International Journal of Applied Science and Technology Application
Vol. 1, No. 1, March 2026, pp. 19-29
ISSN: xxxx-xxxx (online), DOI: 10.62712/ijapset.v1i1.3

Urban Vegetation Cover Prediction Using Sentinel-2 NDVI and Random Forest: A Brief Narrative Review

Muhammad Hasanuddin[1*], Abil Alwi Prayoga[2], Supiyandi[3]

[1]Master's Degree Student in Information Technology, Master of Information Technology, Universitas Pembangunan Panca Budi, Medan, Indonesia
[2,3]Faculty of Computational Science and Digital Intelligence, Information Technology, Universitas Pembangunan Panca Budi, Medan, Indonesia
Author's Email: [1*]muhammadhasanuddin.skom@gmail.com, [2]abilabil1752@gmail.com, [3]supiyandi.mkom@gmail.com

Article Info

Article history:

Received March 5, 2026
Revised March 5, 2026
Accepted March 8, 2026

Keywords:

NDVI, Sentinel-2, Google Earth Engine, Random Forest, urban vegetation cover, remote sensing.

ABSTRACT

A predictive model of urban vegetation cover is developed by integrating remote sensing technology, cloud computing, and machine learning algorithms. The study used the Normalized Difference Vegetation Index (NDVI), calculated from Sentinel-2 satellite imagery and analyzed in Google Earth Engine (GEE), to monitor vegetation conditions at a wide spatial scale. The research approach uses quantitative methods, including spatial analysis based on satellite imagery and predictive modeling with the Random Forest algorithm. The research process includes acquiring Sentinel-2 Level-2A images, pre-processing them with cloud masking and atmospheric correction, calculating NDVI values, and developing vegetation prediction models using machine learning methods. The results showed that the Random Forest model predicted vegetation cover with high accuracy, as indicated by a Coefficient of Determination (R^2) of 0.85 and a Root Mean Square Error (RMSE) of 0.045. The resulting vegetation distribution map shows significant variations in vegetation density between natural vegetation areas, agricultural land, and built-up areas. The findings of this study show that integrating NDVI from Sentinel-2, Google Earth Engine, and the Random Forest algorithm is an effective approach for monitoring and predicting urban vegetation cover. The results of this study make a methodological contribution to the development of remote sensing-based geospatial analysis and provide a scientific basis for sustainable urban planning and green open space management in urban areas.

Corresponding Author:

Muhammad Hasanuddin,
Master's Degree Student in Information Technology, Master of Information Technology,
Universitas Pembangunan Panca Budi,
Jalan Gatot Subroto Km 4.5, Medan, Indonesia
Email: muhammadhasanuddin.skom@gmail.com

1. INTRODUCTION

The development of urbanization in the 21st century has become one of the most significant socio-ecological transformation processes affecting spatial structures and environmental conditions across different regions of the world. Rapid urban growth is often accompanied by the expansion of built-up areas, infrastructure development, and the conversion of land that previously functioned as green open space or natural vegetation. This phenomenon causes significant changes in land-cover structure and directly impacts on the sustainability of urban ecosystems. Studies have shown that the loss of urban vegetation can increase urban surface temperatures, reduce carbon sequestration capacity, and worsen the overall quality of the urban environment [1], [2], [3]. Therefore, urban vegetation monitoring is an important component of sustainable urban planning and urban environmental management. Urban vegetation not only serves as an aesthetic element in the urban landscape but also plays important ecological roles, such as regulating microclimates, controlling air pollution, and protecting against soil degradation. In this context, the existence of urban green space is an important indicator of the quality of the urban environment. However, the dynamics of vegetation change are often difficult to monitor conventionally due to the large urban area and the rapid pace of change. Therefore, approaches based on remote sensing technology and geospatial analysis are becoming increasingly important alternatives for systematically and sustainably understanding vegetation change.

Remote sensing technology enables the observation of vegetation conditions at large scales and over long periods through analysis of satellite imagery. One of the most widely used indicators in vegetation monitoring is the Normalized Difference Vegetation Index (NDVI), which is calculated from the difference in spectral reflectance between the red and near-infrared bands in satellite imagery. NDVI has been shown to be able to effectively describe vegetation density and plant health in a variety of ecosystem conditions, including urban ecosystems [3], [4], [5]. A high NDVI value usually indicates healthy, dense vegetation, while a low value indicates sparse vegetation or non-vegetated areas such as buildings and roads. The advantage of NDVI lies in its ability to detect vegetation changes quickly and consistently over time. In various urban studies, NDVI is used to map the distribution of urban vegetation, analyze changes in land cover, and evaluate vegetation health conditions on a wide spatial scale [1], [2]. In addition, NDVI can also be combined with other vegetation indices, such as NDMI, to provide a more comprehensive interpretation of vegetation conditions and plant moisture in urban environments [5]. Therefore, NDVI is the primary indicator in research on vegetation dynamics and changes in urban ecosystems.

The development of satellite technology has significantly improved the quality of data used in vegetation analysis. One of the satellite sensors widely used in vegetation research is Sentinel-2, developed by the European Space Agency as part of the Copernicus program. Sentinel-2 provides multispectral imagery with a spatial resolution of up to 10 meters and has a relatively high observation frequency thanks to its configuration of two operational satellites. These advantages allow for the analysis of vegetation changes on a more detailed scale compared to previous-generation satellite sensors [2]. With higher spatial resolution, Sentinel-2 can identify vegetation variations at the scale of urban landscapes, including urban green spaces, street vegetation, and small patches of vegetation scattered between buildings. Several studies have shown that Sentinel-2 data is very effective for monitoring urban vegetation dynamics as well as detecting changes in plant phenology in urban contexts [3], [6]. In addition, the combination of Sentinel-2 data with other satellite data sources, such as Landsat or PlanetScope, allows for a more comprehensive temporal analysis of changes in urban vegetation [1]. With these characteristics, Sentinel-2 is one of the primary data sources for remote sensing research focused on monitoring vegetation and land-cover changes in urban areas.

