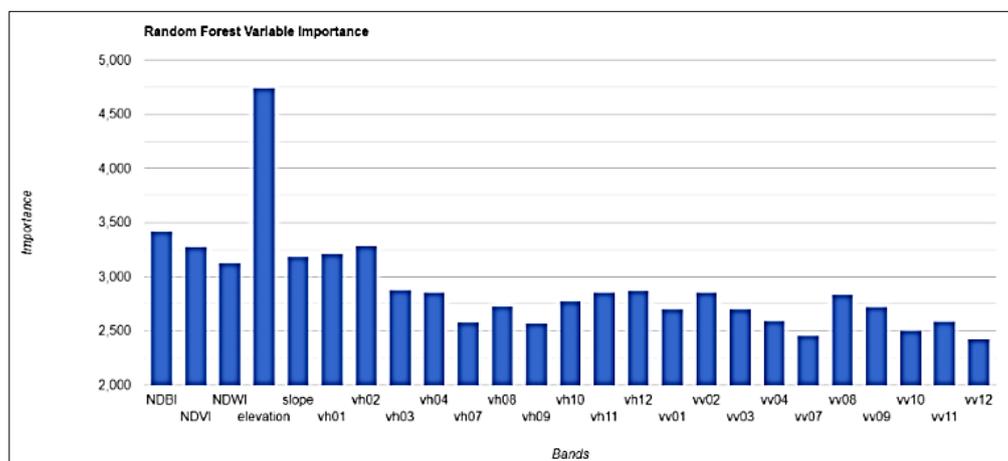


Figure 11. Random Forest Variable Importance of Paddy Field Classification In 2021.

In the image portraying the analysis of the classification results in 2021, Figure 9 illustrates that elevation remains the variable with the most influence in the classification of rice fields. Elevation variables continue to be one of the most important variables in conducting land assistance for rice plants. Conversely, these variables are not the main variables applied to examine paddy field conversion from 2017 to 2021. However, the rice variable was used to identify rice fields and non-

paddy fields in the study area. Concerning the factors that affect below elevation, 2021 is different from the variables that exerted an influence in 2017. The variable is the NDBI or Normalised Different Build-up Index, which is an index that uses the NIR and SWIR bands to highlight the area of built-up land. The reason for the high influence of NDBI on the image in 2021 is due to the construction of Yogyakarta International Airport in the Temon area. The existence of an airport covering a substantial area can impact the interaction between radar pulses and the Earth's surface.

In the research conducted by Kilbride *et al.* (2023), land mapping of paddy fields was also conducted using a combination of optical and SAR data, with Landsat-8 OLI being utilised as the optical data source. In that study, parameter tuning for the number of trees in the Random Forest classification was performed, resulting in the most advantageous models with 80 trees (overall accuracy 92.10%) and 50 trees (overall accuracy 92.47%). The evaluation of variable importance revealed that VH polarisation and NDVI were the two most influential variables, highlighting the significance of the multi-sensor combination in paddy field identification.

Furthermore, in this current study, the most important variables, apart from elevation, varied in each year of classification. In 2017, the principal variables were VH04 (April) radar polarisation and NDWI, while in 2020, they were VH02 (February) radar polarisation and NDBI. Despite the differences, both years demonstrated the importance of two top variables from different sensors, indicating the crucial role of the multi-sensor combination in paddy field identification in this research.

3.5. Rice-Field Land-Conversion

When comparing productive rice land in Kulon Progo Regency between 2017 and 2021, in general, the overall trend indicates a decline in productive rice land in 2021, in comparison to the levels detected in 2017. The information presented is consistent with the content depicted in Figure 12. Figure 12 confirms the rice-field land-conversion from 2017 to 2021. This map shows particular rice-field conversions to non-rice-field that occurred between 2017 and 2021. As illustrated in Figure 2, this conversion occurred close to the newly constructed airport and developed area. These conversions predominantly occurred at the east and southwest part of the study area. Nonetheless, the unconverted rice-fields comprise a larger proportion of area than the converted area. The distribution of unconverted areas extends from the northeast to the southern coast of the study area. In 2017, paddy fields in the southern part of Kulon Progo Regency were established along the southern coast. Similarly, in 2017, the northern part of the Kulon Progo Regency exhibited paddy fields within its territory that remained partially covered. Nevertheless, in 2021, the land cover in the form of rice fields had disappeared.

