Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

Analysis of Land Cover Change Using Support Vector Machine Algorithm on Satellite Imagery (Case Study: Mining Area of IWIP).

Efraim Julian Soma*, Winsy Christo Deilan Weku, Rillya Arundaa, Mahardika Inra Takaendengan
Information System, Universitas Sam Ratulangi, Indonesia

Abstract: Unsupervised mining activities and the lack of proper management of the environment can bring very serious ecological effects. One of the worst is large-scale land cover change. This paper investigates the land cover changes in the mining area of PT Indonesia Weda Bay Industrial Park (IWIP), Weda Bay Nickel (WBN) area, and measures the performance of the Support Vector Machine technique of land cover classification using satellite imagery. This research is a quantitative study based on a remote sensing technique and spatial analysis in a GIS environment. Multi-temporal satellite images, Landsat-7 (2010), Landsat-8 (2014), and Sentinel-2 (2019 and 2025) were processed to land cover classes (vegetation, open land, cloud) based on the spectral indices NDVI, NDBI, and BSI. Through the post-classification comparison method, land cover changes were detected. The overall change from 2010 to 2025 is quite significant. The vegetation cover decreased from 91.2% to 84.3%, while open land increased from less than 1% to 11.4%. The greatest changes were observed in 2019 and 2025, suggesting that mining activities and related industry development have speed up A lot. The Support Vector Machine algorithm reached excellent classification accuracy over the different periods (97.47%, 96.67%, 89.52%, 95.12% in the cases of 2010 2014 2019 and 2025 respectively), and the results were supported by the high Kappa agreement scores (ranging from 0.84 to 0.96). The low difference between training and testing accuracy indicates a strong generalization without any sign of the model getting overfitted to the training data. The work highlights that Support Vector Machine algorithm, is a very good and reliable method for detecting land cover change in mining areas, where patterns are both complex as well as dynamic.
Keywords: Land Cover; Support Vector Machine; Remote Sensing; Satellite Imagery; Geographic Information System;

1. INTRODUCTION

Land is an essential part of human life and is utilized for various activities, including agriculture, industry, and mining. Development and regional growth can lead to changes in land cover, which may have both positive and negative impacts. Therefore, proper monitoring and planning are necessary to support sustainable land use [1]. Mining is also an important sector that contributes significantly to Indonesia's economy, as the country has abundant natural resource potential [2]

Efraim Julian Soma: *Corresponding Author

Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : **https://doi.org/10.58602/itsecs.v4i4.388**
Page 188-200

Mining is one of the industries that makes a significant economic contribution to Indonesia. Indonesia is a country with vast resource potential [3].

Monitoring land-cover changes can be conducted using remote sensing technology. In satellite image classification, the Support Vector Machine (SVM) algorithm is widely used because it can handle high-dimensional data with limited training samples while generally providing high classification accuracy. Therefore, applying SVM to land-cover change analysis in mining areas is expected to produce more accurate and reliable results [4].

The Support Vector Machine (SVM) algorithm is frequently employed in satellite image classification—particularly for land-cover change studies involving limited datasets—due to its effectiveness. As a machine learning method, it is capable of handling high-dimensional data despite having a restricted training set and typically yields high classification accuracy. These advantages suggest that utilizing SVM for land-cover change analysis in mining areas will likely produce more reliable results [5].

The primary contribution of this study is to provide a comprehensive analysis of land-cover dynamics in a large-scale nickel mining area (IWIP) using the Support Vector Machine (SVM) algorithm on multi-temporal satellite imagery. The study also establishes an empirical baseline regarding the reliability of SVM in handling image data affected by significant cloud cover in the equatorial region, serving as a strategic reference for sustainable environmental monitoring efforts in mining areas.

2. RESEARCH METHODOLOGY

2.1. Research location

The study area is the IWIP mining area, located in Lelief Village, Weda Tengah District, North Maluku Regency. Four land-cover classes were identified in the study area: vegetation, bare land, built-up land, and clouds, with the cloud class included only if a small amount of cloud cover is present in the imagery of the study area.

