ZIBELINE INTERNATIONAL™ PUBLISHING

ISSN: 2521-0920 (Print)
ISSN: 2521-0602 (Online)
CODEN: MJGAAN

Malaysian Journal of Geosciences (MJG)

DOI: http://doi.org/10.26480/mjg.01.2024.82.92

CrossMark

RESEARCH ARTICLE

GEOSPATIAL CROP YIELD MODELLING IN FUTA FARM

Elufisan Paul Tomilola[a], Aroge Sunday Kayode[b]*, Odeyemi Felix Gbenga[b], Titilade Adedeji Samuel[c], Ajetunmobi Ridwan[d]

[a] Department of Surveying and Geoinformatics, Ajayi Crowther University Oyo, Oyo State
[b] Department of Surveying and Geoinformatics, Federal Polytechnic Ede
[c] Department of Surveying and Geoinformatics, Federal School of Surveying, Oyo
[d]Department of Surveying and Geoinformatics, Federal University of Technology Akure, Ondo State.
**Corresponding author email: kayodearoge2000@gmail.com*

ARTICLE DETAILS

Article History:

Received 10 February 2024
Revised 18 March 2024
Accepted 24 April 2024
Available online 28 April 2024

ABSTRACT

This research aims to geospatially model crop yield in a FUTA farm to enhance productivity and management practices. Primary and secondary data were collected for seven planting seasons (2014-2021), including maize seeds, NPK fertilizers, urea used, harvesting dates, and yield measurements. GNSS observation was used to determine planting boundaries, while satellite imagery and climate records were used for secondary data. The study examined the vegetation indices (NDVI and GCI) of the farm between 2012 and 2022. Then, Artificial Neural Network (ANN) was used to model crop yield in the study area using the primary and secondary data and the NDVI, and GCI values. Result showed that Vegetative indices (NDVI and GCI) showed an increase between 2014 and 2016, while between 2017 and 2019, there was a decrease. In 2021, there was an increase in vegetative indices values, indicating healthier crops. The decline in crop yield between 2017 and 2019 was not coincidental, but it is believed that the decline in crop health is responsible for the corresponding reduction. The ANN model had a regression coefficient of 0.73282, and the coefficient of determination was 0.5176. The maximum and minimum crop yield values were 24.7 and 25.26 in 2016, and 5 and 4.01 in 2018, respectively. It was observed from that the minimum value of difference is -9.883708757 while the maximum value of difference is 1.451557122. The root means square error (RMSE) and the Mean Absolute Error (MAE) are 0.4296 and 0.2947, respectively. Modelled crop yield values were close to actual yield values, except for 2017 when a large difference was observed due to herdsmen invasion into the school farm. Since, the model showed close correlations with actual yield values, making it a recommended model for predicting crop yield in the study area.

KEYWORDS

Crop Yield, NDVI, GCI, Artificial Neural Network (ANN)

1. INTRODUCTION

Agriculture plays a crucial role in ensuring food security and supporting global economies. Efficient crop management practices are essential for optimizing agricultural productivity, reducing waste, and mitigating the environmental impact of farming activities. One of the major goals of agricultural production is to accomplish maximum yield at a low cost (Hassan and Goheer, 2021). Crop yield modeling is a vital component of agricultural research and planning. It involves the analysis of various factors that influence crop productivity, such as soil characteristics, climate variables, topography, and management practices. These factors operate at different spatial and temporal scales (Bendre et al., 2015; Bose et al., 2016).

Climate change is posing new challenges to crop-related concerns, including food insecurity, supply stability, and economic planning (Fan et al., 2022). The yields of important crops are likewise gradually impacted by climatic changes (Zhao et al., 2017). According to crop production is susceptible and sensitive to changes in climatic parameters such as temperature, precipitation, soil moisture, and many other factors (Ortiz-Bobea et al., 2018). All of these quickly shifting elements could affect annual crop yields as the earth warms up; climate change has already slowed the rise of agricultural output by 21% (Ortiz-Bobea et al., 2021). Topographic features, such as slope and aspect, influence water drainage and soil moisture availability can also affect crop development and yield potential (Fan, et al., 2022). Additionally, management practices, such as irrigation, fertilization, and pest control, impact crop productivity. Understanding the combined effects of these factors and their spatial distribution is crucial for effective farm management and sustainable agricultural production.

Traditionally, crop yield assessments have been carried out using field surveys, which can be time-consuming, labor-intensive and limited in spatial coverage (Zhou et al., 2017). However, the integration of geospatial technologies has revolutionized the data collection process by allowing for the collection of large-scale, high-resolution data across different spatial and temporal scales. Remote sensing techniques enable the acquisition of continuous and high-resolution data, allowing for the monitoring of crop growth, health, and stress levels (Wang et al., 2015). Satellite imagery, with its wide coverage and frequent revisit times, provides valuable information on vegetation indices, such as Normalized Difference Vegetation Index (NDVI), which can be correlated with crop yields (Balogun et al., 2023; Zhou et al., 2017). Aerial photographs, captured at a finer scale, offer detailed information on crop phenology and growth patterns (Bolton, and Friedl, 2013).

Geographic Information Systems (GIS) also provide a powerful framework for integrating and analyzing diverse datasets. By overlaying crop yield data with spatial layers representing soil properties, climate variables,

Quick Response Code	Access this article online	
	Website: www.myjgeosc.com	DOI: 10.26480/mjg.01.2024.82.92

Cite The Article: Elufisan Paul Tomilola[a], Aroge Sunday Kayode[b]*, Odeyemi Felix Gbenga[b], Titilade Adedeji Samuel (2024). Geospatial Crop Yield Modelling In Futa Farm. *Malaysian Journal of Geosciences, 8(1): 82-92.*

topography, and management practices, researchers can identify spatial patterns and relationships (Sagan et al., 2021). GIS-based spatial interpolation techniques, such as kriging or inverse distance weighting, allow for the estimation of crop yields at unsampled locations, providing a comprehensive understanding of the spatial variability of crop productivity within the FUTA farm (Balogun et al., 2023).

The application of advanced modeling techniques, including machine learning algorithms like ANN, GNN, CNN etc, and statistical analyses, further enhances the accuracy and predictive capabilities of geospatial crop yield models (Manjula and Djodiltachoumy, 2017; Shahhosseini et al., 2021; Tiwari and Shukla 2019; Satorras et al., 2017; Kipf and Welling, 2016; Wang et al., 2015; Bendre, 2015; Sekulić et al., 2020). By training the models on historical crop yield data and relevant environmental variables, the study develops a robust model capable of predicting crop yields across the farm. These models can then be validated using independent field measurements to assess their accuracy and reliability.

This research aims to explore the analysis of multi-temporal remote sensing data for crop yield modeling with a view to enhance crop management practices and productivity in FUTA farm. In this research, the

spatial and temporal variability of crop yields within the farm were assessed to develop a robust geospatial crop yield model using advanced modeling techniques; and evaluate the performance of the crop yield model and validate its accuracy through field measurements.

1.1 Study Area

The research focused specifically on the FUTA farm, located in Akure, Nigeria. The study area encompasses the entire farm, including its diverse crop cultivation and experimental plots. The study area is located between 7º18′0″ N, 5º8′ 45″ E and 7º17′ 45″ N, 5º9′0″ E in Akure as shown in Figure 1. In terms of land cover, Akure is a rapidly growing city. Urban development in the city has converted agricultural and forestry land into built-up areas, such as residential, commercial, and industrial areas. The study area consists of 10 plots which covers an area of 45.895 hectares. The area of each of the 10 plots are 5.854 hectares, 9.499 hectares, 2.958 hectares, 3.539 hectares, 5.428 hectares, 5.134 hectares, 0.891 hectares, 6.479 hectares, 5.459 hectares and 0.654 hectares respectively.

