

Research article

Integration of Texture and PCA Information from Sentinel-1 SAR Data for Land Cover-Analysis using Random Forest Classifier Method in Sidoarjo Regency, Indonesia

Filsa Bioresita*, Trisya Sayyidah Fithri Larastika, Muhammad Taufik, Noorlaila Hayati, Chelsea Alfarella Putri Taslyanto

Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Geomatics Engineering Building, ITS Campus, Sukolilo, Surabaya 60111, Indonesia

*Correspondence: filsa.bioresita@its.ac.id

Citation:

Bioresita, F., Larastika, T. S. F., Taufik, M., Hayati, N., & Taslyanto, C. A. P. (2025). Integration of Texture and PCA Information from Sentinel-1 SAR Data for Land Cover-Analysis using Random Forest Classifier Method in Sidoarjo Regency, Indonesia. *Forum Geografi*. 39(1), 38-52.

Article history:

Received: 30 July 2024
Revised: 13 January 2025
Accepted: 12 March 2025
Published: 13 March 2025

Abstract

Land cover has an important role in modelling to spatially analyse natural phenomena that occur on the earth's surface. The identification of land cover can also be used to determine the availability of green space and the percentage of built-up land in an area. Through this information, it can help the government to formulate policies related to development planning in an area. Currently, land cover identification can be done with remote sensing technology, generally using optical imagery. However, there are obstacles when using optical imagery, namely, if the cloud cover in an area is thick enough, it will affect the accuracy of the land cover results. To anticipate this, land cover identification can be done using active or radar imagery, one of which is the Sentinel-1 GRD image. The active image is not influenced by clouds and can record information without being constrained by weather both during the day and night. Sentinel-1 GRD data contains backscattering information that can be extracted using texture analysis and Principal Component Analysis (PCA). The Random Forest classifier was employed early in this study to analyze Sentinel-1 data, enabling classification using various inputs. Land cover classification from several inputs, namely, sigma, gamma, and beta from backscattering data, resulted in overall accuracy of 86.154%, 87.692%, and 86.154%.

Keywords: Random Forest Classifier; Sentinel-1 GRD; Land Cover

1. Introduction

Land cover refers to the physical features of the earth's surface (Regasa *et al.*, 2021). Land cover is a description of the appearance of physical and biological cover on the earth's surface that is formed naturally such as rivers and hills or man-made such as rice fields, gardens, and buildings (Prasetyo, 2023). This land cover has an important role in modelling to spatially analyse natural phenomena that occur on the earth's surface (Liang, 2008). The identification of land cover can also be used to determine the availability of green spaces and the percentage of built-up land in an area (Huda *et al.*, 2022). Through this information, it can help the government to formulate policies related to development planning in an area.

Mapping land cover and its changes is crucial for understanding the current state of the environment (Šćepanović *et al.*, 2021). These changes in LC can result from human activities or climate variations at a regional level. The identification of land cover in Sidoarjo Regency is necessary to monitor the influence of Lapindo mudflow. Sidoarjo Regency is synonymous with Lapindo mudflow, which has been spewing for 18 years ago. The Lapindo mudflow then flowed into the Porong River. From the disposal of the mud, there was a sedimentation process that formed at the mouth of the Porong River (Pryambodo *et al.*, 2021). The result of this sedimentation process is a new landmass that is now called Lusi Island. This continuous sedimentation process is not impossible if a delta or landmass is created in the future. if there will be a delta or new land (Bagaskara *et al.*, 2017). Therefore, it is necessary to monitor land cover in Sidoarjo Regency regularly.

Land cover can be identified using remote sensing technology. Remote sensing was chosen because this method can be used to obtain information on an object or phenomenon on the earth's surface without making direct contact with the object. This remote sensing approach can use both optical imagery and radar imagery. The use of optical imagery in Indonesia is currently, still the main choice to support mapping in various sectors, because the visualization of optical images is easier to interpret by the human eye because it uses hue, colour, and original appearance (Fadlin *et al.*, 2021). Meanwhile, radar or SAR images are only visualized in black and white. However, the advantage that optical images do not have is that radar images in their recording are not affected by clouds and can record information day and night (Fadlin *et al.*, 2021). One of the SAR images that is freely available is the Sentinel-1 image from the European Space Agency (ESA).

This study investigates land cover classification using the Random Forest classifier on Sentinel-1 imagery with input features including PCA, texture, and backscattering, addressing gaps in prior



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

research. Previous studies utilized Sentinel-1 imagery with varying preprocessing steps and classification methods. Most land cover classification methods are currently, carried out by combining polarimetric to get a combination of polarimetric with the best accuracy test as was done in research (Makinde & Oyelade, 2018) using the SVM (Support Vector Machine) method. Another land cover method that can be used is the Random Forest Classifier, several studies using this method such as research (Ardha, 2021) observed land cover changes in the Citarum watershed. Antara *et al.* (2021) achieved 96.90% accuracy using Random Forest with dual-polarimetric composite images. Dahhani *et al.* (2022) classified agricultural areas in Kaffrine with multiple classifiers, highlighting Random Forest's superior performance (84% accuracy) and PCA's role in data dimensionality reduction. Chulafak *et al.* (2018) focused on texture-based land use classification with Neural Networks, achieving 77% accuracy, emphasizing entropy features. This study integrates PCA, texture, and backscattering features to enhance land cover classification using Random Forest, aiming to improve accuracy and efficiency. In that study, classification using random forest was done directly and without adding inputs.

Random forest is an advanced decision tree model that generates several trees for regression and classification investigations. Overfitting is conceivable in traditional DT models, which rely on a single tree for explanatory power. The random forest model was built expressly to address this issue (Sadia *et al.*, 2023). The Random Forest classifier is chosen because it is capable of handling a variety of input types, namely backscattering data, texture features, and PCA results. In contrast to SVM, which works best with fewer input variables and linear separability, Random Forest can analyse more complicated and multidimensional data, making it more suitable for land cover classification in this study. Furthermore, Random Forest is less sensitive to overfitting and is well-known for its great accuracy across a wide range of data inputs. Random forest involves generating multiple trees and then combining their results through a majority vote, where each tree has an equal weight (Tariq *et al.*, 2023). The Random Forest (RF) machine learning algorithm employs 100 estimators. The data is separated into two sets: 70% is used for training, and 30% is reserved for testing (Haydar *et al.*, 2024). Random forest utilizes the bootstrap aggregation method to generate multiple samples from the original dataset and applies decision tree modeling to these samples (Wu *et al.*, 2021). Every decision tree is built independently without pruning, and each node is split using a subset of randomly chosen variables from the number of features specified by the user. This approach allows for the creation of trees with high variance and low bias, resulting in a forest composed of multiple trees (Ntree), or estimators, as defined by the user (Adugna *et al.*, 2022). The random forest method allows users to use various inputs in the classification process. One of them is by adding the input of PCA (Principal Component Analysis) results and texture results from backscattering data. All of this information will be input into the Random Forest classification process. This research was conducted by utilizing backscattering information, namely, sigma, beta, and gamma. Each of the backscattering information will be extracted PCA (Principal Component Analysis) and texture information to be classified.

PCA is frequently utilized as a linear, unsupervised method for feature extraction to decrease the dimensionality of remote sensing data (Uddin *et al.*, 2021). Principal component analysis is a technique to identify redundancies and differences in a stack of bands and convert them into new uncorrelated bands. This PCA algorithm will produce a value that can be called the Principal Component (PC). The PC value comes from a linear combination of the values before the reduction (Pujianto *et al.*, 2019). The principle of PCA is to minimize data dimensionality while preserving the highest possible variance. The objective of PCA is to determine an optimal position that maximizes information variance and efficiently reduces the dimensionality of vector features (Salih Hasan & Abdulazeez, 2021). By applying PCA, redundant and unnecessary features are eliminated, making the remaining features more prominent and organized within a new space known as the principal component (Almaiah *et al.*, 2022). The first principal component (PC1) represents the direction with the highest variance. The subsequent principal components (PC2, PC3, etc.) are orthogonal to each preceding one and capture the maximum remaining variance (Dharani & Sreenivasulu, 2021).

Texture extraction aims to calculate local spatial patterns in an image, providing beneficial details for determining land cover types. Homogeneity, energy, maximum probability, entropy, and mean are key characteristics to determine land cover patterns. Homogeneity reveals areas that are identical, whereas entropy reflects the region's complexity. These characteristics are especially important for distinguishing between developed areas and natural land cover. Texture extraction is extracting metric derivatives that describe local spatial patterns from gray-scale images quantitatively (Braun, 2020). The objective of texture extraction is to characterize texture using a mathematical model that enables the identification of distinct textures and objects in an image (Aouat *et al.*, 2021). Texture extraction approaches can be categorized by statistical properties,

mathematical models, geometry (structural approach), and signal processing (Najafi Khanbebin & Mehrdad, 2021). One method that can be used in extracting the texture feature is GLCM. The GLCM is a two-dimensional array where both dimensions are equal to the number of gray levels in the image (Aggarwal, 2022). GLCM is a technique for image analysis that uses statistical methods to investigate pixel patterns within a gray-scale matrix, also referred to as the gray-scale spatial co-occurrence matrix. As a classification approach, its final stage involves training a classifier. The main purpose of GLCM is to extract texture features from images. It creates a gray-scale co-occurrence matrix and then derives statistical functions from this matrix based on the specified values and spatial relationships of the image's texture (Althubiti *et al.*, 2022). The statistical characteristics of GLCM include maximum probability (maximum value), entropy (degree of randomness), energy (degree of uniformity), correlation (linear relationship of gray levels), contrast (content of local variation), and homogeneity (homogeneity of pixels) (Septiarini & Wardoyo, 2015).