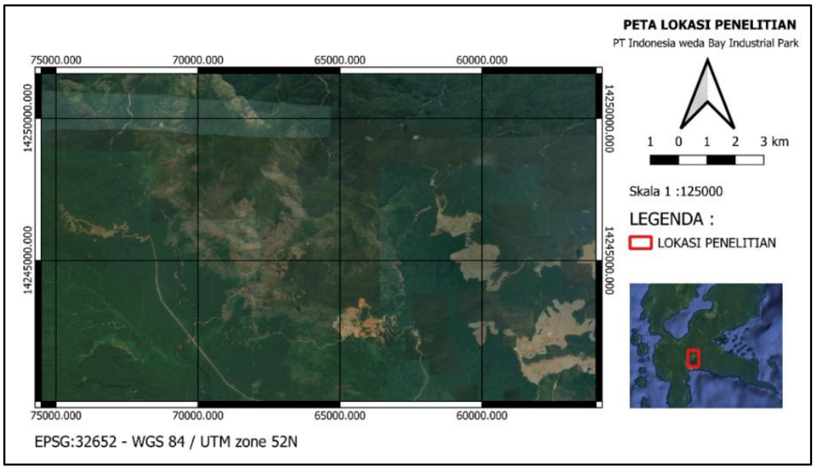

Figure 1. Research location

2.2. Research data and data sources

The data used in this study consist of Landsat 7, Landsat 8, and Sentinel-2 satellite imagery. Landsat 7 and Landsat 8 imagery can be downloaded free of charge from the official United States Geological Survey (USGS) EarthExplorer website, while Sentinel-2 imagery can be accessed free of charge through the ESA Copernicus Open Access Browser. The types of satellite imagery used in this study are presented in Table 1.

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

Table 1. Types of satelite imagery

Number	Year	Image Types	Resolution
1	2010	Landsat 7	30x30m
2	2014	Landsat 8	30x30m
3	2019	Sentinel 2	10x10m
4	2025	Sentinel 2	10x10m

Two types of data were used for this study: raster data consisting of satellite imagery—incorporating various spectral bands and spectral indices—and shapefile data containing multipoint training samples (points representing specific, pre-determined land cover classes). The dataset was then randomly split into 70% for training and 30% for testing. An SVM model was trained using the randomly selected 70% of the data, enabling it to recognize the characteristics of each land cover class based on the spectral values of individual pixels. The number of points for each class is presented in Table 2.

Table 2. Distribution of the number of points per class and splitting data

Year	Training data				Testing data			
	Vegetation	Cloud	Bareland	Built-up	Vegetation	Cloud	Bareland	Built-up
2010	68	30	48	-	32	20	27	-
2014	68	46	51	-	32	29	29	-
2019	60	63	70	-	40	37	28	-
2025	58	65	72	32	42	35	28	18

After the data is split into training and testing sets, the training data is used to teach the model to recognize each land cover class, while the testing data is used to verify whether the model can recognize new unseen data.

2.3. Research workflow

The research begins by defining the Area of Interest (AOI), followed by downloading the satellite imagery used in the study: Landsat 7 and Landsat 8 for 2010 and 2014, and Sentinel-2 for 2019 and 2025. The imagery is then pre-processed through geometric correction, true-color visualization, and clipping using the AOI shapefile to ensure that the analysis covers only the study area. Next, spectral indices are calculated to support the collection of multipoint training samples representing three land-cover classes: vegetation, bare land, and built-up land. The clipped imagery, spectral indices, and training samples are then used as inputs for SVM classification for each study year. The classification results are subsequently evaluated using accuracy metrics to identify potential overfitting or underfitting. If the model performs well without significant overfitting or underfitting, the results are visualized. If either condition occurs, the training samples are re-collected and the classification process is repeated. The research workflow is presented in Figure 2.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : **https://doi.org/10.58602/itsecs.v4i4.388**
Page 188-200

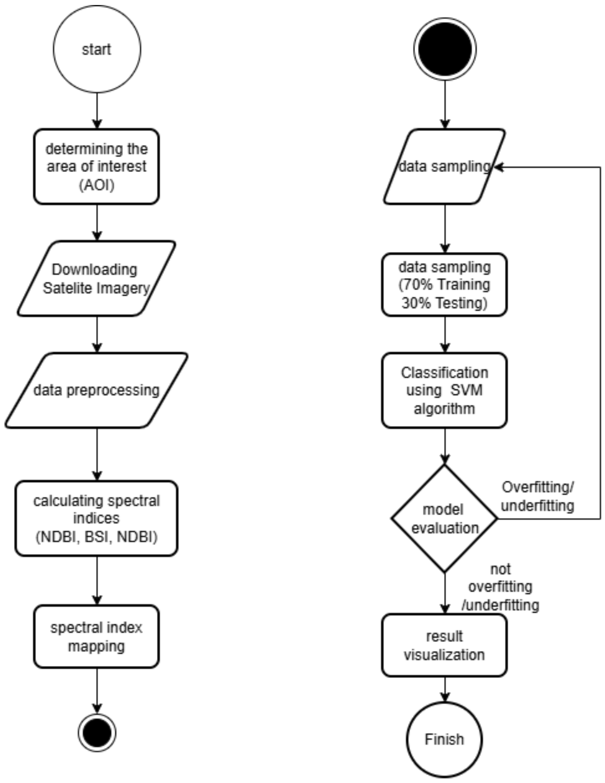

Figure 2. Research workflow

2.4. Satelite imagery

Satellite imagery consists of digital images of the Earth's surface compiled from spectral data acquired by sensors mounted on specialized satellites; these images are available for all regions of the world from various sources, including both commercial and government entities [6].