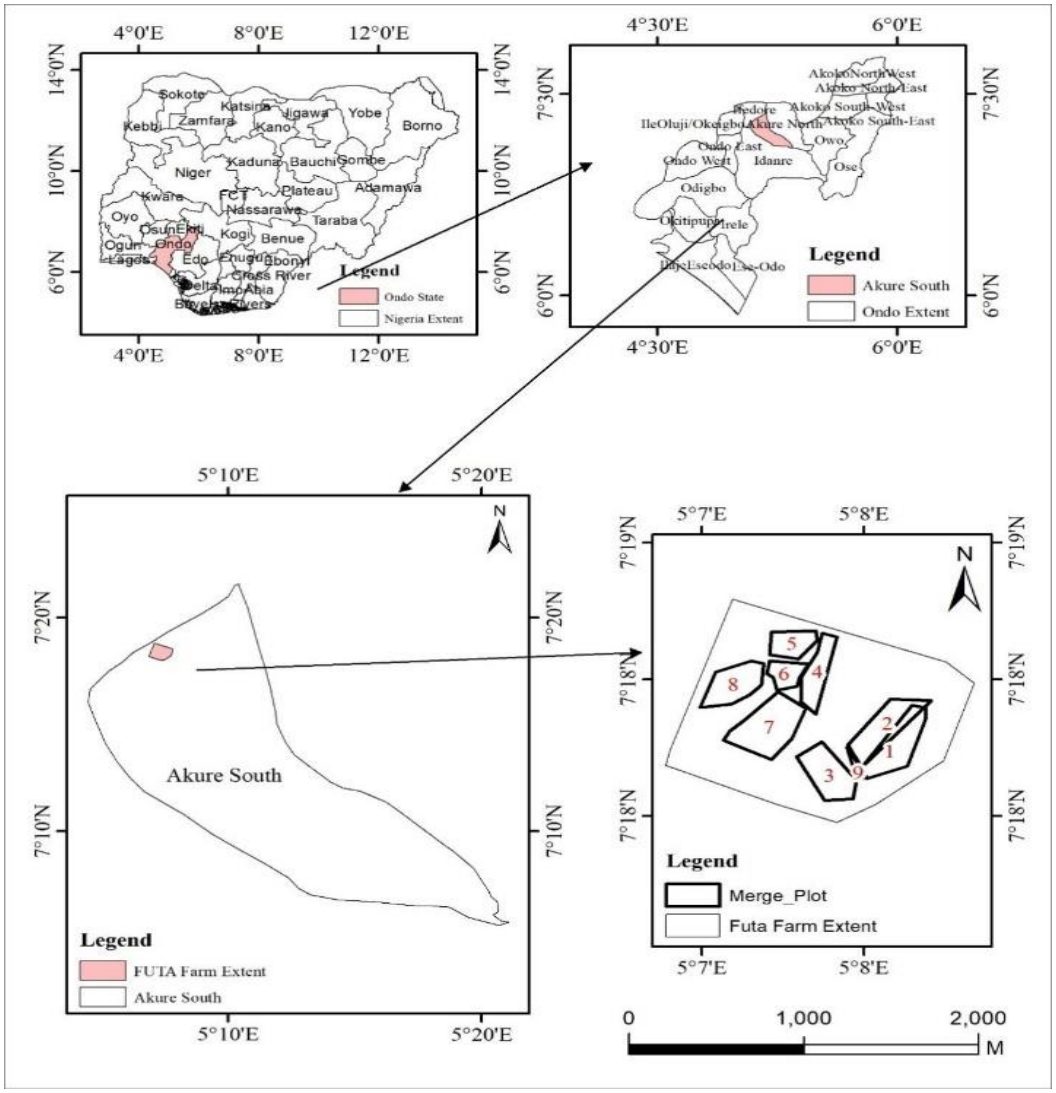

Figure 1: Map of the Study area (Source: Authors' Work, 2024)

2. METHODS

This research used field surveys and satellite imagery to gather data for geospatial crop yield modeling. GNSS observation was used to determine farmland plot boundaries, while secondary data was obtained from satellite imagery, climate records, and SRTM. Crop yield measurements were collected from various plots within the FUTA farm. The data were integrated and analyzed using ArcGIS, with satellite imagery subjected to preprocessing for accuracy. Vegetation indices like the Normalized Difference Vegetation Index were calculated to quantify crop growth. Machine learning algorithms, such as artificial neural networks, were employed to develop a geospatial crop yield model. The model's performance was evaluated using independent field measurements and statistical metrics.

2.1 Data Collection

The study involves the collection of both primary and secondary sources of data. The primary source of data includes GNSS data of each plot for designating the boundaries of the cultivated areas. The secondary data were gathered from existing datasets, such as satellite imagery (Landsat) and climate data. Landsat satellite imagery was acquired with appropriate spectral bands and temporal coverage to capture crop growth dynamics throughout the growing seasons. Climatic data was acquired for the study area from world weather online website. Also, data on crop yield samples were collected from the department of Crop, Soil & Pest management, Federal University of Technology, Akure. The process by which these data were collected are explained below. Table 1 show the summary of the source, data type, and output of the data acquired.

Table 1: Primary Data, Source, Mode of acquisition, Resolution, format and usefulness

S/N	Data	Source	Resolution	Period of Acquisition	Application
1	Crop Yield in tonnes	CSP, Department, FUTA	-	2012 - 2022	Ground truth data for validation
2	Farmland Plots boundary data	Field Observation	-	2022	To understand the area used for plantation within the study area alongside their effects in determining the yield of crops within such area.
3	Landsat 7, 8, 9 Imagery OLI/TIRS	USGS through Earth Explorer	30	2012 – 2021	Vegetative indices (NDVI and GCI)
5	Temperature	NASA Power		2012 – 2021	Input for crop yield modelling
6	Rainfall	NASA Power		2012 – 2022	Input for crop yield modelling
7.	Shuttle Radar Topographic Mission (SRTM)	Earth explorer	30m		To obtain the slope and aspect of the terrain which were used as Input for crop yield modelling.

Source: Author's Work

2.2 Data Processing

2.2.1 Image Processing

The digital numbers (DN) in the imagery were Converted to radiance values to ensure consistency across different images. Then the atmospheric effects such as scattering and absorption was corrected to improve the accuracy of the imagery. Rectification was done to the imageries to correct for distortions caused by terrain and sensor characteristics. multiple images mosaic to create a seamless composite for each year.

2.2.2 Remote Sensing Indices

Vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) Green Chlorophyll Vegetation Index (GCI) was calculated, using satellite imagery to quantify crop growth and vegetation vigor.

The Normalized Difference Vegetation Index (NDVI) uses Red and Near-Infrared bands of Landsat images to determine the state of health of vegetative properties within the area. The appropriate bands (typically Landsat bands 4 and 5) were selected for NDVI calculation. Hence, it is calculated in ArcGIS 10.8 environment using the expression shown in Equation 1

$$\text{NDVI}=\frac{NIR-R}{NIR+R} \quad (1)$$

Where:

NIR is the Near Infrared band of the Landsat series (Band 5 for Landsat 8 and Band 4 for Landsat 7); R is the Red Band (Band 4 for Landsat 8 and Band 3 for Landsat 7)

The GCI's ability to determine the chlorophyll content of a particular vegetated area, was used to determine the growth stage and health of crops in the study area. GCI is calculated using blue, green, and red band data. The appropriate Landsat bands (typically bands 1, 2, and 3) was selected for GCI calculation. The formula in Equation 2 was used to derive the GCI of the study area for this study.

$$\text{GCI} = \frac{NIR}{GREEN} - 1 \quad (2)$$

Where: Green is Band 3 in Landsat 8 and Band 2 for Landsat 7

These indices are used as an indicator for crop yield in the study area and it was validated by crop yield values of each year collected from CSP department of FUTA.

2.2.3 Temperature and Rainfall

The study used NASA Power's temperature and rainfall data to analyze the Study area's temperature distribution in degrees Celsius and precipitation distribution. Data was preprocessed to remove missing values and outliers, and average values for 2012-2022 were calculated using Microsoft Excel's average function.

2.2.4 Slope

The study area's slope was obtained using SRTM (DEM) data, clipped using ArcGIS 10.7 software, and analyzed using the slope tool. The degree rise option was selected for visual representation. The output was categorized into Low, Medium, and High based on slope values.

2.2.5 Aspect

The Aspect of the area was also acquired from the DEM of the study area. The processing is similar to the processing of the slope variable. Still, it involves using the Aspect tool instead of the Slope tool to get the Aspect representation of the study area. The final result gives the direction of the slope (Aspect) with the study area.

2.2.6 Geospatial Crop Yield Modeling

This study developed a geospatial crop yield model using Artificial Neural Networks (ANN) to establish relationships between vegetation indices and factors influencing crop yield. The model was trained in MATLAB using NDVI, GCI, and average values for each factor from 2012-2022. Crop yield values were the target, while inputs included maize seed, NPK, urea, GCI, NDVI Average Temperature, and precipitation. 70% of the datasets were used for training, and 30% for validating and testing.

The training datasets were used to learn the ANN, adjusting weights and biases based on the relationship between input and target datasets. Over-fitting can cause weak mapping. Validation datasets were used to guide and stop the training process. The developed ANN crop yield model was evaluated using standard coefficients like root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R-squared). A lower RMSE indicates better fit between modelled and actual crop yield values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)}{n}} \quad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \quad (4)$$

Where,

y_i = Model NDVI and GCI values for each year; x_i = GIS-based NDVI and GCI values for each year; $\bar{x}$= Mean of Observed GIS-based NDVI and GCI values; n = number of years

3. RESULT

3.1 Normalized Difference Vegetative Index (NDVI) changes in the past eleven (11) years

The various changes in the eleven (11) years were explained within the study to understand the various changes in the health of crops planted within the study area. Vegetated areas typically have higher NDVI values, while non-vegetated areas (e.g., open fields or urban areas) exhibit lower NDVI values. In 2012, the NDVI change shown in figure 2 showed moderate health, with high values observed in most planting plots except for plot 8. In 2013, all plots shown in figure 3 had low NDVI values, indicating little health. In 2014, the NDVI change shown in figure 4 showed little health, with values ranging from 0.09 to 0.309, indicating little health. In 2015, the NDVI change shown in figure 5 showed little health, with values ranging from 0.079 to 0.345.