Land cover identification in Sidoarjo Regency is critical for monitoring the consequences of the Lapindo mudflow. The Lapindo mud eruption, which have been persistent for the past 18 years, flows into the Porong River, resulting in sedimentation at the river's estuary. This sedimentation formed new land, presently known as Lusi Island. Continuous sedimentation may eventually result in new deltas or other landforms. Therefore, periodic monitoring of land cover in Sidoarjo Regency is essential.

The research problems addressed in this project focus on the extraction and analysis of information from Sentinel-1 imagery. Specifically, the study examines the extraction of backscattering information, including σ_0 (sigma nought), γ_0 (gamma nought), and β_0 (beta nought), as well as Principal Component Analysis (PCA) and texture information from the imagery. Additionally, the research explores the application of the Random Forest classifier method to identify land cover using a combination of input data, such as σ_0 , γ_0 , β_0 , PCA, and texture information, derived from Sentinel-1 imagery.

2. Research Methods

2.1 Study Area

This land-cover research using the Random Forest method was conducted in Sidoarjo Regency. Sidoarjo Regency is located between 112.5° to 112.9° east and 7.3° to 7.5° south with an area of 719.34 km². Sidoarjo regency borders Surabaya city and Gresik regency to the north, Pasuruan regency to the south. To the west is Mojokerto regency and to the east is the Madura Strait.

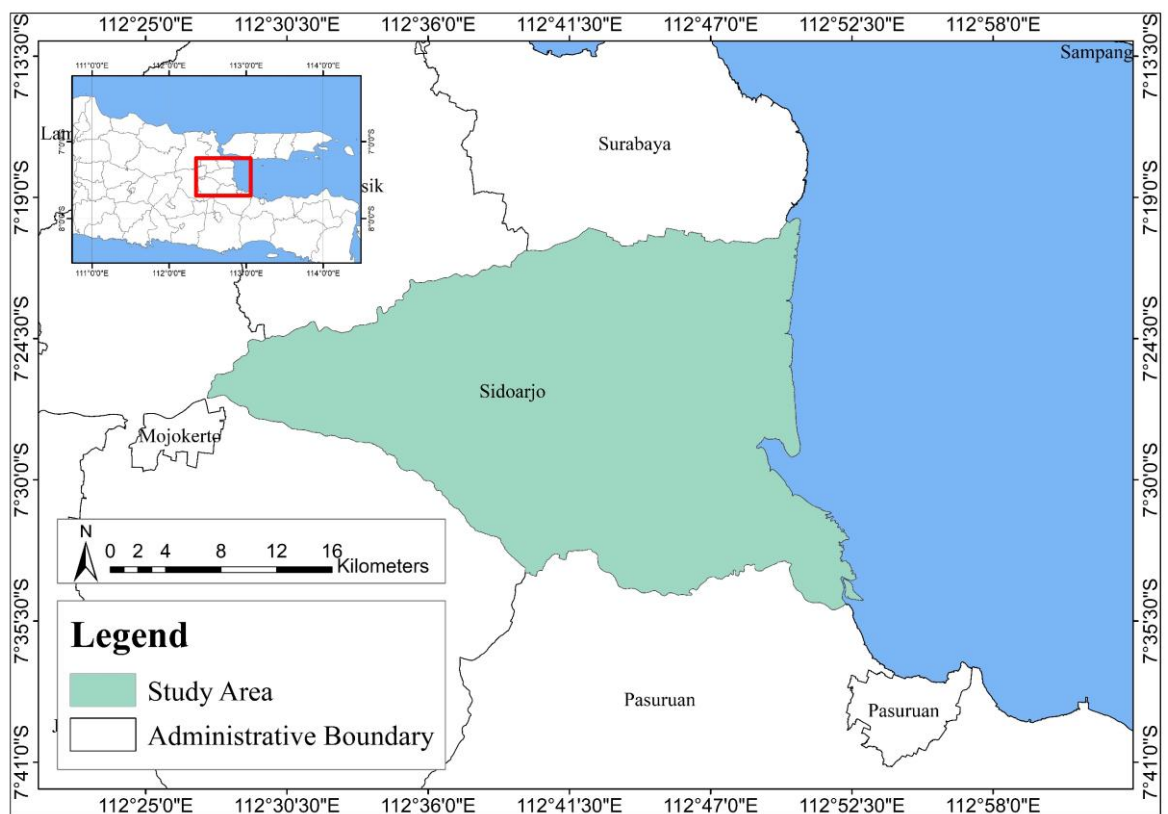


Figure 1. Study Area.

2.2 Pre-Processing Data

The data required is Sentinel-GRD SAR image data dated January 14, 2024 and April 19, 2024. The coordinates of the area of interest (AOI) of the research area, in this case Sidoarjo, and ground truth data are also needed for validation. The Sentinel-1 SAR GRD imagery from January 14, 2024, represents the rainy season, while the imagery from April 19, 2024, represents the dry season.

The preprocessing step comprises of many essential steps that prepare SAR images for accurate classification. Image cropping guarantees that only the areas of interest (AOI) are analyzed. Coregistration to unify the two images obtained from the same trajectory. Radiometric calibration corrects the raw SAR image by modifying pixel values to reflect true radar backscatter from the surface, ensuring that the data appropriately represents the reflectivity of objects in the scene. Speckle filtering is implemented to decrease noise in the image, which improves the clarity of land cover elements.

In addition, terrain correction is carried out to eliminate any topographic distortions and project the image into a consistent reference system. The last step is to convert the dB scale to achieve a more normal distribution of values, using the log function which functions so that the pixel value becomes a logarithmic scale and produces a higher contrast.

2.3 Texture and PCA Extraction

Image texture is a derived metric that quantitatively describes the local spatial pattern of a gray-scale image. The image performed as input to the texture extraction is the January 14, 2024 image of VH and VV polarization. The texture extraction is performed from the Sentinel-1 image by extracting Homogeneity, Energy, Maximum Probability, Entropy and Mean information. Meanwhile, PCA is used to identify redundancies and differences in a stack of bands and transform them into new uncorrelated bands (principal components) into PCA1, PCA2, PCA3 and PCA4. The images used in this PCA extraction input are both images dated 19 January 2024 and 14 April 2024 of VH and VV polarization.

2.4 Classification of Land Cover

Digitization of training areas can be done by manually digitizing by looking at optical images. The data division between training data and testing data is 70:30 with a total of 260 training data and testing data obtained from field validation. Land cover classification is performed using the random forest classifier method using three inputs namely, PCA, texture, and backscattering. The results of this classification consist of 5 land cover classes namely, water bodies, built-up land, open land, non-agricultural vegetation, and agricultural vegetation.

2.5 Accuracy Test

The accuracy test is used to determine the level of accuracy of land cover classification results in Sidoarjo. This accuracy test is carried out by digitizing land cover samples that match field conditions (ground truth). Ground truth is used to perform calculations using the confusion matrix. This accuracy test includes overall accuracy, user accuracy, producer accuracy, and kappa coefficient. The accuracy test formula according to Islami (2022) (Equation 1, 2, 3, and 4).

$$\text{Overall accuracy} = \frac{\text{Number of true pixels (diagonal)}}{\text{Number of reference pixels}} \times 100 \quad (1)$$

$$\text{User accuracy} = \frac{\text{Number of correct pixels per class}}{\text{Number of reference pixels per class (total rows)}} \times 100 \quad (2)$$

$$\text{Producer accuracy} = \frac{\text{Number of correct pixels per class}}{\text{Number of reference pixels of each class (total columns)}} \times 100 \quad (3)$$

$$\text{Kappa coefficient} = \frac{(TS \times TCS) - \sum(\text{total columns} \times \text{total rows})}{TS^2 - \sum(\text{total columns} - \text{total rows})} \times 100 \quad (4)$$

The overall stages of data processing are summarized in the flow chart in Figure 2.