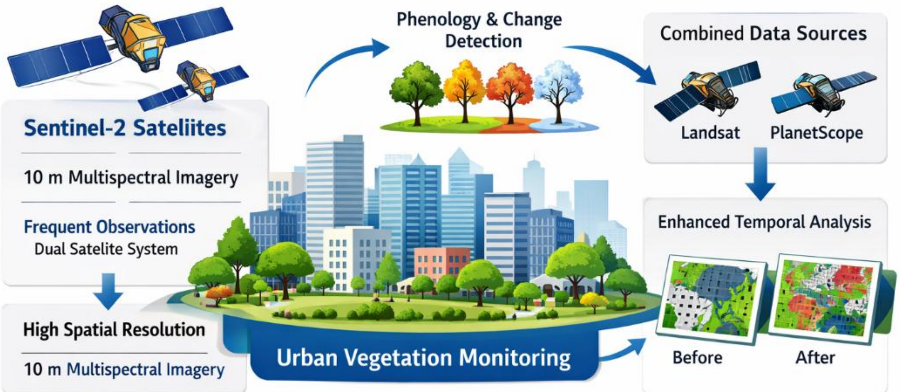

Figure 1. Key Data Source for Urban Vegetation Research

Although satellite data are increasingly abundant, a major challenge in remote sensing research lies in efficiently processing and analyzing data at large scale. Satellite image analysis often involves large volumes of data, especially when research is conducted over long periods or across large areas. In this context, the availability of cloud computing platforms such as Google Earth Engine (GEE) is an important innovation in geospatial analysis. GEE allows researchers to access global satellite data catalogs and perform image processing directly through cloud computing infrastructure without the need for high-capacity hardware. The platform provides a wide range of analytical tools for efficient image processing, vegetation index calculation, land cover classification, and environmental change analysis [7]. Several studies show that GEE can accelerate large-scale geospatial analysis and increase research reproducibility by enabling the entire analysis to be carried out in an integrated manner within a single computing environment [8], [9]. In the context of urban vegetation monitoring, GEE enables systematic analysis of NDVI time-series as well as facilitates the integration of various machine learning algorithms for remote sensing data-driven environmental modeling.

In addition to advances in computing platforms, machine learning significantly improves the accuracy of satellite image analysis. Machine learning algorithms allow the identification of complex patterns in geospatial data that are difficult to detect using conventional statistical methods. One widely used algorithm for land cover classification is Random Forest (RF). This method is an ensemble learning algorithm that builds several decision trees and combines their predictions to produce more stable and accurate predictions. Random Forest has the advantage of handling multiband data and can handle non-linear relationships between variables, making it well-suited for high-dimensional satellite image analysis. Studies have shown that Random Forest can achieve high accuracy in Sentinel-2 image-based land cover classification, even reaching an AUC of around 0.89 in a LULC classification study in the Bhutan region [10]. In addition, this algorithm is also used in various studies to predict vegetation index, plant biomass estimation, as well as NDVI time-series reconstruction when optical imagery is disturbed by clouds [11], [12], [13]. With this capability, Random Forest is among the most effective algorithms for machine-learning-based remote sensing analysis.

Although various studies have demonstrated the effectiveness of using NDVI, Sentinel-2, Google Earth Engine, and Random Forest for vegetation monitoring, most studies still focus on descriptive analyses of land-cover change or static classification. Research that comprehensively integrates these four components to build a predictive model of urban vegetation cover remains relatively limited, especially in urban areas of developing countries. In addition, differences in urban environmental characteristics, such as building density, vegetation variation, and local climatic conditions, often make vegetation analysis results context-specific and cannot be generalized directly from one region to another [4], [5], [14]. Some studies have also highlighted that NDVI threshold values for distinguishing vegetation from non-vegetation can vary across regions, as they are influenced by vegetation types and local environmental conditions [14]. In addition, differences in atmospheric correction methods during satellite image pre-processing can affect NDVI calculations, so a consistent methodological approach is needed to ensure the reliability of the analysis results [15]. These limitations suggest that further research is needed to develop a more robust, context-aware vegetation analysis model for urban environments.

In this context, this study offers novelty by integrating remote sensing, cloud computing, and machine learning to more accurately predict urban vegetation cover. In contrast to previous studies that generally use NDVI as a static indicator, this study combines Sentinel-2-based NDVI analysis with the Google Earth Engine platform and the Random Forest algorithm to develop a vegetation prediction model using geospatial data. This integration enables vegetation analysis to be carried out more efficiently at a wide spatial scale while producing predictive models that describe vegetation dynamics more comprehensively. In addition, this study seeks to integrate vegetation analysis with an urban environmental systems approach that emphasizes the relationship between vegetation change and land-use dynamics in urban areas. With this approach, this research is expected to make a methodological contribution to the development of a remote sensing-based urban vegetation monitoring model and make an empirical contribution in understanding vegetation dynamics in the context of sustainable urban development. Based on the background description, literature review, and identification of research gaps that have been presented, the main problem formulation in this study is: how can the integration of Google Earth Engine, NDVI Sentinel-2, and the Random Forest algorithm be used to accurately and effectively predict urban vegetation cover in support of monitoring urban ecosystem dynamics? The formulation of this problem is the basis for developing a research methodology that integrates remote sensing analysis with machine learning to produce vegetation prediction models that support environmental planning and green open space management in urban areas. Thus, this research is expected not only to make an academic contribution to remote sensing and geospatial analysis but also to provide practical implications for sustainable urban planning and future urban environment management.