In 2016, the NDVI values ranged between 0.065 and 0.385 as shown in figure 6, suggesting a bit of health of crops. In 2017, the NDVI value ranged between 0.082 and 0.293 as shown in figure 7, indicating little health of crop. In 2018, the NDVI ranged between 0.078 and 0.283 as shown in figure 8, with crops in Plot 2 showing more health than crops planted in other plots. In 2019, crops planted in Plot 1, 3, and 4 had the least health, with NDVI values ranging from 0.166 to 0.183. In 2020, the NDVI ranged between 0.07 and 0.252 as shown in figure 10, with most crops having poor health. In 2021, the NDVI values ranged between 0.06 and 0.389 as shown in figure 11, with most crops having poor health. In 2022, the NDVI ranged between 0.06 and 0.4, indicating poor crop health.

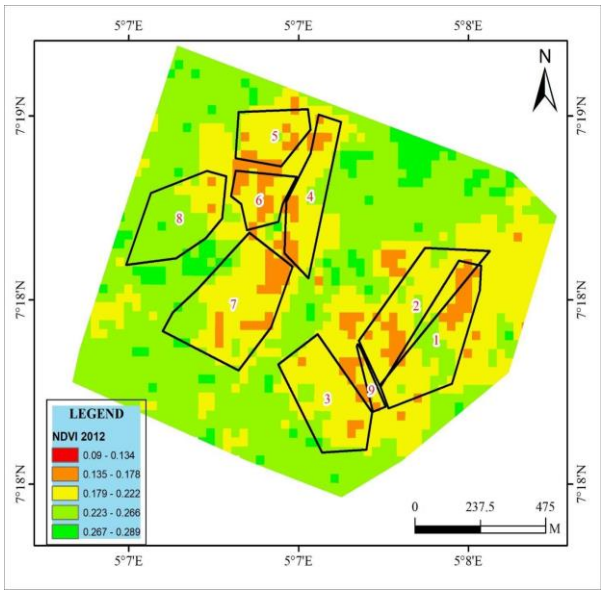

Figure 2: NDVI changes for the year 2012

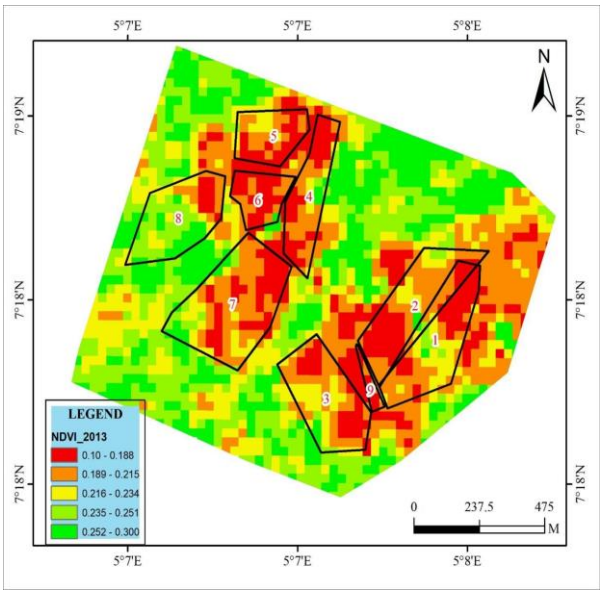

Figure 3: NDVI changes for the year 2013

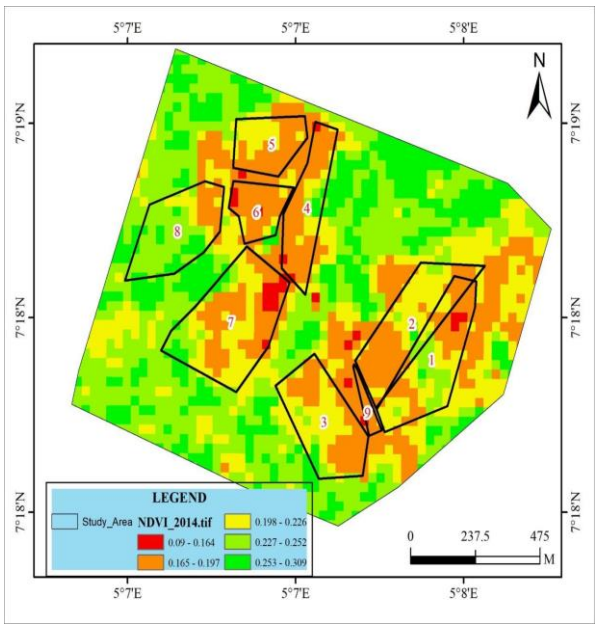

Figure 4: NDVI changes for the year 2014

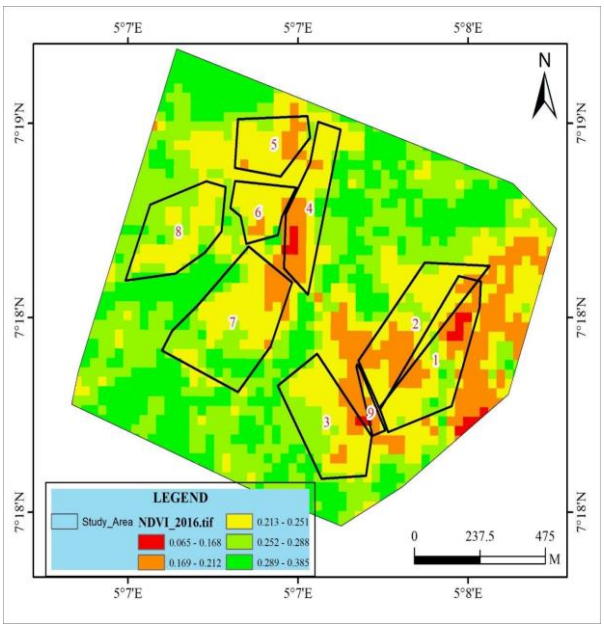

Figure 5: NDVI changes for the year 2015

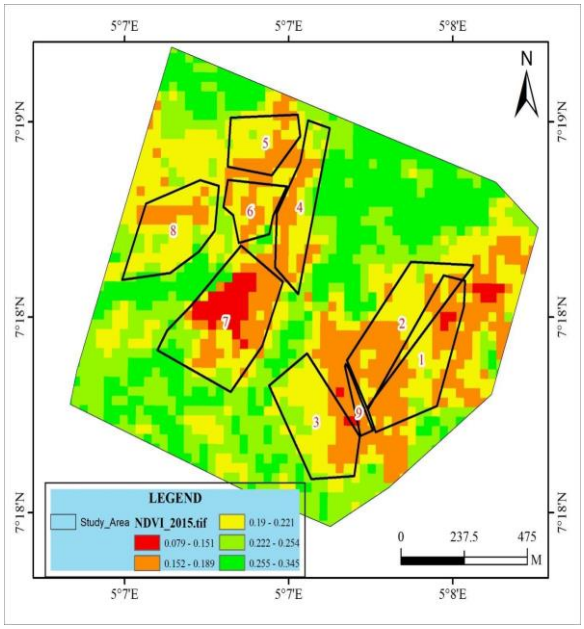

Figure 6: NDVI change for the year 2016

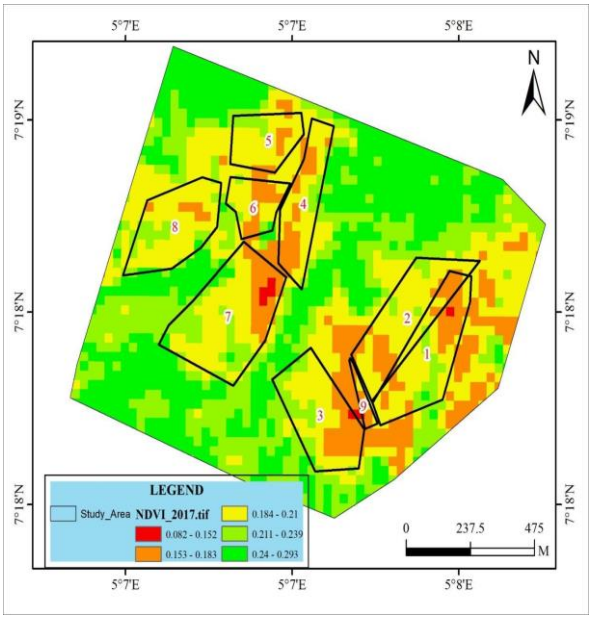

Figure 7: NDVI change for the year 2017

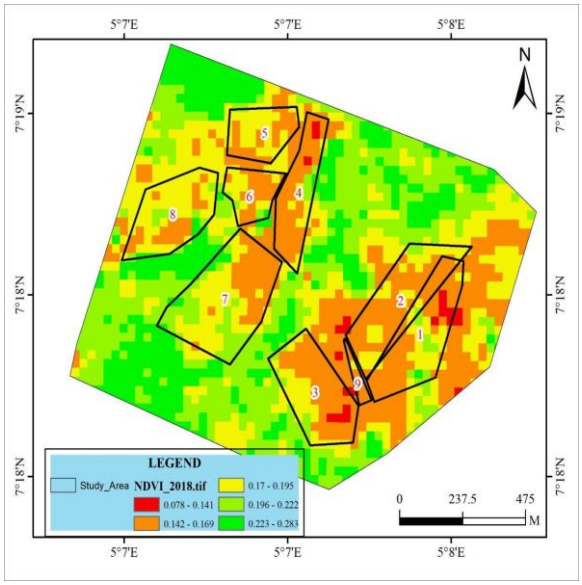

Figure 8: NDVI change for the year 2018

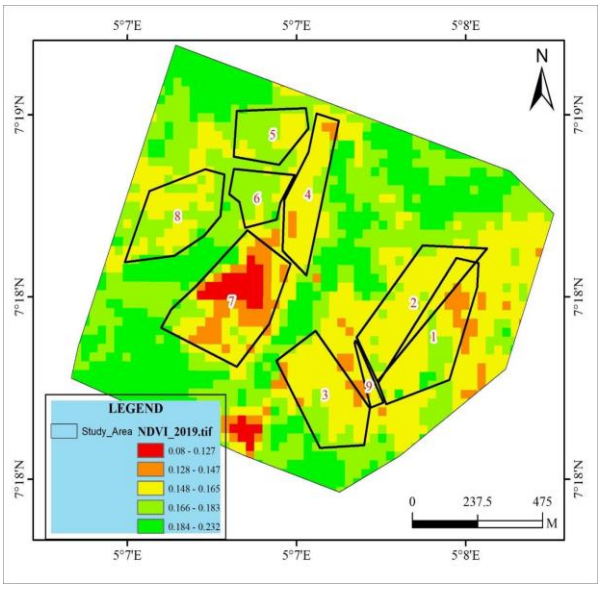

Figure 9: NDVI change for the year 2019

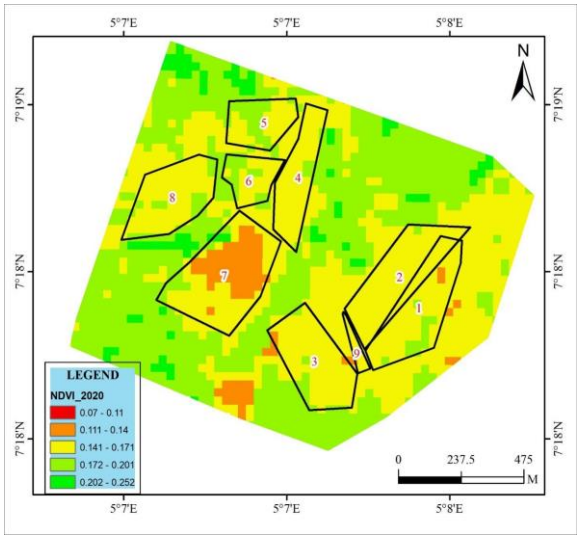

Figure 10: NDVI Change for the year 2020

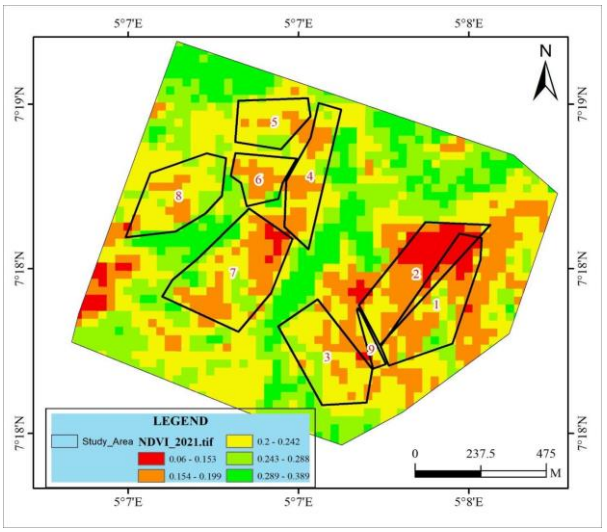

Figure 11: NDVI Change for the year 2021

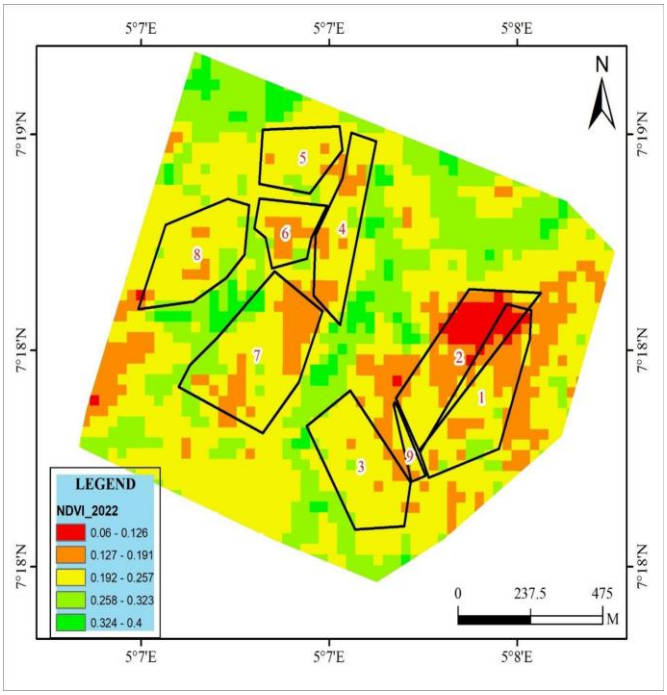

Figure 12: NDVI changes for the year 2022

Table 2 present the Average NDVI value of the study area between 2012 and 2022. The highest average NDVI value of 0.50 was recorded in 2022 while the lowest average NDVI value of 0.13 was recorded in 2020.

Table 2: Average NDVI value of the study area