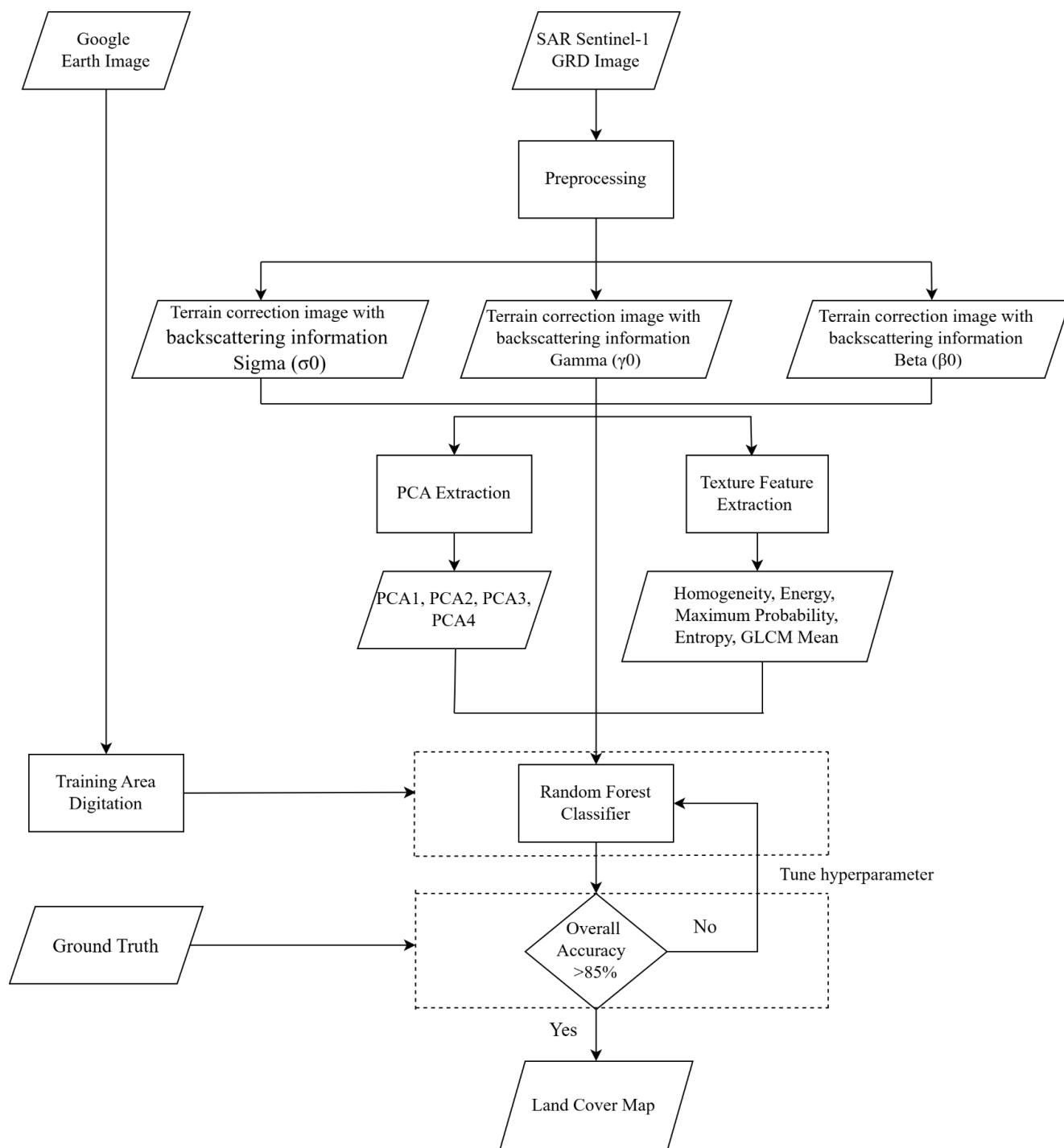


Figure 2. Flow Diagram of Research Method.

3. Results and Discussion

3.1. Extraction of Backscattering Information

The backscattering used in this study is sigma nought, gamma nought, and beta nought. Table 1 shows the results of backscattering information extraction on the image dated January 14, 2024. From the visualization results of the backscattering information in the Table 1, it is found that between sigma, gamma, and beta there is no significant difference. This can be seen in the max and min values of the image. The polarization of the Sentinel-1 image is divided into two polarizations, namely, VH and VV polarization. In the image dated January 14, 2024, the results of the extraction of backscattering information between VH and VV polarizations have the same colour, namely, black for water bodies and white for land. However, what distinguishes between VV and VH is that VV polarization tends to have a darker hue compared to VH polarization. When viewed from the max and min values in Table 2, VV polarization tends to have a value greater than VH polarization, this value is thought to affect the colour that tends to be darker in VV polarization.

Table 1. Backscattering Information Extraction on January 14, 2024 Image.

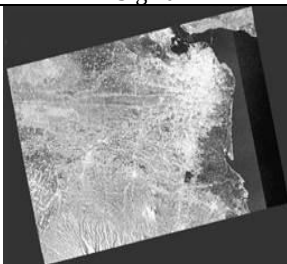
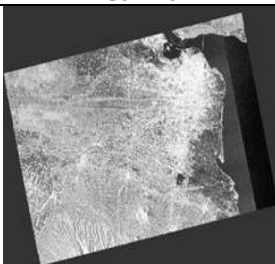
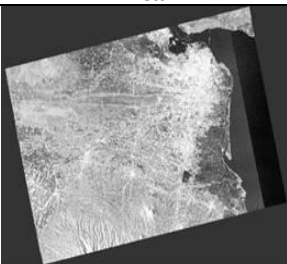
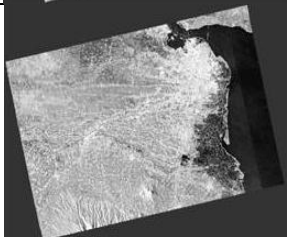
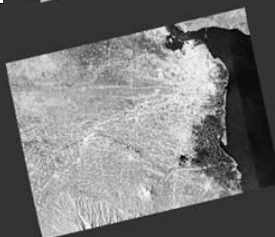
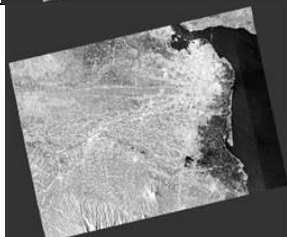
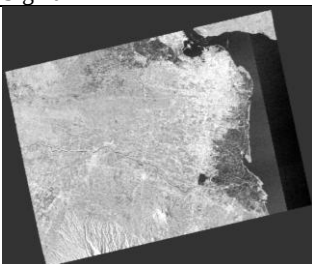
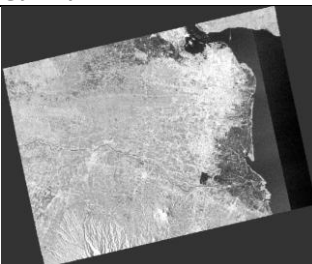
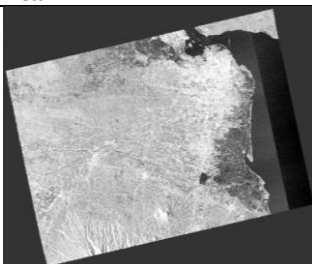
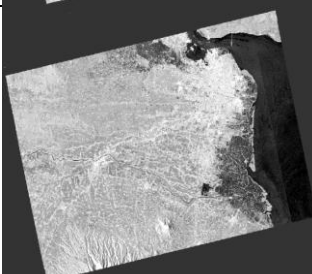
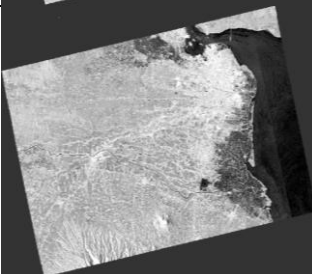
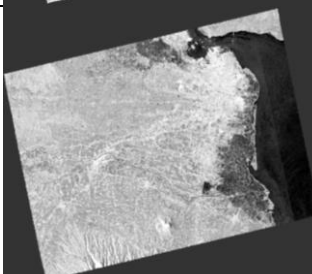
	Sigma	Gamma	Beta
VH			
VV			

Table 2. Range of Value of Backscattering Extracted Image.

No	Image Processing	Range of Value		
		Sigma	Gamma	Beta
1.	January 14, 2024 VH Polarization	-26.630 to 7.759	-24.394 to 10.114	-25.676 to 8.654
2.	January 14, 2024 VV Polarization	-22.473 to 18.968	-21.512 to 19.818	-24.394 to 10.114
3.	April 19, 2024 VH Polarization	-26.277 to 9.011	-25.317 to 9.906	-24.043 to 11.366
4.	April 19, 2024 VV Polarization	-23.478 to 21.450	-22.538 to 22.373	-21.206 to 23.754

While the results of backscattering information extraction on the image dated April 19, 2024 are presented in Table 3. In the image dated April 19, 2024, between VH and VV polarizations tend to have no difference. The VH and VV polarizations for sigma, gamma, and beta have values that are not much different when viewed from Table 2, which can affect the visual appearance of the image which also does not show any difference. However, when compared between VH and VV, VV polarization tends to have a larger value which affects the visualization of VV which has a darker hue and has a more visible texture.

Table 3. Backscattering Information Extraction on April 19, 2024 Image.

	Sigma	Gamma	Beta
VH			
VV			

3.2. Extraction of Principal Component Analysis (PCA)

PCA extraction resulted in PCA1, PCA2, PCA3, and PCA4 as shown in Table 4. From the extraction of principal component analysis (PCA), the results of PCA1, PCA2, PCA3, PCA4, and PCA response tend to be the same between sigma, gamma, and beta and do not have significant differences. This can be seen from the max min value in PCA1, which is presented in Table 5. In this PCA1, between land and water, can still be clearly distinguished, black colour for water and gray colour for land, between residential areas and vegetation, can still be clearly distinguished.

Unlike the results of PCA1, the results of PCA2, PCA3, and PCA4 do not have clear boundaries between land and water bodies, as well as residential areas and vegetation cannot be identified properly. PCA2 is grouped based on pixel values from the range of -14.612 to 13.366 for sigma, -14.588 to 13.406 for gamma, and -14.651 to 13.287 for beta. The PCA3 is grouped by pixel values from the range of -15.313 to 20,070 for sigma, -15.339 to 20.060 for gamma, and -15.269 to 20.090 for beta. PCA4 results have pixel values in the range of -10.009 to 6.075 for sigma, -10.002 to 6.077 for gamma, and -10.019 to 6.072 for beta.

