P5 – AI-Enabled Forest Cover Change Detection using Multi-temporal NDVI and Machine Learning

Project Report Submitted in fulfilment of the

Requirements for the Award of the Internship of

Winter Training Program on Space Science

Technology

(Subject Name: Advanced drone Technology/ Remote sensing and GIS/ Astronomy and Astrophysics)

By

Student Name: Amuktamalyada

Institute Name: Keshav Memorial Engineering College

Institute Roll no.: 245523733193

Enrolment no.: 559113

Under the Supervision of

Ms. Alisha Sinha

India Space Academy,

Department of Space Education, India Space Week

TABLE OF CONTENTS

AI-ENABLED FOREST COVER CHANGE DETECTION

1. INTRODUCTION

Forests are essential components of the Earth's ecosystem, playing a vital role in maintaining ecological balance, regulating climate, conserving biodiversity, and supporting human livelihoods. However, rapid urbanization, agricultural expansion, mining, and infrastructure development have led to noticeable changes in forest cover across many regions of India. Monitoring these changes is crucial for sustainable forest management and environmental planning.

Traditional forest monitoring methods rely on field surveys, which are time-consuming and limited in spatial coverage. Remote sensing technology provides an efficient alternative by enabling large-scale and long-term observation of vegetation dynamics. Among various vegetation indices, the Normalized Difference Vegetation Index (NDVI) is widely used to assess vegetation health and density. NDVI is calculated using the Near Infrared (NIR) and Red spectral bands, where healthy vegetation shows high reflectance in NIR and low reflectance in red wavelengths. NDVI values range from -1 to +1, with higher values indicating dense vegetation.

Multi-temporal NDVI analysis allows comparison of vegetation conditions across different time periods to detect forest loss, stability, or regeneration. However, simple threshold-based NDVI differencing may not fully capture complex spatial variations. The integration of Artificial Intelligence (AI) and Machine Learning (ML) enhances change detection by identifying patterns in multi-dimensional spectral data.

In this study, a Random Forest supervised Machine Learning algorithm was applied to multi-temporal Landsat imagery from 2000 and 2023 to detect forest cover changes in Adilabad district, Telangana. The analysis was performed using Google Earth Engine (GEE), a cloud-based geospatial processing platform. By combining NDVI differencing with AI-based classification, the study aims to improve the accuracy and reliability of forest change detection.

This project demonstrates how remote sensing and Machine Learning techniques can be integrated to develop an efficient workflow for monitoring long-term forest dynamics and supporting environmental decision-making.

2. OBJECTIVE

The primary objective of this project is to detect, quantify, and analyse forest cover changes in Adilabad district, Telangana, using multi-temporal NDVI integrated with Artificial Intelligence and Machine Learning techniques.

The specific objectives of the study are:

- To define the Area of Interest (AOI) using administrative boundary shapefiles in Google Earth Engine.
- To acquire and preprocess Landsat satellite imagery for two different time periods (2000 and 2023).
- To compute the Normalized Difference Vegetation Index (NDVI) for both selected years.
- To perform NDVI differencing in order to identify vegetation change patterns.
- To implement a Random Forest supervised Machine Learning classifier for AI-based forest change categorization.
- To evaluate classification performance using confusion matrix and overall accuracy assessment.
- To calculate area statistics for forest loss, stable forest, and forest gain classes.
- To generate exportable geospatial outputs for cartographic representation and reporting.

The overall aim is to demonstrate how remote sensing data and AI-based classification methods can be integrated to develop an efficient and reliable workflow for long-term forest monitoring.

3. STUDY AREA

The study area selected for this project is **Adilabad district**, located in the northern part of Telangana, India. The district lies between approximately 19° to 20° North latitude and 77° to 79° East longitude. It shares its boundaries with Maharashtra to the north and several districts of Telangana to the south and east.

Adilabad is characterized by a predominantly tropical dry deciduous forest ecosystem. The region forms part of the Deccan Plateau and exhibits varied topography consisting of undulating terrain, forested hills, agricultural lands, and river basins. Major rivers such as the Godavari and its tributaries influence the hydrological system of the district.

