Jurnal Geografi Gea 25 (2) (2025) 160-175

Jurnal Geografi Gea

Journal homepage: https://ejournal.upi.edu/index.php/gea

Identification of Rice Fields Using The Random Forest Method Based on Spectral and Physical Parameters (Case Study: Bandung Regency)

Jania Sutia Rahayu[1], Albertus Delia[1], Riantini Virtriana[1]*

[1]Geodesy and Geomatics Engineering, Institut Teknologi Bandung, Bandung, Indonesia

Correspondence e-mail: *srjania@gmail.com

ABSTRACT

This research aims to identify paddy fields in Bandung Regency using the Random Forest method by incorporating spectral and physical parameters. Typically, mapping paddy fields requires several satellite images over different periods due to rice's planting cycle. To reduce time and cost, this study optimizes Random Forest classification with single-time satellite images, adding physical characteristics of paddy fields. Data includes NDVI, NDWI, LST, SMI, BSI, SAVI, slope, elevation, soil pH, total nitrogen, and clay content. Results show that combining these parameters enhances classification accuracy, achieving 91% overall accuracy and a Kappa of 0.89. This approach is both effective and efficient for mapping paddy fields, crucial for agricultural management and food security, and supports sustainable monitoring of paddy field distribution in Bandung Regency. By integrating spectral parameters (NDVI, NDWI, LST, SMI, BSI, SAVI) and physical characteristics (slope, elevation, soil pH, total nitrogen, and clay content), the Random Forest method significantly improves the classification accuracy of paddy fields compared to using only spectral data. The classification results indicated a substantial improvement, with the overall accuracy reaching 91% and a Kappa coefficient of 0.89. This methodological approach not only demonstrates its effectiveness and efficiency but also plays a vital role in agricultural management and food security. It provides a sustainable solution for monitoring paddy field distribution in Bandung Regency.

ARTICLE INFO

Article History:

Submitted/Received 26 July 2024
First Revised 29 November 2024
Accepted 09 October 2025
First Available online 31 October 2025
Publication Date 31 October 2025

Keyword:

Paddy fields
Bandung Regency
Random Forest method
Spectral parameters
Physical parameters

1. INTRODUCTION

Indonesia has approximately 40.5 million hectares of wetlands that play a crucial role in supporting life systems, such as being a source of water, a source of food, preserving biodiversity, and controlling the global climate. One of the most important types of wetlands in Indonesia is rice fields. Rice fields are wetlands that produce rice, which is the staple food for the Indonesian people.

The intensive increase in population exerts pressure on the availability of natural and agricultural resources, particularly rice fields. The limited availability of rice fields is a factor in the food crisis caused by land degradation and land (Ridwana et al., 2022). One of the efforts to maintain food security is to optimize land use (Mulyani et al., 2018) and preserve rice fields. Therefore, land monitoring is essential for decision-making processes in developing land use policies that will support sustainable food security (Feizizadeh et al., 2013) especially for rice fields.

Bandung Regency is one of the regencies in West Java Province with a high population density of 2,136 people/km² (Central Statistics Agency of Bandung Regency, 2023). Agricultural land use dominates land use in Bandung Regency, accounting for 78.37% of the total area, with 23.25% used for rice fields (West Java Province Communication, Informatics and Statistics Office, 2018). Rice production in Bandung Regency shows a fluctuating trend. According to the Bandung Regency Agricultural Profile data in 2023, rice production in Bandung Regency decreased from approximately 700,000 tons in 2018 to 600,000 tons in 2022. This is due to Bandung Regency being an urban buffer area, resulting in a shift in economic activities from agriculture to non-agriculture (Dzikrillah et al., 2017).

In order to monitor rice fields, it is necessary to update information on the distribution of rice fields continuously through mapping, which requires the identification of rice fields. Identifying rice fields requires a lot of supporting information, such as soil type, climate conditions, and topography. This makes direct data collection in the field time-consuming, labor-intensive, and costly. Machine learning methods can be used to enhance efficient and effective automatic detection capabilities according to the given classification model (Fauzi et al., 2020). One of the machine learning methods that can be used in classification is the Random Forest method (Ridwana et al., 2022) .

The Random Forest algorithm consists of several decision trees where each tree provides a calculation for the most dominant class unit to classify (specific classes according to the input training data (Noi Phan et al., 2020), The concept of the decision tree algorithm provides an understanding of the classification process carried out through developed classification rules and is often used for multi-source data integration (Sari et al., 2022). Mapping using the Random Forest algorithm can provide accurate and efficient results, which are important for agricultural management and food security (P. Wei et al., 2023; Ridwana et al., 2024).