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
NDVI	0.20	0.21	0.20	0.21	0.23	0.19	0.18	0.16	0.13	0.22	0.50

The curve of NDVI shown in figure 13 was plotted using the average values of the NDVI change for each year under study shows that there was a decline in the plant's health from 2017 down to 2019. Although it is quite hard to understand the causes of this decline in health, one of the major things that is believed will be responsible for its decline is the land usage rate. A critical look at the curve shows a sharp increase in the crop's health after the crop has been left to rest, regained its nutrients (due to COVID-19), and then prepared for plantation in the year 2021. This indicates that the break experienced by the land as a result of COVID-19 is needed and also responsible for the increase in the health status of the crops cultivated within the study area.

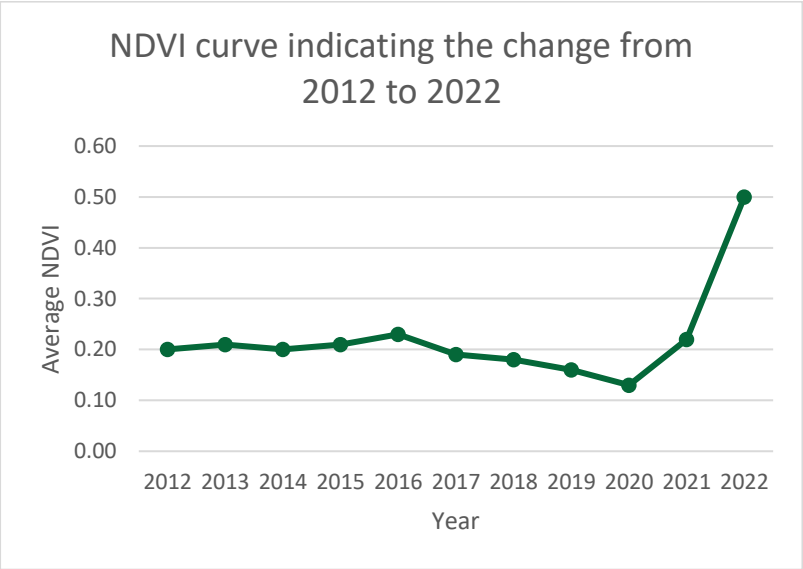

Figure 13: NDVI curve indicating the change from 2012 to 2022

3.1 Green Chlorophyll Vegetation Index (GCI) Changes

The study examines the chlorophyll content of crops planted in a specific area during different periods. In 2012, the crops were found to be moderately healthy, with values ranging from 0.203 to 0.549 as shown in figure 14. In 2013, figure 15 showed that the whole farm land was moderately healthy within the period of this study. In 2014, the crops were found to be healthy, with values ranging from 0.211 to 0.749 as shown in figure 16. However, plot 8 was expected to yield better due to its high GCI value. In 2015, the crops were healthy, with values ranging from 0.134 to 0.883 as shown in figure 17. Most plots had moderate chlorophyll content, except for plot 7 which had low content. In 2016, the crops were very healthy, with values ranging from 0.194 to 1.060 as shown in figure 18. However, some plots had low chlorophyll content, possibly due to early crop growth.