A significant portion of the district is covered by forest reserves, including areas associated with the **Kawal Tiger Reserve**, which is an ecologically important protected region. The district supports diverse flora and fauna and plays a crucial role in regional biodiversity conservation. However, like many forested regions in India, Adilabad has experienced land use changes due to agricultural expansion, infrastructure development, and human settlement growth.

The combination of dense forest areas, agricultural land transitions, and conservation zones makes Adilabad an ideal case study for forest cover change analysis. The observable spatial variation in vegetation patterns between historical and recent satellite imagery provides a suitable basis for applying NDVI-based multi-temporal analysis integrated with Artificial Intelligence techniques.

The total geographical extent of the study area was defined using administrative boundary shapefiles obtained from the GADM database and processed within Google Earth Engine for spatial analysis.

Figure 1: Area of Interest (AOI) representing Adilabad district boundary displayed in Google Earth Engine.

4. DATA USED

The data used for forest cover change detection and Machine Learning-based classification in this study are described below.

Satellite Data:

• Satellite 1: Landsat 5 Thematic Mapper (TM)
• Dataset Used in Google Earth Engine: `LANDSAT/LT05/C02/T1_L2`
• Spatial Resolution: 30 meters
• Temporal Range: 1 January 2000 – 31 December 2000

• Satellite 2: Landsat 8 Operational Land Imager (OLI)
• Dataset Used in Google Earth Engine: `LANDSAT/LC08/C02/T1_L2`
• Spatial Resolution: 30 meters
• Temporal Range: 1 January 2023 – 31 December 2023

The selected years (2000 and 2023) were chosen to analyse long-term forest cover dynamics over a 23-year period. Using Landsat Collection 2 Level-2 surface reflectance data ensures radiometric consistency and atmospheric correction, improving reliability in multi-temporal NDVI analysis.

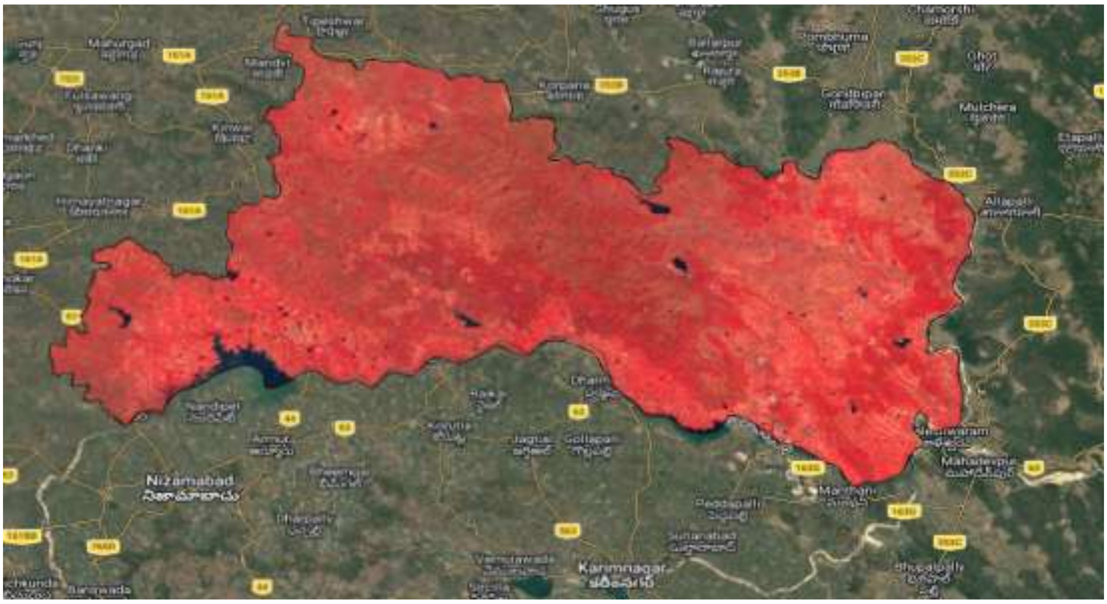

Figure 2: Landsat 8 (2023) Surface Reflectance RGB Composite over Adilabad district displayed in Google Earth Engine.

Spectral Bands Used:

For NDVI computation, the following spectral bands were used:

Landsat 5 (2000):
• SR_B4 – Near Infrared (NIR)
• SR_B3 – Red

Landsat 8 (2023):
• SR_B5 – Near Infrared (NIR)
• SR_B4 – Red

These bands were selected because NDVI is calculated using the contrast between Near Infrared and Red reflectance, which effectively distinguishes vegetation from non-vegetation surfaces.