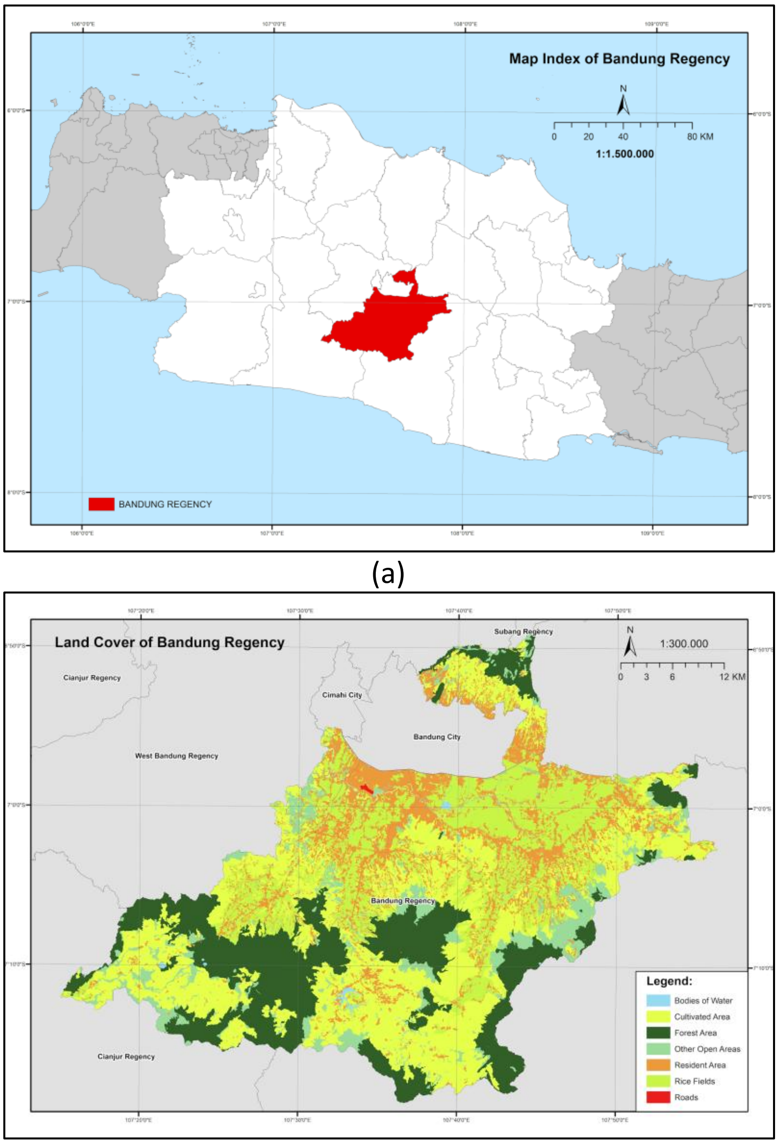
Mapping rice fields requires several satellite images over different periods because rice plants have their own planting period. Therefore, to minimize time and cost, this study aims to optimize the classification using the Random Forest method to identify rice fields using single-time satellite imagery by adding physical characteristics of rice fields. The combination of rice field classification based on spectral values and other physical parameters can significantly improve compared to spectral values. The developed model is expected to improve the accuracy of rice field mapping classification and provide faster and easier identification processes, supporting the monitoring efforts of rice field distribution in Bandung Regency.

2. METHODS

2.1 Study Area

The study area is located in Bandung Regency as explain in **Figure 1(a)**, geographically situated at 07°22’ - 108°50’ East Longitude and 6°41’ - 7°19’ South Latitude, with a total area of 1,762.4 km². The administrative boundaries of Bandung Regency are as follows (Central Statistics Agency of Bandung Regency, 2023): the north area is mostly bordered by Bandung City; Part of West Bandung Regency, and Cimahi City; the south area it borders Cianjur Regency and Garut Regency; the east area it borders Sumedang Regency and Garut Regency.

The land cover of Bandung Regency in **Figure 1(b)** consists of protected areas, agricultural cultivation areas, non-agricultural areas, and other areas. Most of the land in Bandung Regency is agricultural land, with a proportion of 78%, and 23% of it is rice fields (West Java Province Communication, Informatics and Statistics Office, 2018). The following are the study areas for this research.



(a)

(b)

Figure 1. (a) Bandung Regency; (b) Land Cover

2.2 Data

This research was conducted to develop a method in mapping using machine learning, specifically the Random Forest classification algorithm. The aim is to achieve a higher accuracy in land mapping classification and a faster and easier identification process, supporting efforts to monitor land distribution. In this rice field mapping, two main parameters were used: spectral data and physical land data, which were collected through digital downloads from various sources and processed on the Google Earth Engine platform. The following data in **Table 1** were used in this study.

Table 1. Data

No	Type	Data	Source	Year	Resolution
1	Spectral Data	Normalized Difference Vegetation Index (NDVI)	Landsat 9	2022	Raster-30m
2		Normalized Difference Water Index (NDWI)	Landsat 9	2022	Raster-30m
3		Land Surface Temperature (LST)	Landsat 9	2022	Raster-30m
4		Soil Moisture Index (SMI)	Landsat 9	2022	Raster-30m
5		Bare Soil Index (BSI)	Landsat 9	2022	Raster-30m
6		Soil-Adjusted Vegetation Index (SAVI)	Landsat 9	2022	Raster-30m
1	Physical Data	Slope	NASA SRTM	-	Raster-30m
2		Elevasi	NASA SRTM	-	Raster-30m
3		Soil pH	SoilGrids	2022	Raster-30m
4		Total Nitrogen	SoilGrids	2022	Raster-30m
5		Clay Soil	SoilGrids	2022	Raster-30m

2.3 Methodology

The use of these two scenarios is to compare the accuracy between mapping paddy fields using only spectral data with mapping that also considers physical data in **Figure 2**. It aims to perform classification optimization using the Random Forest method on the Google Earth Engine platform, and the results are tested using cross tabulation for accuracy. The following are the steps taken:

2.3.1 Literature Study and Data Collection

The research begins with a literature study to understand relevant concepts and techniques, followed by the collection of spectral data (NDVI, NDWI, BSI, SAVI, SMI, LST) and physical data (slope, elevation, soil pH, total nitrogen, clay content).

2.3.2 Paddy Field Characteristic Factors

The data collected were grouped into spectral variables and physical variables.

2.3.3 Classification Reference

Land use data was used to differentiate between paddy fields and non-rice fields.

2.3.4 Dataset

A combination of spectral and physical variables are used as explanatory variables and training features. In this step, The data collected is split into two sets: training data and testing data. The training data is utilized to develop the classification model, while the testing data is used to assess the model's accuracy. Although there isn't a strict rule for this division, a common practice is to allocate 70% of the data for training and 30% for testing.

