A Crop Area Mapping Procedure Using High Spatial Resolution Imagery with Hyper-Temporal NDVI Data.

A Case Study of Cocoa in Ghana

PAULINA ANSAA ASANTE
February, 2017

SUPERVISORS:
Dr.ir. C.A.J.M (Kees) de Bie
Ir. L.M. van Leeuwen - de Leeuw (Louise)

A Crop Area Mapping Procedure Using High Spatial Resolution Imagery with Hyper-Temporal NDVI Data.

A Case Study of Cocoa in Ghana

PAULINA ANSAA ASANTE
Enschede, The Netherlands, February, 2017

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-Information Science and Earth Observation.

Specialization: Natural Resources Management

SUPERVISORS:
Dr.ir. C.A.J.M. (Kees) de Bie
Ir. L.M. van Leeuwen - de Leeuw (Louise)

THESIS ASSESSMENT BOARD:
Prof. Andy Nelson (Chair, Department of Natural Resources, ITC, University of Twente,)
Dr. Ir. B.G.J.S Sonneveld (External Examiner, SOW-VU, Centre for World Food Studies)

ABSTRACT

The availability of accurate and timely crop area information is fundamental to any agricultural crop monitoring program. Remote sensing by itself or integrated with GIS and ground surveys is proved to offer unique capabilities to improve crop area information provision. However, it is still not widely operational for mapping crops. This is mainly because the existing procedures rely on remotely sensed data which have either; high spatial resolution but with often low temporal resolution unable to derive crop phenology. Or a high temporal but coarse spatial resolution unable to resolve crop fields. To overcome this trade-off issues the primary objective of this research was to investigate the applicability of integrating these two types of datasets in a procedure to map cocoa area in Ghana.

MODIS hyper-temporal NDVI data and Sentinel-2 high spatial resolution image were acquired in combination with other datasets to produce cocoa area maps based on the existing approaches and integration. For the hyper-temporal NDVI approach procedure, clustered NDVI cocoa classes for Ghana were identified based on NDVI-ecological strata and combined statistically with COCOBOD area statistics. Exact cocoa cropping intensities were estimated using field data for Goaso district. For the high spatial resolution procedure, a single-date Sentinel-2 high resolution image was stratified and linked to field data to specify cocoa cropping intensity within the map classes. Two cocoa maps were created based on these procedures. To integrate, the two cocoa maps were combine and cocoa cropping intensities estimated using simple arithmetic.

To compare the mapping performance of the existing procedures used separately and the integrated, generated cocoa maps were correlated with actual cocoa field polygons. The results of the correlation analysis indicated that, the cocoa map generated using hyper-temporal NDVI alone recorded the lowest with Adjusted R^2 of 25% showing a low relationship and the high spatial resolution recorded an Adjusted R^2 of 68% which indicates a relatively high relationship. The integrated product cocoa map was at middle with an Adjusted R^2 value of 38% which indicates moderate relationship with the actual cocoa polygons. These results indicate that combining hyper-temporal and high spatial resolution data is feasible however the performance of the integrated product depends very much on the ability of the underlying approaches to distinguish crops from other cover types well. The procedure provides a cheap, rapid and efficient method for mapping crops from a small to large scale.

Keywords: Hyper-temporal NDVI, High spatial resolution, crop area, cocoa

ACKNOWLEDGEMENTS

I am really grateful to the NFP for funding my MSc studies and stay in the Netherlands. Also like to appreciate the efforts of the administrative staff of the ITC faculty and Hotel for making all the necessary arrangements for my stay and smooth 18 months' journey.

To the ITC teaching staff, I am thankful for the knowledge and training I have acquired from you as your student. I regard this knowledge and training as a seed sown on fertile soil which will definitely with God's blessings become a fruitful tree.

My supervisors, first to Dr. ir. Kees de Bie, I truly appreciate your committed efforts towards keeping me on track with this whole thesis. I particularly appreciate your patience and motivation in the times I was really stressed with the analysis and your constructive criticism which has made me to be more critical and focused. To Ir. Louise van Leeuwen-de Leeuw, my second supervisor, who suggested and supported vehemently the initial idea of pursuing this particular research and saw through its implementation. You two have been true mentors to me.