2. RESEARCH METHOD

This study adopts a positivist paradigm that emphasizes objective measurement, hypothesis testing, and causal analysis through a quantitative approach grounded in empirical data. This paradigm assumes that changes in urban vegetation cover are a reality that can be objectively measured using spatial and spectral indicators derived from satellite imagery. With this approach, vegetation change is understood as an ecological phenomenon that can be analyzed through mathematical and statistical models to generate scientific generalizations regarding vegetation dynamics in urban areas. This study integrates remote sensing technology, cloud computing, and machine learning to develop an NDVI-based model for predicting urban vegetation cover using Sentinel-2 imagery processed in Google Earth Engine and analyzed with the Random Forest algorithm. This approach allows research to be conducted systematically and measurably and can be replicated by other researchers in different regional contexts with similar methodological procedures.

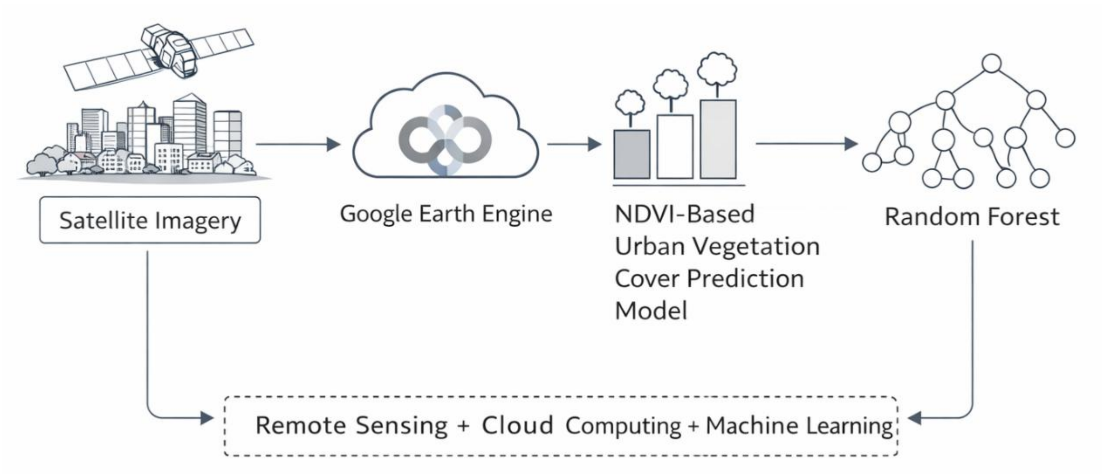

Figure 2. NVDI-Based Urban Vegetation Prediction Approach

The type of research used is explanatory quantitative research with a spatial-predictive approach based on remote sensing. Explanatory research was conducted to examine the relationship between vegetation cover change and urban land-use dynamics using vegetation indicators derived from satellite imagery. A spatial approach is used to map vegetation distribution and identify patterns of vegetation change within a given area. Meanwhile, a predictive approach is used to build a model that projects vegetation cover conditions based on patterns identified in remote sensing data. The analysis was carried out using Sentinel-2 Level-2A imagery, which has a spatial resolution of 10 meters and provides multispectral data for calculating the NDVI vegetation index. The image data are accessed through the Google Earth Engine (GEE) platform, which enables efficient processing of geospatial data at large scales. The use of GEE provides advantages in terms of global access to satellite data and cloud-based image processing capabilities that can accelerate remote sensing data analysis [7], [8], [9].

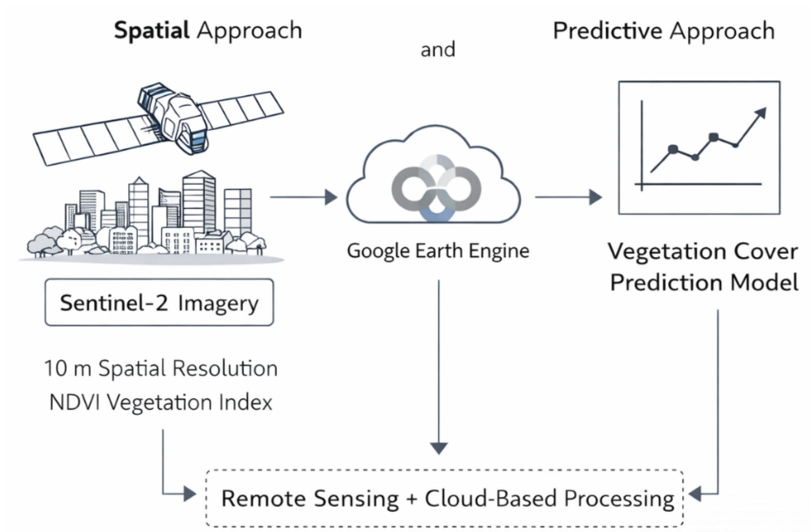

Figure 3. Explanatory Quantitative Research with A Spatial-Predictive Approach Based on Remote Sensing

The research location is in urban areas experiencing urbanization pressure and significant land-use changes that affect vegetation conditions. The research area was analyzed using official administrative boundaries, which served as study areas for satellite image processing. The research object in this study is the urban vegetation cover, identified using NDVI values derived from Sentinel-2 imagery. The main variables in this study are NDVI, a dependent variable representing vegetation conditions, and several independent variables derived from the spectral characteristics of Sentinel-2 imagery, such as red-band reflectance and near-infrared (NIR). NDVI is calculated using the standard remote sensing formula, namely:

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + \mathrm{Red})} \quad (1)$$

NDVI values range from −1 to 1, where higher values indicate healthier, denser vegetation. The vegetation classification is then carried out by dividing the NDVI into several vegetation-density categories to facilitate the spatial interpretation of urban vegetation cover conditions.

