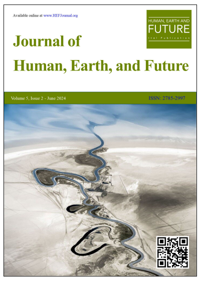
Available online at www.HEFJournal.org

Journal of
Human, Earth, and Future

ISSN: 2785-2997

Vol. 5, No. 2, June, 2024



Towards Improving Sustainable Water Management in Geothermal Fields: SVM and RF Land Use Monitoring

Widya Utama [1*], Rista Fitri Indriani [1], Maman Hermana [2], Ira Mutiara Anjasmara [3], Sherly Ardhya Garini [4], Dhea Pratama Novian Putra [3]

[1] *Department of Geophysical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, 60117, Indonesia.*

[2] *Department of Geosciences, Universiti Teknologi Petronas, Seri Iskandar, 32610, Malaysia.*

[3] *Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember , Surabaya, 60117, Indonesia.*

[4] *Department of Informatics, Institut Teknologi Sepuluh Nopember , Surabaya, 60117, Indonesia*.

Received 22 February 2024; Revised 13 May 2024; Accepted 19 May 2024; Published 01 June 2024

Abstract

The management and monitoring of land use in geothermal fields are crucial for the sustainable utilization of water resources, as well as for striking a balance between the production of renewable energy and the preservation of the environment. This study primarily compared Support Vector Machine (SVM) and Random Forest (RF) machine learning methods, using satellite imagery from Landsat 8 and Sentinel 2 between 2021 and 2023, to monitor land use in the Patuha geothermal area. The objective is to improve sustainable water management practices by accurately categorizing different land cover types. This comparative analysis assessed the efficacy of these techniques in upholding water sustainability in geothermal regions. This study examined the application of SVM and RF machine learning techniques, with particular emphasis on parameter refinement and model assessment, to enhance land use classification accuracy. By employing Kernlab and e1071 for algorithm comparison, the research sought to produce a precise Land Use Model Map, which underscores the significance of advanced analytical techniques in environmental management. This approach was of utmost importance in improving land use monitoring and reinforcing sustainable practices. The comparative evaluation of SVM and RF methods for land use classification demonstrates the superiority of RF in terms of accuracy, stability, and precision, particularly in intricate urban settings, hence establishing it as the preferred model for tasks demanding high reliability. The application of SVM and RF for monitoring land use in geothermal areas is in alignment with Sustainable Development Goals (SDGs) 6 and 15, as it fosters sustainable water management and the conservation of ecosystems.

Keywords: Geothermal; Land Use; Random Forest; Sustainability; Support Vector Machine; Water.

1. Introduction

Effective water resource management is crucial at geothermal sites to balance energy production with environmental stewardship. Geothermal energy production involves extracting heat from deep within the earth and often needs significant amounts of water for tasks like stimulating reservoirs, extracting heat, and generating electricity [1]. Chandrasekharam et al. (2020) emphasize the critical role of water resource management in geothermal energy production [2]. Furthermore, adept water management promotes sustainable development and improves water

* Corresponding author: widya@geofisika.its.ac.id

http://dx.doi.org/10. 28991/HEF-2024-05-02-06

security, especially in water-scarce countries such as MENA and Sub-Saharan Africa. As a result, a careful water resources management strategy is required to limit impacts on critical groundwater systems, ensuring the sustainability of the geothermal industry and the health of local ecosystems [3–6].

Efficient land use management is crucial for sustainably preserving water resources in geothermal areas [7]. Land use management can support the regulation of surface water flow. Surface water flow must be properly regulated to ensure water quality and availability, meet geothermal energy production requirements, and support local ecosystems. Thus, integrating land use management and water sustainability strategies is crucial to supporting the productivity and sustainability of geothermal operations. Cheng et al. (2022) underlined the importance of varied land uses, such as agriculture and urban growth, for water quality and sustainability. It emphasizes the significance of smart land use management to avoid negative impacts on water resources, which is crucial for sustaining sustainable geothermal output [8]. It is crucial to preserve water supplies for future generations while also maintaining the hydrological balance required by the geothermal industry and neighboring ecosystems. As a result, incorporating this land use strategy into geothermal development plans is crucial, emphasizing the critical relationship between land preservation and the conservation of water resources [9, 10].

Land use monitoring and management heavily rely on satellite imagery to obtain accurate spatial information crucial for effective planning [11]. Satellite systems like Landsat 8 and Sentinel 2 are invaluable [12–14]. Landsat 8's Operational Land Imager (OLI) is widely employed in supervised land cover categorization in plantations, as proven in Anua et al.'s (2021) research, highlighting its usefulness in recognizing and managing distinct land cover types for sustainable land use planning [15]. Sentinel 2's multispectral instruments have a broad range of applications; according to Phiri et al. (2020), they are crucial in monitoring agricultural, urban, and areas prone to natural hazards [16]. From the Sentinel-2 image data, an approach was made to the NDVI, NDBI, and NDWI spectral indices. In the study conducted by Hu et al. (2023), the Normalized Difference Vegetation Index (NDVI) is proposed as a viable method for delineating vegetation areas, such as agriculture and forests [17]. Muhaimin et al. (2020) utilized NDBI to accurately map the density of built-up areas, indicating varying levels of built-up area density through different Normalized Difference Built-up Index (NDBI) values [18]. Additionally, Patil et al. (2024) demonstrated the effective use of the Normalized Difference Water Index (NDWI) in differentiating water bodies by harnessing specific spectral bands from Sentinel 2 satellite imagery. This research underscored the significance of NDWI in precise water body mapping and classification for land use monitoring [19]. Integrating spectral indices such as NDVI, NDBI, and NDWI is important in accurately identifying and distinguishing agricultural land, urban zones, and water bodies, thus significantly increasing the precision in monitoring and managing these land resources [20, 21]. This approach improves the accuracy of land use maps and actively contributes to the sustainable management of vital environmental resources, thereby supporting broader water and geothermal resource sustainability goals [22, 23].

Machine Learning (ML) methods such as Support Vector Machine (SVM) and Random Forest (RF) algorithms have been adopted to develop models efficiently in geothermal land use [24, 25]. Xie et al. (2021) employed a Support Vector Machine (SVM) in their study to assess and monitor land use alterations on the Crozon Peninsula in Brittany, France. This investigation enhances comprehension of the frequency and repercussions of such modifications, demonstrating SVM's capacity to map and oversee land use change dynamics effectively. Consequently, this research provides valuable insights concerning land use adaptation and its environmental implications. Furthermore, Xie et al. (2021) also highlighted the utility of the Random Forest (RF) approach in modeling land use for groundwater surveillance in geothermal zones [26]. In a related study by Chen et al. (2022), RF was explicitly selected due to its proficiency in handling extensive datasets and complex relationships, rendering it an efficient instrument for managing geothermal and water resources. The amalgamation of these two algorithms plays a pivotal role in enhancing the precision and efficacy of monitoring and supervising water and geothermal resources [27].

Refinement of parameters in Machine Learning (ML) algorithms such as Support Vector Machine (SVM) and Random Forest (RF) is crucial for enhancing model efficiency and accuracy, particularly in monitoring land use in geothermal fields. For SVM, parameter tuning involves adjusting the Cost C, which balances model complexity against misclassifications; the Hyperparameter (sigma), which determines the Gaussian kernel's width in non-linear SVMs; and the Number of Support Vectors, which affects computational costs when increased. In geothermal land monitoring, the precise setting of these parameters ensures that the SVM can effectively handle the unique land cover characteristics typical of geothermal areas [28–31]. Meanwhile, tuning in the RF algorithm includes optimizing the Training Error, mtry (the number of variables considered at each node split), and the number of trees, significantly influencing the algorithm's ability to predict and classify complex geothermal land use patterns accurately. Adjusting these parameters ensures optimally that RF models are accurate and efficient in real-world applications [32–34].

Parameter adaptation in Machine Learning (ML) is critical for improving the performance and precision of models, especially when analyzing land usage in geothermal fields imaging data. The assessment of feature importance (VarImp) is significant in this context because it helps identify the most important variables that discriminate between different types of land uses, such as possible geothermal sites, conservation areas, or extraction zones. Understanding which features most substantially determine the decision boundaries between these land use classifications can help guide SVM parameter adjustment. Similarly, in the RF algorithm, recognizing the most influential features in the decision tree's construction enhances land use classification accuracy. Thus, comprehending the influence of these variables through feature importance is essential for making informed adjustments to the model parameters, thereby ensuring more accurate and effective monitoring of land use in geothermal environments [35, 36].

Assessing land use models is essential for evaluating their effectiveness in geothermal field applications. Various metrics, such as the confusion matrix, Overall Accuracy (ACC), Kappa Coefficient (KC), Sensitivity, and Specificity, play key roles in this evaluation [37, 38]. The confusion matrix offers a detailed view of the model's classification accuracy by showing true positive and negative classifications, facilitating the calculation of other important metrics like ACC, which indicates the overall proportion of correct classifications. The KC provides insight into how the model performs better than random classification, which is essential for understanding model reliability in complex geothermal landscapes. Sensitivity and Specificity are crucial for assessing the model's accuracy in identifying geothermal-related land uses specifically, with Sensitivity focusing on correctly identifying relevant land use classes and Specificity on accurately ignoring non-relevant classes. Utilizing these metrics gives a comprehensive insight into the performance of SVM and RF models in managing and monitoring land use in geothermal areas [39].

Evaluating machine learning algorithms through software packages like Kernlab and e1071 is crucial in assessing their effectiveness for specific applications, such as monitoring land use in geothermal fields. These tools provide comprehensive frameworks for performance evaluation, parameter tuning, and comparative analysis, which are vital in a geothermal context where data characteristics can be complex and highly variable. Using Kernlab and e1071, researchers can fine-tune models to optimize accuracy and resource efficiency, which is crucial for effective management and monitoring of geothermal land use. This detailed analysis helps determine the most suitable machine learning model to handle the unique challenges of thermal variations and land cover classifications in geothermal areas [40].

This article introduces a comprehensive study on enhancing sustainable water management in geothermal fields through advanced land use monitoring techniques, leveraging Support Vector Machine (SVM) and Random Forest (RF) algorithms. The approach precisely classifies various land cover types, supporting conservation efforts and promoting responsible water use within geothermal projects. This research is strongly aligned with the Sustainable Development Goals, including SDG 6, which focuses on guaranteeing universal access to potable water and sustainable management of water resources, and SDG 15, which stresses the conservation and sustainable use of terrestrial ecosystems. The article's methodology part will go over the data sources analyzed, the methodologies employed, the model assessment criteria, and a comparison of the applicable machine learning approaches. It highlights the utilization of training data derived from NDVI (Normalized Difference Vegetation Index), NDBI (Normalized Difference Built-up Index), and NDWI (Normalized Difference Water Index) as foundational elements for these models. Furthermore, this paper discusses implementing these models using E1071 and Kernlab tools, chosen for their precision, in determining the best method based on accuracy and kappa values. The results and discussions will delve into the techniques and methods applied, ultimately interpreting how these land use outcomes support the preservation of sustainable water resources in geothermal fields and contribute crucially to groundwater conservation and the advancement of SDGs 6 and 15. This introduction sets the stage for a detailed exploration aimed to understand and enhance land use strategies to underpin water sustainability initiatives, thereby fostering a sustainable future [19, 41, 42].

2. Methods

2.1. Study Area

The research site is within the Patuha geothermal field in the Bandung Regency of Indonesia, covering 25 km^2. The geothermal fields spread from the Rancabali to Pasirjambu sub-districts, encompassing an area of 351.9 km^2. Figure 1 depicts the map of the research area.

2.2. Data

The study utilized Landsat 8 and Sentinel-2 Surface Reflectance image data, and administrative areas. The Landsat 8 image data radiometric correction using ENVI 5.3 was then used as input data in Rstudio for the Machine Learning (ML) stage [43]. Sentinel-2 Surface Reflectance image data is processed through Google Earth Engine, so there were three data obtained, namely, Normalized Difference Forest Index (NDVI), Normalized Difference Built-up Index (NDBI), and Normalized Difference Water Index (NDWI). These datasets served as reference points for training data and were further processed using ArcMap 10.8, resulting in training data points in shapefile format. The ML method stage can be continued by using the corrected Landsat-8 image data and training data points as input data in Rstudio [22, 44, 45].

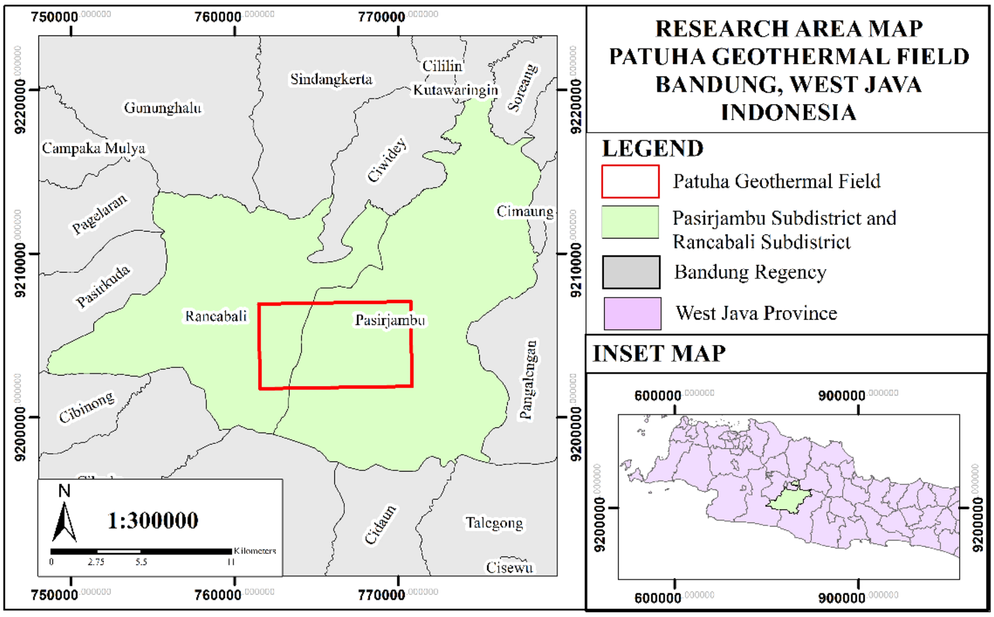


Figure 1. Research Area Map-Patuha Geothermal Field

2.3. Training Data

Training data was acquired using Normalized Difference Forest Index (NDVI), Normalized Difference Built-up Index (NDBI), and Normalized Difference Water Index (NDWI) data [22, 23, 46]. The Normalized Difference Forest Index (NDVI) is a valuable tool for analyzing land use, particularly in forest areas and agricultural plantations. NDVI is calculated using near-infrared (NIR) and visible (red) light, providing information on Forest density in the area. The formula of NDVI = (NIR-RED)/(NIR+RED) [22]. The Normalized Difference Built-up Index (NDBI) represents a specialized remote sensing tool widely used to analyze built-up or urban areas in land use studies. NDBI utilizes the spectral reflectance properties of built-up materials, which are distinct from natural land cover types. Built-up regions generally reflect more in the short-wave infrared (SWIR) band than genuine covers. On the other hand, these areas reflect less in the near-infrared (NIR) band. NDBI leverages this difference using the formula NDBI = (SWIR - NIR) / (SWIR + NIR) [23]. The Normalized Difference Water Index (NDWI) derived from remote sensing is often used to monitor and analyze water bodies associated with land use. NDWI is calculated using specific bands from satellite data that measure reflectance in the electromagnetic spectrum's green and near-infrared (NIR) portions. The formula for NDWI is typically (Green - NIR) / (Green + NIR), with variations depending on the satellite data used. NDWI distinguishes water bodies from other land surfaces due to water's absorbance and low radiation from visible to infrared wavelengths [47]. The training data was determined based on threshold data presented in Table 1.