This thesis work will not have been successful without the support of Institutions like COCOBOD' CHED, COCOALIFE who provided me with baseline information and data needed for the analysis. I really appreciate particularly the efforts of this individuals. Am grateful to Dr. Francis Baah (Executive Director, CHED), Mr. Eric Bani (Deputy Manager, M&E) for giving attention and direction to getting the necessary assistance from CHED. To Mr. Faruk Kwansah Nyame (Manager, Cartography), Mr. Enock Tettey (Deputy Manager, Cartography), a heartfelt thanks for the prompt and expert assistance rendered.
And to all staff of CHED Cartography Unit not forgetting Kaizer and in Goaso office, George, and Yaaba. To CocoaLife, the very first organization to respond and provide the requested data for analysis, my gratitude.

I honestly cannot put in this limited space the whole list of names I am truly thankful for support and assistance during this past 18 months of MSc work. But I like to acknowledge a few whose contributions has made this thesis a success. To Kingsley Kwako Amoako (MOFA), Lilian Lucy Lartey (COCOBOD), George Owusu (CERSGIS) past students of ITC who really supported in data acquisition from the relevant institutions in Ghana. To Marian (Kenya), John Reuben (Tanzania), Grace Simon (Nigeria), Benjamin Appiah (Ghana), Jefferson Okojie (Nigeria), Emma (Ghana), Ekow (Ghana), Kelvin (Tanzania), Marijana (Serbia), Lucas (Argentina), Fetene (Ethiopia), Yuvenal Mtui (Tanzania), Weicheng (China) my truly wonderful multinational classmates who gave ear and suggestions to my cocoa and hyper-temporal stories. And to Justin Yieri (Ghana) and E.E. Agbor (Thailand) my dear expert thesis editors' thanks for your time and corrections.

To my dear parents whose occupation is the very core motivation for this thesis and siblings I truly acknowledge you for the continual support and prayers. I feel the same towards my family in ICF, Enchede and Grace Center AG, Madina, Accra.

My heart is full of joy and gratitude as I remember how you been an immense blessing to me.

I thank the Almighty God for all of you.

God zegene je! (God bless you!)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer)
AVHRR	Advanced Very High Resolution Radiometer
CERSGIS	Centre for Remote Sensing and Geographic Information Systems
COCOBOD	Ghana Cocoa Board
GIS	Geographic Information Systems
GPS	Global Positioning System
HCA	Hierarchical Cluster Analysis
ISODATA	Iterative Self-Organizing Data Analysis Technique
LP DAAC	Land Process Distributed Active Archive Centre
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
SPOT	Satellite Pour l'Observation de la Terre
TOA	Top of Atmosphere
VIs	Vegetation Indices

1. INTRODUCTION

1.1. Background & Justification

The availability of accurate and timely crop area information is fundamental to any agricultural crop monitoring program. Crop area maps provide information on where what is grown, when and how much, a basic requirement to effective crop monitoring (FAO, 2000; Khan et al., 2010; Thomson, 2003). This kind of information is in increasing demand by Governments, international and local agricultural organizations because it has significant agricultural, economic, policy and environmental implications (Leeuw et al., 2010).

Remote sensing by itself or integrated with Geographic Information Systems (GIS) and ground surveys offer unique capabilities for providing accurate and timely information on crop growing areas (Liu et al., 2014). For over three decades, this technology has proven relevant for effective crop identification and mapping across different scales (Badhwar, 1984; Crapper, 1980). Several studies have demonstrated and successfully implemented remote sensing based procedures to derive crop area maps for single or multiple crop types (Durgun et al, 2016; Foerster et al, 2012; Gumma et al, 2011; Vaudour et al., 2015). Some of the studies however noted challenges in identifying certain crop types like cocoa from remote sensing imagery (Boaitey, 2016; Tesema, 2015) and hence crop area mapping . This was is as a result of observed similarity in spectral reflectance properties between the cocoa trees and that of natural forest trees. Others factors identified have impact on crop identification and area mapping includes; spatial resolution of image and target's field size and spatial heterogeneity within the observed area (Buechel et al, 1989; de Bie et al., 2012)

Generally, in remote sensing based crop area mapping, existing procedures have either used phenology-based classification of hyper-temporal datasets (de Bie et al., 2008; Inglada et al., 2015; Khan et al., 2010; Wardlow & Egbert, 2008) or mono-temporal image classification of medium to high spatial resolution imagery (Conrad et al., 2010; Li et al, 2015; Singha et al, 2016). Studies that presented the phenology-based classification procedures utilizes the phenological features derived from hyper-temporal datasets like MODIS, AVHRR, MERIS and SPOT Vegetation to discriminate and map crops. This method is termed hyper-temporal based crop area mapping in this study.