The data used in this study included satellite imagery and supporting spatial data. The main data are Sentinel-2 Level-2A imagery downloaded via Google Earth Engine and cover the observation period defined in the study. The Sentinel-2 image was chosen because it has a sufficiently high spatial resolution for urban vegetation analysis and has a relatively high frequency of observation that allows the analysis of vegetation changes over a certain time span [2], [6]. In addition, this study also uses additional spatial data in the form of administrative boundaries of the research area, which are used as clipping boundaries in the image analysis process. All data is collected digitally through the Google Earth Engine platform, which provides direct access to a wide range of global remote sensing datasets. The data collection technique comprises the stages of satellite image acquisition, image selection based on the research period, and downloading datasets relevant to the research objectives.

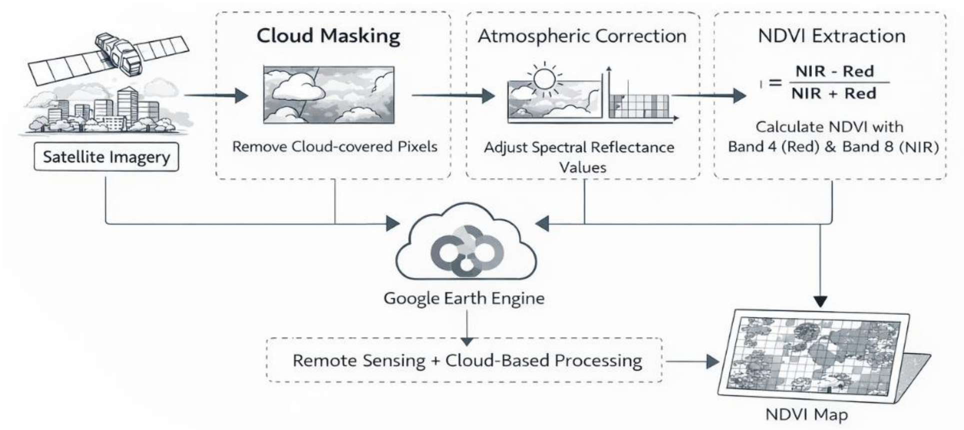

Figure 4. Satellite Imagery Processing Procedure

The data processing procedure begins with pre-processing satellite imagery to ensure the quality of the data used in the analysis. The first stage is the cloud masking process, which removes cloud-covered pixels and their shadows, which can interfere with vegetation index calculations. This process is carried out using the cloud mask function available in the Sentinel-2 dataset on Google Earth Engine. After the cloud masking process is complete, the image is then subjected to atmospheric correction to ensure the spectral reflectance values used in the NDVI calculation are corrected for atmospheric effects. The next stage is to extract NDVI values from the red (Band 4) and near-infrared (Band 8) spectral bands. This process is carried out automatically using programming scripts in Google Earth Engine to produce NDVI maps that represent vegetation conditions in the research area. The NDVI results were then analyzed to assess vegetation distribution and changes in vegetation density across the study area. The next stage of data analysis uses a machine-learning approach with a Random Forest (RF) algorithm to build a vegetation-cover prediction model. Random Forest was chosen because it is effective with multivariable data and can identify non-linear relationships among variables in geospatial datasets. The model training process was carried out using sample data obtained through visual interpretation of satellite imagery and reference to available field data. The sample data were then divided into two groups: training and validation.

The Random Forest model was trained on the training data to study the relationships between spectral variables from satellite imagery and vegetation cover conditions. After the training process is complete, the model is used to predict NDVI values and classify vegetation cover within the study area. This process produces a vegetation prediction map that can be used to analyze vegetation distribution spatially. To ensure the reliability of the resulting model, this study evaluated its accuracy using several statistical metrics commonly used in remote sensing and machine learning analyses. The model evaluation was conducted by comparing the NDVI predictions generated by the Random Forest model with the available validation data. Some of the evaluation metrics used include Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Confusion Matrix for classification analysis. The R^2 value is used to measure the strength of the relationship between the predicted value and the observation value of the NDVI, while the RMSE is used to measure the rate of model prediction error. A confusion matrix is used to evaluate vegetation cover classification accuracy by comparing the model's classification results with reference data. Through this evaluation process, the resulting prediction model can be assessed for its accuracy and ability to represent the actual vegetation conditions. By following a systematic methodological procedure ranging from data acquisition, image pre-processing, NDVI calculation, Random Forest modeling, to model accuracy evaluation, this study produces a methodological framework that can be replicated by other researchers in different regional contexts. The integration of remote sensing technology, Google Earth Engine's cloud computing platform, and machine learning algorithms provides an efficient and accurate analysis approach in monitoring and predicting changes in urban vegetation cover. This approach not only makes a methodological contribution to geospatial research but also provides a scientific basis for the development of satellite data-based urban environmental monitoring systems that can support decision-making in sustainable urban planning.