Table 1. Training Data for Land Use Based on NDVI, NDBI, and NDWI [22, 23, 46]

Land Use Type	NDVI	NDBI	NDWI
Forest	>0.7	-	-
Agriculture and Plantation	0.65-0.7	-	-
Built-up	-	>0.2	-
Water Bodies	-	-	>0.5

2.4. Workflow Scenario

The workflow diagram (Figure 2) presents a methodical process for producing a land use model map. It starts with sourcing data from Landsat 8 OLI/TIRS Level 1 for the years 2021 to 2023, which undergoes radiometric calibration and preprocessing, including Dark Subtraction (DOS) and subsetting by Region of Interest (ROI) [48, 49]. The training and testing data were obtained from Sentinel 2 Surface Reflectance data for the same years. The focus was on indices such as NDVI for agriculture and Forest, NDBI for built-up areas, and NDWI for water bodies [22, 23, 46, 47]. The data was split into 70:30 for training and testing. Parameter tuning is conducted for machine learning (ML), and feature importance is assessed for both methods [50, 51]. The SVM algorithm uses the parameters cost c, hyperparameter sigma, number of support vectors, and training error. In contrast, the RF algorithm uses the parameters mtry, number of trees, number of variables tried at each split, and OOB estimate of error rate [52–55]. The model is then evaluated using a confusion matrix and Cohen's Kappa to assess accuracy [56, 57]. The map produced by the land use model is informed by algorithm evaluation metrics, such as Kernlab and e1071 [58].

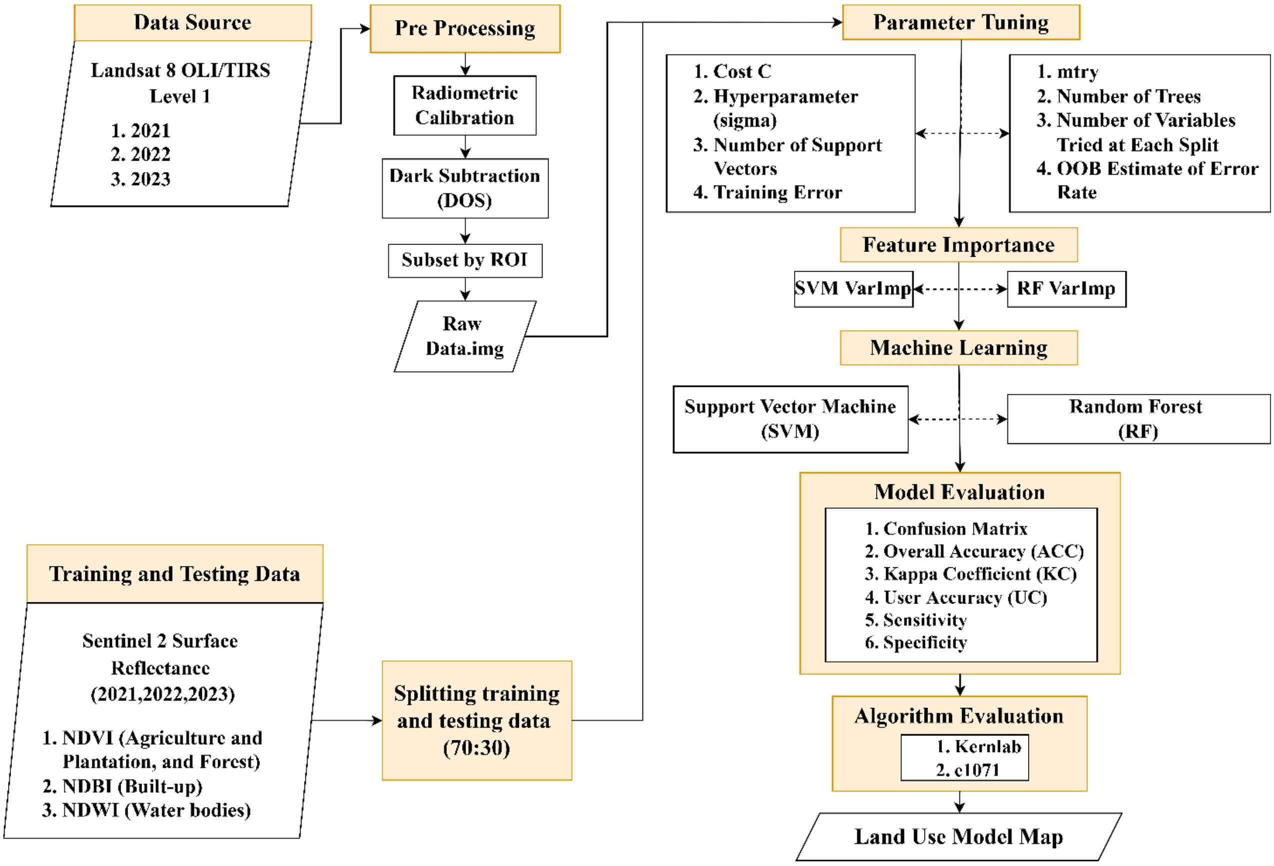


Figure 2. Workflow Diagram

2.5. Parameter Tuning

The optimization of models' performance through the adjustment of hyperparameters is essential in parameter tuning for ML models. SVM incorporates parameters such as Cost C, a regularization parameter that balances the need to model the training data accurately and keep the model simple for good performance on unseen data [50]. The sigma hyperparameter determines the kernel coefficients and influences the impact of a single training example on defining the decision boundary [59]. The number of support vectors indicates the examples used to define the decision boundary [60]. Training error reflects the model's performance on the training dataset. A low training error indicates that the model has learned well from the training data [61]. RF uses parameters such as 'mtry,' which determines the number of features considered when searching for the best split at each tree node and the number of trees, which increases the accuracy of the model to a certain extent [62]. The 'Number of Variables Attempted at Each Split' parameter guides the algorithm to consider many variables in the decision process, while the Out-of-Bag (OOB) error rate estimate provides an unbiased picture of the model's accuracy [51].

2.6. Feature Importance

The Variable Importance (VarImp) measure in machine learning algorithms ranks the importance of different predictors in the model. For SVM, the VarImp might be determined based on the weights assigned to the features in the separating hyperplane. Larger weights indicate greater importance. In RF, VarImp is typically assessed by the decrease in model accuracy or increase in node purity when a feature is permuted, which reflects its contribution to the predictive power of the model. Features with high VarImp scores are more crucial for the model's decision-making process [63, 64].

2.7. Support Vector Machine (SVM)

SVM is an advanced and adaptable supervised machine learning technique that is commonly used for classification and regression applications. At its heart, SVM seeks the ideal hyperplane that optimally separates data points of distinct classes in the feature space [65]. For data that can be separated linearly, this hyperplane is a line (in two dimensions) or a plane (in three dimensions) that maximizes the margin between the two classes. In situations when data cannot be separated linearly, SVM employs a method known as the kernel trick. This strategy entails mapping the data onto a higher dimensional space, where a hyperplane may effectively separate [66, 67]. Figure 3 illustrates the concept of SVM.

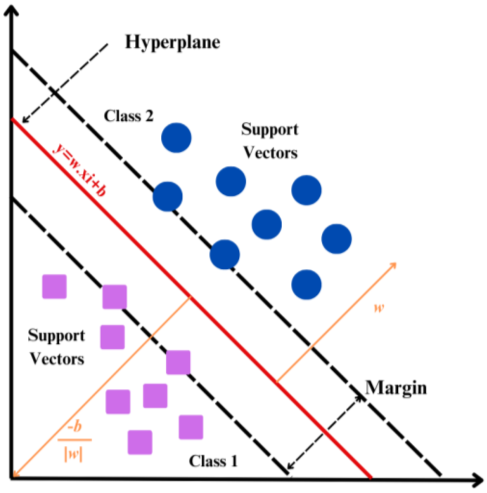


Figure 3. Concept of SVM Algorithm [68]

2.8. Random Forest (RF)

RF is an ML algorithm known for its versatility and accuracy in handling categorization and regression tasks. It belongs to the ensemble learning family, which integrates predictions from multiple algorithms to enhance overall predictive accuracy. The algorithm is based on decision trees and employs a technique known as bootstrap aggregation (or bagging) to create a forest of trees. Each tree is built from a random selection of data and characteristics. This method adds variation to the trees and decreases the possibility of overfitting, which is a typical issue in individual decision trees. During prediction, RF aggregates the outcomes of all trees, usually by majority voting for classification or averaging for regression, to arrive at a final prediction. This method enhances forecast accuracy and gives insight into the significance of variables, making it useful for a variety of applications, notably in the energy and industrial sectors. Despite its strengths, RF may be less intuitive to interpret than a single decision tree and require substantial computational resources, mainly when dealing with large datasets. Figure 4 illustrates the concept of RF [51, 69].

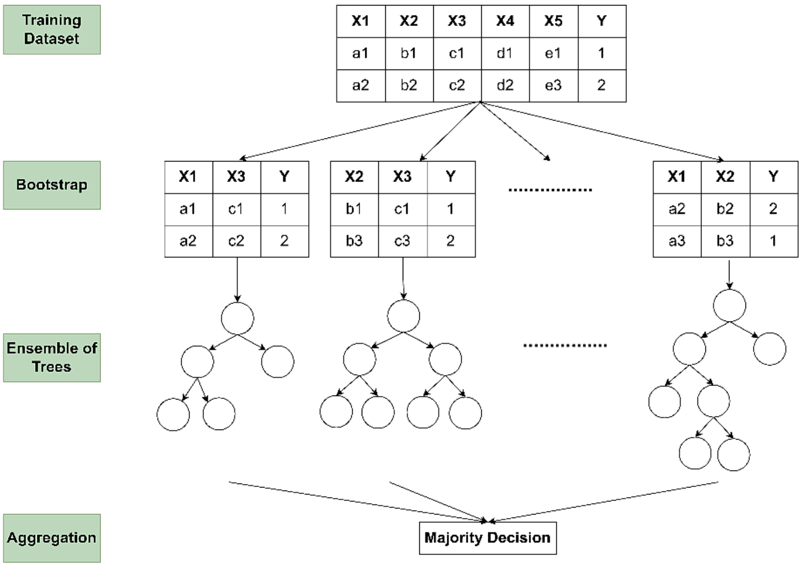


Figure 4. Concept of RF Algorithm [68]

2.9. Model Evaluation and Accuracy Assessment

Model evaluation methods are critical for determining the efficacy of ML in land use prediction models. These methods shed light on how good these models are in classifying and predicting various land uses. RF and SVM are well-known for their ability to handle complicated, multivariate datasets such as those used in land use studies. However, the accuracy and reliability of predictions depend heavily on the appropriate choice and application of model evaluation techniques. These methods determine the extent to which models accurately represent real-world phenomena they predict [51, 70].

The confusion matrix (CM) is an essential evaluation tool in ML, particularly for classification models such as SVM and RF A. The table compares the actual and predicted classifications to assess a model's performance. The matrix is divided into four segments: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). While TP and TN represent the correct predictions for the positive and negative classes, respectively, FP (Type I errors) and FN (Type II errors) represent the errors made by the model [58, 71]. Table 2 is the confusion matrix table.

Table 2. Confusion Matrix (CM) [72]

Actual/Predicted	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

CM is an important technique for evaluating the performance of classification models. Overall Accuracy (ACC) is a measure of the model's overall efficacy across all classes. It is derived as the ratio of accurate predictions (TP and TN) to total predictions produced. This metric provides a holistic view of the model's performance, encompassing its success rate across all categories [23, 73, 74]. In contrast, User Accuracy (UC), which is sometimes confused with precision in binary classifications, focuses on the model's predictions for a given class. It is calculated for each class by dividing the number of correct predictions for that class by the total number of instances predicted as belonging to that class. This metric is especially useful for assessing how effectively the model recognizes each class, emphasizing its precision in differentiating one class from others [75, 76]. The Kappa Coefficient (KC) is a statistical measure that compares observed and expected accuracy, which would be the accuracy if predictions were made randomly. It accounts for the possibility of a correct prediction occurring by chance, providing a more robust view of the model's performance, especially in cases where class distribution is imbalanced [77, 78]. Sensitivity and specificity are essential metrics for evaluating binary classification models. Sensitivity indicates the proportion of real positive instances detected by the model, whereas specificity measures the proportion of actual negative cases. There is often a trade-off between sensitivity and specificity, and choosing the optimal balance depends on the specific needs of the application domain [79]. The formula of ACC, UC, KC, Sensitivity, and Specificity are presented in Equations 1 to 6 [58, 80]:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$UC = \frac{TP_i}{TP_i + FP_i} \tag{2}$$

$$KC = \frac{ACC - P_e}{1 - P_e} \tag{3}$$

$$P_e = \sum \frac{(row\ totals * coloumn\ totals)}{total\ number\ of\ observations^2} \tag{4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

2.10. Algorithm Evaluation

R’s Kernlab and e1071 packages are essential for comparing SVM and RF algorithms. Kernlab provides a variety of kernel-based methods, especially for SVM, enabling a nuanced approach to model training and fine-tuning. It can handle various data types and structures, making it versatile for complex model comparisons. The e1071 package provides SVM capabilities and diverse ML tools that help compare different algorithms. It facilitates model training, parameter optimization, and validation across other methods. Using these packages, one can thoroughly evaluate the performance of SVM compared to RF. It includes analyzing accuracy, precision, and the capability to perform generalization from training data to unseen data. Such a comparison can reveal each method's inherent strengths and limitations, helping to select the most appropriate algorithm for a given dataset or problem [81, 82].

3. Results and Discussion

3.1. Training Data

The training data used in this study are NDVI, NDBI, and NDWI data. NDVI data is used for land use, especially for forests, agriculture, and plantations. NDBI data is used on land use, mainly built-up. At the same time, NDWI data is used for land use, especially water bodies. Figure 5 shows the NDVI, NDBI, and NDWI data overlayed with the training data points for each land use type. Based on Table 3, the training data used for forests with NDVI values greater than 0.7. NDVI is calculated using visible and near-infrared light reflected by vegetation. This index is vital for monitoring plant health because healthy plants absorb visible light and reflect near-infrared light [83, 84]. Research by Xue et al. (2021) shows that high NDVI values usually indicate dense and healthy vegetation, generally found in forest areas, while moderate NDVI values may indicate agricultural areas where vegetation density varies. In this study, the forest NDVI value interval ranged from 0.79 to 0.84 [85]. In this study, the forest NDVI value was greater than 0.7. These values indicate robust plant health and play a pivotal role in maintaining hydrological cycles and regulating surface water resources, which are crucial for sustaining geothermal activities. Forests act as natural reservoirs storing water, crucial for recharging groundwater and managing surface water flow, ultimately impacting the stability of geothermal processes [86]. The study conducted by Xue et al. (2021) shows that the NDVI values for agricultural land usually range from 0.69 to 0.75 [85]. In agricultural and plantation regions, the authors observed NDVI values ranging from 0.65 to 0.7. These moderate values represent the many crop kinds and growth phases seen in these locations and are required for the implementation of precision agricultural techniques. Precision agriculture optimizes water consumption and improves overall water resource stability in geothermal fields and local ecosystems by tailoring irrigation and fertilization tactics based on NDVI data [86].