2.5. Land Cover

Land cover refers to the observable physical condition of the Earth's surface, encompassing how humans manage, utilize, and treat specific types of land cover—whether for production, modification, or maintenance purposes [7]. Information regarding land cover plays a vital role across various sectors, including agriculture, mining, and forestry, as well as in land utilization for other needs. With population growth and rapid technological advancement, many areas have undergone changes in land use to meet human requirements—ranging from the provision of clothing and food to the development of residential areas. Understanding the dynamics of land cover change is crucial, as it provides insight into current conditions and helps predict future land states within a given region [8].

2.6. Remote sensing

According to USGS [9], remote sensing is a technique used to obtain and monitor information about objects, areas, or phenomena on the Earth’s surface without direct contact, by using sensors mounted on satellites or aircraft to record reflected or emitted energy. This technology can be applied for various purposes, such as Earth surface mapping, seafloor observation, and monitoring changes in ocean temperature [10]. Thus, remote sensing can be defined as the science and art of obtaining information about objects or phenomena through the analysis of data recorded from a distance [11].

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

2.7. Support vector machine (SVM)

Support Vector Machine (SVM) is a machine learning algorithm that classifies data by finding the optimal separating boundary, or hyperplane, that maximizes the margin between different classes. Developed by Vladimir N. Vapnik and his colleagues in the 1990s, SVM was formally introduced in 1995 through the paper "Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing" [12].

The SVM classification process begins with input data, followed by the selection of support vectors, which are the data points that most strongly influence the hyperplane. When data cannot be separated linearly, a kernel function, such as linear, polynomial, RBF, or sigmoid, can be used to transform the data into a higher-dimensional space. In this study, pixel-based image classification is performed using the Radial Basis Function (RBF) kernel in ENVI, as illustrated in Figure 3.

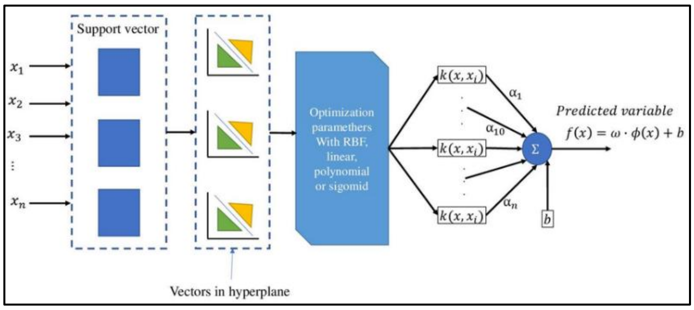

Figure 3. Support Vector Machine Architecture

Before model training, the image data are standardized using StandardScaler to ensure that all input features have a mean of zero and a standard deviation of one. This step is important because SVM is sensitive to differences in feature scales, and standardization prevents variables with larger numerical ranges from dominating the classification process [13]. To obtain the optimal SVM parameters, GridSearchCV is applied to systematically evaluate predefined combinations of hyperparameters using cross-validation. For the RBF kernel, the main parameters optimized are C, which controls the trade-off between classification errors and the hyperplane margin, and gamma, which determines the influence of individual training samples on the decision boundary. This approach helps improve model generalization and reduce the risk of overfitting [14]. SVM is widely used in remote sensing because it can provide high classification accuracy with relatively limited training data. It is also effective for high-dimensional datasets, requires relatively low computational resources, and maintains stable performance in complex feature spaces [15].

2.8. Model Evaluation

Model evaluation is the process of assessing the performance of a machine learning model using specific evaluation metrics. This step is essential to ensure that the model does not simply memorize the training data but can also generalize well to previously unseen data [16].

Several metrics can be used to evaluate classification models, including accuracy, confusion matrix, and K-fold cross-validation. K-fold cross-validation evaluates model performance by dividing the dataset into k equal subsets (folds). One fold is used for validation while the remaining folds are used for training. This process is repeated k times so that each subset is used as validation data once, providing a more reliable estimate of model performance and reducing the risk of overfitting [17]. An illustration of the confusion matrix can be seen in Table 3.