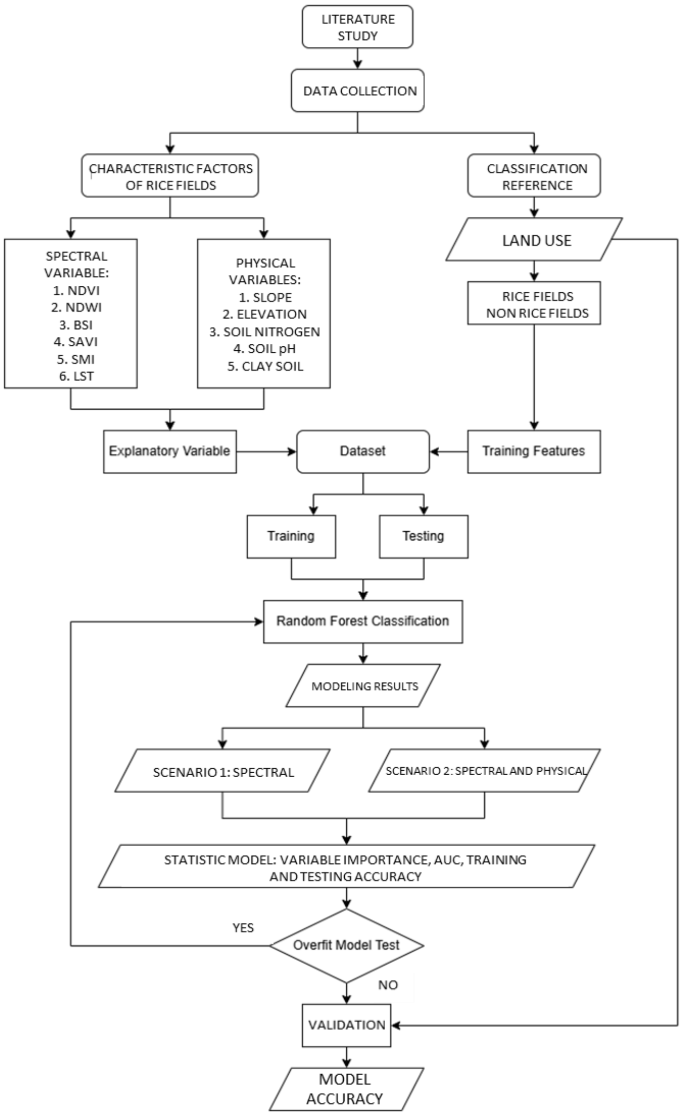


Figure 2. Research Flowchart

2.3.5 *Random Forest Classification*

This data was then used to train the Random Forest model in two scenarios, such as Scenario 1: using only spectral data and Scenario 2 : using spectral data supported by physical data.

2.3.6 *Modeling Results*

The results from both scenarios were compared to see the improvement in classification accuracy.

2.3.7 *Model Statistics*

Model evaluation is done by looking at variable importance, AUC, and accuracy of training and testing data.

2.3.8 *Model Overfit Test and Validation*

The model is tested for overfitting and then validated to ensure accuracy on data that has never been seen before.

2.3.9 Model Accuracy

The final result shows the classification accuracy which is compared between the two scenarios, ensuring the methodology used is effective and efficient.

2.3.10 Data Collection

Data collection was carried out through the process of downloading from various sources digitally, and data processing was conducted on the Google Earth Engine platform. GEE is highly effective for tracking vegetation cover, offering ready-to-use results and being quite easy to implement (Aldiansyah et al., 2021). The spectral data used include:

1) Normalized Difference Vegetation Index (NDVI) to understand the distribution and condition of vegetation.
2) Normalized Difference Water Index (NDWI) to understand the distribution and condition of water resources.
3) Land Surface Temperature (LST) to understand the distribution of surface temperature.
4) Soil Moisture Index (SMI) to understand the distribution and condition of soil moisture.
5) Bare Soil Index (BSI) to identify areas of bare or non-vegetated soil.
6) Soil-Adjusted Vegetation Index (SAVI) to understand the distribution and condition of vegetation.
7) Slope plays a role in determining the topographic characteristics of land, affecting the speed and volume of surface runoff and erosion.
8) Elevation is the height of an object relative to a specific point. Low elevation areas have the potential to be waterlogged due to surface water movement.
9) Soil pH is an indicator of soil fertility.
10) Nitrogen is an essential nutrient needed by plants for growth.
11) Clay soil has physical properties that affect water retention and soil structure (Cleophas et al., 2022; Medhina et al., 2023).

2.3.11 Sample Data Determination

The classification using the Random Forest algorithm requires sample data for training and testing (Breiman, 2001). The population size used is taken from the total pixels with the same resolution as the model results, which have a resolution of 100 meters. To determine the sample size (n) needed for spatial analysis in the context of machine learning modeling with a 100-meter resolution in Bandung Regency, which has 715,766 pixels, the Slovin's formula with a confidence level of 95% and a margin of error of 5% is used. The Slovin's formula is as follows:

$$n = \frac{N}{1 + (N\,x\,e^2)}$$

Description:

n = minimum sample size
N = population size = 715.766
e = margin of error = 0,05

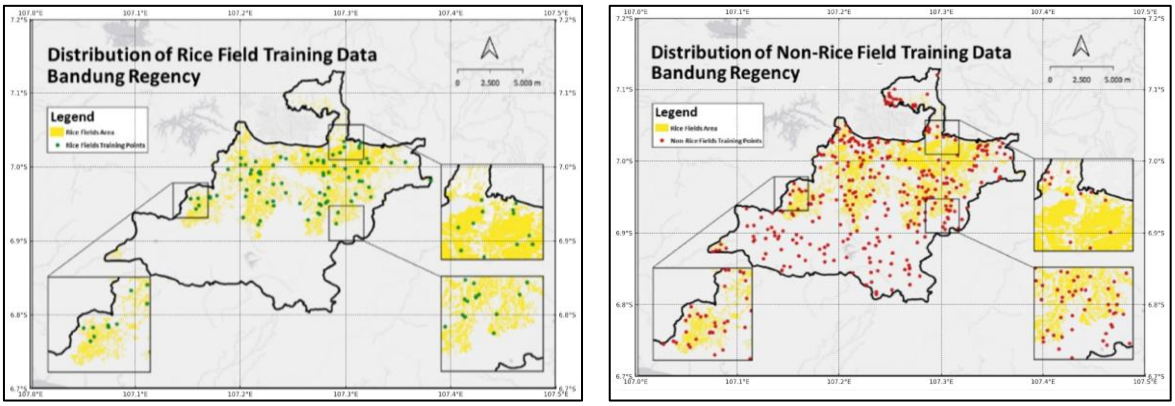
Then:

$$n = \frac{715.766}{1 + 715.766 \times 0{,}05^2} = \frac{715.766}{1790{,}415} = 400$$

DOI: https://doi.org/10.17509/gea.v25i2.85964
p-ISSN 1412-0313 e- ISSN 2549-7529

Based on the above calculation, with a margin of error of 5%, the required sample size is approximately 400 pixels. The classification in this study differentiates between rice and non-rice fields, thus the sample points need to be divided according to the area proportions. The area proportion between rice fields and non-rice fields is 1:5, resulting in 80 sample points for rice fields and 320 for non-rice fields (Schulthess et al., 2023).