The Normalized Difference Built-up Index (NDBI) is calculated to highlight urban land cover by analyzing the reflection properties of built-up materials in the short-wave infrared spectrum [87]. As Muhaimin et al. (2020) demonstrated, NDBI effectively maps urban density variations, which are crucial for urban planning and environmental management [18]. In this study, areas with an NDBI value greater than 0.2 were classified as built-up, indicating significant urban development. Urban expansion typically increases impervious surfaces, contributing to higher runoff and reduced groundwater recharge. These changes have profound implications on both surface and sub-surface water systems, which are essential for the efficiency and sustainability of geothermal energy operations. The ability of NDBI to detect urban materials such as concrete and asphalt makes it an essential tool for monitoring urban sprawl and guiding strategic water resource management. By integrating NDBI data into urban development plans, planners can more effectively balance urban growth with environmental conservation, enhancing the sustainability of local and broader ecosystems and ensuring that geothermal operations are supported by adequate water management practices [88, 89].

Normalized Difference Water Index (NDWI) is widely used in remote sensing to identify and monitor water features such as lakes, rivers, and reservoirs. NDWI utilizes green and near-infrared light to enhance the presence of water in satellite images [90]. Patil et al. (2024) demonstrated the effective use of the Normalized Difference Water Index (NDWI) in differentiating water bodies by harnessing specific spectral bands from Sentinel 2 satellite imagery. This research underscored the significance of NDWI in precise water body mapping and classification for land use monitoring [19]. In this research, water bodies present distinct signatures on this index, with values greater than 0.5 typically indicating water bodies. This high sensitivity to the presence of water makes NDWI particularly valuable for water resource management, aiding in the detection and monitoring of water bodies, which is essential for environmental conservation and planning [91, 92].

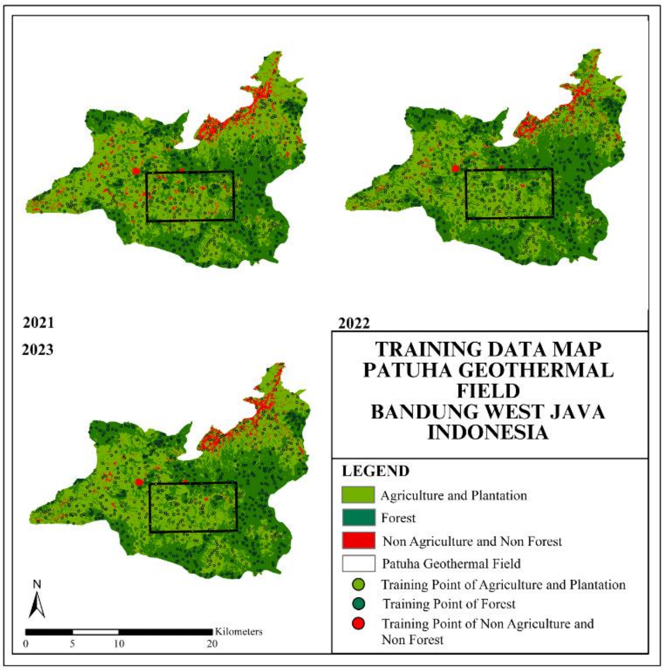


Figure 5. Training Data

Table 3. Total Training Data based on Land Use Type

Land Use Type	Interval Values			Total Training Data (per year)
	NDVI	NDBI	NDWI	
Forest	>0.7	-	-	350
Agriculture and Plantation	0.2-0.7	-	-	325
Built-up	-	>0.2	-	275
Water Bodies	-	-	>0.5	250

3.2. Parameter Tuning

Parameter optimization for the Support Vector Machine (SVM), as shown in (Table 4), from 2021 to 2023, demonstrates a strategic balance in model tuning. The consistent setting of the Cost C parameter at 1 indicates a targeted equilibrium between the model's complexity and its fitting accuracy to the training data. Concurrently, the sigma hyperparameter has been adjusted downward from 2.11 to 1.41, suggesting a refined influence of the kernel on the model's decision boundary, potentially enhancing the precision of classifications. The count of support vectors defining this boundary has varied from 272 to 230 and then slightly increased to 240. This variability indicates adjustments in the model's complexity and adaptability to new data patterns. The training error rates have remained low and stable, underscoring the model's robust performance across different datasets. Understanding these parameters' roles and impacts helps apply SVM more effectively in complex scenarios, where model accuracy and generalizability are crucial [60, 61].

Table 4. Parameter Tuning of SVM

Parameter Tuning	2021	2022	2023
Cost C	1	1	1
Hyperparameter (sigma)	2.11	1.23	1.41
Number of Support Vectors	272	230	240
Training error	0.06	0.07	0.07

In applying the Random Forest (RF) method, detailed in Table 5, the 'mtry' parameter, which specifies the number of features considered at each tree split, has varied from 2 to 7 and reverted to 2. This fluctuation represents adaptive strategies to enhance model accuracy by exploring different complexities in feature selection. Despite these changes, the number of trees has been maintained at 500 to ensure the stability and robustness of predictions. Correspondingly, the Out-of-Bag (OOB) error rate, serving as an internal validation measure to estimate prediction error, peaked at 7.5% before reducing to 6.79%. These oscillations in the OOB error rate indicate the model's sensitivity to changes in 'mtry' and give information about the best configuration for certain data sets. Such thorough parameter adjustment is critical for applications that need high model accuracy, such as ecological modeling or disease prediction. Understanding the implications of these settings can guide practitioners in refining RF models to achieve higher precision in diverse analytical contexts [51, 62].

Table 5. Parameter Tuning of RF

Parameter Tuning	2021	2022	2023
mtry	2	7	2
Number of Trees	500	500	500
Number of Variables Tried at Each Split	2	7	2
OOB Estimate of Error Rate	6.56	7.5	6.79

3.3. Feature Importance

The varImp function in R is a crucial tool for evaluating the importance of predictor variables in machine learning models, particularly Random Forest (RF) and Support Vector Machines (SVM). For RF models, varImp assesses each variable's importance based on the impact of its permutation on model accuracy and node purity—a measure of how well each split in a decision tree classifies the data. This process helps identify which features significantly influence the model's predictive accuracy. In the context of SVM, determining variable importance is more nuanced due to the model's structure. Techniques such as calculating the squared coefficients of linear SVMs or employing Recursive Feature Elimination (RFE) are utilized. RFE, for instance, iteratively removes features to assess their impact on model performance, providing a systematic approach to feature selection. Given the diversity of methods and their dependence on specific model types and R packages, thoroughly understanding of how these variable importance assessments are implemented is paramount for effectively interpreting their outcomes in practical applications [93–98].

The SVM algorithm's performance in land use classification for water bodies has shown notable variations in the importance of spectral features over the years. The classification heavily relied on bands 6 (SWIR1), 1 (Coastal/Aerosol), and 7 (SWIR2)—with Band 6 being particularly sensitive to moisture content, thus distinguishing water from land by capturing the higher absorption rates in the SWIR spectrum. Band 1 was instrumental in identifying fine particulate matter in water, distinguishing between clear and turbid water conditions, while Band 7 enhanced water detection by penetrating atmospheric obstructions like haze. The focus shifted in 2022 to include bands 6, 5 (NIR), and 7. Band 5's sensitivity to water absorption helped further differentiate water bodies from vegetation and urban areas, a crucial factor in precise land mapping for sustainable water resource management in geothermal fields. By 2023, the emphasis on bands 5, 3 (Green), and 6 emphasized the changing technological capabilities and environmental monitoring requirements; Band 3's capacity to detect water's absorption of green light made it a critical feature in differentiating aquatic from terrestrial surfaces. These changes reflect significant advancements in remote sensing technology and highlight the critical role of spectral analysis in supporting the conservation and sustainable management of water resources within geothermal fields, aligning with broader environmental monitoring and sustainability goals [99, 100].

Leveraging Landsat 8 imagery for land use classifications is pivotal in forest analysis, employing specific bands known for their unique spectral responses to effectively assess forest health and density. This dependency has developed over the years, reflecting advances in remote sensing technologies as well as shifting environmental monitoring requirements. In 2021, Bands 4 (Red), 3 (Green), and 5 (NIR) were useful for recognizing healthy vegetation, defining forest borders, and estimating forest density. The focus shifted in 2022 to include Band 2 (Blue) for its ability to detect shadows and moisture, enhancing differentiation between forested and non-forested areas, while Band 3 continued to highlight healthy vegetation and Band 1 (Coastal/Aerosol) became crucial for evaluating atmospheric effects on forest visibility. By 2023, the importance of Band 4 in chlorophyll absorption remained critical for vegetation health identification; Band 2 was pivotal in providing textural details due to its moisture and shadow sensitivity, and Band 1 was essential for atmospheric corrections to improve forest health assessments. This expanding understanding is vital for establishing adaptive forest management techniques that respond to changing climatic circumstances and enable sustainable water resource management in geothermal fields, thus improving conservation efforts and maintaining ecological balance [101, 102].

Over recent years, Landsat 8 imagery has been instrumental in enhancing land use classifications for built-up areas, providing vital data for urban planning and environmental sustainability. In 2021, bands 6 (SWIR1), 7 (SWIR2), and 4 (Red) were pivotal for assessing urban buildings because to their sensitivity to urban materials and their ability to differentiate them from natural landscapes. The SWIR bands, sensitive to moisture, contrast sharply with the reflective properties of built-up areas, while the Red band, absorbing chlorophyll, helps to separate vegetation from artificial structures. In 2022, the focus shifted to bands 6, 5 (NIR) and 7, which collectively improved the differentiation of urban from natural areas by capturing variances in vegetation presence and moisture content. By 2023, bands 5 (NIR), 2 (Blue), and 1 (Coastal/Aerosol) were highlighted for their unique capabilities to refine urban analysis further. NIR and Blue bands were crucial for discerning built-up areas from vegetation and enhancing shadow differentiation, which is important for detailed urban texture mapping. The coastal/aerosol band played a significant role in atmospheric correction, crucial for clarifying urban features. These spectral insights aid in precise urban mapping and support sustainable urban planning, which is crucial for managing water resources effectively in geothermal fields. Such planning helps mitigate the impact of urban expansion on local hydrology, which is essential for conserving water and maintaining the integrity of geothermal operations [103, 104].

Landsat 8 imagery plays a pivotal role in the detailed monitoring of agricultural lands, utilizing specific bands tailored to discern various aspects of agricultural health and land management over different years. In 2021, bands 1 (Coastal/Aerosol), 2 (Blue), and 4 (Red) were crucial, with Band 1 enhancing atmospheric corrections for clearer feature detection, Band 2 distinguishing between water bodies and crop types by sensing soil characteristics, and Band 4 identifying healthy vegetation through chlorophyll absorption. These bands collectively improved the detection of variations within agricultural landscapes. The focus in 2022 shifted to bands 7 (SWIR2), 1, and 3 (Green), where Band 7's ability to detect moisture content became invaluable for assessing soil health and irrigation needs. Band 1 continued to provide atmospheric corrections, and Band 3 highlighted plant vigor by detecting reflected green light. By 2023, the emphasis on bands 2, 3, and 1 for agriculture and plantation underscored their roles in enhancing soil and crop residue contrast, assessing plant health, and providing atmospheric corrections. These spectral analyses aid in precise crop condition assessments and contribute to water conservation strategies crucial in geothermal fields. Enhanced monitoring promotes optimum irrigation methods, minimizing wasteful water usage and contributing to the balance of local hydrological cycles, therefore maintaining the integrity and sustainability of both agricultural and geothermal resources (see Figure 6) [64, 105].

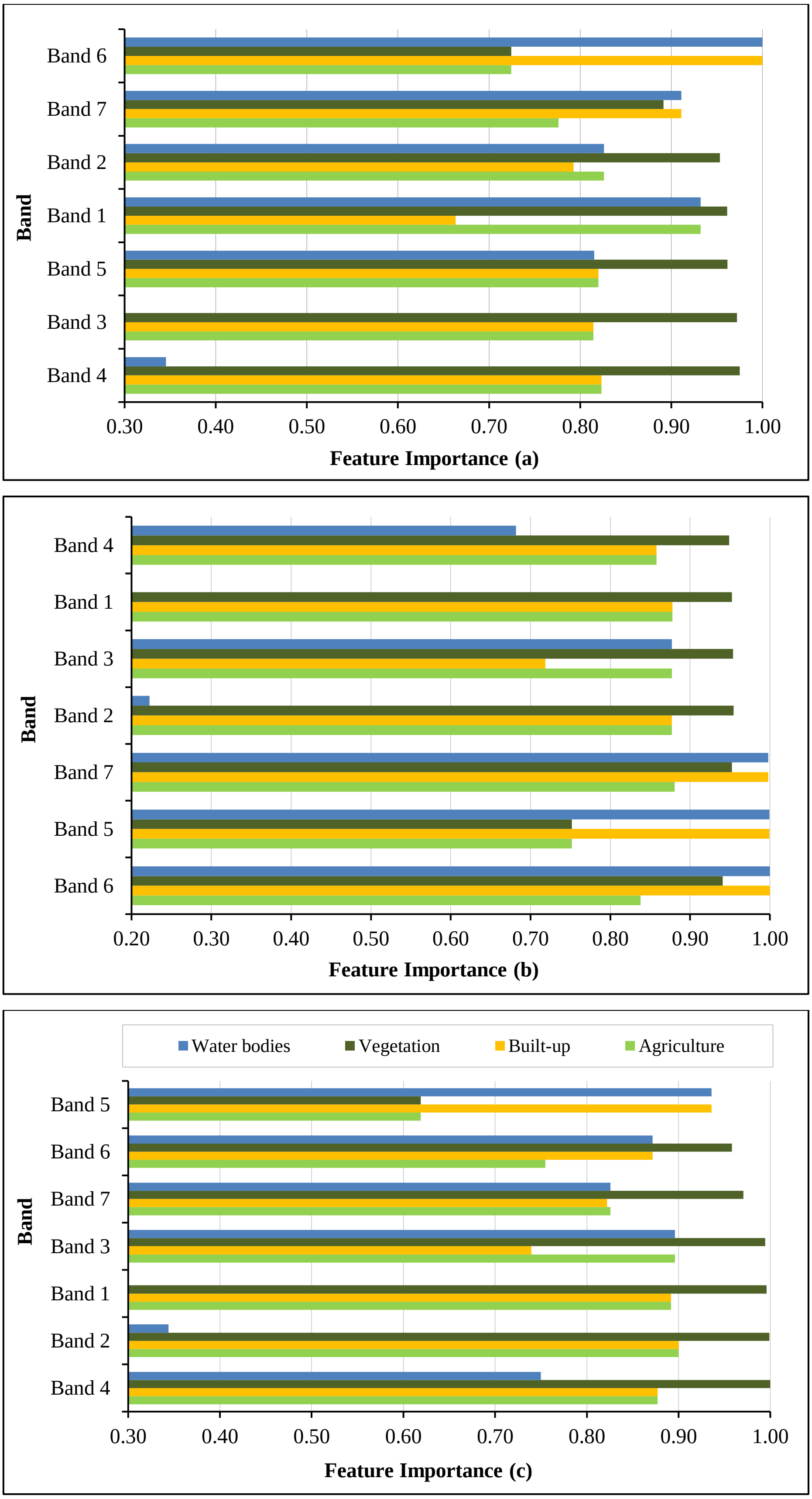


Figure 6. Feature Importance of SVM Model (a) 2021 (b) 2022 (c) 2023

The utilization of the Random Forest (RF) technique for land use classification within geothermal areas leverages the unique spectral sensitivities of Landsat bands 7 (shortwave infrared 2), 3 (green), and 6 (shortwave infrared 1), which have proven particularly effective by 2023. Band 7's sensitivity to moisture variations is crucial for detecting subtle changes in soil moisture that may indicate geothermal activity (Figure 7). At the same time, Band 3 enhances the visibility of vegetation, aiding in the differentiation of land covered by natural flora from geothermally altered terrains. Band 6 excels at penetrating particles in the atmosphere, resulting in sharper views of the ground surface, which is critical for accurate mapping in steam-rich geothermal areas. Together, these bands create a powerful toolkit for recognizing geothermal characteristics, which contributes greatly to water resource conservation by allowing for accurate monitoring and regulation of land use in vulnerable locations. This tailored approach ensures the sustainability of geothermal operations and protects the surrounding ecosystem by mitigating adverse environmental impacts [93–96].