Changes in land use typically involves a shift from agricultural land to non-agricultural land. Land changes in the southern part of Kulon Progo Regency are possible owing to the construction of the new airport in the area. The existence of an airport can increase accessibility to transportation, resulting in an influx of people in the vicinity of the airport. People who arrive in the area will seek to meet their essential needs, such as clothing, food and accommodation. The productive land which was originally rice fields was subsequently converted into a housing complex or hotel which is used as accommodation for airport users.

The conversion of paddy fields will affect food production in Kulon Progo Regency. The significant conversion which is not accompanied by an increase in the production of the remaining paddy fields can ultimately threaten food security. In addition, the high land conversion rate can also affect the condition of the surrounding environment where the level of possible pollution will be higher. As a consequence, this can have an impact on the level of productivity in relation to the remaining paddy fields.

Table 3. Comparison Table of Rice Field Conversion From 2017 to 2021.

2017 Classification	2021 Classification	Pixel Sum	Area [metres^2]
Rice Field	Rice Field	708363	70836300
Non-Rice Field	Rice Field	51387	51387
Rice Field	Non-Rice Field	82666	8266600
Non-Rice Field	Non-Rice Field	4891747	489174700

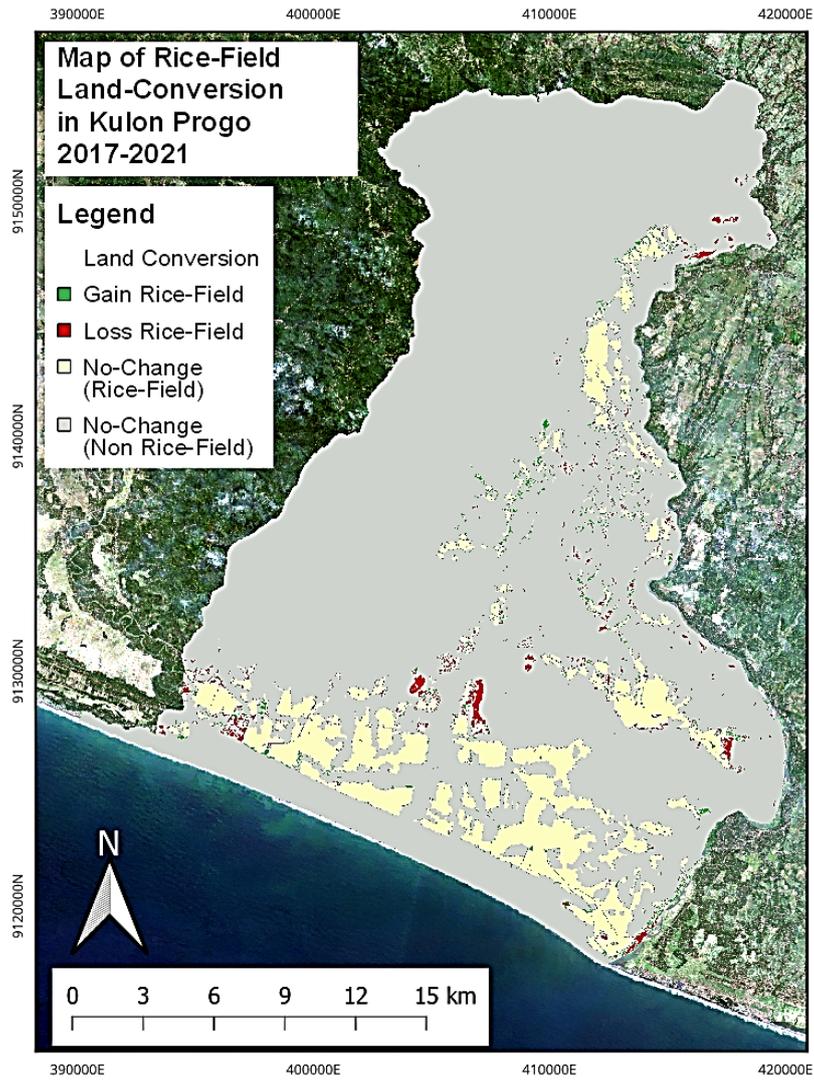


Figure 12. Rice Field Conversion Map in Kulon Progo From 2017 to 2021.