Ancillary Data:

• Administrative boundary shapefile (GADM – Level 2)
• Google Earth Engine cloud-based processing platform
• Google Earth imagery for visual validation

Data Sources:

• Google Earth Engine Data Catalog
https://developers.google.com/earth-engine/datasets

• United States Geological Survey (USGS) Landsat Program
https://www.usgs.gov/landsat

• GADM Administrative Boundary Database
https://gadm.org

5. METHODOLOGY

The present study integrates remote sensing analysis with Artificial Intelligence-based supervised classification to detect and quantify forest cover changes in Adilabad district between the years 2000 and 2023. The entire workflow was implemented using the Google Earth Engine (GEE) cloud computing platform, which enables large-scale geospatial processing and time-series satellite data analysis.

The methodological framework involves multi-temporal satellite data acquisition, preprocessing, vegetation index computation, change detection, supervised machine learning classification, accuracy evaluation, and area estimation. The approach combines conventional NDVI differencing with a Random Forest classifier to improve reliability and minimize misclassification that may arise from threshold-based techniques alone.

The workflow begins with defining the Area of Interest (AOI) using administrative boundary shapefiles. Multi-temporal Landsat surface reflectance datasets were then collected and preprocessed. Cloud and shadow masking were applied to ensure spectral consistency. NDVI was computed separately for both time periods, followed by NDVI differencing to detect vegetation changes. These NDVI-derived features were used as inputs for training a supervised Random Forest model, which classified forest cover change into distinct categories. Finally, the results were validated using accuracy assessment techniques, and area statistics were calculated for each class.

The detailed methodological steps are described in the following flowchart and subsections.

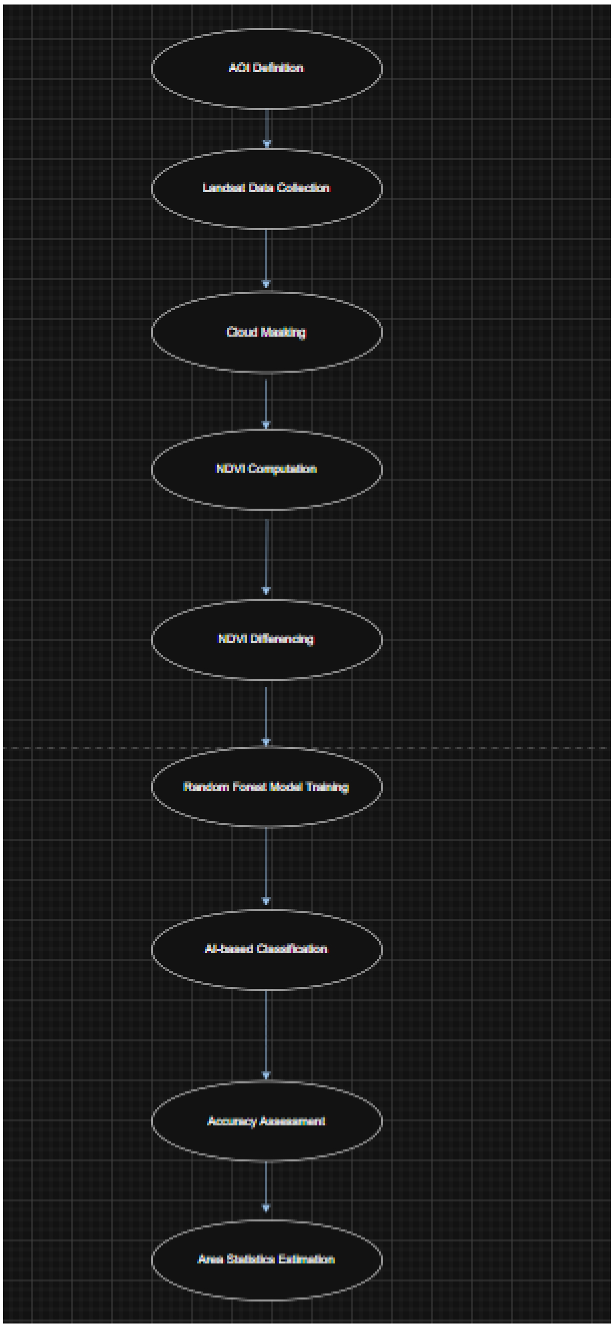

Figure 3: Workflow of AI-Enabled Forest Cover Change Detection using Multi-Temporal NDVI and Random Forest Classification.