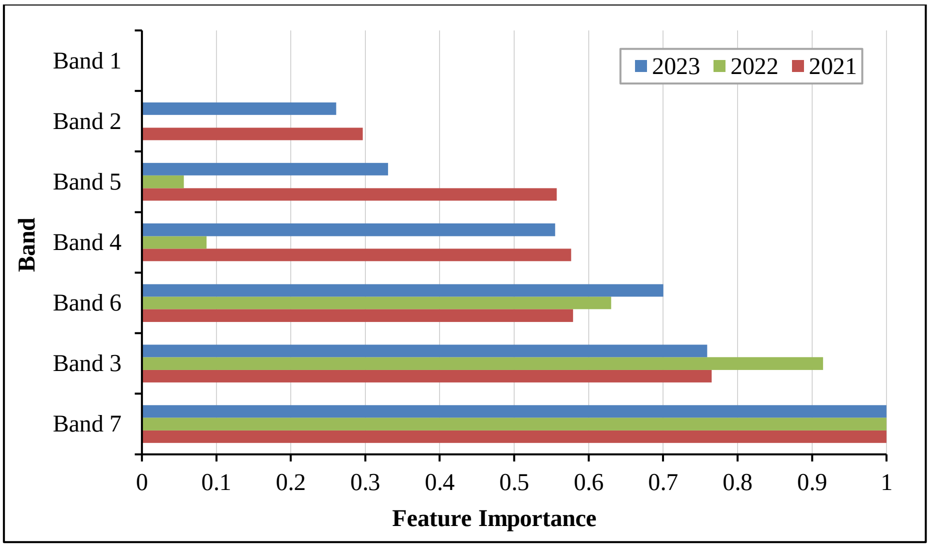


Figure 7. Feature Importance of RF Model

These technologies provide critical data that enable the conservation and sustainable management of natural and human-altered landscapes, ranging from monitoring agricultural vitality and forest health to recognizing built-up areas and water bodies. In geothermal fields, the precise application of these spectral analyses helps manage water resources effectively, ensuring the sustainability of energy production while protecting ecological balance. By continuously adapting to technological advancements and environmental changes, these spectral bands enhance the understanding of the earth's surface and empower policymakers and conservationists to make informed decisions aimed at long-term sustainability. Each application, whether focused on agriculture, water bodies, urban features, or geothermal monitoring, contributes to a holistic approach to environmental stewardship, underlining the critical integration of advanced remote sensing technology with strategic land management practices [25, 79].

3.4. Land Use Model Using SVM Algorithm Method

Using the SVM approach to monitor land use changes between 2021 and 2023, the authors detect major consequences for water sustainability in geothermal areas (Figure 8 and Table 6). The decrease in agricultural and plantation areas from 46.85% to 45.54% reflects water usage and runoff changes, crucial for managing aquifer depletion and maintaining the thermal balance essential for geothermal energy production. Concurrently, a growth in built-up areas means more impermeable surfaces, which reduces groundwater recharge and may contaminate runoff, affecting water quality, which is crucial to geothermal operations. The growth in forested areas from 4.84% to 6.26% enhances ecological conditions that stabilize groundwater levels and provide natural cooling, which benefits geothermal systems. Furthermore, fluctuations in water bodies increasing to 52.66% in 2022 and then decreasing to 48.04% in 2023 highlight thermal and hydrological dynamics shifts that directly impact geothermal production. These findings highlight the need of integrated land and water management methods that balance agricultural, urban, and conservation initiatives with the operating requirements of geothermal energy systems. Adapting to these changes with enhanced remote sensing techniques, such as those given by the SVM approach, stakeholders may assure the sustainability of geothermal resources while conserving the surrounding ecology [106–109].

Table 6. Result of SVM Model

Land Use	2021		2022		2023	
	Area (km^2)	Percentage of Area (%)	Area (km^2)	Percentage of Area (%)	Area (km^2)	Percentage of Area (%)
Agriculture and Plantation	164.81	46.85	148.51	42.21	160.20	45.54
Built-up	17.03	4.84	17.48	4.97	22.01	6.26
Forest	169.39	48.15	185.26	52.66	169.01	48.04
Water Bodies	0.57	0.16	0.56	0.16	0.58	0.16

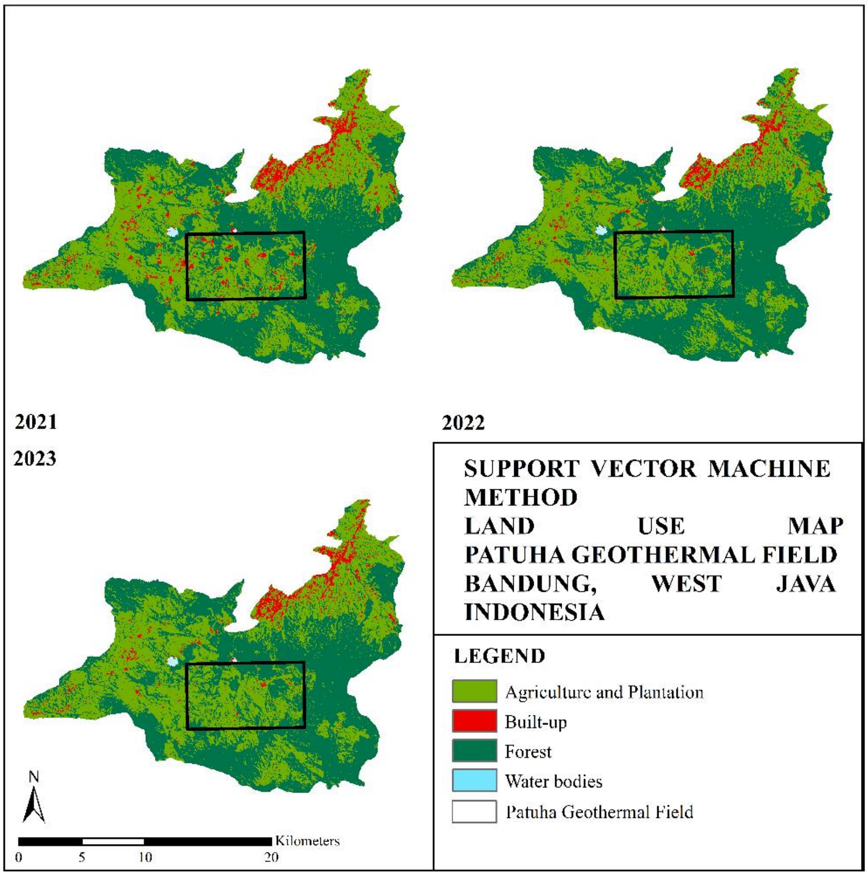


Figure 8. SVM Model

3.5. Land Use Model Using RF Algorithm Method

In the study detailed in Figure 9, the RF algorithm provides a nuanced classification of land use changes from 2021 to 2023, revealing significant fluctuations across various categories, as shown in Table 7. Agricultural and plantation areas saw a marked decrease of 7.77% from 50.93% in 2021 to 43.23% in 2022, followed by a recovery to 47.45% in 2023, reflecting shifts in land management practices that potentially impact water usage and availability in geothermal areas. Conversely, built-up land consistently increased, rising from 4.56% in 2021 to 6% in 2023, suggesting an expansion of urban areas that typically reduces permeable surfaces, thereby affecting groundwater recharge and increasing surface runoff — factors critical to maintaining the thermal stability of nearby geothermal resources. Forest land exhibited significant fluctuations, increasing by 7.47% in 2022 before dropping by 5.39% in 2023; forests play a crucial role in regulating hydrological cycles and enhancing precipitation infiltration. Thus, their expansion or reduction can directly influence the water balance necessary for geothermal operations. The proportion of land occupied by water bodies decreased slightly from 0.21% in 2021 to 0.17% in 2022 before stabilizing in 2023, a change that, while small, is significant in geothermal contexts where even minor variations in water bodies can affect geothermal plants cooling and operational efficiency. These dynamics underscore the need for integrated land and water management strategies that address the demands of land use and ensure the sustainability of water resources vital for geothermal energy production [46, 110, 111].

Table 7. Result of RF Model

Land Use	2021		2022		2023	
	Area (km^2)	Percentage of Area (%)	Area (km^2)	Percentage of Area (%)	Area (km^2)	Percentage of Area (%)
Agriculture and Plantation	179.17	50.93	152.10	43.23	166.94	47.45
Built-up	16.05	4.56	16.99	4.83	21.09	6.00
Forest	155.84	44.30	182.13	51.77	163.17	46.38
Water bodies	0.75	0.21	0.60	0.17	0.60	0.17

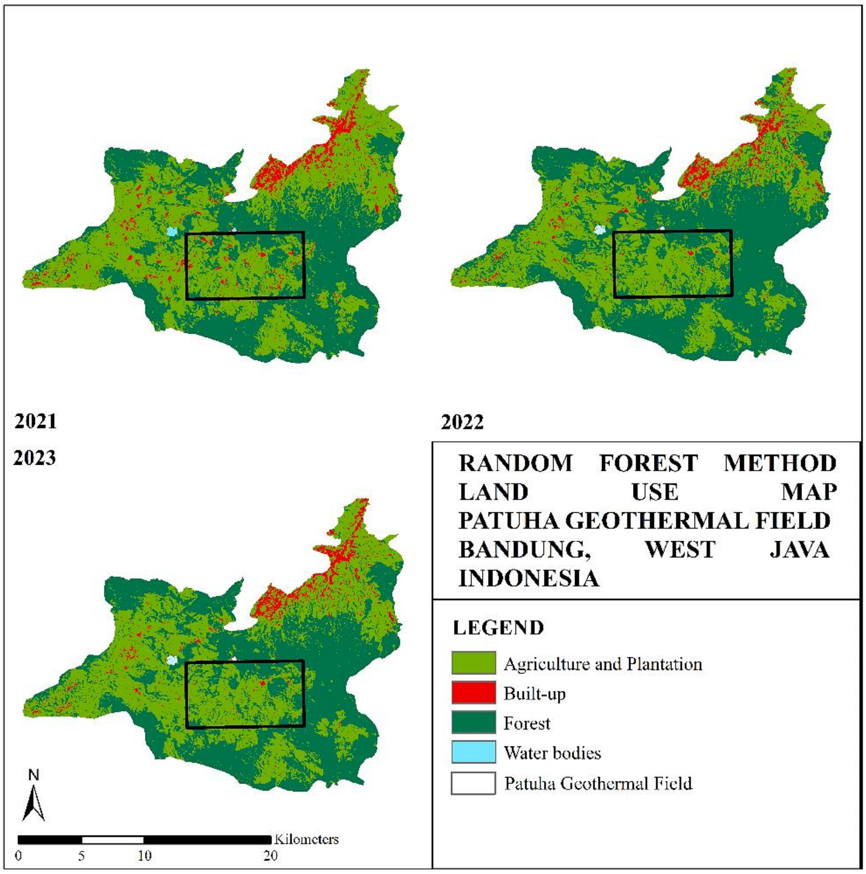


Figure 9. RF Model

3.6. Model Evaluation and Accuracy Assessment

The primary focus of the study is to evaluate the precision of SVM and RF models in classifying land use, which is crucial for effective environmental management in geothermal fields. The confusion matrices (CM) for both models give a thorough breakdown of the expected classifications compared to actual land use observations. For example, the 2021 CM for the SVM model accurately predicts 82 occurrences for agricultural and plantation, 68 for built-up areas, 98 for forest, and 70 for water bodies. Similarly, the RF model demonstrated a comparable accuracy with 84 predictions for agriculture and plantation, 68 for built-up, 96 for forest, and 70 for water bodies. These findings are critical in determining the models' performance in identifying unique land uses, which is required for monitoring and controlling land use changes that influence water sustainability in geothermal areas. Precise classification aids in accurately mapping water bodies and vegetated areas, which are critical for maintaining the hydrological balance and ensuring the availability of adequate water resources for geothermal operations. The detailed results of these classifications are further elaborated in Table 8, illustrating the significant role of accurate land use mapping in supporting sustainable geothermal practices [37, 71, 106].

Machine learning methodologies, as quantified by their respective accuracy (ACC) scores, are crucial for assessing the performance of land use classification models. These scores indicate each model's ability to correctly identify and classify diverse land cover types—ranging from agricultural areas and urban developments to forests and water bodies. Over the period from 2021 to 2023, ACC scores for the SVM method were 0.89, 0.94, and 0.92, while the RF method consistently demonstrated higher accuracy with scores of 0.89, 0.95, and 0.95. The RF model's higher performance, as seen by comparing green bars peaking at an ACC of 0.95, demonstrates its increased competence in comprehensive and dependable land use categorization. Such precision in land use mapping is particularly significant for managing water resources in geothermal fields. Accurately detecting and monitoring land cover types influence water conservation strategies and the sustainable governance of natural resources. By ensuring accurate land cover classification, both models, mainly the RF, contribute to effective environmental surveillance and urban planning, which are critical in maintaining the ecological balance and supporting water sustainability in geothermal operations. These findings, detailed in Table 9, highlight the efficacy of these models and their practical implications for improving geothermal field management through precise land use monitoring [38, 112, 113].

The Kappa Coefficient (KC) is an important performance parameter in assessing land use classification models because it measures the degree of agreement between observed and anticipated classifications while controlling for the possibility of chance agreement. From 2021 to 2023, the SVM method recorded KC scores of 0.85, 0.92, and 0.9, while the RF method showed superior performance with scores of 0.85, 0.93, and 0.93. These results, detailed in Table 9, highlight the consistently higher accuracy of the RF method compared to the SVM method across the evaluated years. This enhanced precision is particularly significant in geothermal fields where accurate land use classification is critical for effective water resource management. Higher KC values indicate a more reliable model for predicting land cover changes directly affecting water runoff patterns, groundwater recharge, and overall water conservation strategies. The RF model's ability to precisely identify land use promotes better informed decisions in environmental planning and management, guaranteeing that actions to conserve water and preserve ecological balance in geothermal fields are based on exact and reliable data [11, 38].