The results of this PCA response have a clear boundary difference between water bodies and land, it can be seen from the colour of the water body represented by white and the land represented by black. The pixel value of the PCA response can be seen in Table 5. From the results of this PCA response, not only water and land are very clear, but the differences between settlements, water bodies, and the topography of mountainous areas are also very clear.

Table 4. Principal Component Analysis (PCA) Extraction.

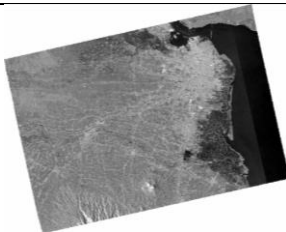
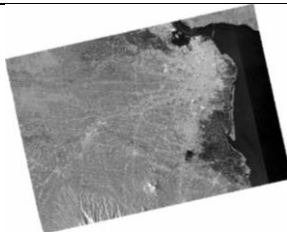
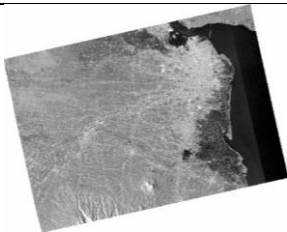
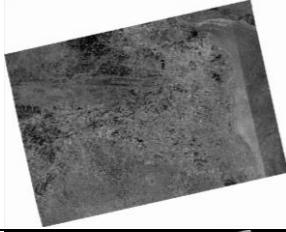
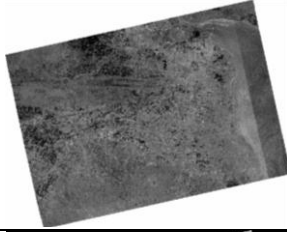
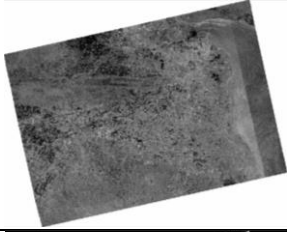
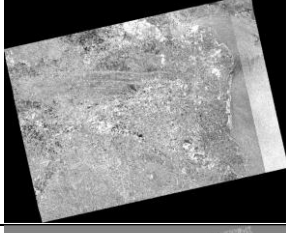
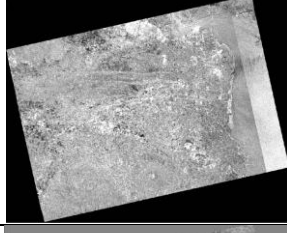
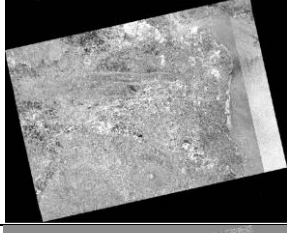
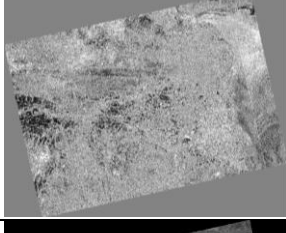
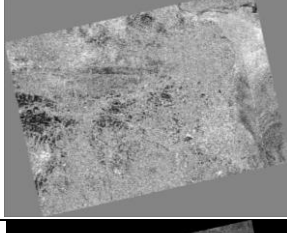
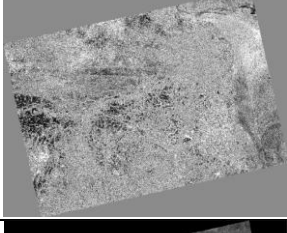
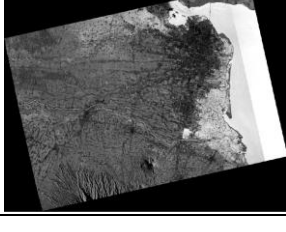
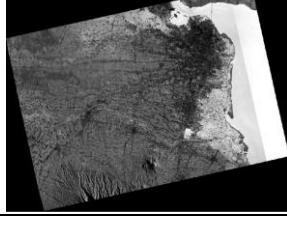
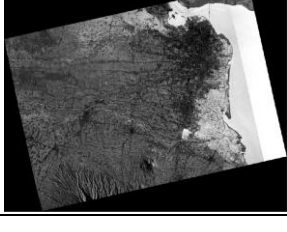
	Sigma	Beta	Gamma
PCA1			
PCA2			
PCA3			
PCA4			
PCA response			

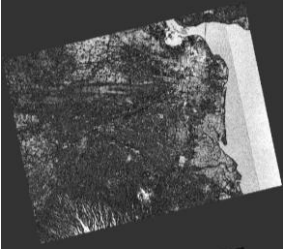
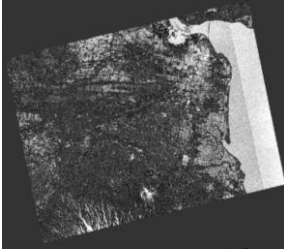
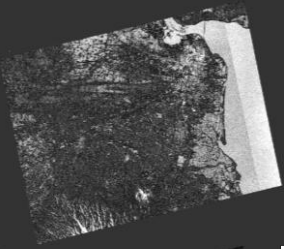
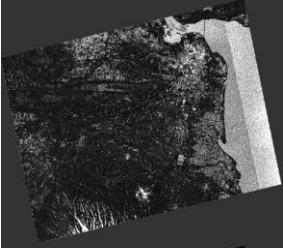
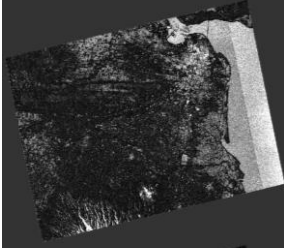
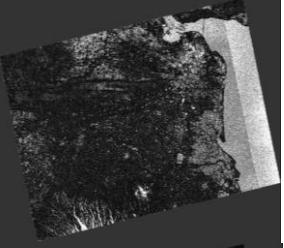
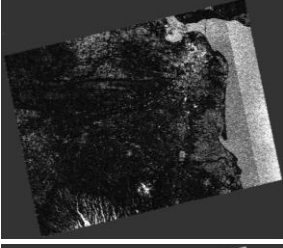
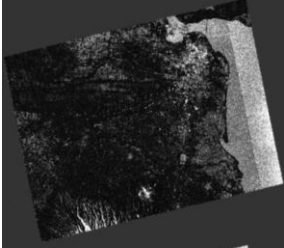
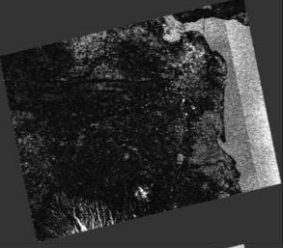
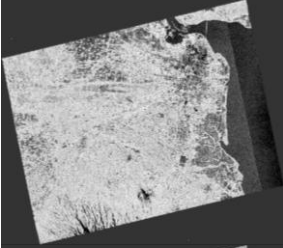
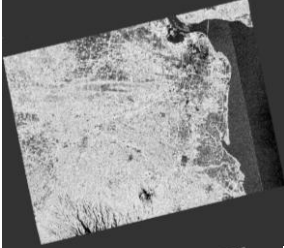
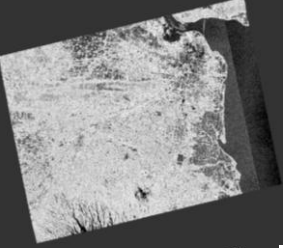
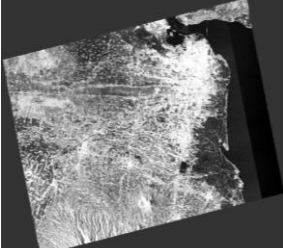
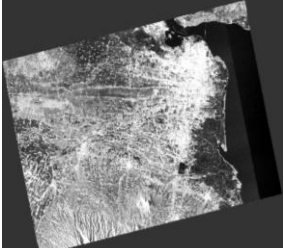
Table 5. Range of Values of PCA Extracted Image.

No	Image Processing	Range of Value		
		Sigma	Gamma	Beta
1.	PCA1	-21.453 to 49.019	-21.198 to 49.145	-21.990 to 48.743
2.	PCA2	-14.612 to 13.366	-14.588 to 13.406	-14.651 to 13.287
3.	PCA3	-14.651 to 13.287	-15.339 to 20.060	15.269 to 20.090
4.	PCA4	-10.009 to 6.075	-10.002 to 6.077	-10.019 to 6.072
5.	PCA response	0 to 46.481	0 to 44.577	0 to 42.041

3.3. Extraction of Texture Information

Image texture extraction is performed for each polarization. Texture extraction on VH polarization is presented in Table 6. The extraction results of homogeneity, energy, maximum probability, entropy and mean with VH polarization between sigma, beta, and gamma have no significant difference. The results of the extraction of homogeneity, energy, and maximum probability in sigma, beta, and gamma tend to be the same with white colour for water bodies while black for land and gray colour for densely populated areas and mountainous topography. This is also supported by the same pixel values for sigma and gamma of 0.014 to 2.000 and a range of 0.015 to 2.000 for beta. The energy extraction results show a black colour that is more intense compared to the black colour in homogeneity. This is indicated by the range of pixel values in energy which tends to be larger, as can be seen in Table 7. This maximum probability result has a less visible texture hue when compared to the energy result.

Table 6. Image Texture Extraction with VH Polarization.