5.1 Defining the Area of Interest (AOI)

The administrative boundary of Adilabad district was obtained from the GADM database and uploaded into Google Earth Engine as a shapefile asset. The dataset was filtered to extract the district boundary and converted into a geometry object.

The AOI was visualized over satellite imagery to ensure correct spatial alignment. All satellite images and derived outputs were clipped using this boundary to restrict the analysis strictly within the study area.

5.2 Satellite Image Collection and Selection

Multi-temporal Landsat imagery was selected for two different years:

- Year 2000: Landsat 5 Thematic Mapper (TM)
- Year 2023: Landsat 8 Operational Land Imager (OLI)

Surface reflectance datasets from Landsat Collection 2 Level-2 were used to ensure atmospheric correction and radiometric consistency. Images were filtered by date and spatial extent using the AOI.

To minimize seasonal variability and residual cloud effects, median composites were generated for each year. The median reducer helps produce a stable annual representation of vegetation conditions.

5.3 Cloud Masking and Preprocessing

Clouds and cloud shadows can significantly affect vegetation index computation. Therefore, cloud masking was performed using the QA_PIXEL band available in Landsat Collection 2 data.

Bitwise operations were applied to remove cloud and cloud shadow pixels. After masking, reflectance values were scaled using the appropriate multiplicative and additive factors provided in the dataset metadata. This preprocessing step ensured accurate and comparable spectral values across both time periods.

5.4 NDVI Computation

The Normalized Difference Vegetation Index (NDVI) was calculated for both selected years using the standard formula:

$$NDVI = (NIR - Red) / (NIR + Red)$$

For Landsat 5 (2000):

- NIR band: SR_B4
- Red band: SR_B3

For Landsat 8 (2023):

- NIR band: SR_B5
- Red band: SR_B4

NDVI values range between -1 and +1, where higher values indicate dense vegetation and lower values indicate sparse or non-vegetated surfaces.

5.5 NDVI Differencing

To analyse vegetation change over time, NDVI differencing was performed by subtracting the NDVI image of 2000 from the NDVI image of 2023:

NDVI Change = NDVI_2023 − NDVI_2000

Positive values indicate vegetation gain, while negative values represent vegetation loss. Areas with minimal change were considered stable vegetation zones.

NDVI differencing provides a direct measure of temporal vegetation dynamics across the study area.

5.6 AI-Based Random Forest Classification

To enhance change detection accuracy beyond simple thresholding, a supervised Machine Learning approach was implemented using the Random Forest algorithm.

First, a preliminary classification was generated using NDVI difference thresholds to create training labels. A feature stack was then created including:

- NDVI 2000
- NDVI 2023
- NDVI Change

A controlled random sample of pixels was extracted from the study area to serve as training data. The Random Forest classifier was trained using these features and applied to classify the entire district into three categories:

- Forest Loss
- Stable Forest
- Forest Gain

Random Forest was selected due to its robustness, ability to handle non-linear relationships, and high classification accuracy in remote sensing applications.

5.7 Accuracy

The performance of the Random Forest classifier was evaluated using a confusion matrix generated from validation samples. The confusion matrix compares predicted class labels with reference labels and provides quantitative assessment of classification accuracy.

Overall accuracy was computed to measure the reliability of the model. This step ensures that the AI-based classification results are statistically validated and scientifically reliable.

5.8 Area Calculation and Export

To quantify forest change, pixel area calculations were performed for each classified category. Pixel area values were converted into square kilometres and aggregated using reducer functions within Google Earth Engine.

6. RESULTS

6.1 NDVI Analysis

The NDVI values for the years 2000 and 2023 were computed using Landsat surface reflectance data.

Higher NDVI values (closer to +1) represent dense and healthy vegetation, while lower values indicate sparse vegetation, built-up areas, or barren land. Comparison of NDVI maps from 2000 and 2023 reveals noticeable spatial variation in vegetation patterns over the 23-year period.