In 2017, the crops were healthy, with values ranging from 0.227 to 0.715 as shown in figure 19. However, some plots had low chlorophyll content, suggesting stress or lack of plantation. figure 20 showed that in 2018, the crops were healthy, but most had low chlorophyll levels, suggesting stress. Figure 21 showed that in 2019, the crops were fairly healthy, with values ranging from 0.197 to 0.546. However, only plot two had very low chlorophyll content, suggesting a built-up area. Other plots showed some increase in chlorophyll content, but these values were low compared to the beginning of the study. Figure 22 showed that in 2020, the crops were found to be very healthy and experiencing little to no stress. However, only plot two had low chlorophyll content, possibly due to an increase after the COVID-19 break. Figure 23 showed that in 2021, the crops were very healthy and experiencing little stress, with values ranging from 0.072 to 1.072. However, only plot two had low chlorophyll content, possibly due to increased chlorophyll content after the COVID-19 pandemic.

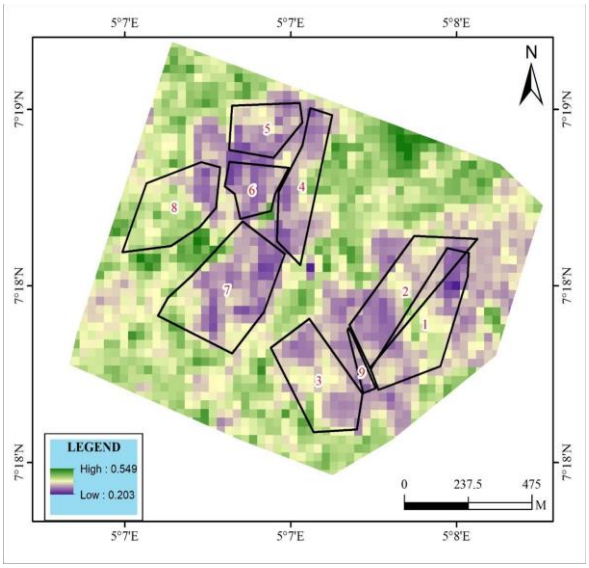

Figure 14: GCI change for the year 2012

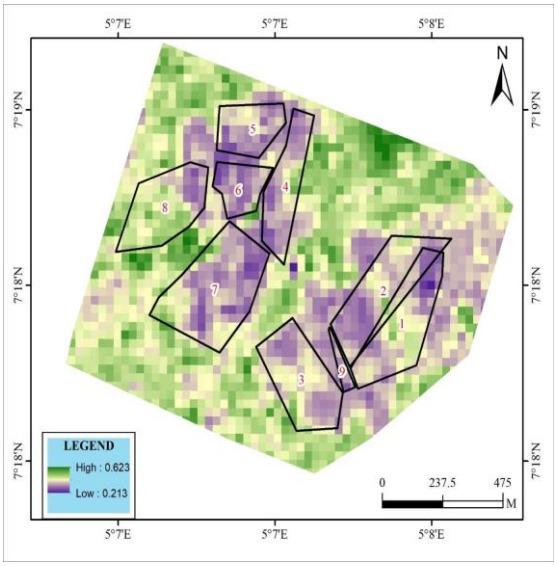

Figure 15: GCI change for the year 2013

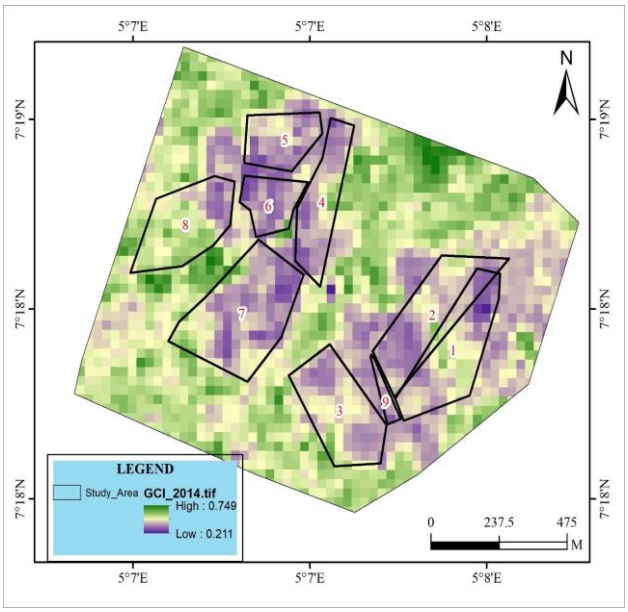

Figure 16: GCI change for the year 2014

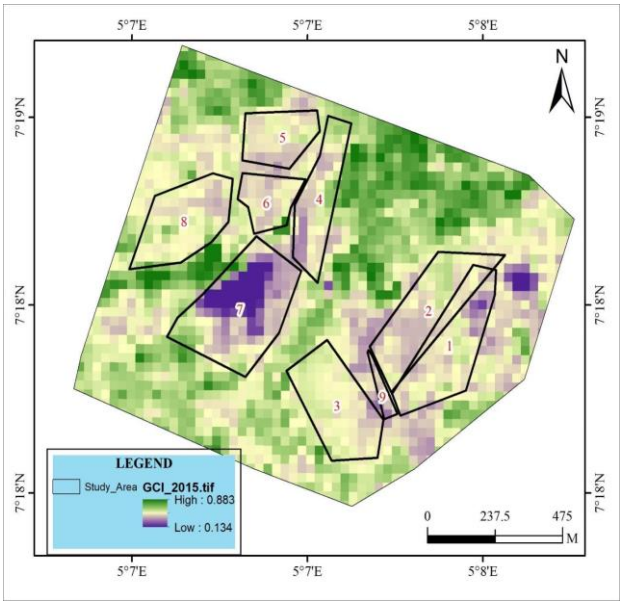

Figure 17: GCI change for the year 2015

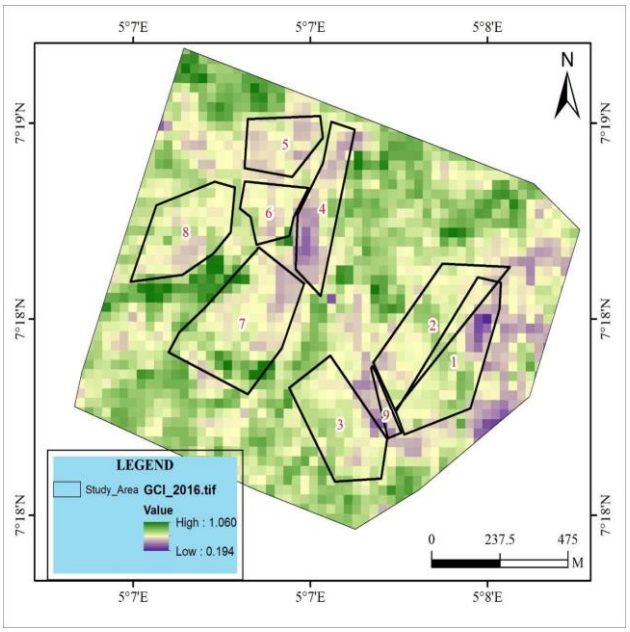

Figure 18: GCI change for the year 2016

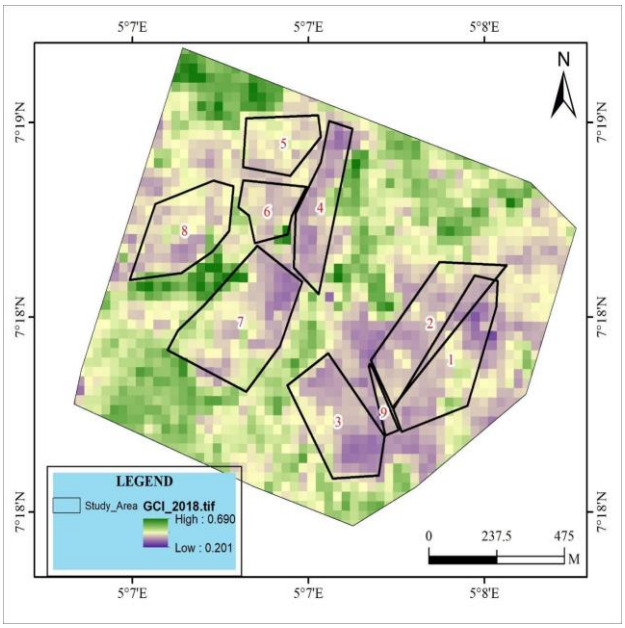

Figure 19: GCI change for the year 2017

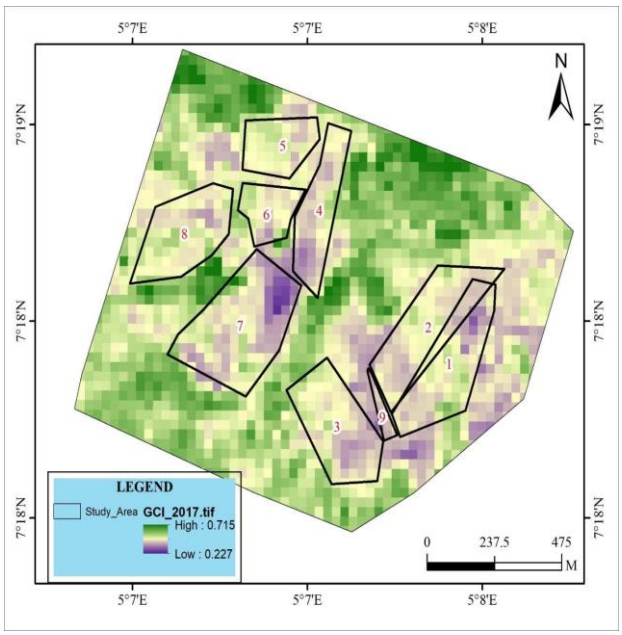