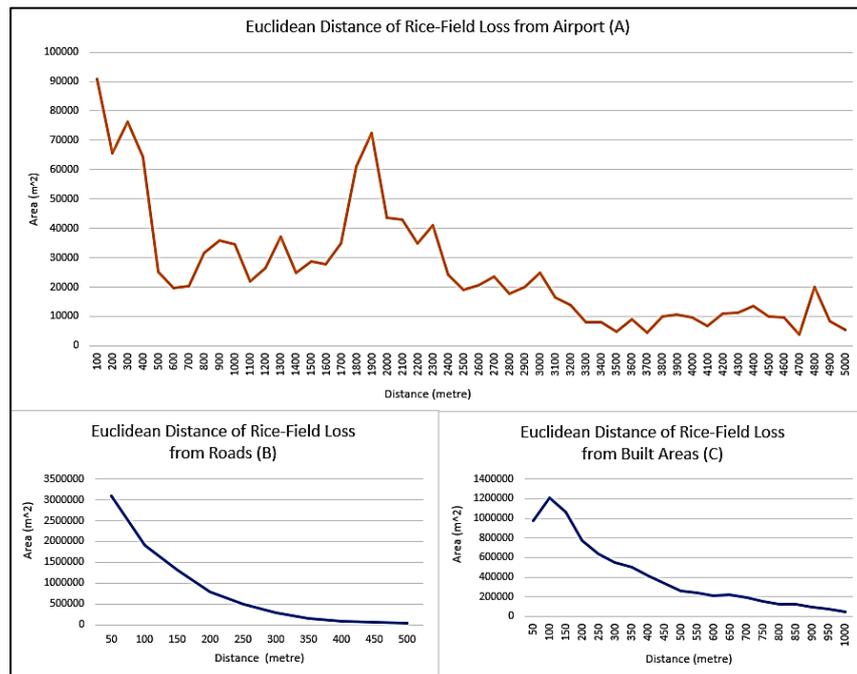


Figure 13. Euclidean Distance of Rice Field Loss (A) from Airport, (B) from Roads, (C) from Built-up Areas.

4. Conclusion

In conclusion, our study highlights the importance of monitoring the conversion of paddy fields in Kulon Progo using a combination of radar Sentinel-1 image, optical Sentinel-2 image, in conjunction with DEM data. Our analysis using multi-sensor and multi-temporal remote sensing data processed using the cloud computing platform generated an accuracy rates of 96.16% (2021) and 95.95% (2017). However, our result also shows that while Random Forest machine learning classification is effective in computing a substantial amount of multi-temporal and multi-sensory data, there are still misclassifications that affect change detection analysis. The misclassifications we identified in our analysis could impact attempts to monitor and manage changes in rice-field land-conversion. Improving the performance of mapping accuracy can be increased by using other more sophisticated methods, for example Deep Learning. However, our study was still able to highlight several of the implications for the management and ecological environment in the study area related to the development of National Strategic Projects of the New International Airport. Our results suggested that it is essential to have accurate and current data as regards changes in productive paddy fields, given that our results determined that roughly 8,266,600 metres² (826.66 hectares) had been converted from paddy fields to non-paddy fields over a period of five years (2017–2021), with most of the conversion occurring close to the new airport and roads. Therefore, efforts and comprehensive assessment and planning should be undertaken to ensure that the management and development of National Strategic Projects and infrastructures are carried out in a way that reduces the impact on the ecological environment and productive paddy fields in Kulon Progo. Additional information regarding the cropping pattern and productivity of the paddy field loss area can underline the future impact in relation to food security.

Acknowledgements

We are very thankful to Pusat Studi Lingkungan Hidup, Universitas Gadjah Mada for the financial support, which made this research possible

Author Contributions

Conceptualisation: Arjasakusuma, Khoirurriqzi, Sasongko, Utami, Irbah; **methodology:** Arjasakusuma, Khoirurriqzi, Sasongko, Utami, Irbah; **investigation:** Khoirurriqzi, Sasongko, Utami, Irbah; **writing—original draft preparation:** Khoirurriqzi, Sasongko, Utami, Irbah; **writing—review and editing:** Arjasakusuma; **visualisation:** Khoirurriqzi, Sasongko. All authors have read and agreed to the published version of the manuscript.