Each sample point is then split into two sets: 70% for training and 30% for testing. This division ensures proper training and testing of the prediction model, aiming for high accuracy in the results. The random point tool is used for the distribution of sample points. The following illustrates in **Figure 3 (a)** and **(b)** for the sample point distribution in this study.



(a) (b)

Figure 3. (a) Distribution of Rice Field Training Data; (b) Distribution of Non-Rice Field Training Data

2.3.12 Classification with the Random Forest Algorithm

There are several input parameter configurations that can be modified on the Google Earth Engine platform used in this study to find the most suitable combination of input parameters for the Random Forest (RF) classifier to achieve the highest classification accuracy. These input parameters include the number of trees, variables per split, minimum leaf population, bag fraction, maximum nodes, and seed. In this study, iterations were performed on the number of trees to find the optimal value and achieve a balance between accuracy and model complexity. Hyperparameter tuning is employed to systematically test different configurations of these parameters to determine the optimal settings that enhance the model’s performance. This process involves adjusting and iterating over various combinations of parameters such as the number of trees and variables per split to find the most effective combination. The goal is to balance the model's accuracy and complexity, ensuring robust performance while avoiding overfitting. Hyperparameter tuning thus plays a crucial role in optimizing the Random Forest classifier for precise classification tasks (Probst et al., 2019; Rimal et al., 2024).

2.3.13 Model Evaluation

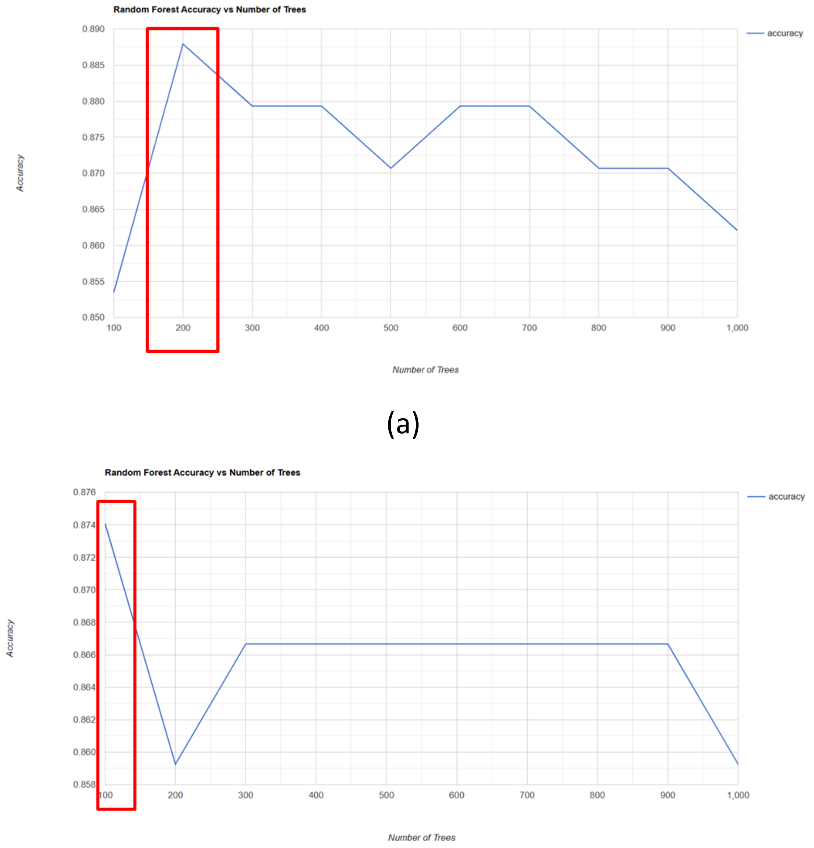
The model evaluation for the two scenarios of rice field identification in Bandung Regency was conducted based on variable importance values and ROC-AUC curves, as well as model validation using cross-tabulation to obtain model accuracy. Variable importance indicates the contribution of each input variable to the model's prediction accuracy. Variable importance can help improve future prediction models by optimizing the use of the most relevant variables.

Cross-tabulation, or contingency table, is a statistical technique used to analyze the relationship between two or more variables. In the context of model validation, cross-tabulation is used to compare the model's prediction results with actual data. By creating a confusion matrix showing the number of true positives, true negatives, false positives, and false negatives, we can calculate metrics such as accuracy, sensitivity, specificity, and others. This approach allows for a comprehensive evaluation of the model's performance and identification of areas that need improvement to enhance prediction accuracy in the future.

3. RESULTS AND DISCUSSION

3.1 Rice Field Classification Results

In the rice field classification study , hyperparameter tuning was used to optimize the Random Forest classifier (Probst et al., 2019; Rimal et al., 2024; Shaharum et al., 2020). For Scenario 1 which is in **Figure 4(a)**, 200 trees were identified as optimal, while in **Figure 4(b)** Scenario 2 used 100 trees. Each split in the decision trees considered four random variables to determine the best split, with a minimum of two samples per leaf to prevent overfitting. Bootstrapping was applied (Kohavi and Edu, 1993), using 50% of the sample data to create diverse trees and reduce variance. A fixed seed ensured reproducibility. These configurations enhanced classification accuracy and efficiency, balancing model complexity and performance.



(a)

(b)

Figure 4. Iteration Results for the Number of Trees in Scenario 1 and Scenario 2

DOI: https://doi.org/10.17509/gea.v25i2.85964
p-ISSN 1412-0313 e- ISSN 2549-7529

In the rice field classification study (Wang et al., 2017; G. Wei et al., 2022), hyperparameter tuning was utilized to optimize the Random Forest classifier. In **Table 2** for Scenario 1, the optimal configuration included 200 trees, while Scenario 2 used 100 trees. Both scenarios employed four variables per split, a minimum leaf population of two, a bag fraction of 0.5, a maximum of 100 nodes, and a fixed seed of 0 for reproducibility. Scenario 1 achieved an overall accuracy of 88.79% and a Kappa of 0.87867, whereas Scenario 2 attained an overall accuracy of 85.18% and a Kappa of 0.8901. These configurations balanced model complexity and performance, demonstrating the effectiveness of hyperparameter tuning.