Table 8. CM Results

Algorithm	Land Use	Agriculture and Plantation	Built-up	Forest	Water bodies
2021					
SVM	Agriculture and Plantation	82	9	7	0
	Built-up	7	68	0	5
	Forest	8	2	98	0
	Water bodies	0	3	0	70
RF	Agriculture and Plantation	84	10	9	2
	Built-up	9	68	0	3
	Forest	4	1	96	0
	Water bodies	0	3	0	70
2022					
SVM	Agriculture and Plantation	86	1	3	0
	Built-up	0	79	0	4
	Forest	11	1	102	1
	Water bodies	0	1	0	70
RF	Agriculture and Plantation	88	1	3	0
	Built-up	2	79	0	4
	Forest	7	0	102	0
	Water bodies	0	2	0	71
2023					
SVM	Agriculture and Plantation	86	3	6	0
	Built-up	3	75	0	3
	Forest	8	2	99	0
	Water bodies	0	2	0	72
RF	Agriculture and Plantation	93	5	4	0
	Built-up	2	73	0	2
	Forest	2	2	101	0
	Water bodies	0	2	0	73

Table 9. ACC and KC Results

Algorithm	Model Evaluation	2021	2022	2023
SVM	ACC	0.89	0.94	0.92
	KC	0.85	0.92	0.90
RF	ACC	0.89	0.95	0.95
	KC	0.85	0.93	0.93

Table 10 illustrates that the SVM and RF models exhibit robust performance across various land use categories, though with distinct strengths. For agriculture and plantation, the SVM model's User's Accuracy (UC) increases from 0.84 to 0.91, reflecting its improved precision in classifying these crucial areas over time. In built-up areas, SVM shows significant progress, starting with a UC of 0.85 and rising to 0.95, indicating enhanced accuracy in urban classification—a key factor in urban water management by reducing impervious surfaces and promoting sustainable land development. SVM also maintains high UC values for forest and water bodies, with the latter reaching an impressive UC of 0.97, which is crucial for accurate water body mapping and directly impacts water conservation strategies in geothermal fields. In contrast, the RF model consistently outperforms in categorizing water categories, with a UC of 0.97, demonstrating its capacity to accurately identify water bodies required for maintaining thermal balance and water quality in geothermal operations. The consistent high performance of RF across all categories ensures dependable land use data, which is essential for comprehensive environmental monitoring and sustainable resource management in geothermal regions. This precise classification is vital for implementing targeted conservation practices, managing hydrological cycles, and supporting geothermal sustainability initiatives [40, 114, 115].

Table 10. UC Results

Algorithm	Land Use	2021	2022	2023
SVM	Agriculture and Plantation	0.84	0.96	0.91
	Built-up	0.85	0.95	0.93
	Forest	0.91	0.89	0.91
	Water bodies	0.96	0.99	0.97
RF	Agriculture and Plantation	0.80	0.96	0.91
	Built-up	0.85	0.93	0.95
	Forest	0.95	0.94	0.96
	Water bodies	0.96	0.97	0.97

The Sensitivity analysis of machine learning models, as shown in Table 11, underscores the robust performance of the RF model in accurately identifying true positives across various land use categories, which is crucial for effective environmental management in geothermal fields. RF displays commendable stability in agriculture and plantation categories, achieving a peak Sensitivity of 0.91. This high Sensitivity ensures that agricultural areas, critical for water management and conservation practices, are accurately monitored. Both models exhibit an upward trend in Sensitivity for built-up areas. RF slightly outperforms SVM with a Sensitivity of 0.96, indicating superior accuracy in detecting urban expansions that potentially impact water runoff and infiltration. In forest land use, RF's Sensitivity reaches 0.97, slightly higher than SVM, which is vital for preserving forest ecosystems that play a significant role in the hydrological cycle affecting geothermal sustainability. Both models show high Sensitivity scores of around 0.97 for water bodies, reflecting their effective detection capabilities crucial for managing water resources in geothermal operations. RF's excellent sensitivity ratings indicate its usefulness in precisely categorizing land use types, hence facilitating precise monitoring and management approaches that contribute to sustainable water usage and conservation measures in geothermal areas [116–118].

The comparative analysis of Specificity between SVM and RF models, as detailed in Table 11, showcases consistently high-performance levels across various land use categories, with Specificity values frequently nearing 1.00. It indicates both models' outstanding ability to correctly identify true negatives, a crucial aspect for accurate land use classification. In the agriculture and plantation category, while SVM started with a Specificity of 0.94, RF initially displayed a slightly lower Specificity but improved significantly over time to match SVM's performance. Both models showed progressive improvements in built-up areas, with RF achieving a Specificity of 0.98, slightly higher than SVM's 0.96. This edge for RF suggests a more reliable exclusion of non-built-up areas, which is critical for urban planning and water runoff management. Both models also had exceptionally high Specificity in the forest and water body categories, with RF peaking at 0.99, suggesting their efficiency in preventing places outside these categories from being incorrectly categorized as such. These high Specificity ratings are especially essential in geothermal sectors, where exact land cover delineation is critical for effective water resource management and promoting sustainable environmental practices. By accurately identifying non-target areas, these models help optimize land management practices and conserve water by preventing unnecessary allocation of resources to incorrectly classified areas [40, 119, 120].

According to the data, the RF model consistently outperformed the SVM in most measures across the research period, with better ACC, KC, and Sensitivity scores. This superior performance is crucial for identifying precisely land use changes that directly impact water sustainability strategies in geothermal operations. For example, RF's high sensitivity in detecting water bodies, as well as its constant UC and specificity in categorizing agricultural and wooded terrain, guarantee that vital water conservation measures may be precisely targeted and implemented [46, 50].

These findings underscore the RF model's robustness in managing complex environmental data, making it a preferred choice for geothermal field management where precise land use classification aids in effectively monitoring and conservation of water resources. By ensuring high accuracy in land classification, the RF model supports sustainable practices by helping to maintain the ecological balance and optimizing land and water use in these sensitive areas [121].

Overall, the detailed evaluation and comparison of SVM and RF models based on the discussed metrics confirm the theoretical approach employed in the study and support the outcome of effective land use monitoring essential for sustaining water resources in geothermal fields. This precision in classification enhances environmental surveillance and urban planning and plays a pivotal role in maintaining the hydrological and thermal stability necessary for the continued viability of geothermal energy production [39].

Table 11. Sensitivity and Specificity Results

Algorithm	Model Evaluation	Land Use	2021	2022	2023
SVM	Sensitivity	Agriculture and Plantation	0.85	0.89	0.89
		Built-up	0.83	0.96	0.91
		Forest	0.93	0.97	0.94
		Water bodies	0.93	0.93	0.96
RF		Agriculture and Plantation	0.87	0.91	0.96
		Built-up	0.83	0.96	0.89
		Forest	0.91	0.97	0.96
		Water bodies	0.93	0.95	0.97
SVM	Specificity	Agriculture and Plantation	0.94	0.98	0.97
		Built-up	0.96	0.99	0.98
		Forest	0.96	0.95	0.96
		Water bodies	0.99	1.00	0.99
RF		Agriculture and Plantation	0.92	0.98	0.97
		Built-up	0.96	0.98	0.99
		Forest	0.98	0.97	0.99
		Water bodies	0.99	0.99	0.99

3.7. Algorithm Evaluation

Algorithm evaluation results, detailed in Table 12, reveal that the RF model, assessed using the e1071 package, consistently outperforms the SVM model evaluated by the Kernlab package across several performance metrics from 2021 to 2023. In the latest assessment of 2023, the RF model achieved a peak accuracy of 0.97, surpassing the SVM's highest accuracy of 0.94. Similarly, the kappa value for RF reached 0.96 in 2023, indicating a more substantial classification agreement than SVM's maximum kappa of 0.95. These results highlight the RF model's better precision and dependability in categorizing land use, making it ideal for jobs that need high accuracy, such as the exact mapping of land cover types in geothermal regions. Accurate land classification is critical for implementing effective water conservation strategies, as it ensures that land use changes—key indicators of water demand and watershed management—are closely monitored. Using the RF model, geothermal energy stakeholders may improve their environmental management methods, ensuring that water resources are allocated optimally and ecological balance is maintained within geothermal zones. This high level of classification accuracy aids in preventing misallocation of resources and contributes to sustainable water management practices essential for the long-term viability of geothermal operations [81, 82].

Table 12. Algorithm Evaluation

Algorithm	Evaluation Algorithm	Evaluation Parameter	2021	2022	2023
SVM (Kernlab packages)	Accuracy	Minimum	0.89	0.88	0.89
		1st Quartile	0.90	0.91	0.91
		Median	0.92	0.92	0.92
		Mean	0.91	0.93	0.92
		3rd Quartile	0.92	0.94	0.93
		Maximum	0.94	0.96	0.93
	Kappa	Minimum	0.85	0.84	0.85
		1st Quartile	0.87	0.89	0.88
		Median	0.89	0.90	0.89
		Mean	0.88	0.90	0.89
		3rd Quartile	0.90	0.92	0.90
		Maximum	0.92	0.95	0.91
RF (e1071 packages)	Accuracy	Minimum	0.89	0.89	0.90
		1st Quartile	0.92	0.91	0.92
		Median	0.94	0.93	0.94
		Mean	0.93	0.93	0.93
		3rd Quartile	0.94	0.93	0.94
		Maximum	0.96	0.95	0.97
	Kappa	Minimum	0.85	0.85	0.86
		1st Quartile	0.90	0.88	0.90
		Median	0.92	0.90	0.91
		Mean	0.91	0.90	0.91
		3rd Quartile	0.92	0.91	0.92
		Maximum	0.95	0.94	0.96

3.8. Discussion

This study stands out for its meticulous execution in classifying land use types using distinct spectral indices—NDVI for vegetation density, NDBI for urban development, and NDWI for water presence, as outlined in Figure 5 and Table 3. High NDVI values (>0.7), indicative of dense forest vegetation, play a critical role in regulating hydrological cycles and surface water resources crucial for geothermal operations. Conversely, agricultural areas exhibit NDVI values ranging from 0.65 to 0.7, reflecting varied crop densities and stages, which are essential for precision agriculture practices that optimize water usage in geothermal fields [83, 84]. Urban areas identified with higher NDBI values (>0.2) highlight increased impervious surfaces contributing to runoff and reduced groundwater recharge, impacting water management in geothermal settings [88, 89]. Additionally, accurately detecting water bodies with NDWI values over 0.5 is crucial for maintaining water availability for geothermal plants [91, 92]. These spectral indices, which have been validated against previous research by Xue et al. (2021) and Muhaimin et al. (2020), not only improve the classifier's accuracy but also help sustainable management by giving specific, actionable data for environmental conservation and geothermal sustainability. This systematic approach ensures that land use monitoring through advanced remote sensing techniques directly supports the conservation efforts necessary for the sustainable operation of geothermal facilities [18, 85].

The performance of both SVM and RF models is intricately tied to parameter tuning, as evidenced by Table 4 and Table 5. In the case of SVM, maintaining a consistent Cost C value ensures stable regularization, while adjusting sigma values allows for sensitivity modifications in the kernel function to capture evolving data characteristics. Variations in support vector numbers over time indicate model development and increased generalization capabilities. In contrast, RF parameter optimization entails adjusting'mtry' values to respond to changing feature spaces and data distributions while keeping the number of trees constant for stability. Variations in out-of-bag error rates reflect year-to-year changes in model robustness. These parameter adjustments are crucial for optimizing land use classification accuracy and ensuring the models' predictive performance accurately reflects changing environmental conditions. Understanding and fine-tuning these characteristics allows practitioners to effectively employ SVM and RF models in various analytical scenarios, facilitating accurate land use monitoring and sustainable water management methods in geothermal areas [60–62].

The examination of feature importance outcomes for machine learning techniques reveals distinct patterns between SVM and RF models, as shown in Figure 6 and Figure 7. While SVM shows annual variability in band selections for different land use types, indicating a flexible approach to capturing relevant spectral characteristics, RF shows greater consistency, particularly for water bodies and built-up areas, implying a reliance on a stable set of significant predictors over time. This variation in feature selection patterns highlights the different tactics used by each algorithm. SVM's dynamic approach may show sensitivity to annual variations in land cover and climatic circumstances, which might affect its performance. In contrast, RF's preference for certain bands annually indicates a stable criterion for feature importance, contributing to its consistent performance. These findings are consistent with earlier studies, emphasizing the necessity of understanding algorithmic variances when evaluating land use categorization results. While SVM may offer flexibility in adapting to changing environmental variables, RF's dependence on robust predictors adds to its consistency across different datasets. Further research could explore ways to leverage the strengths of both algorithms for improved land use monitoring and water resource management in geothermal areas [93–98].

The comparison between SVM and RF methods in land use classification reveals nuanced trends and spatial variations, as outlined in Tables 6 and 7. Notably, agricultural and plantation areas show divergent trends between SVM and RF, with SVM demonstrating a decline followed by an increase. At the same time, RF indicates a general decrease followed by an increase. In contrast, both methods show consistent increases in built-up areas, albeit RF shows a more pronounced rise. Fluctuations in forestland are noted for both SVM and RF, with RF exhibiting a larger increase followed by a decline. In contrast, water bodies show rather constant patterns with minimal fluctuations in area for both techniques. These differences highlight the divergent methods of SVM and RF to evaluating land use changes over time. The disparities in results might be attributed to each model's unique algorithms and techniques. SVM's boundary-based classification may excel in less complex or well-separated data, while RF's ensemble method may handle overlapping classes and complex structures more effectively. Variations in sensitivity to outliers, capacity to handle non-linear connections, and quantity of features can all have an impact on performance. Furthermore, discrepancies in model-specific parameters and their calibration for the dataset might cause variances in classification results. Understanding these nuances is crucial for effective land use monitoring and water resource management in geothermal areas, where each land use type plays a distinct role in influencing hydrological cycles and groundwater recharge. Further research could explore the specific impacts of each land use type on water sustainability in geothermal fields, informing targeted conservation and management strategies [46, 110, 111].

The comparison between SVM and RF models across various metrics underscores the RF model's consistent outperformance, with higher ACC, KC, and Sensitivity scores throughout the study period. This superiority is pivotal for accurately identifying land use changes that directly influence water sustainability in geothermal operations [38, 112, 113]. In detecting water bodies and maintaining high UC and Specificity for agricultural and forested lands, RF's robust performance ensures precise targeting and implementation of water conservation measures [11, 38]. These findings highlight RF's efficacy in managing complex environmental data, making it a preferred choice for geothermal field management where accurate land use classification supports effective water resource monitoring and conservation [40, 114, 115]. Furthermore, the extensive evaluation of SVM and RF models verifies the theoretical approach used in the work, underlining the need of accurate land use monitoring for preserving water supplies in geothermal areas. The superior performance of RF in terms of ACC, KC, UC, Sensitivity, and Specificity further validates its efficacy in accurately classifying land use types, crucial for maintaining hydrological and thermal stability in geothermal energy production [116–118]. This comparative analysis contributes to advancing the understanding of machine learning algorithms' performance in land use classification and provides valuable insights for implementing sustainable water management strategies in geothermal areas [40, 119-121].