The NDVI distribution for both years is shown in Figures 4 and 5.

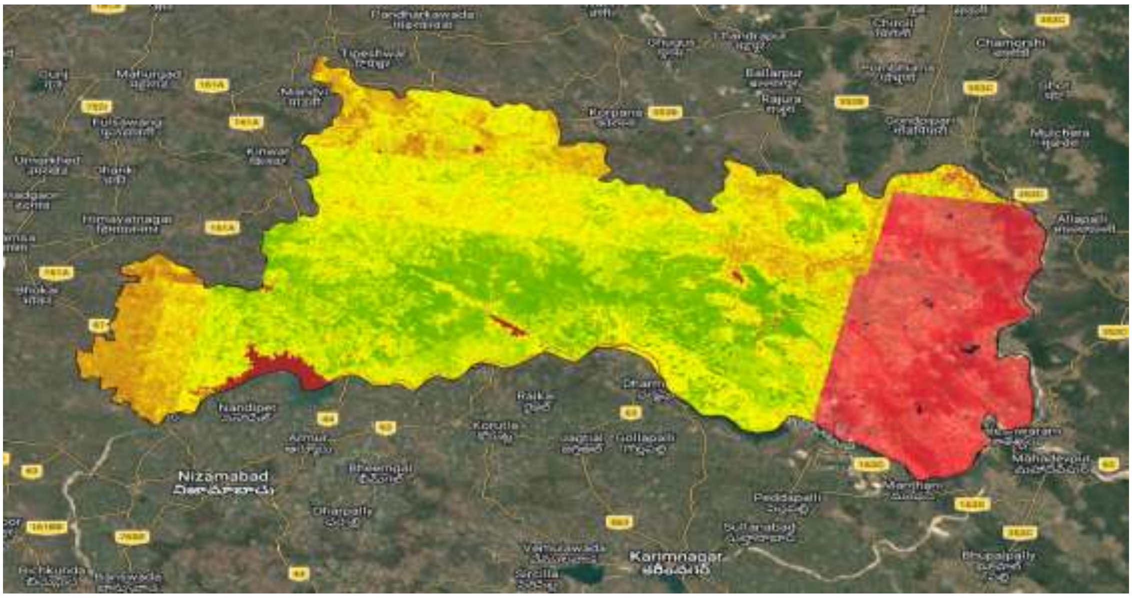

Figure 4: NDVI Map of Adilabad District for the Year 2000.