Table 2. Optimal Input Parameter Combinations

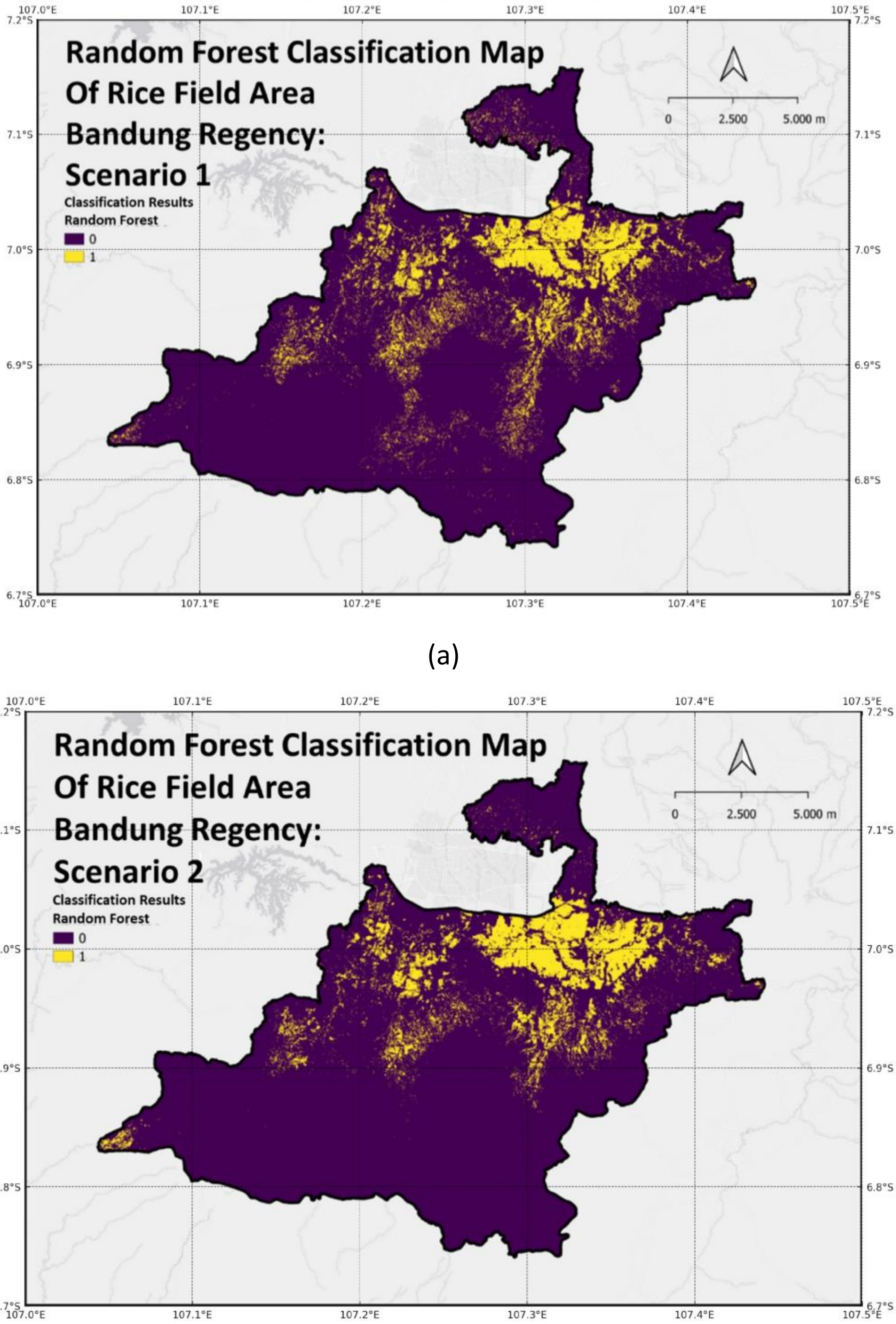
Parameter	Scenario 1	Scenario 2
Number of Trees	200	100
Variable per Split	4	4
Min. Leaf Population	2	2
Bag Fraction	0.5	0.5
Max Nodes	100	100
Seed	0	0
Overall Accuracy	0.8879	0.8518
Kappa	0.87867	0.8901

Source: Data Analysis (2025)

Both scenarios utilized four variables per split, a minimum leaf population of two, a bag fraction of 0.5, a maximum of 100 nodes, and a fixed seed of 0. Although both scenarios performed well, Scenario 2 may be more efficient in terms of computation time due to the lower number of trees. **Figure 5** illustrate the classification map results of rice fields using data from both scenarios.

The differences in performance and efficiency between Scenario 1 and Scenario 2 in the rice field classification study can be attributed to the number of trees used in the Random Forest classifier. Scenario 1 (a), with 200 trees, provides higher overall accuracy and a good Kappa value, indicating a robust model with strong predictive power. However, it requires more computational resources and time. Scenario 2 (b), with 100 trees, achieves slightly lower accuracy but still maintains a good Kappa value, making it more computationally efficient while delivering acceptable performance. This balance between accuracy and efficiency is a result of hyperparameter tuning, which optimizes model parameters to achieve the best possible results with available resources.

The maps in **Figure 5** below show the classification results of rice fields (in yellow) and non-rice fields (in purple) based on spectral parameters. The yellow color on the map indicates areas identified as rice fields. The purple color indicates non-rice areas. The maps reveal that most rice fields are located in the central and northern parts of Bandung Regency, with a fairly wide distribution. The concentration of rice fields is high in certain areas, especially in the northern part, which may reflect areas with intensive agricultural practices. The southern part of Bandung Regency shows fewer areas identified as rice fields, which may reflect different land use or topographic conditions less suitable for rice farming.