Figure 20: GCI Change for the year 2018

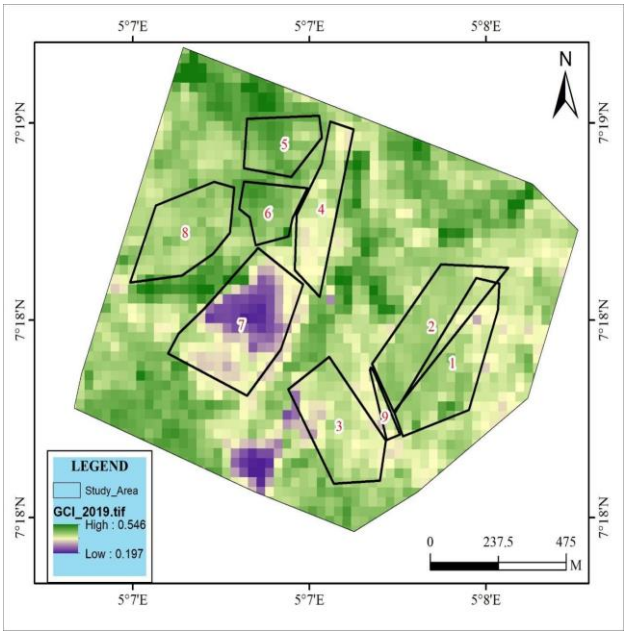

Figure 21: GCI Change for the year 2019

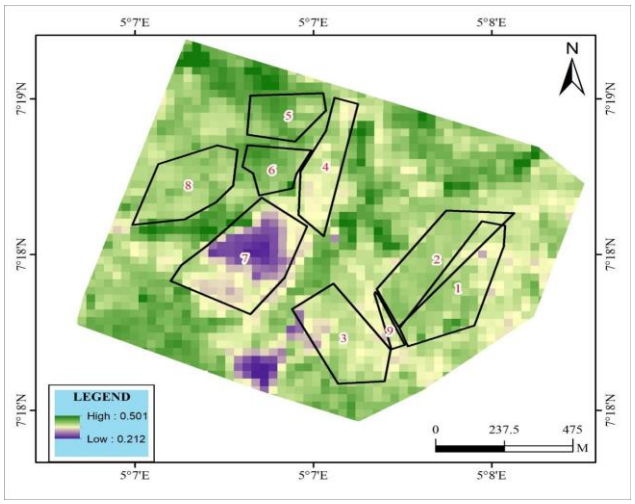

Figure 22: GCI Change for the year 2020

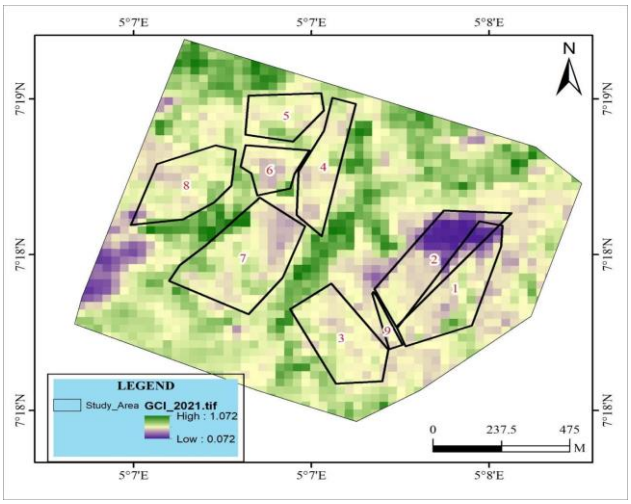

Figure 23: GCI Change for the year 2021

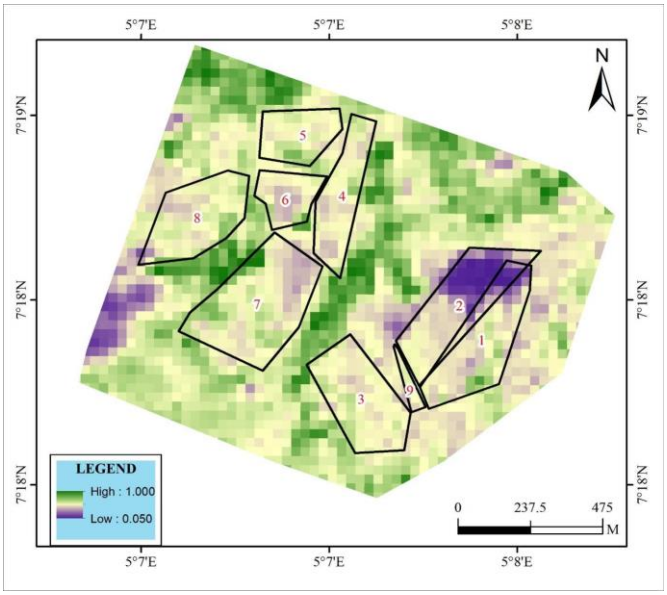

Figure 24: GCI Change for the year 2022

Table 3 presents the Average GCI value of the study area between 2012 and 2022. The highest average NDVI value of 0.63 was recorded in 2016 while the lowest average GCI value of 0.32 was also recorded in 2020.

Table 3: Average GCI value of the study area

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
GCI	0.49	0.50	0.48	0.51	0.63	0.47	0.45	0.37	0.32	0.57	0.56

Figure 25 shows that plant health declined from 2017 to 2020 due to changes in chlorophyll content. Land utilization rate is believed to contribute to this decline. However, when land was given time to rest and regain nutrients due to COVID-19, it improved health. The continuous plantation of crops is believed to be responsible for the deterioration in chlorophyll content, affecting crop health and yield. Other factors may also be involved.

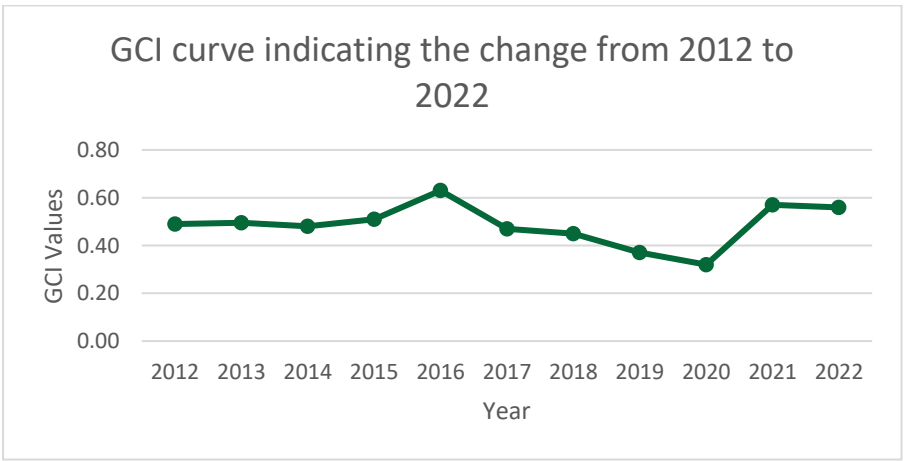

Figure 25: GCI Curve showing the changes in the Chlorophyll content of the crop planted within the study area between 2012 and 2022.

3.2 Average Temperature

Table 4 presents the Average Temperature of the study area between year 2012 and 2022. The highest average temperature of 25.13°C occurred in 2016 while the lowest average temperature of 24.08°C occurs in 2022.