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

Table 3. illustration of confusion matrix

		A	**B**
Actual	A	9	1
	B	3	7
		Predicted	

As illustrated in Table 3, the model correctly predicted Class A nine times (True Positive) and made one incorrect prediction (False Negative); for Class B, the model correctly detected the class seven times (False Positive) and misclassified it three times (True Negative). A confusion matrix is a table used to evaluate classification performance by comparing the model's predicted classes with the actual classes. It shows the number of correct and incorrect predictions for each class [18]. Precision (User's Accuracy) is used to determine the proportion of correctly classified samples within a predicted class. Recall (Producer's Accuracy) measures the model's ability to correctly identify samples belonging to their actual classes, indicating its sensitivity in classification. Cohen's Kappa is also used to assess the agreement between the classification results and the actual data while accounting for the possibility of agreement occurring by chance. The metric formulas for each model evaluation can be found in Table 4.

Table 4. Model evaluation matrix formula

Evaluation matrix	**Formula**
Overall Accuracy	$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$
User's Accuracy (precision)	$Precision = \frac{TP}{TP + FP} \quad (2)$
Producer's Accuracy (Recall)	$Recall = \frac{TP}{TP + FN} \quad (3)$
Cohen's kappa	$K = \frac{OA + P_e}{1 - P_e} \quad (4)$

Information:

TP : True Positife
TN : True Negative
FP : False Positife
FN : False Negative
OA : Overall Accuracy

The kappa coefficient incorporates the p_e component—representing accuracy expected by chance based on class distributions. Factoring in p_e makes the kappa coefficient a more objective measure of agreement than overall accuracy, particularly for datasets with imbalanced classes [19].

Efraim Julian Soma: *Corresponding Author

2.9. Spectral Index

A spectral index can be understood as a quantitative method for utilizing spectral reflectance values to identify or map specific objects from satellite imagery, whether multispectral or hyperspectral [20].

Several spectral indices are employed in this study—namely the Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), and Bare Soil Index (BSI)—each serving a specific function: NDVI detects vegetation, BSI identifies bare soil, and NDBI detects built-up areas.

The equations for each index are presented in Table 5.

Table 5. Spectral index formula

Spectral Index	Formula
Normalized Difference Vegetation Index (NDVI),	$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + \mathrm{Red})} \quad (5)$$
Bare Soil Index (BSI)	$$\mathrm{BSI} = \frac{(\mathrm{SWIR} + \mathrm{Red}) - (\mathrm{NIR} + \mathrm{Blue})}{(\mathrm{SWIR} + \mathrm{Red}) + (\mathrm{NIR} + \mathrm{Blue})} \quad (6)$$
Normalized Difference Built-up Index (NDBI)	$$\mathrm{NDBI} = \frac{\mathrm{SWIR} - \mathrm{NIR}}{\mathrm{SWIR} + \mathrm{NIR}} \quad (7)$$

Information:
NIR : Near Infrared Band
Red : Red Reflective Band
Blue : Blue Reflective Band
SWIR: Shortwave Infrared Reflective Band

Spectral index operate by combining light reflectance values from two or more satellite image wavelength bands to highlight the characteristics of specific features on the Earth's surface. The purpose of using these spectral indices is to enhance the accuracy of sample collection for each classes to be classified by the model.

3. RESULT AND DISCUSSIONS

3.1. Research Results

This study utilizes satellite imagery from different sources: Landsat-7 ETM (2010) and Landsat-8 (2014), accessible via https://earthexplorer.usgs.gov, and Sentinel-2 (2019 and 2025), accessible via https://www.copernicus.eu/en. Prior to classification using the Support Vector Machine (SVM) algorithm, the acquired satellite imagery underwent preprocessing to enhance image quality and convert the imagery to True Color. However, the 2010 Landsat-7 ETM imagery exhibited a Scan Line Corrector (SLC) error, resulting in the appearance of data gaps (blank lines). To address this, a gap mask—downloadable from the USGS website alongside the Landsat 7 imagery—was used to rectify the SLC error in the 2010 dataset.

Following geometric correction to enhance resolution and convert the image to True Color, the image was clipped to the Area of Interest (AOI). This process utilized a pre-prepared AOI shapefile and the "clip raster by mask layer" function in QGIS to restrict the image to the specific study area, the PT IWIP mining site. The pre-processed satellite image showing land cover at the PT IWIP mining site is presented in Figure 4.