3. RESULTS AND DISCUSSION

3.1. General Description of Research Locations and Data Informants

This research was carried out in an urban area in Batu Bara Regency, North Sumatra Province, Indonesia, which is one of the coastal areas with a high rate of urbanization in recent years. This area has geographical characteristics as a coastal area directly adjacent to the Strait of Malacca, with a combination of land cover comprising residential areas, industrial areas, agricultural land, and natural vegetation, such as mangroves and other vegetation. Based on the administrative map of the research area shown in the source article, the Batu Bara Regency comprises several sub-districts, with the administrative center in Lima Puluh District. Spatial analysis of this region shows that development pressures are concentrated in coastal areas and adjacent to economic centers and major transportation networks. The number of data informants in this study was not human respondents but spatial data units, Sentinel-2 satellite image pixels, analyzed through the Google Earth Engine platform. The main data analyzed came from the 2024 Sentinel-2A Level-2A image with a spatial resolution of 10 meters. Each image pixel represents the spectral conditions of the Earth's surface that are used to calculate the Normalized Difference Vegetation Index (NDVI) value. During the analysis, the image pixels used as training and validation samples were selected through digital interpretation of satellite imagery using spatial sampling procedures. Each pixel sample is then labeled as a vegetation class based on its spectral characteristics, and the classification model is trained using the Random Forest algorithm. Therefore, the unit of analysis in this study is vegetation-pixel data, which represents the condition of land cover in the research area.

Based on satellite image processing, the study area exhibited significant variations in vegetation cover across coastal, agricultural, and residential areas. Dense vegetation areas are generally located in mangrove forests or natural vegetation, while areas with moderate vegetation are mostly found on agricultural land and in urban green areas. Meanwhile, areas with low vegetation are dominated by built-up areas such as settlements, industrial estates, and transportation infrastructure. This variation is clearly visible in the NDVI map produced by processing Sentinel-2 imagery in Google Earth Engine.

3.2. Results of Satellite Image Processing and Vegetation Cover Identification

The initial stage of analysis was carried out using Sentinel-2 imagery acquired over the research area. The results of the initial image acquisition showed that several areas were covered by clouds, thereby affecting the quality of the imagery used in vegetation analysis. Therefore, a cloud-masking process is applied to remove cloud-covered pixels and their shadows, which can cause errors in NDVI calculations.

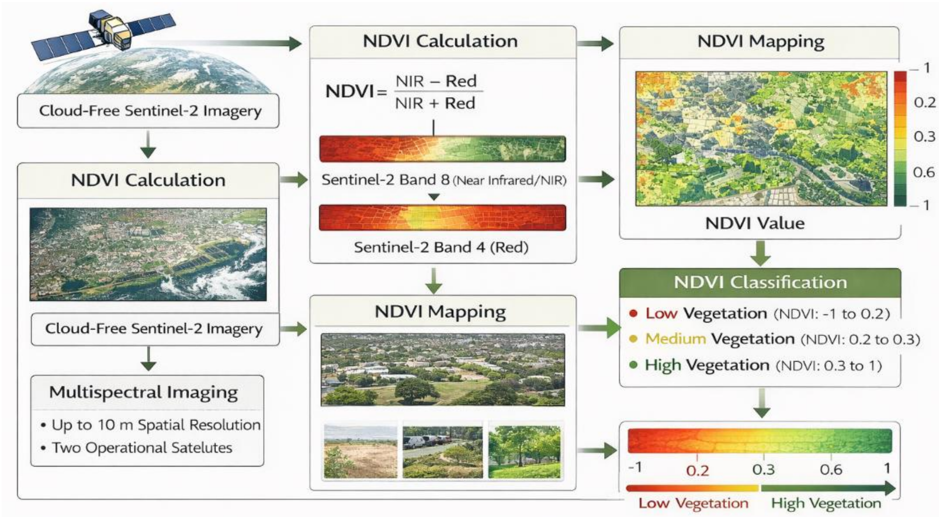

Figure 5. NDVI Calculation and Classification Workflow

The results of this process produce cloud-free images that more clearly display the conditions of Earth's surface. After the cloud masking process is completed, the next step is to calculate NDVI values using the Sentinel-2 image's Band 8 (Near Infrared/NIR) and Band 4 (Red) spectral bands. The NDVI value is calculated for each image pixel using the standard NDVI formula, resulting in a distribution of vegetation values across the study area. The NDVI results are then visualized as a map showing variations in vegetation density, with different colors indicating different vegetation densities. Areas with high NDVI values are shown in green, indicating healthy, dense vegetation, while areas with low NDVI values are shown in red or yellow, indicating sparse vegetation or non-vegetation. Based on the NDVI analysis, vegetation index values in the study area ranged from −1 to 1, with variations indicating different vegetation densities. To facilitate interpretation of the results, NDVI values are classified into three main categories: low vegetation, medium vegetation, and high vegetation. This classification is based on the NDVI interpretation standard commonly used in remote sensing analysis.

Table 1. Classification of Vegetation Density Based on NDVI

Classes	Rentang Nilai NDVI	NDVI Value Range
1	−1 until 0.32	Low Vegetation
2	0.32 until 0.42	Temperate Vegetation
3	0.42 until 1	High Vegetation

The classification results showed that most urban areas in the study area were classified as low- to medium-vegetation, especially those that had experienced land conversion to built-up areas. Areas with high vegetation tend to be found in places with natural vegetation cover, such as mangrove areas or relatively large green areas.

3.3. Results of Vegetation Prediction Modeling Using Random Forest

Once the NDVI map is obtained, the next step is to develop a vegetation cover prediction model using the Random Forest algorithm. This model is trained in sample data derived from the interpretation of satellite imagery and available reference datasets. During model training, several spectral variables from Sentinel-2 imagery were used as inputs to predict NDVI values and vegetation classes. The model training process produces a predictive model that identifies the relationships between the spectral variables of satellite imagery and vegetation cover conditions in the research area. Once the model is trained, validation is performed on the validation dataset to evaluate its accuracy. The evaluation results showed that the Random Forest model produced NDVI predictions with high accuracy.