	Sigma	Gamma	Beta
Homogeneity			
Energy			
Maximum probability			
Entropy			
Mean			

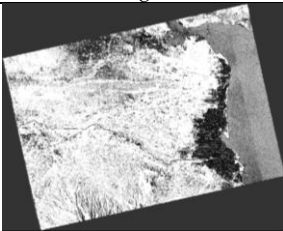
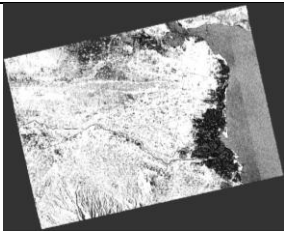
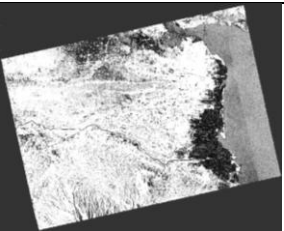
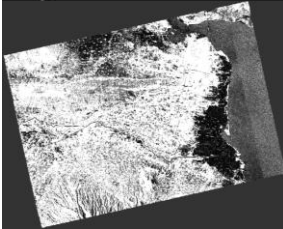
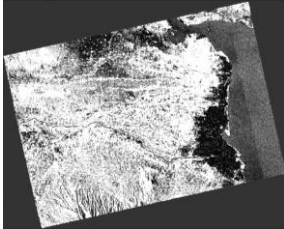
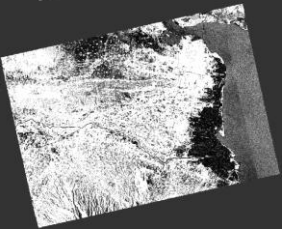
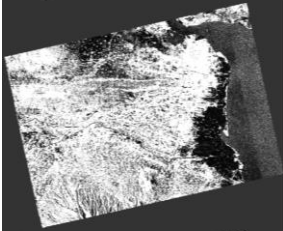
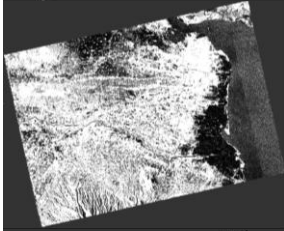
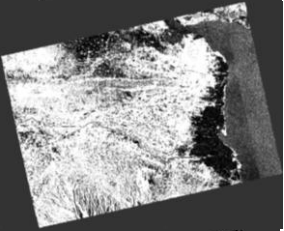
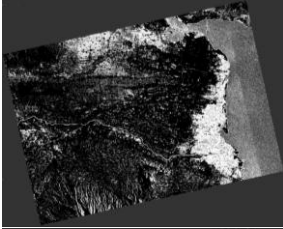
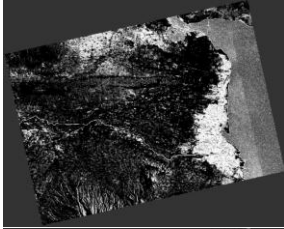
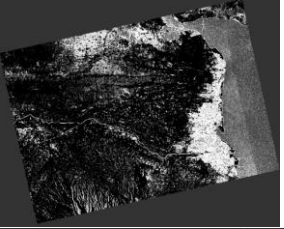
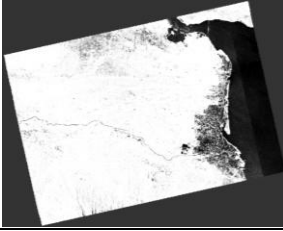
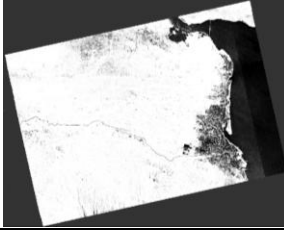
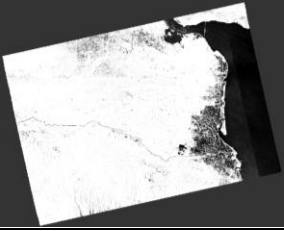
The results of entropy and mean extraction using VH polarization have the opposite colour compared to the extraction of homogeneity, energy and maximum probability. From the entropy and mean extraction results, the water body is solid black and the land is white. However, the mean extraction has a very intense black colour and a clear texture when compared to entropy. The results of entropy and mean extraction using VH polarization between sigma, gamma, and beta tend to be the same and have no significant difference.

Table 7. Range of Texture Extraction Result Values.

No	Texture Extraction Results	Range of Value		
		Sigma	Gamma	Beta
1.	Homogeneity VH polarization	0.014 to 2.000	0.014 to 2.000	0.015 to 2.000
2.	Energy VH polarization	0.137 to 2.000	0.138 to 2.000	0.136 to 2.000
3.	Maximum probability VH polarization	0.021 to 2.000	0.014 to 2.000	0.014 to 2.000
4.	Entropy VH polarization	-1.386 to 9.475	-1.386 to 9.428	-1.386 to 9.487
5.	GLCM Mean VH polarization	0.000 to 62.000	0.000 to 62.000	0.000 to 62.000
6.	Homogeneity VV polarization	0.069 to 2.000	0.104 to 2.000	0.068 to 2.000
7.	Energy VV polarization	0.142 to 2.000	0.139 to 2.000	0.144 to 2.000
8.	Maximum probability VV polarization	0.021 to 2.000	0.021 to 2.000	0.021 to 2.000
9.	Entropy VV polarization	-1.386 to 9.373	-1.386 to 9.410	-1.386 to 9.299
10.	GLCM Mean VV polarization	4.171 to 62.000	4.250 to 62.000	4.107 to 62.000

The results of the extraction of homogeneity, energy, maximum probability, entropy and mean using VV polarization have the opposite colour to when using VH polarization. In the extraction of homogeneity, energy and maximum probability using VV polarization, the black colour indicates water areas, while the white colour is for land. Between sigma, gamma, and beta in each homogeneity, energy and maximum probability extraction using VV polarization, there is no significant difference. Energy extraction results tend to be more textured compared to homogeneity extraction. While the maximum probability extraction results obtained a more textured appearance than energy extraction. The maximum probability extraction results have the same value for sigma, gamma, and beta in the range of 0.021 to 2.000.

Table 8. Image Texture Extraction with VV Polarization.

	Sigma	Gamma	Beta
<i>Homogeneity</i>			
<i>Energy</i>			
<i>Maximum probability</i>			
<i>Entropy</i>			
<i>Mean</i>			

The entropy texture extraction in this VV polarization produces black land and white water. While in the mean extraction, in this VV polarization, the water is shown in black, while the land is white. However, it can be seen that this land area has no texture or can be said to be dominant with white colour. This is supported by the very high max pixel value. The range of pixel values from the mean extraction results is 4.171 to 62.000 for sigma, 4.250 to 62.000 for gamma, and 4.107 to 62.000 for beta. The results of texture information extraction using VV polarization are presented in Table 8.

3.4. Land Cover

Land cover is obtained by performing classification using random forest and input by conducting image analysis using principal component analysis (PCA) techniques and texture extraction using the GLCM method. Table 9 shows that the results of land cover processing using sigma nought, gamma nought, and beta nought. Land cover is divided into 5 classes, namely, water body class, built-up land, open land, non-agricultural vegetation, and agricultural vegetation. From the land cover processing above, there are differences in area from each processing using sigma, gamma, and beta. Figure 3, 4 and 5 show the land cover classification map of Sidoarjo using sigma, beta and gamma backscattering information.

Table 9. Land Cover Area According to Backscattering Information.

Class	Sigma		Gamma		Beta	
	Area (km ²)	Percentage	Area (km ²)	Percentage	Area (km ²)	Percentage
Water body	204.098	28.07%	205.064	28.20%	201.971	27.78%
Built-up land	117.088	16.10%	120.654	16.59%	120.008	16.50%
Open land	123.128	16.93%	125.034	17.20%	126.177	17.35%
Non-agricultural vegetation	152.512	20.98%	151.886	20.89%	152.871	21.02%
Agricultural vegetation	130.302	17.92%	124.489	17.12%	126.1	17.34%
Total	727.128	100.00%	727.127	100.00%	727.127	100.00%

In the water body class, the largest area obtained from processing land cover using gamma backscattering information is 205.064 km² and the smallest area obtained from processing using beta backscattering information is 201.971 km². The difference between the largest area and the smallest area in the water body class is 3.093 km². While in the built-up land class, the largest area obtained from processing using gamma backscattering information is 120.654 km² and the smallest area from processing using sigma backscattering information is 117.088 km².