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

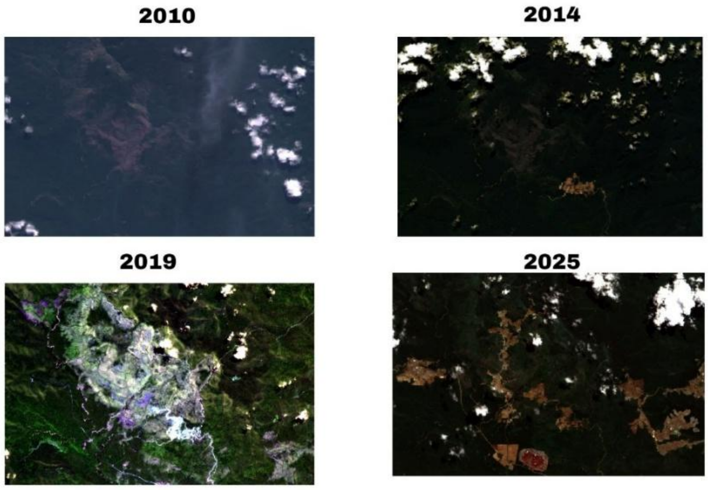

Figure 4. Image visualization of the study area

Figure 4 displays corrected and clipped imagery from 2010, 2014, 2019, and 2025. In 2010, the study area was still dominated by vegetation; by 2014, open land resulting from mining activities had begun to appear. Mining activity within the study area—causing land cover changes—was evident by 2019. Finally, the 2025 imagery clearly reveals a drastic transformation compared to previous years, as mining operations expanded, resulting in a larger area of open land.

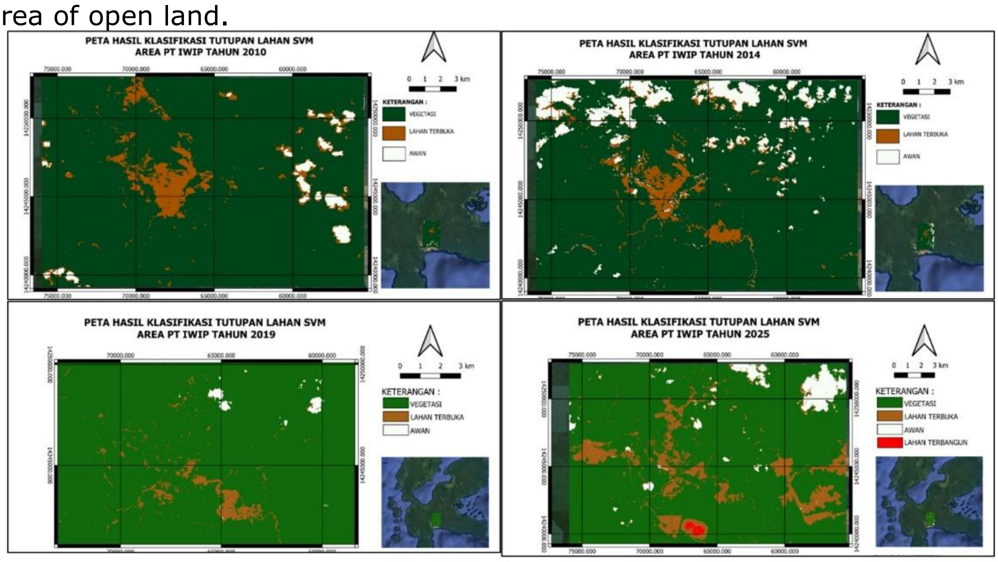

Figure 5. Land cover classification maps using SVM

Based on Figure 5, the study area was still dominated by vegetation in each observation year, represented by the green color. In addition, bare land, represented by the brown color, was found and distributed across several parts of the study area. In 2010, the presence of bare land was still relatively limited. In 2014, bare land became more widely distributed in several locations. In 2019, bare land was still mainly found in the central part of the study area. Meanwhile, in 2025, the land cover condition became more diverse with the appearance of the built-up land class, represented by the red color. This indicates that changes in land cover occurred in the PT IWIP area during the 2010–2025 period. Visually, these changes can be seen from several areas of vegetation that were

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

converted into bare land, as well as the development of built-up areas. The results of the land cover changes can be seen in the table 6.

Table 6. Land cover change on IWIP mining area

Class	2010 (Landsat 7)		2014 (Landsat 8)		2019 (Sentinel 2)		2025 (Sentinel 2)	
	Area (ha)	%	Area (ha)	%	Area (ha)	%	Area (ha)	%
Vegetation	24.661,71	91,2	22.952,34	84,9	25.526,51	94,1	22.856,56	84,3
Cloud	565,20	2,1	2.662,83	9,9	974,06	3,6	10.088,24	4
Bare land	1.804,95	6,7	1.416,69	5,2	618,95	2,3	3.099,52	11,4
Built-up	0	0	0	0	0	0	75,2	0,3
Total	27.031,86	100	27.031,86	100	27.119,52	100	27.119,52	100

Based on Table 5, land-cover composition in the PT Indonesia Weda Bay Industrial Park (IWIP) area changed considerably between 2010 and 2025. The SVM classification identified four main classes: vegetation, clouds, bare land, and built-up land. Vegetation remained the dominant class throughout the study period, covering 24,661.71 ha (91.2%) in 2010, decreasing to 22,952.34 ha (84.9%) in 2014, increasing to 25,526.51 ha (94.1%) in 2019, and declining again to 22,856.56 ha (84.3%) in 2025.