Table 4: Average Temperature of the study area

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Average Temperature	24.26	24.35	24.72	24.74	25.13	24.80	24.48	24.64	24.44	24.67	24.08

Figure 26 shows the average temperature of the study area from 2012 to 2022. The figure 4.27 shows annual fluctuations in temperature, with some years experiencing slightly higher or lower temperatures compared to the adjacent years. For example, 2016 has the highest temperature, while 2022 has the lowest. Overall, the temperatures appear relatively consistent, with most values clustering around the mid 24°C range. There are no extreme outliers or drastic fluctuations, suggesting a degree of stability in the recorded temperatures over the years.

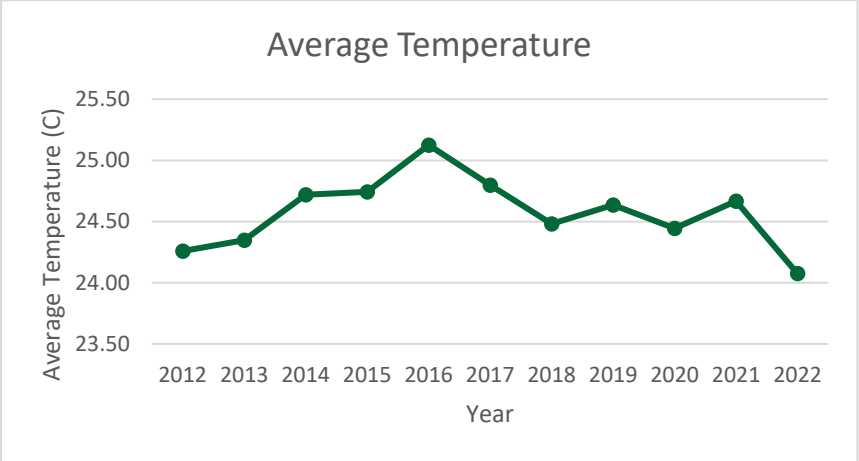

Figure 26: Average temperature of the study area from 2012 to 2022

3.3 Average Rainfall

Table 5 presents the Average Rainfall of the study area between year 2012 and 2022. The highest average Rainfall of 7.56mm occurred in 2021 while the lowest average temperature of 3.07mm occurs in 2013 and 2015.

Table 5: Average Rainfall of the study area

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Average Rainfall (mm)	3.95	3.07	3.51	3.07	3.95	3.96	4.83	5.27	4.39	7.56	5.78

Figure 27 shows the average rainfall of the study area from 2012 to 2022. The average rainfall values vary from a low value of 3.07 mm in 2013 and 2015 to a high value of 7.56 mm in 2021. Figure 27 indicates annual fluctuations in rainfall, with some years experiencing higher or lower rainfall compared to the adjacent years. The year 2021 stands out with significantly higher rainfall compared to the other years, suggesting a potential anomaly or specific weather event during that year. Overall, there seems to be an increasing trend in rainfall from 2012 to 2022, with a noticeable rise in the latter years, particularly in 2019, 2021, and 2022. While there is variability in rainfall, the data doesn't exhibit extreme outliers or drastic fluctuations, indicating a degree of consistency in the recorded rainfall over the years.

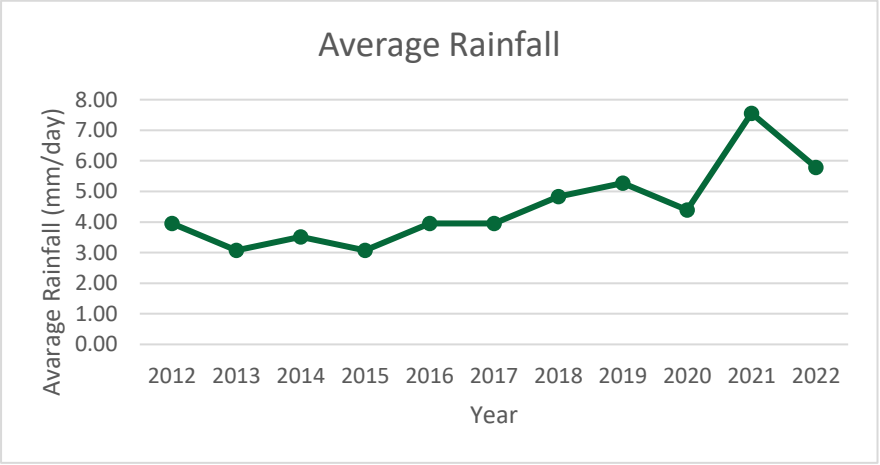

Figure 27: Average rainfall of the study area from 2012 to 2022

3.4 Slope

The value of the slope of the study area is classified into 3 as shown in table 4.6. The three classes are Low, Medium and High with a range of values of 0 - 5.37^0, 5.38^0 - 15.83^{0}and 15.84^0 - 72.06^0 respectively.

Table 6: Slope of the study area

Label	Low	Medium	Medium
Slope Value	0 - 5.37^0	5.38^0 - 15.83^0	15.84^0 - 72.06^0

The slope of the area represented in Figure 28 shows the topographical representation of the study area. It can be observed that the majority of the study area falls within the low to medium slope. This represent that majority of the crop planted within this area have a possibility of being affected by erosion. Furthermore, runoff, erosion and flashflood are some of the effects susceptible to crops within this area.

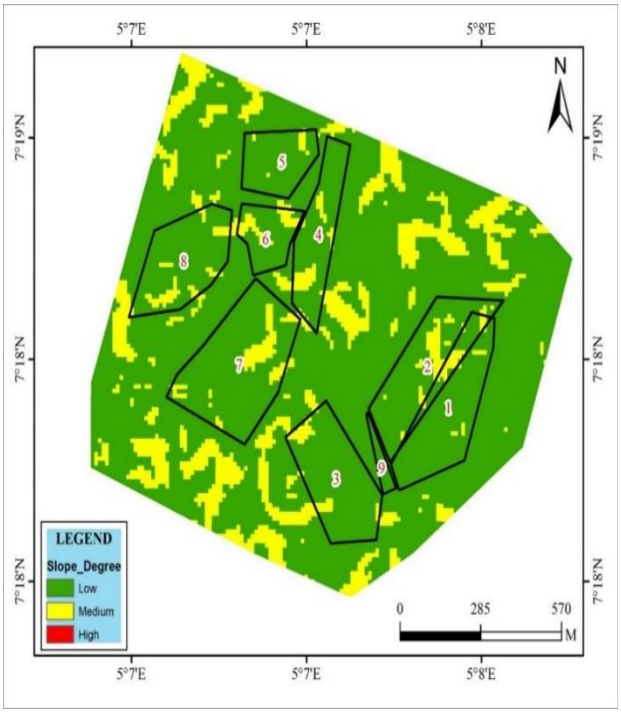

Figure 28: Slope of the study area

The aspect within the study area represented in Figure 29 demonstrated the direction of the slope within the study area. It was observed that majority of the slope within this area are titling towards the Northwest region of the study area. This further suggest that crops planted within this area are at great risk of being affected by flood, runoff or erosion during a heavy down pour of rainfall.

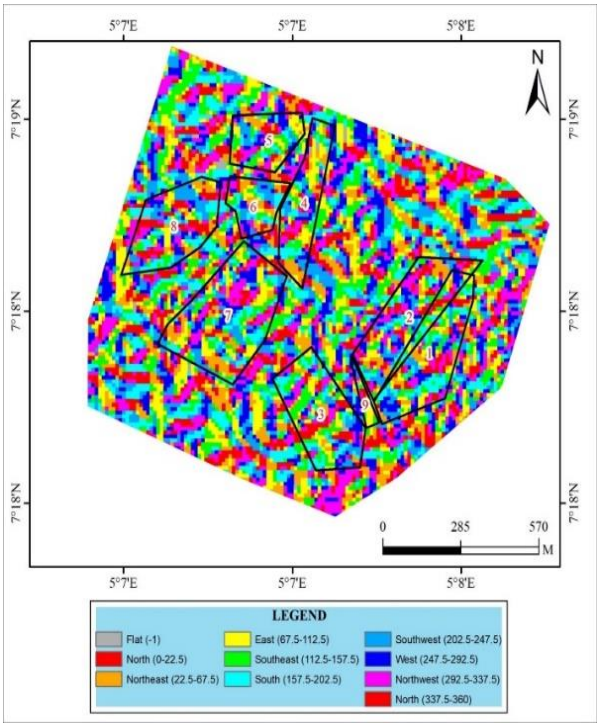

Figure 29: Aspect of the study area

The crop yield record collected from the CSP, department, of FUTA together with climatic data, NDVI and GCI values formed the dataset for artificial Neural Network (ANN) which was used for modelling Crop Yield for the study area. The overall Regression coefficient R for the trained ANN model is 0.73282. The value for coefficient of determination R^2 as computed by MATLAB is 0.5176.