In comparing the performance of SVM and RF algorithms in monitoring land use, a clear trend of improvement in accuracy metrics over time emerges from various studies. Specifically, Rana et al.’s (2020) research reported ACC values of 0.64 for SVM and 0.7 for RF, whereas the current findings show a significant increase to 0.94 and 0.95, respectively [41]. Similarly, Pandit et al., 2024, highlight the robustness of RF with a top ACC of 0.9 and a Kappa Coefficient of 0.9, which the authors have found to increase to 0.95 and 0.93, respectively, in this study [122]. Zhao et al. (2024) also noted a range in User Accuracy from 0.6 to 0.9 for both algorithms, which the data further refines to 0.84 to 0.99. These comparisons not only highlight the advancements in algorithmic precision, but also underline the superiority of RF in the context of land use monitoring, continuously obtaining greater accuracy across varied investigations [42].

The evaluation of algorithm performance, as depicted in Table 12, underscores the superiority of the RF model over the SVM model for land use classification tasks in geothermal fields. Utilizing the e1071 package for RF evaluation and the Kernlab package for SVM assessment, the RF model consistently outperforms SVM across multiple metrics from 2021 to 2023. In 2023 specifically, RF achieved a peak accuracy of 0.97, surpassing SVM's highest accuracy of 0.94, while maintaining a higher kappa value throughout, reaching 0.96 compared to SVM's maximum of 0.95. These results emphasize the RF model's precision and reliability in accurately classifying land use types. It is particularly suitable for tasks demanding high accuracy, such as mapping land cover types in geothermal regions. Accurate land classification is essential for implementing effective water conservation strategies, as it enables the close monitoring of land use changes, which are crucial indicators of water demand and watershed management. Utilizing the RF model, geothermal energy stakeholders may improve their environmental management methods, optimize water resource allocation, and preserve ecological balance within geothermal zones. This high level of classification accuracy not only prevents misallocation of resources but also contributes to sustainable water management practices vital for the long-term sustainability of geothermal operations [81, 82].

The study concludes that the RF algorithm is the best alternative for land use categorization in geothermal zones, outperforming SVM in a variety of land use categories. Its capacity to handle difficult categorization tasks demonstrates its usefulness for precisely monitoring land use trends and their consequences for groundwater sustainability. Given the enormous influence of land use on groundwater management, particularly in recharge areas, enacting policies that prioritize their conservation is crucial. Protecting and expanding critical recharge areas, controlling urban growth, and encouraging water-efficient agricultural practices are essential measures indicated by RF classification results. Continued RF utilization for monitoring can help detect changes in land use patterns over time and enable the development of adaptive management strategies. Moreover, aligning groundwater conservation efforts with SDGs 6 and 15 underscores their contribution to broader environmental and societal objectives. In order to enhance outcomes, future studies could explore integrating additional remote sensing techniques or algorithms and assess the socioeconomic impacts of recommended land use policies. Such endeavors would help to increase the understanding and implementation of sustainable groundwater management strategies in geothermal areas. These efforts are essential for forecasting land use trends, predicting water balance calculations, and monitoring subsurface fluids by integrating surface and subsurface data over the next five years, thus contributing to the sustainable management of geothermal resources and environmental conservation [11, 38, 40, 109].

4. Conclusions

In this work, the authors evaluated the usefulness of SVM and RF algorithms for monitoring land usage in geothermal zones, which is critical for groundwater conservation and advancing SDGs 6 (Clean Water and Sanitation) and 15 (Life on Land). Key findings include:

- The strategic selection of training data using spectral indices—such as high NDVI for forests, varied NDVI for agriculture, high NDBI for urban areas, and high NDWI for water bodies—is essential for the precise classification of land use, providing a robust foundation for distinguishing between different land cover types.
- Optimal tuning of model parameters significantly enhances the performance of both SVM and RF models. SVM benefits from stable regularization and adaptive kernel sensitivity, while RF excels with a consistent ensemble size and dynamic feature selection, both pivotal for accurate land use classification.
- The feature significance in land use classification shows distinct trends between SVM and RF. SVM adjusts annually to changes in spectral attributes. In contrast, RF shows a preference for specific bands, particularly in water bodies and built-up areas, highlighting RF's advantage in handling static feature sets.
- Disparities in land use classification trends between SVM and RF were observed; SVM showed irregular patterns in agricultural areas. RF displayed more uniform trends and outperformed SVM in urban settings, suggesting RF's superior ability to manage complex data structures.
- Overall, RF models surpass SVM in accuracy, stability, and precision in land use classification. The ensemble approach of RF, which amalgamates multiple decision trees, evidently enhances its capacity to interpret complex patterns, thus achieving higher accuracy in various land use categories.
- Comprehensive analysis from 2021 to 2023 consistently demonstrated RF's superiority over SVM in terms of both accuracy and kappa coefficient. Consequently, for high-stakes classification tasks requiring precision and reliability, the RF model, particularly with the e1071 package, is recommended over the SVM model assessed with the Kernlab package.

Finally, the RF algorithm's outstanding performance in classifying land use inside geothermal regions demonstrates its vital role in encouraging sustainable groundwater management. The RF model contributes significantly to the SDGs for clean water and biodiversity by directing the implementation of land use policies that support groundwater replenishment as well as ensuring the sustainable management and conservation of groundwater resources.

4.1. Limitations and Post-Study Research Plan

While this research lays a solid foundation for managing geothermal land use through the integration of satellite imagery and machine learning techniques, it is important to acknowledge certain limitations and outline future steps. Despite its contributions, the study may need help in accurately forecasting land use trends and predicting water balance calculations over the next five years due to the dynamic nature of environmental factors. Furthermore, monitoring subsurface fluids by combining surface and subsurface data may be complicated and need additional methodological refinement. Moving forward, continued utilization of RF for monitoring can facilitate the detection of changes in land usage patterns and contribute to the development of adaptive management strategies. Aligning groundwater conservation efforts with SDGs 6 and 15 underscores their broader environmental and societal significance. In order to improve the outcomes, future studies should explore integrating additional remote sensing techniques or algorithms and assessing the socioeconomic impacts of recommended land use policies. These efforts are crucial for developing the understanding and implementation of sustainable groundwater management practices in geothermal regions, therefore contributing to the sustainable management of geothermal resources and environmental preservation.

5. Declarations

5.1. Author Contributions

Conceptualization, W.U. and I.M.A.; methodology, R.F.I.; software, R.F.I.; validation, R.F.I.; formal analysis, R.F.I.; investigation, R.F.I.; resources, R.F.I.; data curation, R.F.I.; writing—original draft preparation, R.F.I.; writing—review and editing, R.F.I.; visualization, S.A.G. and D.P.N.P.; supervision, W.U., M.H., and I.M.A.; project administration, S.A.G. and D.P.N.P.; funding acquisition, W.U. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The satellite imagery data utilized in this study consists of Landsat 8 and Sentinel 2 data. Landsat 8 OLI TIRS data can be obtained from the publicly accessible website https://earthexplorer.usgs.gov/. Similarly, Sentinel 2 data can be accessed from the European Union/ESA/Copernicus/SentinelHub platform.

5.3. Funding

The researchers received financial support from Institut Teknologi Sepuluh Nopember (ITS) as part of the World Class University 2023 Outbound Researcher Mobility (ORM) Scheme to aid in the conduct, writing, and dissemination of this article.

5.4. Acknowledgements

The authors express their gratitude to Institut Teknologi Sepuluh Nopember (ITS) for their financial assistance provided through the World Class University 2023 program with the Outbound Researcher Mobility (ORM) Scheme. The authors also acknowledge Universiti Teknologi Petronas (UTP) for their significant contribution as a research partner. Furthermore, the authors extend their appreciation to all individuals who contributed to the successful completion of this research project.

5.5. Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

6. References

[1] Brahim, F. Ben, Boughariou, E., Makni, J., & Bouri, S. (2020). Evaluation of groundwater hydrogeochemical characteristics and delineation of geothermal potentialities using multi criteria decision analysis: Case of Tozeur region, Tunisia. Applied Geochemistry, 113(August 2019), 104504. doi:10.1016/j.apgeochem.2019.104504.

[2] Chandrasekharam, D., Lashin, A., Al Arifi, N., Al-Bassam, A. M., & Chandrasekhar, V. (2020). Geothermal energy for sustainable water resources management. International Journal of Green Energy, 17(1), 1–12. doi:10.1080/15435075.2019.1685998.

[3] Łukasiewicz, E., & Shamoushaki, M. (2022). Heating potential of undeveloped geothermal water intakes in Poland in the context of sustainable development and air protection. Water Resources and Industry, 27, 100175. doi:10.1016/j.wri.2022.100175.

[4] Solano-Olivares, K., Santoyo, E., & Santoyo-Castelazo, E. (2024). Integrated sustainability assessment framework for geothermal energy technologies: A literature review and a new proposal of sustainability indicators for Mexico. Renewable and Sustainable Energy Reviews, 192, 114231. doi:10.1016/j.rser.2023.114231.

[5] Iqbal, M., Al-Hassan, M. A., Herdianita, N. R., & Juliarka, B. R. (2023). Determining recharge area in ULUBELU geothermal field, LAMPUNG, Indonesia using stable isotope data. Applied Geochemistry, 156, 105763. doi:10.1016/j.apgeochem.2023.105763.

[6] Zhang, Y., Xiao, Y., Yang, H., Wang, S., Wang, L., Qi, Z., Han, J., Hao, Q., Hu, W., & Wang, J. (2024). Hydrogeochemical and isotopic insights into the genesis and mixing behaviors of geothermal water in a faults-controlled geothermal field on Tibetan Plateau. Journal of Cleaner Production, 442, 140980. doi:10.1016/j.jclepro.2024.140980.

[7] Ali, B., Hedayati-Dezfooli, M., & Gamil, A. (2023). Sustainability assessment of alternative energy power generation pathways through the development of impact indicators for water, land, GHG emissions, and cost. Renewable and Sustainable Energy Reviews, 171, 113030. doi:10.1016/j.rser.2022.113030.

[8] Cheng, C., Zhang, F., Shi, J., & Kung, H. Te. (2022). What is the relationship between land use and surface water quality? A review and prospects from remote sensing perspective. Environmental Science and Pollution Research, 29(38), 56887–56907. doi:10.1007/s11356-022-21348-x.

[9] van der Laan, E., Nunes, J. P., Dias, L. F., Carvalho, S., & Mendonça dos Santos, F. (2023). Assessing the climate change adaptability of sustainable land management practices regarding water availability and quality: A case study in the Sorraia catchment, Portugal. Science of the Total Environment, 897, 165438. doi:10.1016/j.scitotenv.2023.165438.

[10] Anthony, T., Shohan, A. A. A., Oludare, A., Alsulamy, S., Kafy, A. Al, & Khedher, K. M. (2024). Spatial analysis of land cover changes for detecting environmental degradation and promoting sustainability. Kuwait Journal of Science, 51(2), 100197. doi:10.1016/j.kjs.2024.100197.

[11] Rash, A., Mustafa, Y., & Hamad, R. (2023). Quantitative assessment of Land use/land cover changes in a developing region using machine learning algorithms: A case study in the Kurdistan Region, Iraq. Heliyon, 9(11), 21253. doi:10.1016/j.heliyon.2023.e21253.

[12] Carrasco, L., O'Neil, A. W., Daniel Morton, R., & Rowland, C. S. (2019). Evaluating combinations of temporally aggregated Sentinel-1, Sentinel-2 and Landsat 8 for land cover mapping with Google Earth Engine. Remote Sensing, 11(3), 288. doi:10.3390/rs11030288.

[13] Mao, L., Li, M., & Shen, W. (2020). Remote sensing applications for monitoring terrestrial protected areas: Progress in the last decade. Sustainability (Switzerland), 12(12), 5016. doi:10.3390/su12125016.

[14] Kayad, A., Sozzi, M., Gatto, S., Marinello, F., & Pirotti, F. (2019). Monitoring within-field variability of corn yield using sentinel-2 and machine learning techniques. Remote Sensing, 11(23), 2873. doi:10.3390/rs11232873.

[15] Anua, S. N., & Wong, W. V. C. (2022). Utilizing Landsat 8 OLI for land cover classification in plantations area. IOP Conference Series: Earth and Environmental Science, 1053(1), 0–6. doi:10.1088/1755-1315/1053/1/012027.

[16] Phiri, D., Simwanda, M., Salekin, S., Ryirenda, V.R., Murayama, Y., Ranagalage, M., Oktaviani, N., Kusuma, H.A., Zhang, T., Su, J., Liu, C., Chen, W.H., Liu, H., Liu, G., Cavur, M., Duzgun, H.S., Kemec, S., Demirkan, D.C., Chairet, R., Ben Salem, Y., Aoun, M., Kiala, Z., Mutanga, O., Odindi, J., Peerbhay, K. (2020). Remote Sensing Sentinel-2 Data for Land Cover / Use Mapping: A Review. Remote Sensing, 12(2291), 1–35. doi:10.3390/rs12142291.

[17] Hu, Y., Raza, A., Syed, N. R., Acharki, S., Ray, R. L., Hussain, S., Dehghanisanij, H., Zubair, M., & Elbeltagi, A. (2023). Land Use/Land Cover Change Detection and NDVI Estimation in Pakistan's Southern Punjab Province. Sustainability (Switzerland), 15(4), 3572. doi:10.3390/su15043572.

[18] Muhaimin, M., Fitriani, D., Adyatma, S., & Arisanty, D. (2022). Mapping Build-Up Area Density Using Normalized Difference Built-Up Index (NDBI) and Urban Index (Ui) Wetland in the City Banjarmasin. IOP Conference Series: Earth and Environmental Science, 1089(1), 012036. doi:10.1088/1755-1315/1089/1/012036.

[19] Patil, P. P., Jagtap, M. P., Khatri, N., Madan, H., Vadduri, A. A., & Patodia, T. (2024). Exploration and advancement of NDDI leveraging NDVI and NDWI in Indian semi-arid regions: A remote sensing-based study. Case Studies in Chemical and Environmental Engineering, 9(December 2023), 100573. doi:10.1016/j.cscee.2023.100573.

[20] Pande, C. B., Moharir, K. N., Varade, A. M., Abdo, H. G., Mulla, S., & Yaseen, Z. M. (2023). Intertwined impacts of urbanization and land cover change on urban climate and agriculture in Aurangabad city (MS), India using google earth engine platform. Journal of Cleaner Production, 422(August), 138541. doi:10.1016/j.jclepro.2023.138541.

[21] Neinavaz, E., Skidmore, A. K., & Darvishzadeh, R. (2020). Effects of prediction accuracy of the proportion of vegetation cover on land surface emissivity and temperature using the NDVI threshold method. International Journal of Applied Earth Observation and Geoinformation, 85, 101984. doi:10.1016/j.jag.2019.101984.

[22] Lastovicka, J., Svec, P., Paluba, D., Kobliuk, N., Svoboda, J., Hladky, R., & Stych, P. (2020). Sentinel-2 data in an evaluation of the impact of the disturbances on forest vegetation. Remote Sensing, 12(12), 1914. doi:10.3390/rs12121914.