In contrast, bare land showed a significant increase in 2025. Its area decreased from 1,804.95 ha (6.7%) in 2010 to 618.95 ha (2.3%) in 2019, before increasing sharply to 3,099.52 ha (11.4%) in 2025. This represents an increase of approximately 2,480.57 ha or 9.1% from 2019 to 2025, which may indicate intensified mining activities, land clearing, infrastructure development, and industrial expansion. The built-up class was not detected in 2010, 2014, or 2019 but appeared in 2025, covering 75.20 ha (0.3%). Its emergence indicates the development of industrial facilities and supporting infrastructure within the IWIP area. Overall, the results demonstrate a notable shift in land-cover composition, particularly the reduction of vegetation and substantial expansion of bare land and built-up areas by 2025.

Land cover changes in the PT IWIP area reveal drastic dynamics, particularly regarding the reduction in vegetation from 91.2% (24,661.71 ha) in 2010 to 84.3% (22,856.56 ha) in 2025. This decline corresponds directly to a surge in open land—rising to 11.4% (3,099.52 ha)—and the emergence of a built-up land category accounting for 0.3% (75.2 ha) by the end of the observation period. The 9.1% surge in open land since 2019 indicates an intensification of large-scale land clearing for open-pit mining activities and the expansion of RKEF smelter infrastructure within the IWIP area. Consistent with previous studies by [5] and [3]—which highlight the impact of mining downstreaming on land-use conversion—the patterns of change at IWIP confirm that the primary drivers of this landscape modification are the policy to increase national nickel production capacity and the surge in investment in integrated industrial estates, which has intensified significantly since 2018–2019.

3.2. Evaluation of support vector machine models

In this case, cross-validation and a confusion matrix are used to assess model performance. For the cross-validation process, the training data is divided into five groups (5-fold); one-fold serves as the validation set, while the remaining four are used for training. This process is repeated five times so that each fold acts as the validation set at

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : https://doi.org/10.58602/itsecs.v4i4.388
Page 188-200

least once. A confusion matrix is then employed to evaluate the model's performance by determining the number of test samples correctly predicted versus those misclassified across each class.

Table 7. Cross-validation test results 2010–2025

Year	Test1	Test2	Test3	Test4	Test5	Mean	STD
2010	96,67	100	100	100	100	99,33	1,33
2014	96,97	100	96,97	100	100	98,79	1,14
2019	82,05	94,87	92,31	89,47	94,74	90,69	4,75
2025	95,65	93,33	100	97,78	97,83	96,92	2,26

Based on Table 6, the results of the five-fold cross-validation show distinct patterns for each year. In 2010, the model achieved an accuracy of 96.67% in the first test and 100% in tests two through five, with a mean of 99.33% and a standard deviation of 0.13 (1.33%); this indicates minimal variation in accuracy across folds, signifying stable and consistent model performance. In 2014, accuracy ranged from 96.97% to 100%, with a mean of 98.79% and a standard deviation of 0.14 (1.14%); this suggests relatively uniform classification results across folds and a high level of model stability that is not dependent on any specific data split. In 2019, the accuracy for each fold was 82.05%, 94.87%, 92.31%, 89.47%, and 94.74%, with a mean (CV Mean) of 90.69% and a standard deviation of 0.47 (4.75%); the higher standard deviation compared to other years indicates greater performance variation across folds, although the mean accuracy exceeding 90% still demonstrates good generalization capability. In 2025, the highest accuracy was 100% (third fold) and the lowest was 93.33% (second fold), with a mean (CV Mean) of 0.96 and a standard deviation of 0.26 (2.26%); despite slight variations in accuracy across folds, all values remained above 90%, indicating that the model possesses excellent classification performance and good generalization capability across various data splits. Based on the evaluation results using the 5-Fold Cross-Validation method, the Support Vector Machine (SVM) model demonstrated good to excellent classification performance across all years of observation. Test results indicate that the average accuracy (CV Mean) was 99.33% (standard deviation: 1.33%) in 2010, 98.79% (standard deviation: 1.48%) in 2014, 90.69% (standard deviation: 4.75%) in 2019, and 96.92% (standard deviation: 2.26%) in 2025.