The trained Crop Yield model was used in predicting the crop yield values of points with known input parameters (**Maize Seed (Kg), NPK (bags), Urea (bags), GCI, NDVI, Average Temperature, Precipitation (mm)**). The results of the prediction are shown in table 4.3. From Table 7, the maximum crop yield value for both the actual crop yield and Modelled Crop is 24.7 and 25.26 respectively which was both experienced in the year 2016. The minimum crop yield value for both the actual crop yield and Modelled Crop is 5 and 4.01 respectively which was both experienced in the year 2018. All Modelled crop yield values had a close difference with the actual yield values expect for year 2017 which showed a large difference. The actual value in 2017 ought not be the 7 tonnes as given by CSP department but the reason for the low Crop Yield that year was because of the invasion of the herdsmen into the school farm. If not for this reason, the expected yield should be around 16 tonnes as predicted by the model.

Table 7: Result of Actual Crop Yield and Modelled Crop Yield

Year	Actual Crop Yield (Tonnes)	Modelled Crop Yield (Tonnes)	Differences in Crop Yield Values
2014	18.7	19.38	0.684131
2015	13.8	12.35	1.451557
2016	24.7	25.26	-0.55685
2017	7	16.89	-9.88371
2018	5	4.01	0.985893
2019	14.8	15.10	-0.30937
2021	14.4	14.36	0.034469

The Actual Crop Yield values were compared with Modelled Crop Yield values to see the discrepancies between the two values. The statistics of the differences in Crop yield values is shown in table 8. It was observed from the that the minimum value of difference is -9.883708757 while the maximum value of difference is 1.451557122. The root means square error (RMSE) and the Mean Absolute Error (MAE) are 0.4296 and 0.2947, respectively. The value for the coefficient of determination R^2 as computed by MATLAB is 0.5176

Table 8: Statistics of the difference in Crop Yield Values

Statistics	Max.	Min.	Mean	RMSE	MAE
Differences in Crop Yield Values	1.451557122	-9.88370876	-1.08484004	0.4296	0.2947

Figure 30 shows the graphical representation of the actual crop values and modelled crop yield values for each planting years. The figure shows that there is no much difference in the crop yield values for years 2014, 2015, 2016, 2018, 2019, 2021. Year 2017 shows a large difference in crop yield values owing to the invasion of herdsmen into the school farm during planting season.

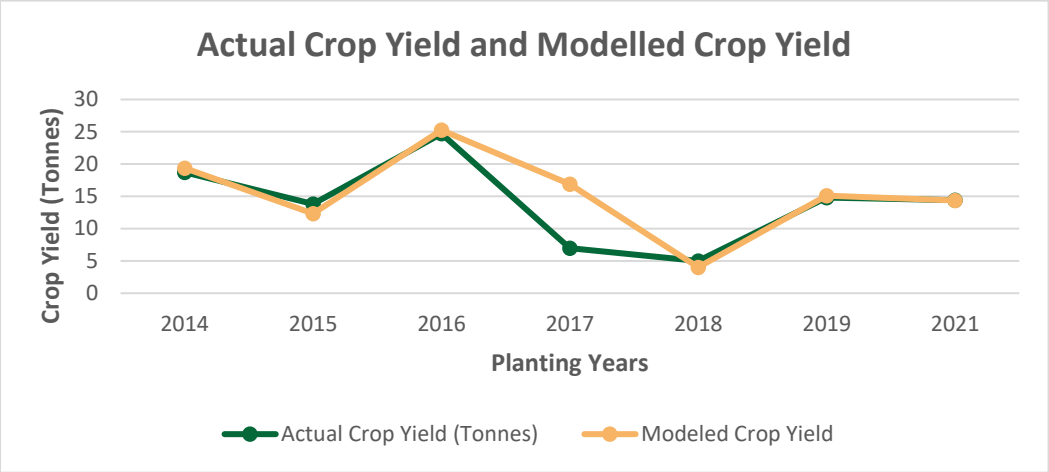

Figure 30: Showing the Actual Crop Yield and the Modelled Crop Yield

4. Conclusion

This research utilized field surveys to gather data for geospatial crop yield modeling, including crop types, planting dates, harvesting dates, management practices, and yield measurements from FUTA farm plots. GNSS observation was used to determine planting boundaries, while secondary data was obtained from satellite imagery and climate records. Landsat 8 satellite imagery captured crop growth dynamics throughout the growing seasons, while climate databases provided data on temperature and precipitation. The study used Artificial Neural Networks (ANN) to model crop yield in the study area. The dataset included crop yield records from the CSP department of FUTA, climatic data, NDVI, and GCI values. The ANN model had a regression coefficient of 0.73282, and the coefficient of determination was 0.5176. The maximum and minimum crop yield values were 24.7 and 25.26 in 2016, and 5 and 4.01 in 2018, respectively. It was observed from that the minimum value of difference is -9.883708757 while the maximum value of difference is 1.451557122. The Root Means Square Error (RMSE) and the Mean Absolute Error (MAE) are 0.4296 and 0.2947, respectively. Modelled crop yield values were close to actual yield values, except for 2017 when a large difference was observed due to herdsmen invasion into the school farm.

References

Balogun, E.T., Abdulla, A., Ajeyomi, A.S., Zullyadini, A.R., Ologun, E.A., Mahir, S., Bushra, M.D., Muhammad, T.R., Olarewaju, T.P., and Olamiju, O.A., 2023. Monitoring and predicting the influences of land use/land cover change on cropland characteristics and drought severity using remote sensing techniques, Environmental and Sustainability Indicators, 18, 100248, ISSN 2665-9727, https://doi.org/10.1016/j.indic.2023.100248.

Bendre, M.R., Thool, R.C., and Thool, V.R., 2015. Big Data in Precision Agriculture: Weather Forecasting for Future Farming. In 1st International Conference on Next Generation Computing Technologies, Pp. 744-750. doi:10.1109/NGCT.2015.7375220

Bolton, D.K., and Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agricultural and Forest Meteorology, 173, Pp. 74–84.

Bose, P., Kasabov, N.K., Bruzzone, L., and Hartono, R.N., 2016. Spiking neural networksfor crop yield estimation based on spatiotemporal analysis of image time series. IEEE Transactions on Geoscience and Remote Sensing, 54 (11), Pp. 6563–6573. doi:10.1109/TGRS.2016.2586602

Fan, J., Bai, J., Li, Z., Ortiz-Bobea, A., and Gomes, C.P., 2022. A GNN-RNN Approach for Harnessing Geospatial and Temporal Information: Application to Crop Yield Prediction. Proceedings of the AAAI Conference on Artificial Intelligence, 36 (11), Pp. 11873-11881. https://doi.org/10.1609/aaai.v36i11.21444

Food and Agriculture Organization (FAO). 2019. Crop yield forecasting. Retrieved from http://www.fao.org/in-action/inventory-fao-geospatial-tools/advanced-fao-geospatial-tools/crop-yield-forecasting/en/

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ, 25, Pp. 295–309.

Kipf, T.N., and Welling, M., 2016. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.s

Manjula, E., and Djodiltachoumy, S., 2017. A Model for Prediction of Crop Yield. International Journal of Computational Intelligence and Informatics, 6 (4), Pp. 2349–6363.

Ortiz-Bobea, A., Ault, T.R., Carrillo, C.M., Chambers, R.G., and Lobell, D.B., 2021. Anthropogenic climate change has slowed global agricultural productivity growth. Nature Climate Change, 11 (4), Pp. 306–312.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Spec, 351, Pp. 309.

Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D.R., Sidike, P., and Fritschi, F.B., 2021. Field-scale crop yield prediction using multi-temporal WorldView-3 and Planet Scope satellite data and deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 174, Pp. 265–281. doi: 10.1016/j.isprsjprs.2021.02.008

Satorras, V.G., and Estrach, J.B., 2018. Few-Shot Learning with Graph Neural Networks. In International Conference on Learning Representations.

Sekulić, A., Milan, K., Gerard, B.M., Heuvelink, M.N., and Branislav, B., 2020. Random Forest Spatial Interpolation. Remote Sensing, 12 (10), Pp. 1687. https://doi.org/10.3390/rs12101687

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific reports, 11 (1), Pp. 1–15.

Tiwari, P., and Shukla, P.K., 2019. A Review on Various Features and Techniques of Crop Yield Prediction Using Geo-Spatial Data. International Journal of Organizational and Collective Intelligence, 9 (1), Pp. 37–50. doi:10.4018/ijoci.2019010103

Wang, H., Huo, Z., Zhou, G., Wu, L., and Feng, H., 2015. Monitoring and Forecasting Winter Wheat Freeze Injury and Yield from Multi-Temporal Remotely Sensed Data. Intelligent Automation & Soft Computing, 22 (2), Pp. 255–260. doi:10.1080/10798587.2015.1095475

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W., 2014. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crops Research, 164, Pp. 178–188. doi: 10.1016/j.fcr.2014.05.001

Zhou, X., Zheng, H. B., Xu, X. Q., He, J.Y., Ge, X.K., Yao, X., and Tian, Y.C., 2017. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 130, Pp. 246–255. doi:10.1016/j.isprsjprs.2017.05.003