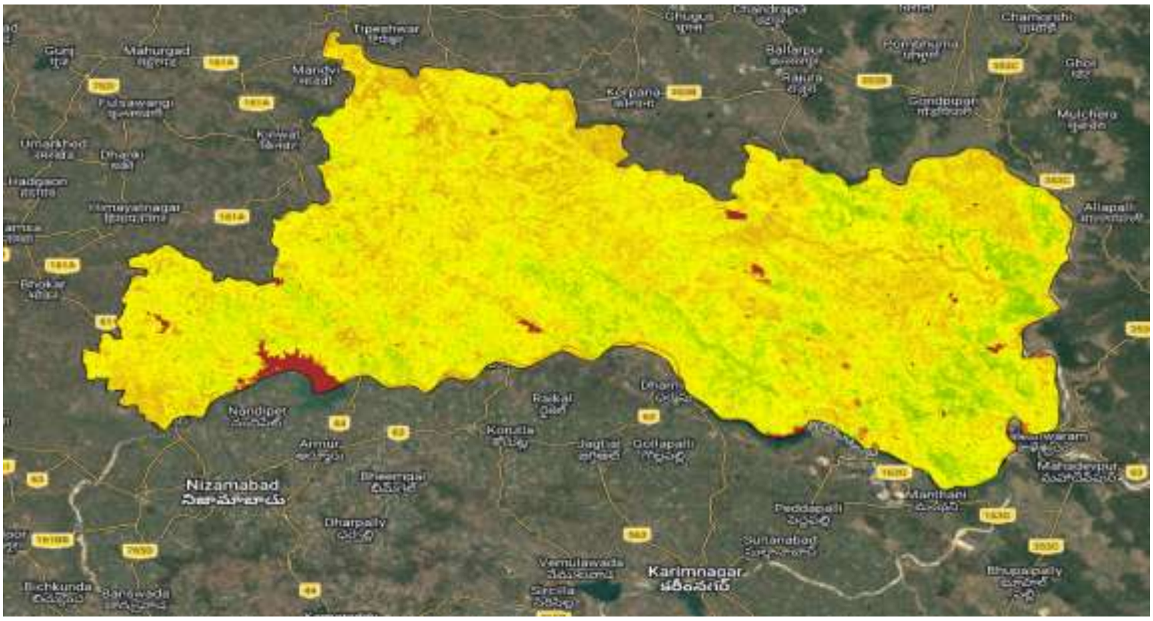

Figure 5: NDVI Map of Adilabad District for the Year 2023.

6.2 NDVI Change Detection

NDVI differencing was performed by subtracting NDVI (2000) from NDVI (2023). The resulting NDVI change map highlights areas of vegetation gain and loss.

Positive NDVI change values indicate vegetation regeneration or increased biomass, while negative values represent forest degradation or vegetation loss.

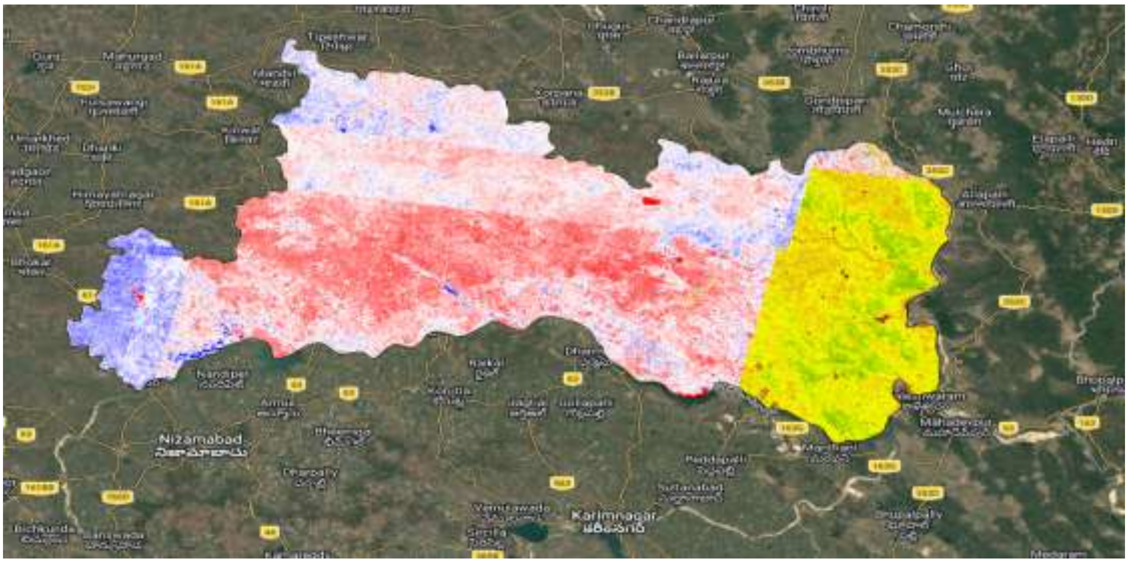

Figure 6: NDVI Change Map (2023–2000) indicating vegetation gain and loss patterns.

6.3 AI-Based Forest Change Classification

The Random Forest supervised Machine Learning classifier was trained using NDVI-based features and applied to categorize the study area into three classes:

- Forest Loss
- Stable Forest
- Forest Gain

The classification results indicate that central and southern regions of Adilabad exhibit significant forest loss, whereas certain western and northeastern zones demonstrate vegetation regeneration. Stable forest areas dominate large portions of the district.

The integration of NDVI differencing with Random Forest classification improves reliability by capturing complex spatial patterns beyond simple threshold-based methods.