References

- Arjasakusuma, Sanjiwana, Sandiaga Swahyu Kusuma, Raihan Rafif, Siti Saringatin, and Pramaditya Wicaksono. (2020.) Combination of Landsat 8 OLI and Sentinel-1 SAR Time-Series Data for Mapping Paddy Fields in Parts of West and Central Java Provinces, Indonesia. *ISPRS International Journal of Geo-Information*, 9(663), 1–17. doi: 10.3390/ijgi9110663.
- Azedou, Ali, Aouatif Amine, Isaya Kisekka, Said Lahssini, Youness Bouziani, and Said Moukrim. (2023). Enhancing Land Cover/Land Use (LCLU) Classification through a Comparative Analysis of Hyperparameters Optimisation Approaches for Deep Neural Network (DNN). *Ecological Informatics* 78 (4), 16–25. doi: 10.1016/j.ecoinf.2023.102333.
- Bahagia, Bahagia, Bambang Hidayana, Rimun Wibowo, and Zuzy Anna. (2020). Local Wisdom to Overcome Covid-19 Pandemic of Urug and Cipatat Kolot Societies in Bogor, West Java, Indonesia. *Forum Geografi*, 34 (2), 146–60. doi: 10.23917/forgeo.v34i2.12366.
- Becker, Alexander, Stefania Russo, Stefano Puliti, Nico Lang, Konrad Schindler, and Jan Dirk Wegner. (2023). Country-Wide Retrieval of Forest Structure from Optical and SAR Satellite Imagery with Deep Ensembles. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195(11), 269–86. doi: 10.1016/j.isprsjprs.2022.11.011.
- Blickensdörfer, Lukas, Marcel Schwieder, Dirk Pflugmacher, Claas Nendel, Stefan Erasmí, and Patrick Hostert. (2022). Mapping of Crop Types and Crop Sequences with Combined Time Series of Sentinel-1, Sentinel-2 and Landsat 8 Data for Germany. *Remote Sensing of Environment*, 269(2022), 112831. doi: 10.1016/j.rse.2021.112831.
- Brown, Dana R. N., Christopher D. Arp, Todd J. Brinkman, Barbara A. Cellarius, Melanie Ingram, Mark E. Miller, and Katie V. Spellman. (2023). Long-Term Change and Geospatial Patterns of River Ice Cover and Navigability in Southcentral Alaska Detected with Remote Sensing. *Arctic, Antarctic and Alpine Research*, 55(1), 99775–7340. doi: 10.1080/15230430.2023.2241279.
- Chen, Bohan, Kevin Miller, Andrea L. Bertozzi, and Jon Schwenk. (2023). Batch Active Learning for Multispectral and Hyperspectral Image Segmentation Using Similarity Graphs. *Communications on Applied Mathematics and Computation*, 1(1), 11–15. doi: 10.1007/s42967-023-00284-8.
- Fu, Bolin, Yiyin Liang, Zhinan Lao, Xidong Sun, Sunzhe Li, Hongchang He, Weiwei Sun, and Donglin Fan. (2023). Quantifying Scattering Characteristics of Mangrove Species from Optuna-Based Optimal Machine Learning Classification Using Multi-Scale Feature Selection and SAR Image Time Series. *International Journal of Applied Earth Observation and Geoinformation*, 122(4), 1569–8432. doi: 10.1016/j.jag.2023.103446.
- Giofandi, Eggy Arya. (2020). Persebaran Fenomena Suhu Tinggi Melalui Kerapatan Vegetasi Dan Pertumbuhan Bangunan Serta Distribusi Suhu Permukaan. *Jurnal Geografi: Media Informasi Pengembangan Dan Profesi Kegeografian*, 17(2), 56–62. doi: 10.15294/jg.v17i2.24486.
- Gupta, Anishi, and Sambhav Kumar Jain. (2022). Conversion from Multi-Spectral Data into SAR Data with Deep Convolution Neural Architecture Using Generative Adversarial Network. *Procedia Computer Science*, 218(2023), 1760–67. doi: 10.1016/j.procs.2023.01.154.
- Hadibasyir, Hamim Zaky, Seftiawan Samsu Rijal, and Dewi Ratna Sari. (2020). Comparison of Land Surface Temperature During and Before the Emergence of Covid-19 Using Modis Imagery in Wuhan City, China. *Forum Geografi*, 34(1), 1–15. doi: 10.23917/forgeo.v34i1.10862.
- He, Xiaoning, Shuangcheng Zhang, Bowei Xue, Tong Zhao, and Tong Wu. (2023). Cross-Modal Change Detection Flood Extraction Based on Convolutional Neural Network. *International Journal of Applied Earth Observation and Geoinformation*, 117(1), 1569–8432. doi: 10.1016/j.jag.2023.103197.
- Hidayati, Iswari Nur, R. Suharyadi, and Projo Danoedoro. (2018). Developing an Extraction Method of Urban Built-Up Area Based on Remote Sensing Imagery Transformation Index. *Forum Geografi* 32(1), 96–108. doi: 10.23917/forgeo.v32i1.5907.
- Hossain, Md Sazzad, Md Asif Haider Khan, Tomiwa Victor Oluwajuwon, Jayanta Biswas, S. M. Rubaiot Abdullah, Md Seikh Sadiul Islam Tanvir, Sirajum Munira, and Md Naif Ahmed Chowdhury. (2023). Spatiotemporal Change Detection of Land Use Land Cover (LULC) in Fashiakhali Wildlife Sanctuary (FKWS) Impact Area, Bangladesh,