(a)

(b)

Figure 5. Rice Field Classification Map (a) Scenario 1 Results; (b) Scenario 2 Results

3.2 Variable Importance

The importance of variables (Gregorutti et al., 2017) differs in each scenario due to different relationships between input variables and classification targets in each scenario. The variable importance for scenario 1 and scenario 2 is as follows in **Figure 6**:

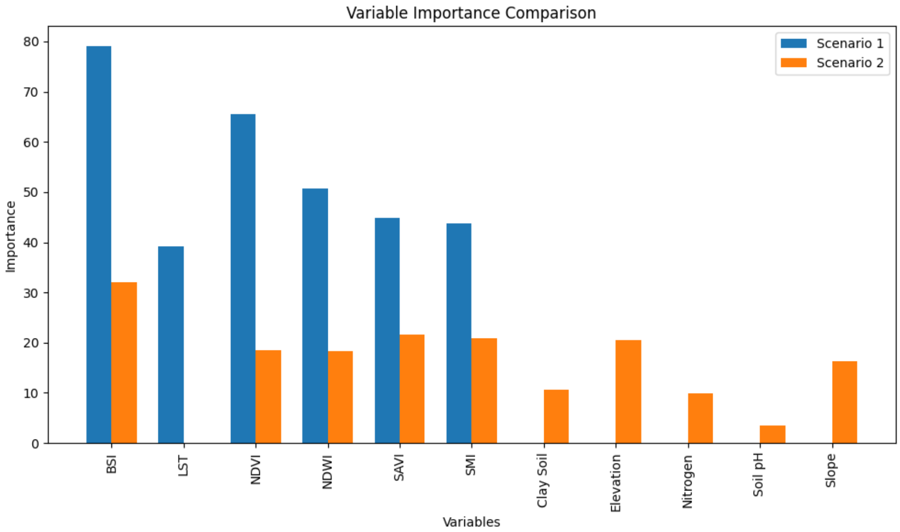


Figure 6. Variable Importance

BSI is the most important variable in classifying rice fields and non-rice fields in both scenarios. BSI (Bare Soil Index) measures the light reflection from bare soil, which is typical for rice fields as they are often flooded or wet. BSI helps identify bare soil areas without vegetation, facilitating rice field mapping and management through remote sensing. The comparison between spectral and physical data in **Table 3** shows that spectral data is important for vegetation health and water content, while physical data is important for understanding the topography of rice fields. The following table shows the differences between spectral and physical data.

Table 3. Differences Between Spectral and Physical Data

Data	Scenario 1	Scenario 2
BSI	High	Low
NDVI	Low-Moderate	High
NDWI	High	Low
SAVI	Moderate	Low
SMI	High	Low
Elevasi	Variative	Variative
Slope	Low	Variative
Clay Soil	Variative	Variative
Nitrogen	Variative	Variative
pH	Variative	Variative

BSI is dominant because rice fields tend to have bare and wet soil, resulting in high BSI values that effectively distinguish rice fields from other land types in remote sensing (Li and Chen, 2014; Nguyen et al., 2021).

3.3 ROC Curve and AUC Value

To assess how well the model performs in classification, further analysis was conducted using ROC curves and AUC values. The **Figure 7** show the ROC curves for the classification results of scenario 1 and scenario 2.

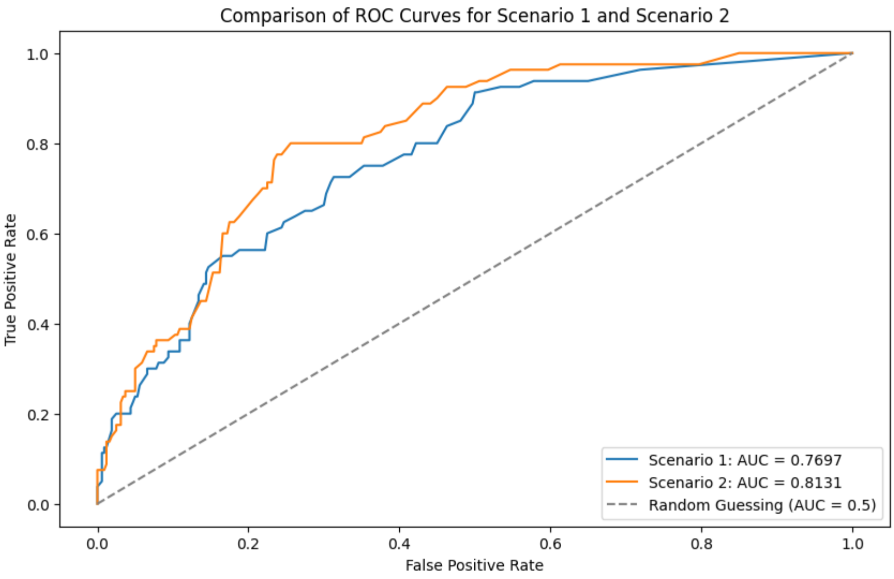


Figure 7. ROC-AUC Curve (a) Scenario 1; (b) Scenario 2

The higher AUC for scenario 2 indicates that the model in scenario 2 has a better ability to distinguish between the two classes than scenario 1. The ROC curve for scenario 2 is closer to the top-left corner than scenario 1. Closer to the top-left corner indicates that the model in scenario 2 has higher sensitivity and a lower false positive rate, indicating better performance in distinguishing between positive and negative classes. Spectral variables provide basic information related to vegetation and soil conditions. Physical variables add additional dimensions such as topography and soil chemical properties. The combination of spectral and physical variables provides more information to the model, enhancing its ability to distinguish between positive and negative classes. Additional information from physical variables reduces overlap between classes, improving classification accuracy. Physical variables increase data complexity and diversity, allowing the model to explore more complex relationships between features and target classes.

3.4 Accuracy Test

The model accuracy test was conducted using cross-tabulation to verify the model's accuracy results and using the grid method with a 100-meter interval. This test was conducted for both model scenarios. **Figure 8** shows an overlay map of the model with the reference rice field map used for accuracy testing.