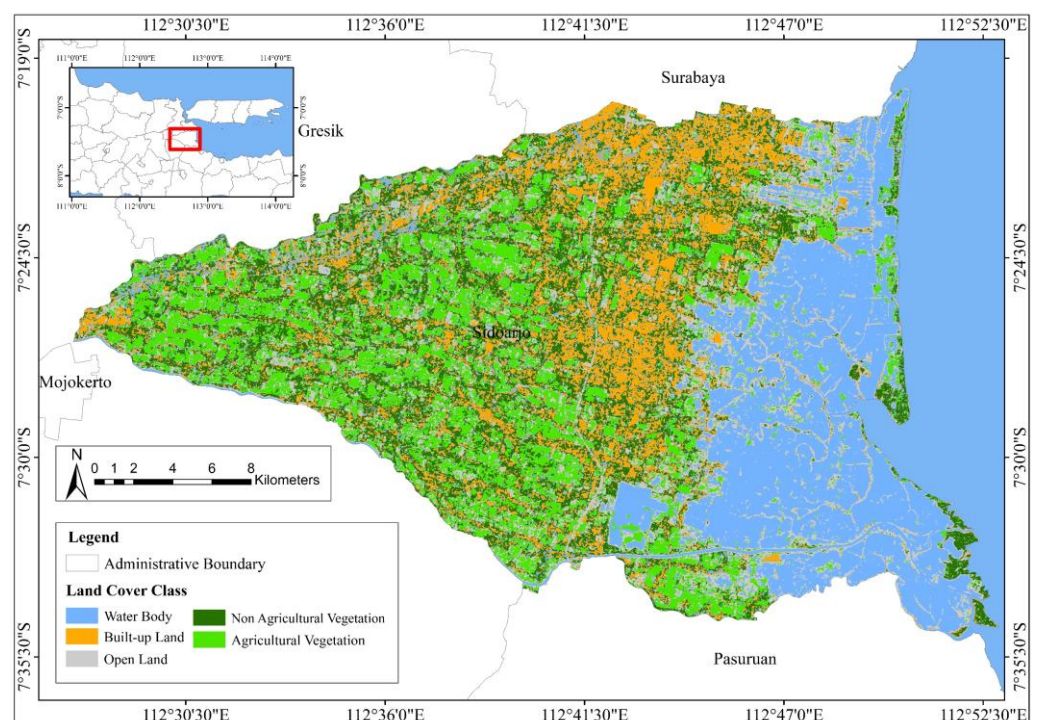


Figure 3. Land Cover Map Using Sigma Backscattering Information.

The open land class area between sigma, gamma, and beta backscattering has a considerable difference in area. The largest open land class area is obtained from processing using beta backscattering of 126.177 km². While the smallest open land area is obtained from processing using sigma backscattering of 123.128 km². The difference between the largest area and the smallest area in the open land class is 3.049 km². The largest area of non-agricultural vegetation class was obtained from processing using beta backscattering of 152.871 km² and the smallest area was obtained from processing using gamma backscattering of 151.886 km². The difference between the largest and smallest area in this non-agricultural vegetation class is 0.985 km². The agricultural vegetation class with the largest area obtained from processing using sigma backscattering is 130.302 km², the smallest area obtained from processing using gamma backscattering is 124.489 km². The difference between the largest and smallest area is 5.813 km².

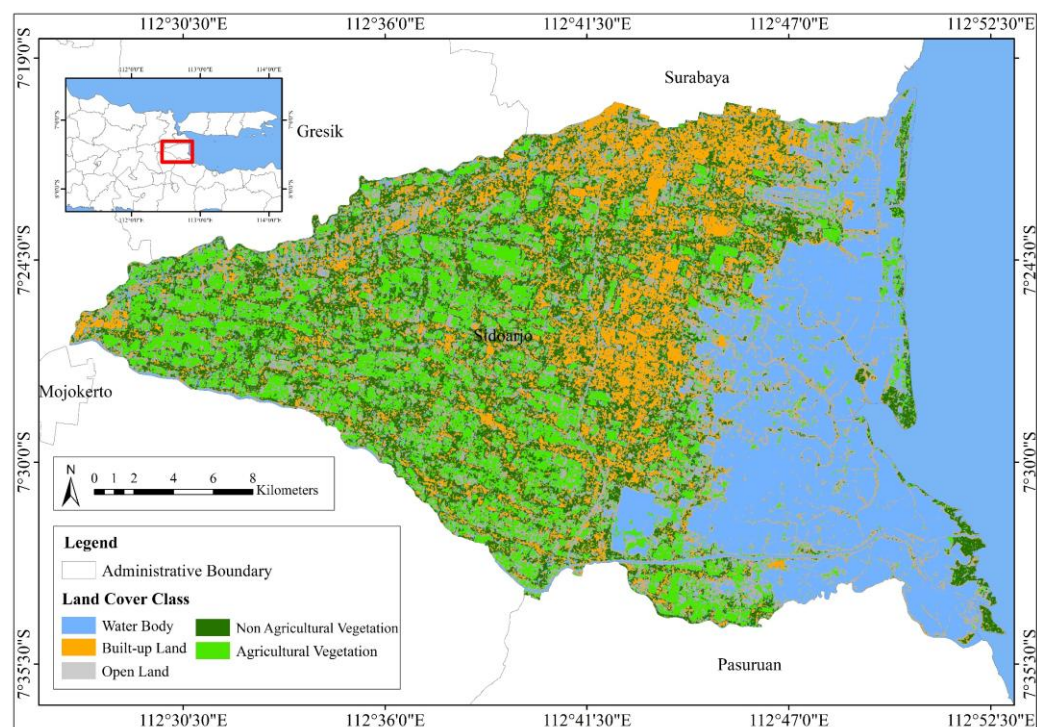


Figure 4. Land Cover Map Using Gamma Backscattering Information.

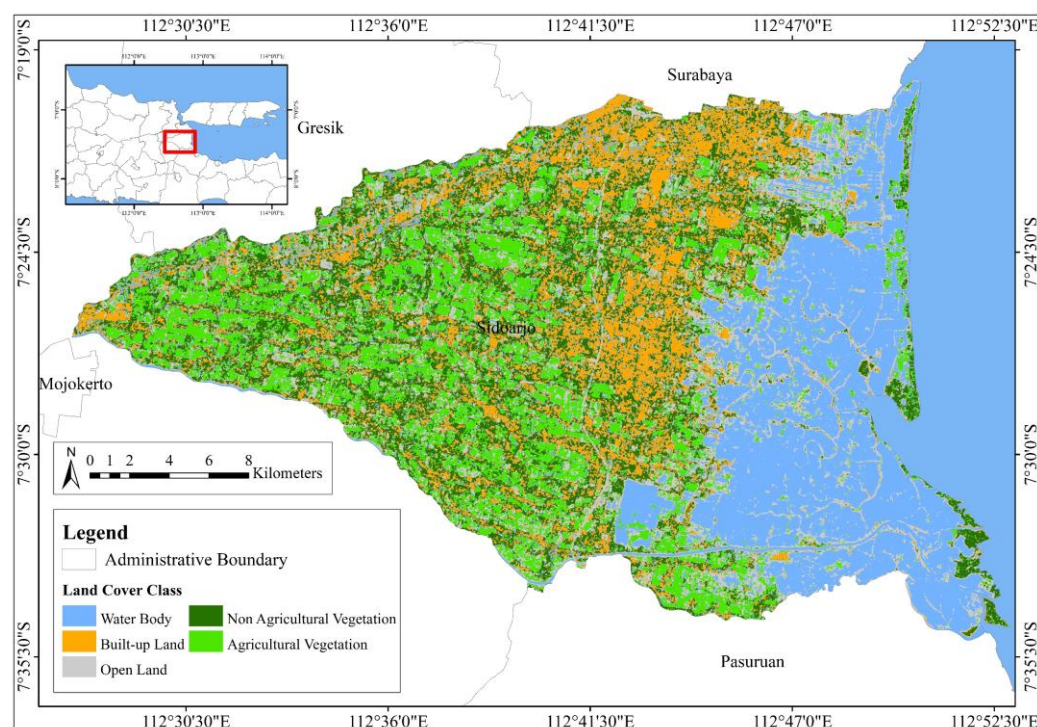


Figure 5. Land Cover Map Using Beta Backscattering Information.

3.5. Accuracy Test

This field validation was conducted directly on April 20, 2024 to April 22, 2024 and was carried out at 65 points spread across Sidoarjo. These validation points represent 5 classes equally. This validation stage was carried out by tagging. The distribution of validation points is presented in Figure 6.

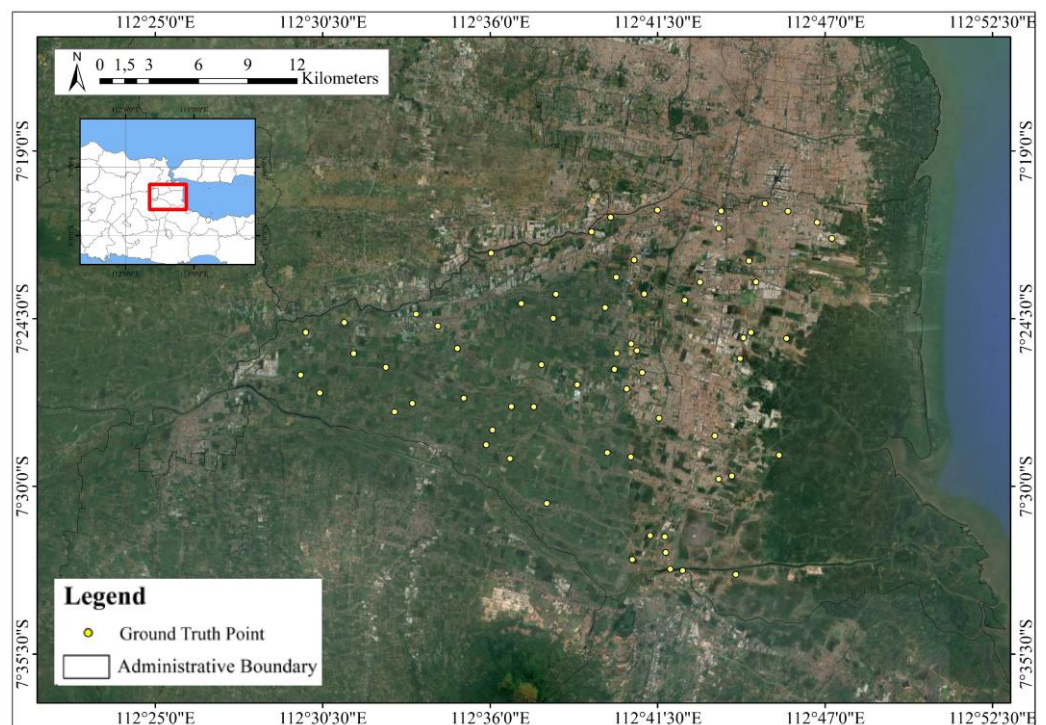


Figure 6. Ground Truth Point Distribution.