- Employing Multispectral Images and GIS. *Modelling Earth Systems and Environment*, 9(3), 151–73. doi: 10.1007/s40808-022-01653-7.
- Jia, Kun, Shunlin Liang, Xiangqin Wei, Yunjun Yao, Yingru Su, Bo Jiang, and Xiaoxia Wang. (2014). Land Cover Classification of Landsat Data with Phenological Features Extracted from Time Series MODIS NDVI Data. *Remote Sensing* 6(11), 11518–32. doi: 10.3390/rs6111518.
- Kilbride, John Burns, Ate Poortinga, Biplav Bhandari, Nyein Soe Thwal, Nguyen Hanh Quyen, Jeff Silverman, Karis Tenneson, David Bell, Matthew Gregory, Robert Kennedy, and David Saah. (2023). A Near Real-Time Mapping of Tropical Forest Disturbance Using SAR and Semantic Segmentation in Google Earth Engine. *Remote Sensing*, 15(5223), 1–20.
- Mahesh Batta. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR)* 9(1), 381–86. doi: 10.21275/ART20203995.
- Maxwell, Aaron E., Timothy A. Warner, and Fang Fang. (2018). Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. *International Journal of Remote Sensing*, 39(9), 2784–2817. doi: 10.1080/01431161.2018.1433343.
- Movchan, Dmytro, Andrii Bilous, Lesia Yelistratova, Alexander Apostolov, and Artur Hodorovsky. (2023). Application of Various Approaches of Multispectral and Radar Data Fusion for Modelling of Aboveground Forest Biomass. *Folia Forestalia Polonica, Series A*, 65(2), 55–67. doi: 10.2478/ffp-2023-0006.
- Onojeghuo, Alex Okiemute, Yuxin Miao, and George Alan Blackburn. (2023). Deep ResU-Net Convolutional Neural Networks Segmentation for Smallholder Paddy Rice Mapping Using Sentinel 1 SAR and Sentinel 2 Optical Imagery. *Remote Sensing* 15(6), 1–20. doi: 10.3390/rs15061517.
- Parelius, Eleonora Jonasova. (2023). A Review of Deep-Learning Methods for Change Detection in Multispectral Remote Sensing Images. *Remote Sensing* 15(8), 65–70. doi: 10.3390/rs15082092.
- Pham, Tien Dat, Nam Thang Ha, Neil Saintilan, Andrew Skidmore, Duong Cao Phan, Nga Nhu Le, Hung Luu Viet, Wataru Takeuchi, and Daniel A. Friess. (2023). Advances in Earth Observation and Machine Learning for Quantifying Blue Carbon. *Earth-Science Reviews* 243(5), 104501. doi: 10.1016/j.earscirev.2023.104501.
- Rafif, Raihan, Sandiaga Swahyu Kusuma, Siti Saringatin, Giara Iman Nanda, Pramaditya Wicaksono, and Sanjiwana Arjasakusuma. (2021). Crop Intensity Mapping Using Dynamic Time Warping and Machine Learning from Multi-Temporal PlanetScope Data. *Land* 10(12), 1–18. doi: 10.3390/land10121384.
- Ratnawati, Luthfiana, and Dwi Ratna Sulistyanningrum. (2020). Penerapan Random Forest Untuk Mengukur Tingkat Keparahatan Penyakit Pada Daun Apel. *Jurnal Sains Dan Seni ITS*, 8(2), 2337–3520. doi: 10.12962/j23373520.v8i2.48517.
- Shaik, Riyaz Uddien, Shoba Periasamy, and Weiping Zeng. (2023). Potential Assessment of PRISMA Hyperspectral Imagery for Remote Sensing Applications. *Remote Sensing*, 15(5), 1–14. doi: 10.3390/rs15051378.