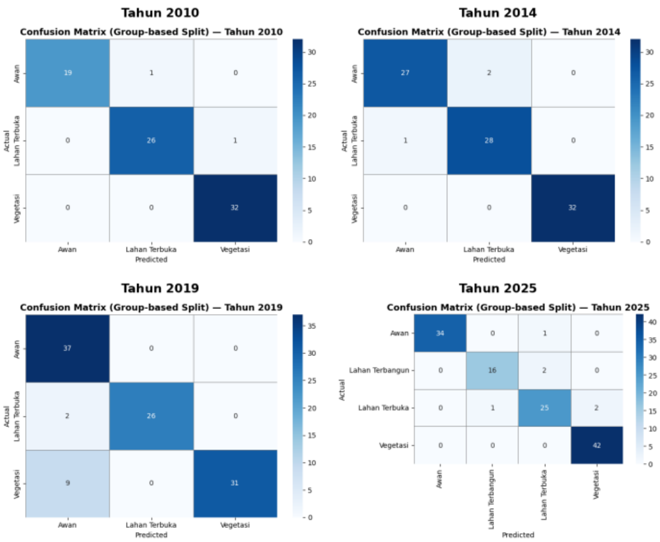

Figure 6. Confusion matrix of SVM classification results for 2010–2025

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : **https://doi.org/10.58602/itsecs.v4i4.388**
Page 188-200

In 2010, out of 79 total test samples, the model correctly predicted 77 (vegetation class: 32/32 correct; open land: 26/27 correct, with 1 sample misclassified as vegetation;clouds: 19/20 correct, with 1 sample misclassified as open land). In 2014, the model correctly predicted 77 out of 80 total test samples; classification errors occurred with 1 open land sample predicted as cloud and 2 cloud samples predicted as open land. In 2019, out of 105 total test samples, 94 were predicted accurately, while 11 were misclassified. The highest error rates occurred in the vegetation class (9 out of 40 samples predicted as cloud), likely due to cloud cover or shadows over vegetated areas, and in the open land class (2 out of 28 samples predicted as cloud). In 2025, with the addition of a built-up land class, 117 out of 123 total test samples were correctly classified, while 6 were misclassified; the built-up land class achieved 88.9% accuracy (16 out of 18 samples, with 2 misclassified as open land), the open land class achieved 89.3% accuracy (25 out of 28 samples, with errors distributed across the built-up land and vegetation classes), and the vegetation class achieved perfect accuracy (42 out of 42 samples).

The performance testing of the SVM algorithm employed a data split of 70% for training and 30% for testing, validated via the 5-fold cross-validation method. Quantitative evaluation metrics included Overall Accuracy, User's Accuracy (Precision), Producer's Accuracy (Recall), and Cohen's Kappa. Test results demonstrated very high prediction accuracy: 97.47% (2010), 96.67% (2014), 89.52% (2019), and 95.12% (2025). In comparison to the baseline approach—conventional classification relying solely on single spectral index thresholds (such as NDVI or BSI)—the SVM method overcomes fundamental limitations in identifying cloud-covered or shadowed areas. SVM proved significantly more effective than the pure spectral index baseline because it simultaneously learns multi-dimensional features (combining band reflectance and indices) to determine the optimal hyperplane for separating land cover classes. The model's low error rate is further evidenced by the minimal gap between training and testing accuracies (ranging from only 1.15% to 2.12%), confirming the model's genuine effectiveness and the absence of overfitting.

4. CONCLUSION

Based on the research objectives—to determine the extent of land cover change and evaluate the effectiveness of the SVM algorithm—the following conclusions can be drawn: Massive land cover change has occurred in the PT IWIP mining area due to the intensification of industrial activities. This is evidenced by a reduction in vegetation cover from 91.2% (24,661.71 ha) in 2010 to 84.3% (22,856.56 ha) in 2025. The most significant changes were observed between 2019 and 2025, characterized by a sharp increase in open land to 11.4% (3,099.52 ha) and the emergence of 75.20 ha of built-up land cover.

The Support Vector Machine (SVM) algorithm demonstrated stability in conducting remote sensing-based land cover change analysis. Quantitative evaluation confirmed the model's ability to classify land cover types with excellent testing accuracy—specifically 97.47% (2010), 96.67% (2014), 89.52% (2019), and 95.12% (2025)—and Cohen's Kappa values ranging from 0.84 to 0.96. The model also demonstrated strong generalization capabilities without overfitting, as evidenced by the minimal difference between training and testing accuracy (<2.2%).

CREDIT AUTHORSHIP CONTRIBUTION

Efraim Julian Soma: Writing – review & editing, Writing – original draft, Validation, Model Implementation, Methodology, Conceptualization. **Winsy Christo**

Efraim Julian Soma: *Corresponding Author

Deilan Weku: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Rillya Arundaa:** Writing – original draft, Model Implementation, Data curation. **Mahardika Inra Takaendengan:** Writing – original draft, Model Implementation, Data curation.

ACKNOWLEDGMENT

The authors thank the USGS and ESA Copernicus Open Access Browser for providing access to the satellite imagery data used in this study. This research received no external funding. The authors declare no conflicts of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The authors used ChatGPT-5.6 to improve the writing clarity of this paper. They reviewed and edited the AI-assisted content and take full responsibility for the final publication.