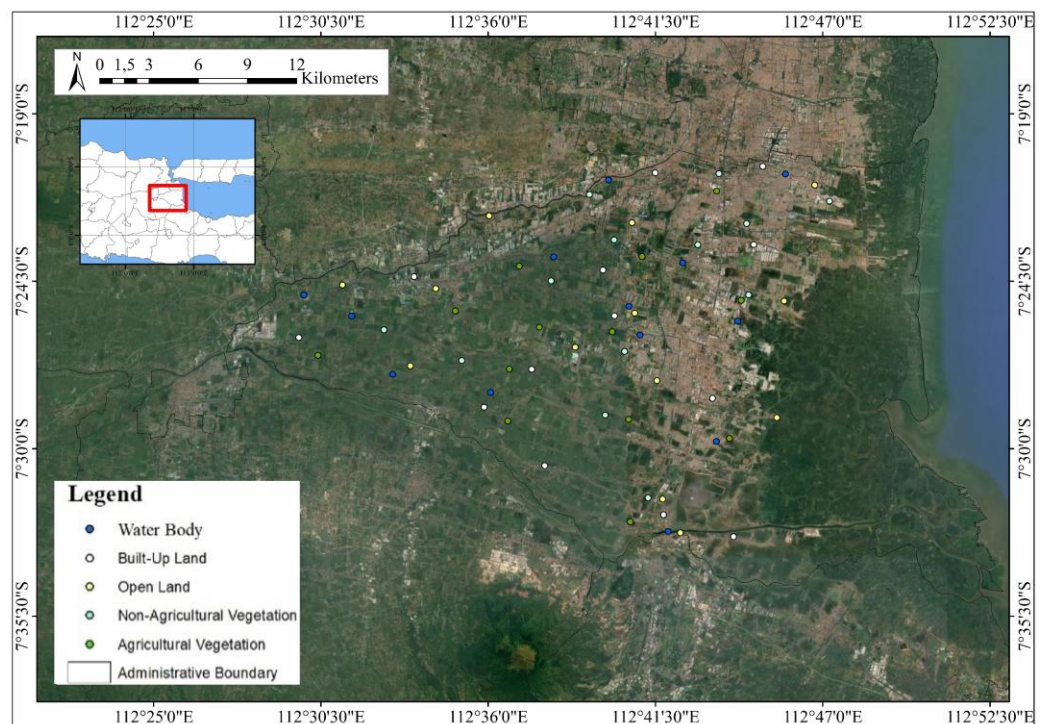


Figure 7. Ground Truth Point Distribution for Each Class.

The accuracy test was conducted using a confusion matrix to calculate overall accuracy, user accuracy, producer accuracy, and kappa coefficient. Confusion matrix was conducted on land cover maps with different backscattering information. The land cover map using sigma backscattering information produces an overall accuracy of 86.154% with user accuracy and producer accuracy for the water body class of 100% and 84.6154%, for built-up land of 84.6154% and

84.6154%, open land of 90% and 69.2308%, non-agricultural vegetation of 70.5882% and 92.3077%, while agricultural vegetation of 92.8571% and 100%. From the calculation with confusion matrix, the kappa coefficient for the land cover map using sigma nought is 0.82692.

In addition to using sigma nought, there is also a land cover map using gamma backscattering information resulting in an overall accuracy of 87.6923% with user accuracy and producer accuracy for water bodies of 100% and 84.6154%, built-up land of 80% and 92.3077%, open land of 100% and 69.2308%, non-agricultural vegetation of 75% and 92.3077%, and agricultural vegetation of 92.8571% and 100%. Kappa coefficient of the land cover map using gamma nought is 0.84615.

Table 10. Accuracy of Sigma Backscattering Information.

	Water Body	Built-Up Land	Open Land	Non-Agricultural Vegetation	Agricultural Vegetation	Total User	User Accuracy
Water Body	11	0	0	0	0	11	100.00%
Built-Up Land	0	11	1	1	0	13	84.62%
Open Land	0	1	9	0	0	10	90.00%
Non-Agricultural Vegetation	2	1	2	12	0	17	70.59%
Agricultural Vegetation	0	0	1	0	13	14	92.86%
Total Producer	13	13	13	13	13	65	
Producer Accuracy	84.62%	84.62%	69.23%	92.31%	100.00%		
Overall Accuracy							86.15%
Kappa Coefficient							0.827

Table 11. Accuracy of Gamma Backscattering Information.

	Water Body	Built-Up Land	Open Land	Non-Agricultural Vegetation	Agricultural Vegetation	Total User	User Accuracy
Water Body	11	0	0	0	0	11	100.00%
Built-Up Land	0	12	2	1	0	15	80.00%
Open Land	0	0	9	0	0	9	100.00%
Non-Agricultural Vegetation	2	1	1	12	0	16	75.00%
Agricultural Vegetation	0	0	1	0	13	14	92.86%
Total Producer	13	13	13	13	13	65	
Producer Accuracy	84.62%	92.31%	69.23%	92.31%	100.00%		
Overall Accuracy							87.69%
Kappa Coefficient							0.846

While the overall accuracy for land cover maps using beta nought is 86.1538% with user accuracy and producer accuracy for water body classes of 84.6154% and 84.6154%, built-up land classes of 91.6667% and 84.6154%, open land classes of 100% and 76.9231%, non-agricultural vegetation classes of 66.6667% and 92.3077%, and agricultural vegetation classes of 100% and 92.3077%. The kappa coefficient was obtained at 0.82692. From the results of the accuracy test calculation above, the greatest accuracy test was obtained on the land cover map using gamma nought. The amount of overall accuracy is proportional to the amount of kappa coefficient. So that the land cover map using gamma nought besides having the greatest overall accuracy, also has the greatest kappa coefficient. Kappa coefficient which is close to 1 and overall accuracy >85% indicates that the land cover map has been classified well.

Gamma backscattering is crucial in land cover categorization since it normalizes pixel values across different scenarios, assuring consistency and comparable results. This normalization method is particularly effective for volume scatterers such as vegetation, making gamma backscattering an excellent choice for precise land cover assessment. The scattering type has a substantial impact on categorization accuracy since the backscatter signal is affected by the physical properties of surface objects. For instance, vegetation exhibits volume scattering due to its complex structure, but urban environments are dominated by double-bounce scattering caused by reflective surfaces such as buildings. Classification errors frequently occur when the scattering type varies among surface characteristics, resulting in misinterpretations in the classification process. Addressing these scattering features is critical to increasing the accuracy of land cover categorization systems.

Table 12. Accuracy of Beta Backscattering Information.

	Water Body	Built-Up Land	Open Land	Non-Agricultural Vegetation	Agricultural Vegetation	Total User	User Accuracy
Water Body	11	1	1	0	0	13	84.62%
Built-Up Land	0	11	0	1	0	12	91.67%
Open Land	0	0	10	0	0	10	100.00%
Non-Agricultural Vegetation	2	1	2	12	1	18	66.67%
Agricultural Vegetation	0	0	0	0	12	12	100.00%
Total Producer	13	13	13	13	13	65	
Producer Accuracy	84.62%	84.62%	76.92%	92.31%		92.31%	
Overall Accuracy							86.15%
Kappa Coefficient							0.827

The efficiency of Sentinel-1 SAR data for land cover monitoring is demonstrated by the classification accuracy for various backscattering types. The practical consequences of these findings are particularly evident in environmental monitoring and regional planning for sustainable development, where precise land cover data is crucial for making decisions.

Acknowledgements

The authors would like to thank The Directorate of Research and Community Service (DRPM) Institut Teknologi Sepuluh Nopember (ITS), Indonesia for supporting the research funding. The authors would also like to thank ESA for providing Sentinel-1 data.

Author Contributions

Conceptualization: Bioresita, F.; **methodology:** Bioresita, F., Larastika, T. S. F.; **investigation:** Taufik, M., Hayati, N.; **writing—original draft preparation:** Bioresita, F., Larastika, T. S. F.; **writing—review and editing:** Larastika, T. S. F., Alfarelia, C.; **visualization:** Larastika, T. S. F.. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

Data availability

Data is available upon Request.

Funding

This research was funded by The Directorate of Research and Community Service (DRPM) Institut Teknologi Sepuluh Nopember (ITS), Indonesia in the final project upgrading program.