Table 2. Prediction Model Evaluation Results

Evaluation Parameters	Value
Coefficient of Determination (R^2)	0.85
Root Mean Square Error (RMSE)	0.045

An R^2 value of 0.85 indicates a strong relationship between the NDVI value predicted by the model and the actual NDVI value obtained from satellite imagery data. Meanwhile, an RMSE value of 0.045 indicates that the model's prediction error rate is relatively low. These results show that the Random Forest model used in this study has a good ability to predict vegetation cover from Sentinel-2 image data.

3.4. Spatial Distribution of Urban Vegetation

The resulting vegetation prediction map from the Random Forest model shows a pattern of vegetation distribution that varies across the study area. Areas with high vegetation are concentrated in places with natural vegetation cover or large green open spaces. On the other hand, areas with low vegetation are generally in urban areas that have undergone intensive development. Some coastal areas show relatively dense mangrove vegetation, as indicated by the high NDVI values on the vegetation map. Meanwhile, areas with moderate NDVI values are generally associated with agricultural land or mixed vegetation with moderate vegetation density. This distribution pattern indicates that vegetation conditions in the study area are influenced by various factors, including land use, development activities, and the geographical characteristics of coastal areas.

3.5. Documentation of Spatial Data Observation Results

During the data analysis process, visual documentation from satellite imagery showed different conditions between images still covered by clouds and those processed into cloud-free images. The results of this observation underscore the importance of image pre-processing to ensure the quality of the data used in vegetation analysis. In addition, the NDVI map derived from Sentinel-2 image processing provides a clear picture of vegetation distribution in the study area. The color variations on the NDVI map indicate differences in vegetation density levels that can serve as a basis for further spatial analysis of urban ecosystem conditions. Overall, the results of this study show that the Google Earth Engine-based approach, NDVI from Sentinel-2, and the Random Forest algorithm can produce accurate vegetation cover maps and provide important spatial information on urban vegetation conditions. The data generated from this analysis provides an empirical picture of vegetation distribution and density in the study area. These results form the basis for further analysis of urban vegetation dynamics and their implications for environmental management and sustainable urban planning.

3.6. Discussion

The results show that integrating NDVI from Sentinel-2, Google Earth Engine (GEE), and the Random Forest algorithm produces a predictive model of urban vegetation cover with high accuracy. The $R^2 = 0.85$ and RMSE = 0.045 indicate that the developed model effectively represents the relationship between the observed and predicted NDVI values. These findings suggest that a remote sensing approach combined with machine learning methods can effectively monitor the dynamics of urban vegetation at broad spatial scales. These results are consistent with previous studies confirming that NDVI is a highly sensitive indicator of vegetation density and plant health across various ecosystems, including urban ecosystems [3], [4], [5]. The use of NDVI as a vegetation indicator enables rapid identification of changes in vegetation cover and long-term monitoring through time-series analysis of satellite images. Therefore, the findings of this study further strengthen the role of NDVI as a key indicator in remote sensing studies focusing on vegetation dynamics in urban areas.

The spatial distribution of vegetation produced in this study shows that areas with high vegetation cover are generally concentrated in areas with natural vegetation cover or large green open spaces. On the other hand, areas with low vegetation are dominated by built-up areas such as settlements, industrial estates, and transportation networks. This pattern is in line with the findings of various studies showing that urban expansion often leads to a decrease in vegetation cover due to the conversion of land to built-up areas [1], [2]. In this context, NDVI from the Sentinel-2 image was shown to detect vegetation characteristics in detail in urban landscapes. The advantages of Sentinel-2, with a spatial resolution of 10 meters, allow the identification of vegetation variations at finer scales than previous-generation satellite sensors. This enables a more accurate analysis of vegetation distribution in urban environments, which generally have complex land-cover mosaics. Thus, the use of Sentinel-2 data in this study contributes to improving the accuracy of urban vegetation mapping and supporting more detailed analysis of land cover changes. The findings of this study also show that the Google Earth Engine platform plays a very important role in facilitating large-scale analysis of remote sensing data. GEE enables efficient processing of satellite imagery on cloud computing infrastructure, allowing image analysis to be performed faster than with conventional approaches. This is in line with the findings of previous research, which stated that GEE provides ease of access to various global remote sensing datasets and provides various analysis tools that support environmental modeling based on satellite imagery [7]. In addition, GEE has been widely used in urban vegetation research to monitor land-cover changes and analyze the dynamics of urban ecosystems [8], [9]. In this study, the use of GEE enables systematic processing of Sentinel-2 imagery from pre-processing through NDVI calculation and vegetation prediction modeling. Thus, the platform not only improves analysis efficiency but also enhances reproducibility, as the entire analysis process can be replicated using the same programming script.

From a methodological perspective, the Random Forest algorithm used in this study showed excellent performance in predicting vegetation cover from satellite imagery. Random Forest is a machine learning algorithm that can handle multidimensional data and non-linear relationships between variables. This capability makes Random Forest very effective at analyzing high-spectral-complexity satellite imagery. The results of this study are consistent with previous studies showing that Random Forest is among the most accurate algorithms for classifying land cover from Sentinel-2 imagery [10]. In the study, Random Forest achieved high accuracy in classifying LULC from satellite imagery. In addition, Random Forest is used in various studies to predict vegetation indices and estimate plant biomass from remote sensing data [11], [12]. Thus, the use of Random Forest in this study further strengthens the empirical evidence of the algorithm's effectiveness for analyzing vegetation from satellite imagery.