[23] Kruasilp, J., Pattanakiat, S., Phutthai, T., Vardhanabindu, P., & Nakmuenwai, P. (2023). Evaluation of Land Use Land Cover Changes in Nan Province, Thailand, Using Multi-Sensor Satellite Data and Google Earth Engine. Environment and Natural Resources Journal, 21(2), 186–197. doi:10.32526/ennrj/21/202200200.

[24] Wolanin, A., Camps-Valls, G., Gómez-Chova, L., Mateo-García, G., van der Tol, C., Zhang, Y., & Guanter, L. (2019). Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. Remote Sensing of Environment, 225, 441–457. doi:10.1016/j.rse.2019.03.002.

[25] Kumar, M., Singh, P., & Singh, P. (2023). Machine learning and GIS-RS-based algorithms for mapping the groundwater potentiality in the Bundelkhand region, India. Ecological Informatics, 74, 101980. doi:10.1016/j.ecoinf.2023.101980.

[26] Xie, G., & Niculescu, S. (2021). Mapping and monitoring of land cover/land use (LCLU) changes in the crozon peninsula (Brittany, France) from 2007 to 2018 by machine learning algorithms (support vector machine, random forest, and convolutional neural network) and by post-classification comparison (PCC). Remote Sensing, 13(19), 3899. doi:10.3390/rs13193899.

[27] Chen, Z., Chang, R., Zhao, W., Li, S., Guo, H., Xiao, K., Wu, L., Hou, D., & Zou, L. (2022). Quantitative prediction and evaluation of geothermal resource areas in the southwest section of the Mid-Spine Belt of Beautiful China. International Journal of Digital Earth, 15(1), 748–769. doi:10.1080/17538947.2022.2061055.

[28] Suthar, G., Kaul, N., Khandelwal, S., & Singh, S. (2024). Predicting land surface temperature and examining its relationship with air pollution and urban parameters in Bengaluru: A machine learning approach. Urban Climate, 53, 101830. doi:10.1016/j.uclim.2024.101830.

[29] Ni, X., Liu, Z., Wang, J., Dong, M., Wang, R., Qi, Z., Xu, H., Jiang, C., Zhang, Q., & Wang, J. (2024). Optimizing the development of contaminated land in China: Exploring machine-learning to identify risk markers. Journal of Hazardous Materials, 465, 133057. doi:10.1016/j.jhazmat.2023.133057.

[30] Radočaj, D., Gašparović, M., Radočaj, P., & Jurišić, M. (2024). Geospatial prediction of total soil carbon in European agricultural land based on deep learning. Science of the Total Environment, 912, 169647. doi:10.1016/j.scitotenv.2023.169647.

[31] Li, L., Liu, L., Peng, Y., Su, Y., Hu, Y., & Zou, R. (2023). Integration of multimodal data for large-scale rapid agricultural land evaluation using machine learning and deep learning approaches. Geoderma, 439, 116696. doi:10.1016/j.geoderma.2023.116696.

[32] Savitha, C., & Talari, R. (2023). Mapping cropland extent using sentinel-2 datasets and machine learning algorithms for an agriculture watershed. Smart Agricultural Technology, 4, 100193. doi:10.1016/j.atech.2023.100193.

[33] Sun, Y., Li, Y., Wang, R., & Ma, R. (2024). Modelling potential land suitability of large-scale wind energy development using explainable machine learning techniques: Applications for China, USA and EU. Energy Conversion and Management, 302, 118131. doi:10.1016/j.enconman.2024.118131.

[34] Tiwari, P., Poudel, K. P., Yang, J., Silva, B., Yang, Y., & McConnell, M. (2023). Marginal agricultural land identification in the Lower Mississippi Alluvial Valley based on remote sensing and machine learning model. International Journal of Applied Earth Observation and Geoinformation, 125, 103568. doi:10.1016/j.jag.2023.103568.

[35] Wang, Y., Zhang, P., Sun, H., Jia, X., Zhang, C., Liu, S., & Shao, M. (2022). Vertical patterns and controlling factors of soil nitrogen in deep profiles on the Loess Plateau of China. Catena, 215, 106318. doi:10.1016/j.catena.2022.106318.

[36] Wang, F., Xia, J., Zou, L., Zhan, C., & Liang, W. (2022). Estimation of time-varying parameter in Budyko framework using long short-term memory network over the Loess Plateau, China. Journal of Hydrology, 607, 127571. doi:10.1016/j.jhydrol.2022.127571.

[37] Indriani, R. F., Anjasmara, I. M., Utama, W., Paramita, E. G. K., & Nainggolan, R. A. O. (2023). Comparative Analysis of Physiograpic Study for Hydrology of Benowo Region, Surabaya. IOP Conference Series: Earth and Environmental Science, 1250(1), 12015. doi:10.1088/1755-1315/1250/1/012015.

[38] Chowdhury, M. S. (2024). Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting. Environmental Challenges, 14, 100800. doi:10.1016/j.envc.2023.100800.

[39] Vu, V. T., Nguyen, H. D., Vu, P. L., Ha, M. C., Bui, V. D., Nguyen, T. O., Hoang, V. H., & Nguyen, T. K. H. (2023). Predicting land use effects on flood susceptibility using machine learning and remote sensing in coastal Vietnam. Water Practice and Technology, 18(6), 1543–1555. doi:10.2166/wpt.2023.088.

[40] Yuh, Y. G., Tracz, W., Matthews, H. D., & Turner, S. E. (2023). Application of machine learning approaches for land cover monitoring in northern Cameroon. Ecological Informatics, 74, 101955. doi:10.1016/j.ecoinf.2022.101955.

[41] Rana, V. K., & Venkata Suryanarayana, T. M. (2020). Performance evaluation of MLE, RF and SVM classification algorithms for watershed scale land use/land cover mapping using sentinel 2 bands. Remote Sensing Applications: Society and Environment, 19, 100351. doi:10.1016/j.rsase.2020.100351.

[42] Zhao, Z., Islam, F., Waseem, L. A., Tariq, A., Nawaz, M., Islam, I. U., Bibi, T., Rehman, N. U., Ahmad, W., Aslam, R. W., Raza, D., & Hatamleh, W. A. (2024). Comparison of Three Machine Learning Algorithms Using Google Earth Engine for Land Use Land Cover Classification. Rangeland Ecology and Management, 92, 129–137. doi:10.1016/j.rama.2023.10.007.

[43] Nasiri, V., Deljouei, A., Moradi, F., Sadeghi, S. M. M., & Borz, S. A. (2022). Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. Remote Sensing, 14(9). doi:10.3390/rs14091977.

[44] Xu, R. (2021). Mapping rural settlements from Landsat and sentinel time series by integrating pixel-and object-based methods. Land, 10(3), 1–18. doi:10.3390/land10030244.

[45] Leder, T. D., Leder, N., & Baučić, M. (2020). Application of satellite imagery and water indices to the hydrography of the Cetina riverbasin (Middle Adriatic). Transactions on Maritime Science, 9(2), 374–384. doi:10.7225/toms.v09.n02.020.

[46] Noi Phan, T., Kuch, V., & Lehnert, L. W. (2020). Land cover classification using google earth engine and random forest classifier-the role of image composition. Remote Sensing, 12(15), 2411. doi:10.3390/RS12152411.

[47] Jiang, W., Ni, Y., Pang, Z., Li, X., Ju, H., He, G., Lv, J., Yang, K., Fu, J., & Qin, X. (2021). An effective water body extraction method with new water index for sentinel-2 imagery. Water (Switzerland), 13(12), 1647. doi:10.3390/w13121647.

[48] Theofanous, N., Chrysafis, I., Mallinis, G., Domakinis, C., Verde, N., & Siahalou, S. (2021). Aboveground biomass estimation in short rotation forest plantations in northern Greece using esa's sentinel medium-high resolution multispectral and radar imaging missions. Forests, 12(7), 902. doi:10.3390/f12070902.

[49] Zhang, C., Huang, C., Li, H., Liu, Q., Li, J., Bridhikitti, A., & Liu, G. (2020). Effect of textural features in remote sensed data on rubber plantation extraction at different levels of spatial resolution. Forests, 11(4), 399. doi:10.3390/F11040399.

[50] Orieschnig, C. A., Belaud, G., Venot, J. P., Massuel, S., & Ogilvie, A. (2021). Input imagery, classifiers, and cloud computing: Insights from multi-temporal LULC mapping in the Cambodian Mekong Delta. European Journal of Remote Sensing, 54(1), 398–416. doi:10.1080/22797254.2021.1948356.

[51] Wang, L., & Zhou, Y. (2023). Combining Multitemporal Sentinel-2A Spectral Imaging and Random Forest to Improve the Accuracy of Soil Organic Matter Estimates in the Plough Layer for Cultivated Land. Agriculture (Switzerland), 13(1), 8. doi:10.3390/agriculture13010008.

[52] Gull, S., Shah, S. R., & Dar, A. M. (2022). Assessing land use/land cover change detection of north-eastern watersheds of Kashmir valley using GIS and remote sensing techniques. Water Practice and Technology, 17(8), 1603–1614. doi:10.2166/wpt.2022.085.

[53] Zhang, R., Tang, X., You, S., Duan, K., Xiang, H., & Luo, H. (2020). A novel feature-level fusion framework using optical and SAR remote sensing images for land use/land cover (LULC) classification in cloudy mountainous area. Applied Sciences (Switzerland), 10(8), 1–24. doi:10.3390/APP10082928.

[54] Ghayour, L., Neshat, A., Paryani, S., Shahabi, H., Shirzadi, A., Chen, W., Al-Ansari, N., Geertsema, M., Amiri, M. P., Gholamnia, M., Dou, J., & Ahmad, A. (2021). Performance evaluation of sentinel-2 and Landsat 8 OLI data for land cover/use classification using a comparison between machine learning algorithms. Remote Sensing, 13(7), 1349. doi:10.3390/rs13071349.

[55] Myroniuk, V., Kutia, M., Sarkissian, A. J., Bilous, A., & Liu, S. (2020). Regional-scale forest mapping over fragmented landscapes using global forest products and Landsat time series classification. Remote Sensing, 12(1), 1–24. doi:10.3390/RS12010187.

[56] Gašparović, M., & Dobrinić, D. (2021). Green infrastructure mapping in urban areas using sentinel-1 imagery. Croatian Journal of Forest Engineering, 42(2), 337–356. doi:10.5552/crojfe.2021.859.

[57] Svoboda, J., Štych, P., Laštovička, J., Paluba, D., & Kobliuk, N. (2022). Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia. Remote Sensing, 14(5), 1189. doi:10.3390/rs14051189.

[58] Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 6308–6325. doi:10.1109/JSTARS.2020.3026724.

[59] Yousefi, S., Mirzaee, S., Almohamad, H., Dughairi, A. A. Al, Gomez, C., Siamian, N., Alrasheedi, M., & Abdo, H. G. (2022). Image Classification and Land Cover Mapping Using Sentinel-2 Imagery: Optimization of SVM Parameters. Land, 11(7), 993. doi:10.3390/land11070993.

[60] Forozan, G., Elmi, M. R., Talebi, A., Mokhtari, M. H., & Shojaei, S. (2020). Temporal-Spatial Simulation of Landscape Variations Using Combined Model of Markov Chain and Automated Cell. KN - Journal of Cartography and Geographic Information, 70(2), 45–53. doi:10.1007/s42489-020-00037-0.

[61] Hosseini, M., McNairn, H., Mitchell, S., Robertson, L. D., Davidson, A., Ahmadian, N., Bhattacharya, A., Borg, E., Conrad, C., Dabrowska-Zielinska, K., De Abelleyra, D., Gurdak, R., Kumar, V., Kussul, N., Mandal, D., Rao, Y. S., Saliendra, N., Shelestov, A., Spengler, D., … Becker-Reshef, I. (2021). A comparison between support vector machine and water cloud model for estimating crop leaf area index. Remote Sensing, 13(7), 1–20. doi:10.3390/rs13071348.

[62] Van Huynh, C., Pham, T. G., Nguyen, L. H. K., Nguyen, H. T., Nguyen, P. T., Le, Q. N. P., Tran, P. T., Nguyen, M. T. H., & Tran, T. T. A. (2022). Application GIS and remote sensing for soil organic carbon mapping in a farm-scale in the hilly area of central Vietnam. Air, Soil and Water Research, 15, 1–11. doi:10.1177/11786221221114777.

[63] Isioye, O. A., Akomolafe, E. A., & Ikwueze, U. H. (2020). Accuracy analysis of sentinel 2A and Landsat 8 OLI+ satellite datasets over kano state (Nigeria) using vegetation spectral indices. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 54(3/W1), 65–72. doi:10.5194/isprs-archives-XLIV-3-W1-2020-65-2020.

[64] Mananze, S., Pôças, I., & Cunha, M. (2020). Mapping and assessing the dynamics of shifting agricultural landscapes using google earth engine cloud computing, a case study in Mozambique. Remote Sensing, 12(8), 1–23. doi:10.3390/RS12081279.

[65] Hikouei, I. S., Kim, S. S., & Mishra, D. R. (2021). Machine-learning classification of soil bulk density in salt marsh environments. Sensors, 21(13), 4408. doi:10.3390/s21134408.

[66] Du, H., Li, M., Xu, Y., & Zhou, C. (2023). An Ensemble Learning Approach for Land Use/Land Cover Classification of Arid Regions for Climate Simulation: A Case Study of Xinjiang, Northwest China. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 2413–2426. doi:10.1109/JSTARS.2023.3247624.

[67] Amindin, A., Pouyan, S., Pourghasemi, H. R., Yousefi, S., & Tiefenbacher, J. P. (2021). Spatial and temporal analysis of urban heat island using Landsat satellite images. Environmental Science and Pollution Research, 28(30), 41439–41450. doi:10.1007/s11356-021-13693-0.

[68] Rani, A., Kumar, N., Kumar, J., & Sinha, N. K. (2022). Machine learning for soil moisture assessment. Deep Learning for Sustainable Agriculture, 143–168. doi:10.1016/B978-0-323-85214-2.00001-X.

[69] Shao, Z., Sumari, N. S., Portnov, A., Ujoh, F., Musakwa, W., & Mandela, P. J. (2021). Urban sprawl and its impact on sustainable urban development: a combination of remote sensing and social media data. Geo-Spatial Information Science, 24(2), 241–255. doi:10.1080/10095020.2020.1787800.

[70] Kibet, R., Olatubara, C. O., Ikporukpo, C. O., & Jebiwott, A. (2021). Land Use Land Cover Changes and Encroachment Issues in Kapkatet Wetland, Kenya. Open Journal of Ecology, 11(07), 493–506. doi:10.4236/oje.2021.117032.

[71] Deur, M., Gašparović, M., & Balenović, I. (2020). Tree species classification in mixed deciduous forests using very high spatial resolution satellite imagery and machine learning methods. Remote Sensing, 12(23), 1–18. doi:10.3390/rs12233926.

[72] Xia, B., Zhang, H., Li, Q., & Li, T. (2015). PETs: A Stable and Accurate Predictor of Protein-Protein Interacting Sites Based on Extremely-Randomized Trees. IEEE Transactions on Nanobioscience, 14(8), 882–893. doi:10.1109/TNB.2015.2491303.