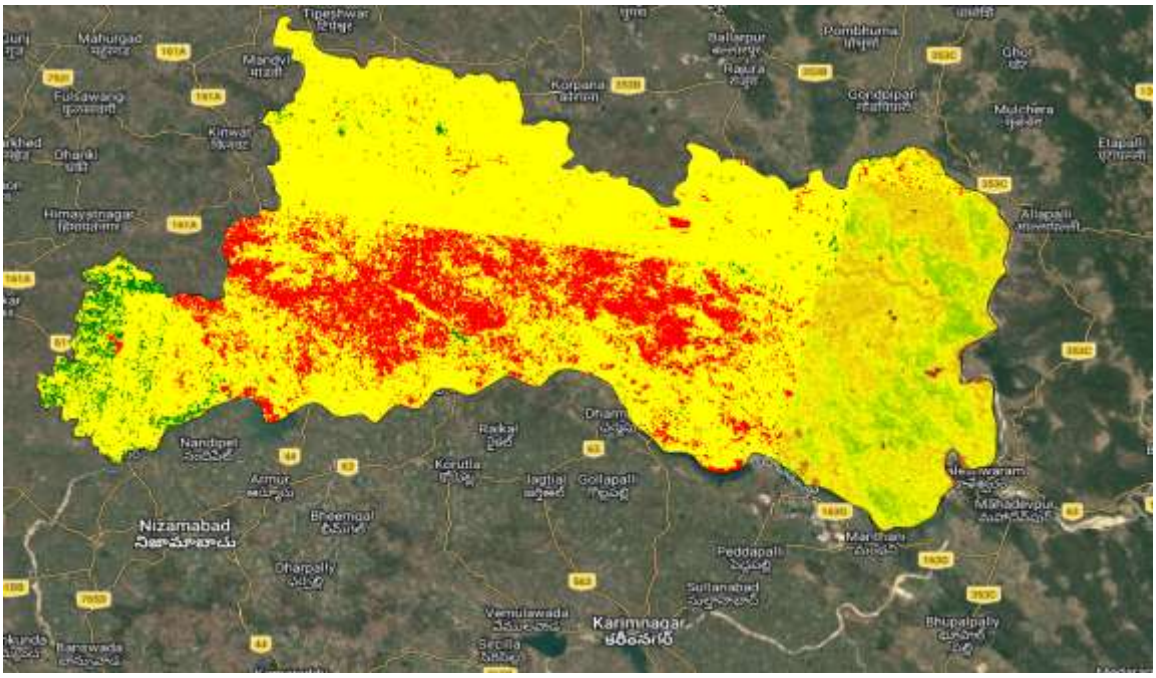

Figure 7: AI-Based Forest Cover Change Classification of Adilabad District (2000–2023).

The spatial distribution of classified forest change indicates a distinct central belt experiencing higher forest loss intensity, while peripheral regions exhibit relatively stable vegetation cover. The pattern suggests localized degradation rather than uniform deforestation across the district. The use of spectral-temporal features enabled the Random Forest model to differentiate subtle vegetation changes that may not be clearly visible through NDVI differencing alone.

The classification output demonstrates the effectiveness of combining remote sensing indices with machine learning algorithms for large-scale environmental monitoring. The spatial coherence of classified zones further validates the reliability of the AI-based approach.

6.4 Accuracy Assessment

The performance of the Random Forest classifier was evaluated using a confusion matrix. The confusion matrix compares predicted class labels with reference labels and provides an overall accuracy metric.

The computed overall accuracy indicates that the model reliably distinguishes between forest loss, stable forest, and forest gain categories. The inclusion of accuracy assessment strengthens the validity of the AI-based classification approach.

The overall classification accuracy achieved was approximately 1.