In addition to providing methodological contributions, the results of this study also have significant implications for urban environmental management. Information on the distribution of urban vegetation derived from NDVI analysis can serve as a basis for green open space planning and sustainable urban environmental management. Urban vegetation plays an important role in improving the quality of the urban environment through various ecosystem functions, such as temperature regulation, air pollutant filtration, and aesthetic enhancement. Therefore, information on vegetation distribution obtained through satellite image analysis can serve as the basis for local governments in designing urban green space management policies. In addition, regular monitoring of vegetation using satellite imagery enables early identification of vegetation degradation that can affect the quality of the urban environment. This research also contributes to the development of a remote sensing-based methodology for urban vegetation monitoring. The integration of NDVI Sentinel-2, Google Earth Engine, and Random Forest demonstrates that a geospatial technology-based approach can produce accurate and efficient vegetation prediction models. This approach enables vegetation analysis at a large scale without the need for intensive field surveys. In addition, the use of cloud computing platforms enables large-scale data processing, allowing vegetation analysis to be carried out more systematically and sustainably. Thus, this research makes an important methodological contribution to the development of an urban vegetation monitoring system based on remote sensing technology.

However, this study also has some limitations to consider. One of the main limitations is the reliance on optical satellite imagery that is susceptible to cloud interference. Cloud conditions can cause the loss of image data, affecting the continuity of vegetation analysis. Several previous studies have shown that a combination of Sentinel-2 optical data and Sentinel-1 radar data can be used to address this problem through a data fusion approach or machine learning-based NDVI reconstruction [11], [16], [17]. In addition, the use of NDVI as a vegetation indicator is limited to very dense vegetation due to potential vegetation index saturation. Therefore, further research may consider using additional vegetation indices, such as NDMI or GRVI, to improve the accuracy of vegetation condition interpretation [4], [5]. In addition, this study still uses satellite imagery as the main data source, so that more intensive field validation can provide additional information on actual vegetation conditions in the field. Overall, the results of this study show that integrating remote sensing technology, cloud computing, and machine learning provides an effective approach to monitoring urban vegetation. This approach not only makes a methodological contribution to geospatial research but also provides practical implications in the management of the urban environment. With increasing pressures of urbanization in various regions of the world, monitoring urban vegetation is becoming increasingly important to ensure the sustainability of the urban environment. Therefore, the approach developed in this study can serve as the basis for a satellite-based urban vegetation monitoring system for sustainable urban planning and future green open space management.

4. CONCLUSION (10 PT)

This study aims to analyze and predict urban vegetation cover through the integration of remote sensing technology, cloud computing, and machine learning algorithms. Specifically, the study implemented an approach that combines NDVI from Sentinel-2 imagery, the Google Earth Engine platform, as well as the Random Forest algorithm to build a predictive model of urban vegetation cover. The results of the study show that the integrative approach is able to produce accurate vegetation mapping and provide a clear picture of the spatial distribution of vegetation in urban areas that experience urbanization pressure. The evaluation values of the model obtained, namely the determination coefficient (R^2) of 0.85 and the Root Mean Square Error (RMSE) value of 0.045, showed that the developed model had a strong ability to predict the NDVI value and vegetation conditions in the study area. The findings of this study also show that the distribution of urban vegetation has a spatial pattern that is influenced by land use dynamics. Areas with high vegetation are generally located in areas that still have natural vegetation cover or relatively large green open spaces, while areas with low vegetation are dominated by built-up areas such as settlements, industrial estates, and transportation infrastructure. Thus, Sentinel-2-based NDVI analysis is able to provide important information about vegetation health conditions and vegetation density in urban environments. In addition, the use of the Google Earth Engine platform has been proven to provide significant efficiency in the process of processing satellite imagery on a large scale, allowing vegetation analysis to be carried out more quickly and systematically. The main contribution of this research lies in the development of a methodological approach that integrates remote sensing, cloud computing, and machine learning in urban vegetation monitoring. This integration not only improves the accuracy of vegetation analysis but also allows for the development of prediction models that can be used to understand vegetation dynamics over a longer period of time. Thus, this research makes a scientific contribution in the field of remote sensing and geospatial analysis, especially in the development of satellite data-based vegetation cover monitoring methods. In addition, the results of this study also have practical implications for sustainable urban planning because information on vegetation distribution can be used as a basis for the management of green open spaces and the control of urban environmental degradation. Although this study shows promising results, some limitations still need to be noted. Reliance on optical satellite imagery such as Sentinel-2 causes vegetation analysis to be susceptible to cloud interference that can reduce data quality. In addition, the use of NDVI as the main indicator of vegetation also has limitations in very dense vegetation conditions due to the potential saturation of vegetation index values. Therefore, further research is recommended to integrate data from radar sensors such as Sentinel-1 or use additional vegetation indexes to improve the accuracy of vegetation analysis. Future research can also expand the scope of analysis by utilizing longer time-series data and integrating field data to improve the validation of remote sensing analysis results. With the development of a more comprehensive methodology, geospatial technology-based urban vegetation monitoring is expected to make a greater contribution in supporting environmental management and sustainable urban development.

5. ACKNOWLEDGEMENTS

Thank you to everyone who helped publish this article..