5. REFERENCES

[1] M. G. Aryatama, A. Sukmono, and F. Hadi, "ANALISIS PERUBAHAN TUTUPAN LAHAN MENGGUNAKAN CITRA SENTINEL-1 MULTITEMPORAL (Studi Lokasi: Kecamatan Singorojo, Kabupaten Kendal)," 2023.

[2] S. Dahlia, A. Luthfia, D. Muhammad, and S. Abfertiawan, "Analisis Perubahan Tutupan Lahan Kawasan Pertambangan Batubara Terhadap Pertumbuhan Penduduk dan Ekonomi: Studi Kasus Kota Ombilin dan Sangatta," 2020.

[3] B. Benedictus Lanang Ido Hernanto and B. Lanang Ido Hernanto, "ANALYSIS OF LAND COVER CHANGE IN MOROWALI USING LANDSAT 8 SATELLITE IMAGERY AND UNSUPERVISED CLASSIFICATION METHOD," *International Journal of Social Science (IJSS)*, vol. 5, pp. 2798–4079, 2025, doi: 10.53625/ijss.v5i1.10997.

[4] M. Webster, "Mengenal Apa Itu Citra Satelit dan Kegunaannya," https://terralogiq.com/citra-satelit/. Accessed: Jul. 12, 2026. [Online]. Available: https://terralogiq.com/citra-satelit/

[5] A. R. Nendrawan and S. Darmawan, "Klasifikasi Tutupan Lahan Berbasis Karakteristik Backscatter Menggunakan Citra Sentinel-1-1 dan Alos Palsar (Studi Kasus : Kecamatan Long Ikis, Kalimantan Timur)," 2021.

[6] M. Duka, F. Lihawa, and S. Rahim, "Perubahan Tutupan Lahan dan Pengaruhnya Terhadap Pola Persebaran Suhu di Kota Gorontalo," *Jambura Geoscience Review*, vol. 2, no. 1, pp. 16–29, Jan. 2020, doi: 10.34312/jgeosrev.v2i1.2682.

[7] USGS, "What is remote sensing and what is it used for?," https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used. Accessed: Aug. 02, 2026. [Online]. Available: https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used

[8] F. fahrid Muhsoni, "Penginderaan-Jauh_Firman-Farid-Muhsoni," 2015.

[9] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *REMOTE SENSING AND IMAGE INTERPRETATION Seventh Edition*, 7th ed., vol. 7. New York.: John Wiley and Sons, 1979.

[10] J. M. Álvarez-Alvarado, J. G. Ríos-Moreno, S. A. Obregón-Biosca, G. Ronquillo-Lomelí, E. Ventura-Ramos, and M. Trejo-Perea, "Hybrid techniques to predict solar

Efraim Julian Soma: *Corresponding Author

Journal of Information Technology, Software Engineering and Computer Science
Website : https://ejournal.techcart-press.com/index.php/itsecs
E-ISSN 2962-0635
P-ISSN 2962-0538
Volume 4, Number 4, October 2026
DOI : **https://doi.org/10.58602/itsecs.v4i4.388**
Page 188-200

radiation using support vector machine and search optimization algorithms: A review," Feb. 01, 2021, *MDPI AG*. doi: 10.3390/app11031044.
[11] Geeksforgeeks, "StandardScaler, MinMaxScaler and RobustScaler techniques - ML," geeksforgeeks. Accessed: Aug. 01, 2026. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/standardscaler-minmaxscaler-and-robustscaler-techniques-ml/
[12] Great learning, "Hyperparameter Tuning with GridSearchCV."
[13] S. A. Salleh *et al.*, "Support Vector Machine (SVM) and Object Based Classification in Earth Linear Features Extraction: A Comparison," *Revue Internationale de Géomatique*, vol. 33, no. 1, pp. 183–199, 2024, doi: 10.32604/rig.2024.050723.
[14] Kemdiktisaintek, "Modul Ajar Evaluasi Model Machine Learning," 2023.
[15] R. Maulid, "Tipe Machine Learning dengan K-Fold Cross Validation." Accessed: Aug. 01, 2026. [Online]. Available: https://dqlab.id/tipe-machine-learning-dengan-k-fold-cross-validation
[16] B. Umar, "Evaluasi Model: Metode dan Metrik Penting dalam Machine Learning."
[17] J. Cohen, "A Coefficient of Agreement for Nominal Scales.," *Educational and Psychological Measurement*, vol. 20, pp. 37–46, 1960.
[18] N. Ahmed, "Application of NDVI in Vegetation Monitoring Using GIS and Remote Sensing in Northern Ethiopian Highlands," 2016. [Online]. Available: https://www.researchgate.net/publication/346099273

Efraim Julian Soma: *Corresponding Author