[73] Vo Quang, A., Delbart, N., Jaffrain, G., Pinet, C., & Moiret, A. (2022). Detection of degraded forests in Guinea, West Africa, based on Sentinel-2 time series by inclusion of moisture-related spectral indices and neighbourhood effect. Remote Sensing of Environment, 281, 113230. doi:10.1016/j.rse.2022.113230.

[74] Jackson, C. M., & Adam, E. (2022). A machine learning approach to mapping canopy gaps in an indigenous tropical submontane forest using WorldView-3 multispectral satellite imagery. Environmental Conservation, 49(4), 255–262. doi:10.1017/S0376892922000339.

[75] Zhang, H., Gorelick, S. M., & Zimba, P. V. (2020). Extracting impervious surface from aerial imagery using semi-automatic sampling and spectral stability. Remote Sensing, 12(3), 506. doi:10.3390/rs12030506.

[76] Kilwenge, R., Adewopo, J., Sun, Z., & Schut, M. (2021). UAV-based mapping of banana land area for village-level decision-support in rwanda. Remote Sensing, 13(24), 1–15. doi:10.3390/rs13244985.

[77] Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., & Mi, J. (2021). GLC_FCS30: Global land-cover product with fine classification system at 30ĝ€¯m using time-series Landsat imagery. Earth System Science Data, 13(6), 2753–2776. doi:10.5194/essd-13-2753-2021.

[78] Wang, S., Jiao, X., Wang, L., Gong, A., Sang, H., Salahou, M. K., & Zhang, L. (2020). Integration of boosted regression trees and cellular automata-Markova model to predict the land use spatial pattern in Hotan Oasis. Sustainability (Switzerland), 12(4), 1396. doi:10.3390/su12041396.

[79] Wiederkehr, N. C., Gama, F. F., Castro, P. B. N., Bispo, P. da C., Balzter, H., Sano, E. E., Liesenberg, V., Santos, J. R., & Mura, J. C. (2020). Discriminating forest successional stages, forest degradation, and land use in central Amazon using ALOS/PALSAR-2 full-polarimetric data. Remote Sensing, 12(21), 1-30. doi:10.3390/rs12213512.

[80] Loukika, K. N., Keesara, V. R., & Sridhar, V. (2021). Analysis of land use and land cover using machine learning algorithms on google earth engine for Munneru river basin, India. Sustainability (Switzerland), 13(24), 13758. doi:10.3390/su132413758.

[81] Rahman, A., Abdullah, H. M., Tanzir, M. T., Hossain, M. J., Khan, B. M., Miah, M. G., & Islam, I. (2020). Performance of different machine learning algorithms on satellite image classification in rural and urban setup. Remote Sensing Applications: Society and Environment, 20, 100410. doi:10.1016/j.rsase.2020.100410.

[82] Yang, H., Wang, P., Chen, A., Ye, Y., Chen, Q., Cui, R., & Zhang, D. (2023). Prediction of phosphorus concentrations in shallow groundwater in intensive agricultural regions based on machine learning. Chemosphere, 313, 137623. doi:10.1016/j.chemosphere.2022.137623.

[83] Hossain, M. S., Khan, M. A. H., Oluwajuwon, T. V., Biswas, J., Rubaiot Abdullah, S. M., Tanvir, M. S. S. I., Munira, S., & Chowdhury, M. N. A. (2023). Spatiotemporal change detection of land use land cover (LULC) in Fashiakhali wildlife sanctuary (FKWS) impact area, Bangladesh, employing multispectral images and GIS. Modeling Earth Systems and Environment, 9(3), 3151–3173. doi:10.1007/s40808-022-01653-7.

[84] Li, Z., Xu, Y., Sun, Y., Wu, M., & Zhao, B. (2020). Urbanization-driven changes in land-climate dynamics: A case study of Haihe River Basin, China. Remote Sensing, 12(17), 2701. doi:10.3390/RS12172701.

[85] Xue, S. Y., Xu, H. Y., Mu, C. C., Wu, T. H., Li, W. P., Zhang, W. X., Streletskaya, I., Grebenets, V., Sokratov, S., Kizyakov, A., & Wu, X. D. (2021). Changes in different land cover areas and NDVI values in northern latitudes from 1982 to 2015. Advances in Climate Change Research, 12(4), 456–465. doi:10.1016/j.accre.2021.04.003.

[86] Oad, V. K., Szymkiewicz, A., Khan, N. A., Ashraf, S., Nawaz, R., Elnashar, A., Saad, S., & Qureshi, A. H. (2023). Time series analysis and impact assessment of the temperature changes on the vegetation and the water availability: A case study of Bakun-Murum Catchment Region in Malaysia. Remote Sensing Applications: Society and Environment, 29(November 2022), 100915. doi:10.1016/j.rsase.2022.100915.

[87] Onyango, D. O., & Opiyo, S. B. (2022). Detection of historical landscape changes in Lake Victoria Basin, Kenya, using remote sensing multi-spectral indices. Watershed Ecology and the Environment, 4, 1–11. doi:10.1016/j.wsee.2021.12.001.

[88] Yasin, M. Y., Abdullah, J., Noor, N. M., Yusoff, M. M., & Noor, N. M. (2022). Landsat observation of urban growth and land use change using NDVI and NDBI analysis. IOP Conference Series: Earth and Environmental Science, 1067(1), 12037. doi:10.1088/1755-1315/1067/1/012037.

[89] Guha, S., & Govil, H. (2021). A long-term monthly analytical study on the relationship of LST with normalized difference spectral indices. European Journal of Remote Sensing, 54(1), 487–511. doi:10.1080/22797254.2021.1965496.

[90] Santecchia, G. S., Revollo Sarmiento, G. N., Genchi, S. A., Vitale, A. J., & Delrieux, C. A. (2023). Assessment of Landsat-8 and Sentinel-2 Water Indices: A Case Study in the Southwest of the Buenos Aires Province (Argentina). Journal of Imaging, 9(9). doi:10.3390/jimaging9090186.

[91] Rani, P. P., Kumar, M. S., & Sireesha, P. V. G. (2021). Mapping of active and empty aquaponds using spectral indices in coastal region of Guntur District, Andhra Pradesh, India. Journal of Environmental Biology, 42(5), 1338–1346. doi:10.22438/jeb/42/5/MRN-1634.

[92] Guha, S., Govil, H., & Besoya, M. (2020). An investigation on seasonal variability between LST and NDWI in an urban environment using Landsat satellite data. Geomatics, Natural Hazards and Risk, 11(1), 1319–1345. doi:10.1080/19475705.2020.1789762.

[93] Rumora, L., Miler, M., & Medak, D. (2020). Impact of various atmospheric corrections on sentinel-2 land cover classification accuracy using machine learning classifiers. ISPRS International Journal of Geo-Information, 9(4), 277. doi:10.3390/ijgi9040277.

[94] Ibrahim, S. (2023). Improving Land Use/Cover Classification Accuracy from Random Forest Feature Importance Selection Based on Synergistic Use of Sentinel Data and Digital Elevation Model in Agriculturally Dominated Landscape. Agriculture (Switzerland), 13(1), 98. doi:10.3390/agriculture13010098.

[95] Dobrinic, D., Gašparovic, M., & Medak, D. (2022). Evaluation of Feature Selection Methods for Vegetation Mapping Using Multitemporal Sentinel Imagery. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43(B3-2022), 485–491. doi:10.5194/isprs-archives-XLIII-B3-2022-485-2022.

[96] Agustin, S., Tjandrasa, H., & Ginardi, R. V. H. (2020). Deep Learning-based Method for Multi-Class Classification of Oil Palm Planted Area on Plant Ages Using Ikonos Panchromatic Imagery. International Journal on Advanced Science, Engineering and Information Technology, 10(6), 2200–2206. doi:10.18517/ijaseit.10.6.12030.

[97] Indriani, R. F., & Utama, W. (2023). Physiographic Study for Hydrology of Benowo Region Surabaya. IOP Conference Series: Earth and Environmental Science, 1127(1), 012019. doi:10.1088/1755-1315/1127/1/012019.

[98] Utama, W., & Indriani, R. F. (2021). Regional Physiographic Study for the Hydrology of Kali Lamong Watershed Area. IOP Conference Series: Earth and Environmental Science, 936(1), 12032. doi:10.1088/1755-1315/936/1/012032.

[99] Psomiadis, E. (2022). Long and Short-Term Coastal Changes Assessment Using Earth Observation Data and GIS Analysis: The Case of Sperchios River Delta. ISPRS International Journal of Geo-Information, 11(1), 61. doi:10.3390/ijgi11010061.

[100] Orynbaikyzy, A., Gessner, U., Mack, B., & Conrad, C. (2020). Crop type classification using fusion of sentinel-1 and sentinel-2 data: Assessing the impact of feature selection, optical data availability, and parcel sizes on the accuracies. Remote Sensing, 12(17), 2779. doi:10.3390/RS12172779.

[101] Boston, T., Van Dijk, A., Larraondo, P. R., & Thackway, R. (2022). Comparing CNNs and Random Forests for Landsat Image Segmentation Trained on a Large Proxy Land Cover Dataset. Remote Sensing, 14(14), 3396. doi:10.3390/rs14143396.

[102] Nallapareddy, A. (2022). Detection and Classification of Vegetation Areas from Red and Near Infrared Bands of Landsat-8 Optical Satellite Image. Applied Computer Science, 18(1), 45–55. doi:10.35784/acs-2022-4.

[103] Pu, D. C., Sun, J. Y., Ding, Q., Zheng, Q., Li, T. T., & Niu, X. F. (2020). Mapping Urban Areas Using Dense Time Series of Landsat Images and Google Earth Engine. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42(3/W10), 403–409. doi:10.5194/isprs-archives-XLII-3-W10-403-2020.

[104] Yan, H., Wang, K., Lin, T., Zhang, G., Sun, C., Hu, X., & Ye, H. (2021). The challenge of the urban compact form: Three-dimensional index construction and urban land surface temperature impacts. Remote Sensing, 13(6), 1–23. doi:10.3390/rs13061067.

[105] Hively, W. D., Lamb, B. T., Daughtry, C. S. T., Serbin, G., Dennison, P., Kokaly, R. F., Wu, Z., & Masek, J. G. (2021). Evaluation of SWIR crop residue bands for the Landsat next mission. Remote Sensing, 13(18), 3718. doi:10.3390/rs13183718.

[106] Mazarire, T. T., Ratshiedana, P. E., Nyamugama, A., Adam, E., & Chirima, G. (2022). Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa. South African Journal of Geomatics, 9(2), 333–347. doi:10.4314/sajg.v9i2.22.

[107] Gyamfi-Ampadu, E., Gebreslasie, M., & Mendoza-Ponce, A. (2020). Mapping natural forest cover using satellite imagery of Nkandla forest reserve, KwaZulu-Natal, South Africa. Remote Sensing Applications: Society and Environment, 18, 100302. doi:10.1016/j.rsase.2020.100302.

[108] Ouma, Y. O., Keitsile, A., Nkwae, B., Odirile, P., Moalafhi, D., & Qi, J. (2023). Urban land-use classification using machine learning classifiers: comparative evaluation and post-classification multi-feature fusion approach. European Journal of Remote Sensing, 56(1), 2173659. doi:10.1080/22797254.2023.2173659.

[109] Hasan, M. A., Mimi, M. B., Voumik, L. C., Esquivias, M. A., & Rashid, M. (2023). Investigating the Interplay of ICT and Agricultural Inputs on Sustainable Agricultural Production: An ARDL Approach. Journal of Human, Earth, and Future, 4(4), 375-390. doi:10.28991/HEF-2023-04-04-01.

[110] Chaves, M. E. D., Picoli, M. C. A., & Sanches, I. D. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. Remote Sensing, 12(18), 3062. doi:10.3390/rs12183062.

[111] Amini, S., Saber, M., Rabiei-Dastjerdi, H., & Homayouni, S. (2022). Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series. Remote Sensing, 14(11), 1–23. doi:10.3390/rs14112654.

[112] Choudhury, B. U., Divyanth, L. G., & Chakraborty, S. (2023). Land use/land cover classification using hyperspectral soil reflectance features in the Eastern Himalayas, India. Catena, 229, 107200. doi:10.1016/j.catena.2023.107200.

[113] Hu, W., Zhang, S., Fu, Y., Jia, G., Yang, R., Shen, S., Li, Y., & Li, G. (2023). Objective diagnosis of machine learning method applicability to land comprehensive carrying capacity evaluation: A case study based on integrated RF and DPSIR models. Ecological Indicators, 151, 110338. doi:10.1016/j.ecolind.2023.110338.

[114] Ebrahimy, H., Mirbagheri, B., Matkan, A. A., & Azadbakht, M. (2022). Effectiveness of the integration of data balancing techniques and tree-based ensemble machine learning algorithms for spatially-explicit land cover accuracy prediction. Remote Sensing Applications: Society and Environment, 27, 100785. doi:10.1016/j.rsase.2022.100785.

[115] Lekka, C., Petropoulos, G. P., & Detsikas, S. E. (2024). Appraisal of EnMAP hyperspectral imagery use in LULC mapping when combined with machine learning pixel-based classifiers. Environmental Modelling and Software, 173, 105956. doi:10.1016/j.envsoft.2024.105956.

[116] Sugumar, R., & Suganya, D. (2023). A multi-spectral image-based high-level classification based on a modified SVM with enhanced PCA and hybrid metaheuristic algorithm. Remote Sensing Applications: Society and Environment, 31, 100984. doi:10.1016/j.rsase.2023.100984.

[117] Yazdian, H., Salmani-Dehaghi, N., & Alijanian, M. (2023). A spatially promoted SVM model for GRACE downscaling: Using ground and satellite-based datasets. Journal of Hydrology, 626, 130214. doi:10.1016/j.jhydrol.2023.130214.

[118] Periasamy, S., Ravi, K. P., & Tansey, K. (2022). Identification of saline landscapes from an integrated SVM approach from a novel 3-D classification schema using Sentinel-1 dual-polarized SAR data. Remote Sensing of Environment, 279, 113144. doi:10.1016/j.rse.2022.113144.

[119] Chundu, M. L., Banda, K., Lyoba, C., Tembo, G., Sichingabula, H. M., & Nyambe, I. A. (2024). Modeling land use/land cover changes using quad hybrid machine learning model in Bangweulu wetland and surrounding areas, Zambia. Environmental Challenges, 14, 100866. doi:10.1016/j.envc.2024.100866.

[120] Cloete, D. N., Shoko, C., Dube, T., & Clarke, S. (2024). Remote sensing-based land use land cover classification for the Heuningnes Catchment, Cape Agulhas, South Africa. Physics and Chemistry of the Earth, 134, 103559. doi:10.1016/j.pce.2024.103559.

[121] Chaturvedi, V., & de Vries, W. T. (2021). Machine Learning Algorithms for Urban Land Use Planning: A Review. Urban Science, 5(3), 68. doi:10.3390/urbansci5030068.

[122] Pandit, S., Shimada, S., & Dube, T. (2024). Comprehensive Analysis of Land Use and Cover Dynamics in Djibouti Using Machine Learning Technique: A Multi-Temporal Assessment from 1990 to 2023. Environmental Challenges, 100920. doi:10.1016/j.envc.2024.100920.