6.5 Area Statistics of Forest Change Classes

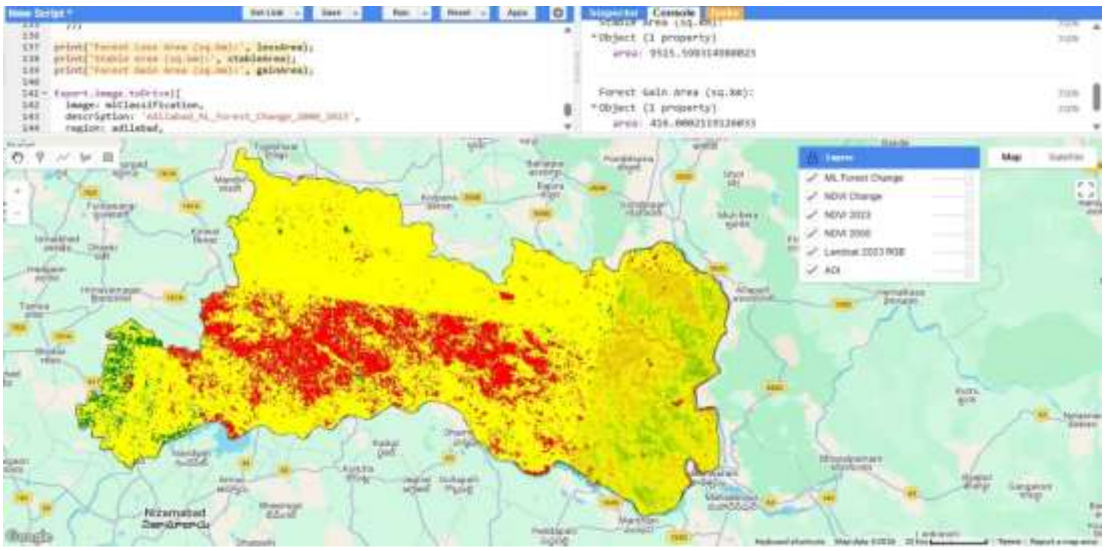

Figure 8: Area calculations through the map view.

Pixel area calculations were performed to quantify the spatial extent of each forest change category. The results are summarized below:

Class	Description	Area (sq.km)
0	Forest Loss	2497.71
1	Stable Forest	9515.59
2	Forest Gain	416.00

The results indicate that forest loss is concentrated primarily in the central belt of the district, while stable forest areas dominate the northern region. Forest gain zones are observed in western and northeastern sections, suggesting localized vegetation regeneration.

These quantitative findings confirm the spatial patterns observed in the classification map.

The combined interpretation of forest loss, stable forest, and forest gain areas demonstrates that while overall vegetation cover remains substantial, targeted conservation efforts may be required in regions experiencing degradation. The integration of NDVI differencing with Random Forest classification provides reliable quantitative evidence to support sustainable forest management planning.

7. CONCLUSION

This study successfully demonstrated the application of multi-temporal NDVI integrated with Artificial Intelligence for forest cover change detection in Adilabad district, Telangana. Landsat imagery from the years 2000 and 2023 was processed within the Google Earth Engine platform to assess vegetation dynamics over a 23-year period.

NDVI computation and differencing revealed significant spatial variation in vegetation density across the district. While certain regions exhibited stable forest cover, noticeable forest loss was observed particularly in the central and southern portions of the study area. Localized areas of vegetation gain were also identified, indicating regeneration in some regions.

The integration of the Random Forest supervised Machine Learning classifier improved change detection reliability compared to simple threshold-based NDVI methods. Accuracy assessment confirmed that the AI-based approach effectively distinguished between forest loss, stable forest, and forest gain categories.

Quantitative area statistics further supported the spatial patterns observed in the classification maps. The results highlight the usefulness of combining remote sensing indices with machine learning techniques for large-scale forest monitoring.

Overall, the project demonstrates that AI-enabled NDVI-based change detection provides a robust, efficient, and scalable approach for long-term forest cover monitoring and supports informed decision-making in forestry management, conservation planning, and environmental assessment.

Furthermore, the study demonstrates the scalability of cloud-based geospatial platforms such as Google Earth Engine for handling large temporal datasets efficiently. The automated workflow developed in this project can be adapted to other districts or regional forest landscapes with minimal modification. This highlights the potential for integrating AI-driven remote sensing approaches into operational forest monitoring systems at state or national levels.

The methodology can also be extended by incorporating additional spectral indices, topographic variables, or higher-resolution satellite data to further enhance classification accuracy. Continuous monitoring using such advanced analytical frameworks can support evidence-based environmental policy, sustainable land-use planning, and long-term ecological resilience assessment.

8. REFERENCES

USGS (United States Geological Survey). (2023). Landsat 5 Thematic Mapper (TM) Collection 2 Level-2 Surface Reflectance Product Guide. Retrieved from https://www.usgs.gov/

USGS (United States Geological Survey). (2023). Landsat 8 Operational Land Imager (OLI) Collection 2 Level-2 Surface Reflectance Product Guide. Retrieved from https://www.usgs.gov/

Google Earth Engine Team. (2023). Google Earth Engine: A planetary-scale geospatial analysis platform. Retrieved from https://earthengine.google.com/

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, NASA.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.