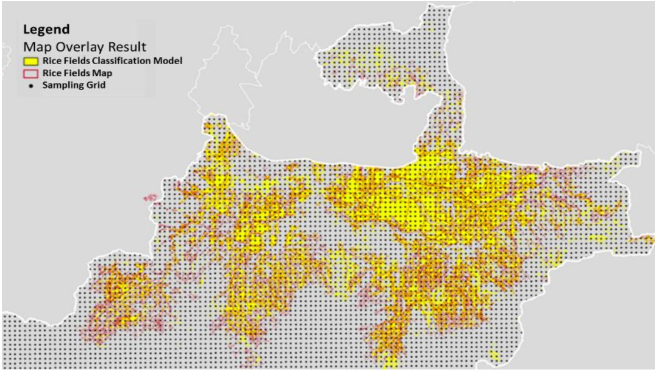
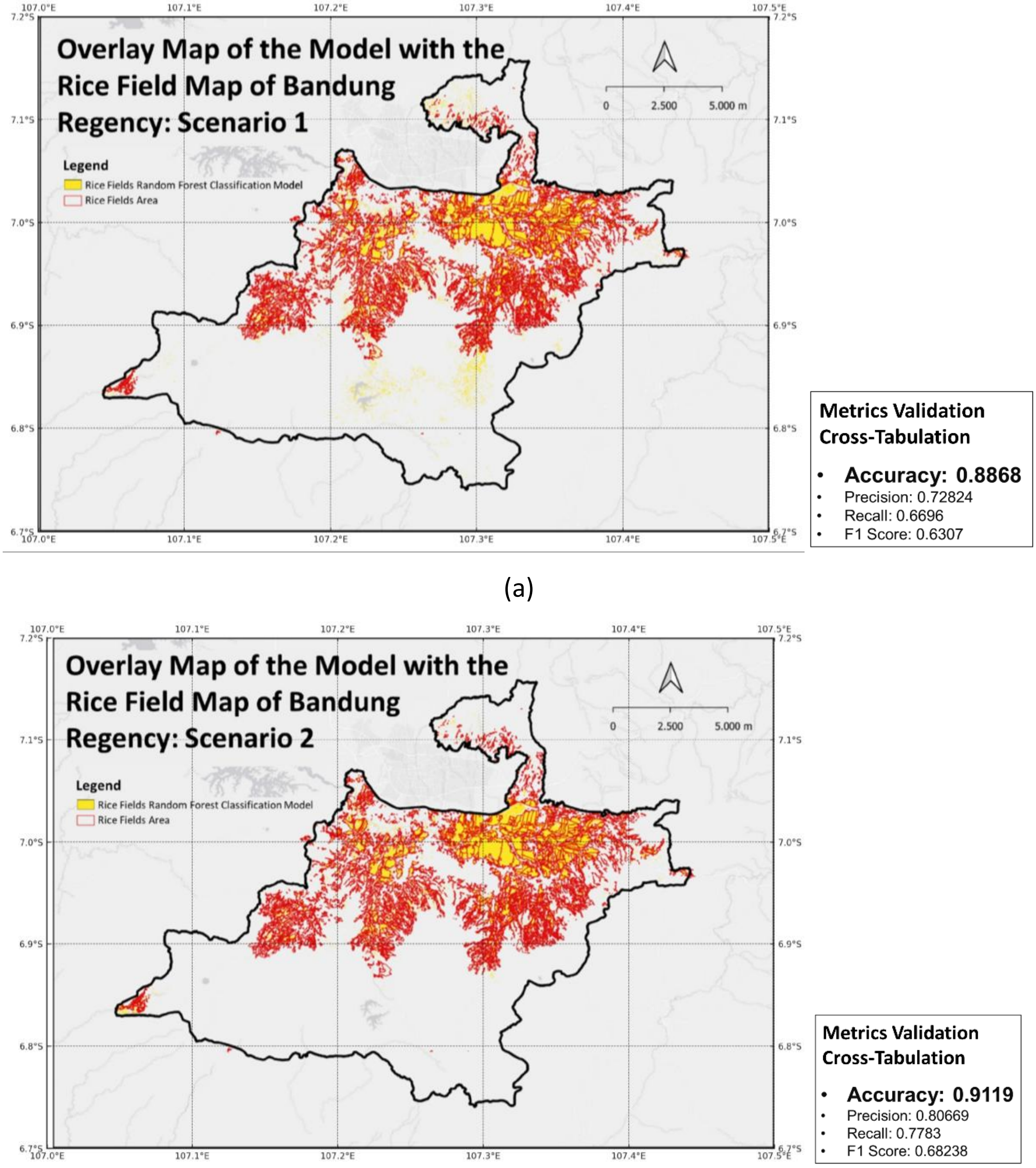


Figure 8. Sampling Grid Overlay Map of the Model with Existing Rice Field Areas for Accuracy Testing

Figure 9 shows the model's overlay results for both scenarios with the existing rice field map of Bandung Regency.



(a)

(b)

Figure 9. Overlay Map of the Model with the Rice Field Map of Bandung Regency
(a) Scenario 1; (b) Scenario 2

In scenario 1, an accuracy of 88% was achieved, meaning that out of every 100 predictions made by the model, 88 predictions were correct according to the actual data. For scenario 2, an accuracy of 91% was achieved, meaning that out of every 100 predictions made by the model, 91 predictions were correct according to the actual data. This shows that overall, the model has a good ability to accurately predict most of the cases tested.

Precision and recall in both scenarios are also quite good, although the lower recall indicates that some rice field areas may not be identified. The F1 Score shows a good balance between precision and recall, ensuring that the model is not only accurate but also comprehensive enough in identifying rice field areas. These results indicate that the model is reliable for mapping and managing rice fields in Bandung Regency, although further improvements may be needed to increase recall. Both scenarios show no significant difference. However, scenario 2 can be considered better in predicting and distinguishing between rice field and non-rice field classes.

4. CONCLUSIONS

The Random Forest method has proven effective in identifying rice fields in Bandung Regency using spectral and physical data, with a high level of accuracy. Both scenarios show that the combination of spectral and physical data results in a more accurate classification model compared to using only spectral data. The analysis shows that the Bare Soil Index (BSI) is the most important variable in the classification process. BSI is very effective in distinguishing rice fields from other land types because rice fields tend to have wet bare soil. Model testing using cross-tabulation shows a high level of accuracy with an AUC value indicating the model's ability to distinguish between positive and negative classes well. The model in scenario 2 with additional physical data shows better results than scenario 1.

5. RECOMMENDATIONS

Enhancing physical data to improve the accuracy of rice field classification models is recommended by adding more physical variables, such as other nutrient contents and more detailed climate information. For further research, it is recommended to use higher resolution data and other classification methods to compare results and further improve accuracy.