- Stathopoulos, N., D. Rozos, and E. Vasileiou. 2023. Water Resources Management in Sperchios River Basin, Using SWOT Analysis. *Journal of the Civil Engineering Forum*, 9(3), 315–28. doi: 10.22146/jcef.7652.
- Tong, Zhonggui, Yuxia Li, Jinglin Zhang, Lei He, and Yushu Gong. (2023). MSFANet: Multiscale Fusion Attention Network for Road Segmentation of Multispectral Remote Sensing Data. *Remote Sensing*, 15(8), 1–23. doi: 10.3390/rs15081978.
- Tripathi, Akshar, Kapil Malik, Arjuman Rafiq Reshi, Md Moniruzzaman, and Reet Kamal Tiwari. (2023). Multi-Temporal SAR Interferometry (MTInSAR)-Based Study of Surface Subsidence and Its Impact on Krishna Godavari (KG) Basin in India: A Support Vector Approach. *Environmental Monitoring and Assessment*, 195(11), 1–17. doi: 10.1007/s10661-023-11896-1.
- Triscowati, Dwi Wahyu, Bagus Sartono, Anang Kurnia, Dede Dirgahayu, and Arie Wahyu Wijayanto. (2020). Classification of Rice-Plant Growth Phase Using Supervised Random Forest Method Based on Landsat-8 Multitemporal Data. *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, 16(2), 187–96. doi: 10.30536/ij.ijreses.2019.v16.a3217.
- Tzepkenlis, Anastasios, Konstantinos Marthoglou, and Nikos Grammalidis. (2023). Efficient Deep Semantic Segmentation for Land Cover Classification Using Sentinel Imagery. *Remote Sensing*, 15(8), 90–95. doi: 10.3390/rs15082027.
- Vinet, Luc, and Alexei Zhedanov. (2020). A ‘missing’ Family of Classical Orthogonal Polynomials. *Jurnal Kesehatan Komunitas Indonesia*, 44(8), 1–12. doi: 10.1088/1751-8113/44/8/085201.
- Wang, Ziyu, Shisong Cao, Mingyi Du, Wen Song, Jinling Quan, and Yang Lv. (2023). Local Climate Zone Classification by Seasonal and Diurnal Satellite Observations: An Integration of Daytime Thermal Infrared Multispectral Imageries and High-Resolution Night-Time Light Data. *Remote Sensing*, 15(10), 1–10. doi: 10.3390/rs15102599.
- Wu, Nitu, Luís Guilherme Teixeira Crusiol, Guixiang Liu, Deji Wuyun, and Guodong Han. (2023). Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries. *Remote Sensing*, 15(3), 1–22. doi: 10.3390/rs15030750.
- Xu, Yiming, Youquan Tan, Amr Abd-Elrahman, Tengfei Fan, and Qingpu Wang. (2023). Incorporation of Fused Remote Sensing Imagery to Enhance Soil Organic Carbon Spatial Prediction in an Agricultural Area in Yellow River Basin, China. *Remote Sensing*, 15(8), 1–16. doi: 10.3390/rs15082017.
- Yusrina, Azizah, Fiky Yosef Suratman, and Dharu Arseno. (2019). Pembentukan Citra Synthetic Aperture Radar (SAR) Menggunakan Metode Backprojection. *E-Proceeding of Engineering*, 6(2), 4268–74.
- Zollini, Sara, Donatella Dominici, Maria Alicandro, Maria Cuevas-González, Eduard Angelats, Francesca Ribas, and Gonzalo Simarro. (2023). New Methodology for Shoreline Extraction Using Optical and Radar (SAR) Satellite Imagery. *Journal of Marine Science and Engineering* 11(3), 1–25. doi: 10.3390/jmse11030627.