4. Conclusion

The extraction of backscattering information between sigma, gamma, and with VV and VH polarization using 2 dates, namely, January 14, 2024 and April 19, 2024, does not produce significant differences. It can be seen from the pixel value of each sigma, gamma, and beta backscattering information, that there is a difference that is not too far. Principal component analysis extraction produces PCA1, PCA2, PCA3, PCA4, and PCA response. Where the results of PCA2, PCA3, and PCA4 do not detect any differences between water and land areas because they have the same colour between water and land. Whereas from the results of PCA1 and PCA response, land and water areas can be clearly seen because of the texture and clear colour differences. Texture extraction using the GLCM method by extracting homogeneity, energy, maximum probability, entropy and mean carried out on January 14, 2024 which represents the rainy season. The results of texture extraction of homogeneity, energy, maximum probability, entropy and mean between sigma, gamma, and beta in VV and VH polarizations do not have a big difference. This can be seen from the difference in the range of sigma, gamma, and beta pixel values that are not too large. Land cover results in 3 maps with different backscattering information. Land cover map using sigma nought produces overall accuracy and kappa coefficient of 86.15385% and 0.826923. Land cover map using gamma nought produces overall accuracy and kappa coefficient of 87.69231% and 0.846154. While with beta nought produces overall accuracy and kappa coefficient of 86.15385% and 0.826923. Gamma backscattering is particularly prosperous for land cover classification due to its ability to normalise pixel values across scenes and its applicability for volume scatterers such as vegetation. This characteristic leads to its improved classification accuracy over sigma and beta backscattering. In order to increase classification accuracy further, future studies might investigate applying deep learning methods or integrating multi-spectral data. Sentinel-1's multiple revisit times could also be used for real time monitoring of changes in land cover.

References

- Adugna, T., Xu, W., & Fan, J. (2022). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sensing*, 14(3), 574. doi: 10.3390/rs14030574
- Aggarwal, A. K. (2022). Learning Texture Features from GLCM for Classification of Brain Tumor MRI Images using Random Forest Classifier. *Wseas Transactions On Signal Processing*, 18, 60–63. doi: 10.37394/232014.2022.18.8
- Almaiah, M. A., Almomani, O., Alsaaidah, A., Al-Otaibi, S., Bani-Hani, N., Hwaitat, A. K. Al, Al-Zahrani, A., Lutfi, A., Awad, A. B., & Aldhyani, T. H. H. (2022). Performance Investigation of Principal Component Analysis for Intrusion Detection System Using Different Support Vector Machine Kernels. *Electronics*, 11(21), 3571. doi: 10.3390/electronics11213571
- Althubiti, S. A., Paul, S., Mohanty, R., Mohanty, S. N., Alenezi, F., & Polat, K. (2022). Ensemble Learning Framework with GLCM Texture Extraction for Early Detection of Lung Cancer on CT Images. *Computational and Mathematical Methods in Medicine*, 2022, 1–14. doi: 10.1155/2022/2733965
- Antara, I. M. O. G., Kusmiyarti, T. B., Suyarto, R., & Wiyanti, W. (2021). Klasifikasi Sawah Dan Non-Sawah Dengan Metode Random Forest Pada Citra Sar Sentinel-1 (Studi Kasus di Kecamatan Kediri, Kabupaten Tabanan, Bali). *Seminar Nasional Geomatika*, 245. doi: 10.24895/SNG.2020.0 0.1140
- Aouat, S., Ait-hammi, I., & Hamouchene, I. (2021). A New Approach for Texture Segmentation Based on the Gray Level Co-occurrence Matrix. *Multimedia Tools and Applications*, 80(16), 24027–24052. doi: 10.1007/s11042-021-10634-4

- Ardha, M. A. G. S. J. A. Y. F.; Y. D. R. Z. D. (2021). Utilization of Sentinel 1-A for Identification Land Use changes in Citarum Watershed. *Journal of Degraded and Mining Lands Management*.
- Bagaskara, D. P., Widada, S., & Rochaddi, S. (2017). Laju Sedimentasi Dan Pergeseran Delta Di Muara Anak Sungai Porong Sidoarjo. *Jurnal Oseanografi*, 6(4).
- Braun, A. (2020). SAR-based Landcover Classification with Sentinel-1 GRD Products. *SkyWatch Space Applications*.
- Chulafak, G. A., Kushardono, D., & Zylshal, N. (2018). Optimasi Parameter Dalam Klasifikasi Spasial Penutup Penggunaan Lahan Menggunakan Data Sentinel Sar (Parameters Optimiza-Tion In Spatial Land Use Land Cover Classification Using Sentinel Sar Data). *Jurnal Penginderaan Jauh Dan Pengolahan Data Citra Digital*, 14(2). doi : 10.30536/j.pjpdcd.1017.v14.a2746
- Dahhani, S., Raji, M., Hakdaoui, M., & Lhissou, R. (2022). Land Cover Mapping Using Sentinel-1 Time-Series Data and Machine-Learning Classifiers in Agricultural Sub Saharan Landscape. *Remote Sensing*, 15(1), 65. doi: 10.3390/rs15010065
- Dharani, M., & Sreenivasulu, G. (2021). Land Use and Land Cover Change Detection by Using Principal Component Analysis and Morphological Operations in Remote Sensing Applications. *International Journal of Computers and Applications*, 43(5), 462–471. doi: 10.1080/1206212X.2019.1578068
- Fadlin, F., Arsyad, M. T., Maricar, F., & Hatta, M. P. (2021). Monitoring Perubahan Penggunaan Lahan Menggunakan Citra Satelit Sentinel 1 Di Das Wanggu Kota Kendari. *In Jurnal Teknik Sumber Daya Air*, 12(2).
- Haydar, M., Sadia, H., & Hossain, M. T. (2024). Data driven forest fire susceptibility mapping in Bangla-desh. *Ecological Indicators*, 166, 112264. doi: 10.1016/j.ecolind.2024.112264
- Huda, H. A. N., Hasyim, A. W., & Parlindungan, J. (2022). Identifikasi Perubahan Tutupan Lahan Di Ko-Ta Batu Menggunakan Metode Penginderaan Jauh. *Planning for Urban Region and Environment Journal (PURE)*, 11(1), 153-160.
- Islami, F. A. T. S. D. W. E. D. D. B. D. (2022). Accuracy Assessment of Land Use Change Analysis Using Google Earth in Sadar Watershed Mojokerto Regency. *IOP Conf. Series: Earth and Environmental Science*, 012091.
- Liang, S. (2008). Advances in Land Remote Sensing. *Springer Netherlands*. doi: 10.1007/978-1-4020-6450-0
- Makinde, E., & Oyelade, O. (2018). Land Cover Mapping Using Sentinel-1 SAR Satellite Imagery of Lagos State for 2017. *The Economy, Sustainable Development, and Energy International Conference*, 1399. doi: 10.3390/proceedings2221399
- Najafi Khanbebin, S., & Mehrdad, V. (2021). Local Improvement Approach and Linear Discriminant Analysis-based Local Binary Pattern for Face Recognition. *Neural Computing and Applications*, 33(13), 7691–7707. doi: 10.1007/s00521-020-05512-3
- Prasetyo, S. Y. J. (2023). *Metode Penelitian Penginderaan Jauh*. Retrieved From www.penerbituwais.com
- Pryambodo, D. G., Prihantono, J., & Akhwady, R. (2021). Model Karakteristik Lapisan Bawah Per-Mukaan Tanah Pulau Lusi Sidoarjo Jawa Timur Menggunakan Metode Geolistrik Untuk Mendukung Wisata Bahari. *Jurnal Kelautan Nasional*, 16(1), 45-54.
- Pujianto, R., Adiwijaya, & Rahmawati, A. A. (2019). Analisis Ekstraksi Fitur Principle Component Analysis pada Klasifikasi Microarray Data Menggunakan Classification And Regression Trees. *E-Proceeding of Engineering*, 6(1).
- Regasa, M., Nones, M., & Adeba, D. (2021). A Review on Land Use and Land Cover Change in Ethiopian Basins. *Land*, 10(6), 585. doi: 10.3390/land10060585
- Sadia, H., Sarkar, S. K., & Haydar, M. (2023). Soil erosion susceptibility mapping in Bangladesh. *Ecological Indicators*, 156, 111182. doi: 10.1016/j.ecolind.2023.111182
- Salih Hasan, B. M., & Abdulazeez, A. M. (2021). A Review of Principal Component Analysis Algorithm for Dimensionality Reduction. *Journal of Soft Computing and Data Mining*, 02(01). doi: 10.30880/jscdm.2021.02.01.003
- Šćepanović, S., Antropov, O., Laurila, P., Rauste, Y., Ignatenko, V., & Praks, J. (2021). Wide-Area Land Cover Mapping With Sentinel-1 Imagery Using Deep Learning Semantic Segmentation Models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 10357–10374. doi: 10.1109/JSTARS.2021.3116094
- Septiarni, A., & Wardoyo, R. (2015). Kompleksitas Algoritma GLCM untuk Ekstraksi Ciri Tekstur pada Penyakit Glaucoma. *Prosiding Seminar Teknik Informatika Dan Sistem Informasi*, 98-102.
- Tariq, A., Yan, J., Gagnon, A. S., Riaz Khan, M., & Mumtaz, F. (2023). Mapping of Cropland, Cropping Patterns and Crop Types by Combining Optical Remote Sensing Images with Decision Tree Classifier and Random Rorest. *Geo-Spatial Information Science*, 26(3), 302–320. doi: 10.1080/10095020.2022.2100287
- Uddin, Md. P., Mamun, Md. Al, & Hossain, Md. A. (2021). PCA-based Feature Reduction for Hyperspectral Remote Sensing Image Classification. *IETE Technical Review*, 38(4), 377–396. doi: 10.1080/02564602.2020.1740615
- Wu, R., Wang, J., Zhang, D., & Wang, S. (2021). Identifying different types of urban land use dynamics using Point-of-interest (POI) and Random Forest algorithm: The case of Huizhou, China. *Cities*, 114, 103202. doi: 10.1016/j.cities.2021.103202