6. REFERENCES

Aldiansyah, S., Mandini Mannesa, M. D., and Supriatna, S. (2021). Monitoring of Vegetation Cover Changes With Geomorphological Forms using Google Earth Engine in Kendari City. *Jurnal Geografi Gea*, *21*(2), 159–170. https://doi.org/10.17509/gea.v21i2.37070

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Cleophas, F., Isidore, F., Musta, B., Mohd Ali, B. N., Mahali, M., Zahari, N. Z., and Bidin, K. (2022). Effect of soil physical properties on soil infiltration rates. *Journal of Physics: Conference Series*, *2314*(1). https://doi.org/10.1088/1742-6596/2314/1/012020

Dzikrillah, G. F., Anwar, S., and Sutjahjo, S. H. (2017). Analisis Keberlanjutan Usahatani Padi Sawah Di Kecamatan Soreang Kabupaten Bandung. *Jurnal Pengelolaan Sumberdaya Alam Dan Lingkungan (Journal of Natural Resources and Environmental Management)*, *7*(2), 107–113. https://doi.org/10.29244/jpsl.7.2.107-113

Fauzi, A., Supriyadi, R., & Maulidah, N. (2020). Detection of breast cancer with feature selection based on Principal Component Analysis and Random Forest. *Jurnal Infortech,* *2*(1), 96-101. https://doi.org/10.31294/infortech.v2i1.8079

Feizizadeh, B., and Blaschke, T. (2013). Land suitability analysis for Tabriz County, Iran: multi-criteria evaluation approach using GIS. *Journal of Environmental Planning and Management*, *56*(1), 1–23. https://doi.org/10.1080/09640568.2011.646964

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, *27*, 659–678.

Kohavi, R., and Edu, S. (1993). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, *2*, 1137–1143.

Li, S., and Chen, X. (2014). A new bare-soil index for rapid mapping developing areas using landsat 8 data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(4), 139–144. https://doi.org/10.5194/isprsarchives-XL-4-139-2014

Medhina, F. M., Harjanti, D. T., and Buwono, S. (2023). Jurnal Geografi Gea Biomass Percentage of Organic Soil Layers in The Rasau Jaya Tiga Village. *Jurnal Geografi GEA*, *23*(2), 144–158.

Mulyani, A., and Agus, F. (2018). The Need and Availability of Reserved Land to Realize Indonesia’s Aspiration as the World’s Food Barn in 2045. *Analisis Kebijakan Pertanian*, *15*(1), 1.

Nguyen, C. T., Chidthaisong, A., Diem, P. K., and Huo, L. Z. (2021). A modified bare soil index to identify bare land features during agricultural fallow-period in southeast asia using landsat 8. *Land*, *10*(3), 1–18. https://doi.org/10.3390/land10030231

Noi Phan, T., Kuch, V., and Lehnert, L. W. (2020). Land cover classification using google earth engine and random forest classifier-the role of image composition. *Remote Sensing*, *12*(15), 1–22. https://doi.org/10.3390/RS12152411

Probst, P., Wright, M. N., and Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(3), 1–19. https://doi.org/10.1002/widm.1301

Ridwana, R., Al Kautsar, A., Saleh, F., Himayah, S., Arrasyid, R., & Pamungkas, T. D. (2022, November). Spatiotemporal monitoring of rice crops in the covid-19 pandemic period for local food security using sentinel 2b imagery case study: tasikmalaya city. In *IOP Conference Series: Earth and Environmental Science*, *1089*(1), 012039. IOP Publishing.

Ridwana, R., Kamal, M., Arjasakusuma, S., & Rabbi, M. F. A. (2024). Evaluation of Machine Learning Models for Mapping Food Crops using Sentinel-2A Imagery in West Java, Indonesia. In *E3S Web of Conferences*, *600*, 03007. EDP Sciences.

Ridwana, R., Kamal, M., Arjasakusuma, S., Sugandi, D., & Sakti, A. D. (2024). Bibliometric computation mapping analysis of publication machine and deep learning for food crops mapping using VOSviewer. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *50*(2), 42-59.

Rimal, Y., Sharma, N., and Alsadoon, A. (2024). The accuracy of machine learning models relies on hyperparameter tuning: student result classification using random forest, randomized search, grid search, bayesian, genetic, and optuna algorithms. *Multimedia Tools and Applications*, *83*, 74349-74364. https://doi.org/10.1007/s11042-024-18426-2

Sari, I. L., Weston, C. J., Newnham, G. J., and Volkova, L. (2022). Developing multi-source indices to discriminate between native tropical forests, oil palm and rubber plantations in indonesia. *Remote Sensing*, *14*(1). https://doi.org/10.3390/rs14010003

Schulthess, U., Rodrigues, F., Taymans, M., Bellemans, N., Bontemps, S., Ortiz-Monasterio, I., Gérard, B., and Defourny, P. (2023). Optimal Sample Size and Composition for Crop Classification with Sen2-Agri's Random Forest Classifier. *Remote Sensing*, *15*(3). https://doi.org/10.3390/rs15030608

Shaharum, N. S. N., Shafri, H. Z. M., Ghani, W. A. W. A. K., Samsatli, S., Al-Habshi, M. M. A., and Yusuf, B. (2020). Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sensing Applications: Society and Environment*, *17*. https://doi.org/10.1016/j.rsase.2020.100287

Wang, Y., Liu, X., Wang, H., Li, Y., Li, Y., Liu, F., Xiao, R., Shen, J., and Wu, J. (2017). Rice agriculture impacts catchment hydrographic patterns and nitrogen export characteristics in subtropical central China: a paired-catchment study. *Environmental Science and Pollution Research*, *24*(18), 15700–15711. https://doi.org/10.1007/s11356-017-9044-z

Wei, G., Cao, J., Xie, H., Xie, H., Yang, Y., Wu, C., Cui, Y., and Luo, Y. (2022). Spatial-Temporal Variation in Paddy Evapotranspiration in Subtropical Climate Regions Based on the SEBAL Model: A Case Study of the Ganfu Plain Irrigation System, Southern China. *Remote Sensing*, *14*(5). https://doi.org/10.3390/rs14051201

Wei, P., Ye, H., Qiao, S., Liu, R., Nie, C., Zhang, B., Song, L., and Huang, S. (2023). Early Crop Mapping Based on Sentinel-2 Time-Series Data and the Random Forest Algorithm. *Remote Sensing*, *15*(13), 1–18. https://doi.org/10.3390/rs15133212