6. REFERENCES (10 PT)

[1] H. Wang and Z.-C. Zeng, "SATELLITE OBSERVATIONS OF URBAN GREENERY PHENOLOGY IN DOWNTOWN BEIJING AT METER TO KILOMETER SCALES," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Jul. 2022, pp. 6268–6271. doi: 10.1109/IGARSS46834.2022.9883366.

[2] J. Dobrzańska, A. Nadolny, R. Kalbarczyk, and M. Ziemiańska, "Urban Resilience and Residential Greenery—The Evidence from Poland," *Sustainability*, vol. 14, no. 18, p. 11317, Sep. 2022, doi: 10.3390/su141811317.

[3] N. Xia *et al.*, "Monitoring and evaluation of vegetation restoration in the Ebinur Lake Wetland National Nature Reserve under lockdown protection," *Front. Plant Sci.*, vol. 15, Apr. 2024, doi: 10.3389/fpls.2024.1332788.

[4] A. Krtalić, D. Linardić, and R. Pernar, "Framework for Spatial and Temporal Monitoring of Urban Forest and Vegetation Conditions: Case Study Zagreb, Croatia," *Sustainability*, vol. 13, no. 11, p. 6055, May 2021, doi: 10.3390/su13116055.

[5] O. Strashok, M. Ziemiańska, and V. Strashok, "Evaluation and Correlation of Normalized Vegetation Index and Moisture Index in Kyiv (2017–2021)," *J. Ecol. Eng.*, vol. 23, no. 9, pp. 212–218, Sep. 2022, doi: 10.12911/22998993/151884.

[6] H. Wang and F.-Y. Gong, "Quantifying City- and Street-Scale Urban Tree Phenology from Landsat-8, Sentinel-2, and PlanetScope Images: A Case Study in Downtown Beijing," *Remote Sens.*, vol. 16, no. 13, p. 2351, Jun. 2024, doi: 10.3390/rs16132351.

[7] M. I. Habibie, "THE APPLICATION OF MACHINE LEARNING USING GOOGLE EARTH ENGINE FOR REMOTE SENSING ANALYSIS," *J. Teknoinfo*, vol. 16, no. 2, p. 233, Jul. 2022, doi: 10.33365/jti.v16i2.1872.

[8] H. Rakuasa and D. A. Sihasale, "Analysis of Vegetation Index in Ambon City Using Sentinel-2 Satellite Image Data with Normalized Difference Vegetation Index (NDVI) Method based on Google Earth Engine," *J. Innov. Inf. Technol. Appl.*, vol. 5, no. 1, pp. 74–82, Jun. 2023, doi: 10.35970/jinita.v5i1.1869.

[9] A. Agustiyara, D. Mutiarin, A. Nurmandi, A. N. Kasiwi, and M. F. Ikhwali, "Mapping Urban Green Spaces in Indonesian Cities Using Remote Sensing Analysis," *Urban Sci.*, vol. 9, no. 2, p. 23, Jan. 2025, doi: 10.3390/urbansci9020023.

[10] K. Tempa, M. Ilunga, A. Agarwal, and Tashi, "Utilizing Sentinel-2 Satellite Imagery for LULC and NDVI Change Dynamics for Gelephu, Bhutan," *Appl. Sci.*, vol. 14, no. 4, p. 1578, Feb. 2024, doi: 10.3390/app14041578.

[11] A. L. de Castro, M. L. Duarte, H. Ewbank, and R. W. Lourenço, "Use of synthetic aperture radar data for the determination of normalized difference vegetation index and normalized difference water index," *J. Appl. Remote Sens.*, vol. 18, no. 01, Feb. 2024, doi: 10.1117/1.JRS.18.014516.

[12] E. Habyarimana and F. S. Baloch, "Machine learning models based on remote and proximal sensing as potential methods for in-season biomass yields prediction in commercial sorghum fields," *PLoS One*, vol. 16, no. 3, p. e0249136, Mar. 2021, doi: 10.1371/journal.pone.0249136.

[13] R. Agrawal, J. D. Mohite, S. A. Sawant, A. Pandit, and S. Pappula, "ESTIMATION OF NDVI FOR CLOUDY PIXELS USING MACHINE LEARNING," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLIII-B3-2, no. B3-2022, pp. 813–818, May 2022, doi: 10.5194/isprs-archives-XLIII-B3-2022-813-2022.

[14] A. D. W. Rahmadana, "Pemanfaatan Foto Udara Multispektral Untuk Sidik Cepat Kerapatan Tutupan Vegetasi Di Wilayah Perkotaan," *Geomedia Maj. Ilm. dan Inf. Kegeografian*, vol. 21, no. 1, pp. 1–9, Jun. 2023, doi: 10.21831/gm.v21i1.35907.

[15] D. Moravec, J. Komárek, S. López-Cuervo Medina, and I. Molina, "Effect of Atmospheric

Corrections on NDVI: Intercomparability of Landsat 8, Sentinel-2, and UAV Sensors," *Remote Sens.*, vol. 13, no. 18, p. 3550, Sep. 2021, doi: 10.3390/rs13183550.

[16] A. B. Ramathilagam, S. Natarajan, and A. Kumar, "SAR2NDVI: a pix2pix generative adversarial network for reconstructing field-level normalized difference vegetation index time series using Sentinel-1 synthetic aperture radar data," *J. Appl. Remote Sens.*, vol. 17, no. 02, May 2023, doi: 10.1117/1.JRS.17.024514.

[17] R. Burger *et al.*, "The Biomass Proxy: Unlocking Global Agricultural Monitoring through Fusion of Sentinel-1 and Sentinel-2," *Remote Sens.*, vol. 16, no. 5, p. 835, Feb. 2024, doi: 10.3390/rs16050835.