The other studies, using mono-temporal image classification procedures utilized mainly medium to high spatial resolution data such as Sentinel-2, Landsat-TM/ETM, ASTER, SPOT-5-HRG to provide a more detailed discrimination of crops at the large scale (e.g. field level) (Conrad et al., 2010; Li et al, 2015; Singha et al, 2016). Crops are differentiated based image spectral and spatial properties. This method is termed high spatial resolution crop area mapping in this study.

Very few studies have integrated the use of the hyper-temporal and high spatial based procedures to map crops.

1.2. Remote Sensing Based Crop Area Mapping Procedures

Hyper-Temporal Based Crop Area Mapping

For about four decades, several techniques to map crop area, using hyper-temporal remote sensing data have been developed (Badhwar, 1984; Price et al., 1994; Smith & Ramey, 1982). These hyper-temporal datasets have generally low/medium (coarse) spatial resolution which limits their ability to map smaller size crop fields. However, the temporal dimension of this data is found particularly relevant in crop area mapping because crops typically exhibit a higher temporal than spatial variability (de Bie et al., 2008).

The utility of the temporal dimension over single-date image analyses for crop area mapping has been rising steadily especially in recent remote sensing developments which allows easy access to increasingly high temporal data (de Bie et al.1, 2008; Inglada et al., 2015; Khan et al., 2010; Wardlow & Egbert, 2008)
The main strength of the hyper-temporal procedure is its ability to capture phenological changes in crop reflectance characteristics. This is used to discriminate crops from other cover types based on their distinct temporal signatures derived through the use of Vegetation Indices (VIs) like the Normalized Difference Vegetation Index (NDVI). Here the relationship between remotely sensed image properties and crop characteristics is described with a VI that is able to quantify for instance the distinct green-up and senescence of crop cycle (Duncan et al., 2015).

It is generally proven that VIs has a correlation with several biophysical variables in crops, like the greenness, photosynthetic activity, canopy structure, biomass and productivity. This makes it a useful indicator for identifying crops, mapping and monitoring when retrieved from hyper-temporal datasets (Duncan et al., 2015; Huete et al., 2002; Pettorelli et al., 2005; Tucker, 1979).

One of the widely used VIs for identifying and mapping crop area is the NDVI derived by dividing the difference between infrared and red reflectance measurements by their sum (de Bie et al., 2008). This index is usually assumed to be broadly indicative of active crop photosynthetic activity (Sarkar & Kafatos, 2004) and therefore associated with the amount of greenness and above-ground dry matter production (Goward & Huemmrich, 1992). Several studies have used temporal NDVI profiles to discriminate croplands from non-croplands and mapped crop specific area and intensity (Asilo et al., 2014; Gumma et al., 2011; Khan et al., 2010; Nguyen et al., 2012; Roumenina et al., 2015).

High Spatial Resolution Based Crop Area Mapping

The use of high spatial resolution procedure is best suited for crop mapping at the field scale (Conrad et al., 2010). Here, crops are identified and distinguished based on the spectral reflectance and spatial resolution properties of the image used. This crop mapping procedure is found to be more useful in mapping crops in regions where small-scale farming is predominant and in areas with isolated patches of crops (Ozdogan et al., 2010). It was however noted, that the ability of a high spatial image to accurately resolve and map crops, requires that; the spatial resolution is equal to or smaller than the size of the target fields (Ozdogan & Woodcock, 2006).

In this procedure, often single or multi-date high spatial resolution data is used to produce the high resolution maps (Jansen & Di Gregorio, 2003). The utility of this method has been demonstrated in several studies(Conrad et al., 2014; Jansen & Di Gregorio, 2003; Vaudour et al., 2015) and have been found to provide a much better account for small crop field sizes than using low/moderate spatial resolution data. However, details are often limited to local scales and its usability for wide area mapping requires large volume image scenes which is relatively expensive and time consuming. Also, this high spatial resolution data is often not available for all areas around the globe

1.3. Problem Statement

Despite the long history and potential of utilizing remote sensing techniques for the provision of accurate and up-to-date crop area maps, it is still not widely operational for mapping crops (Liu et al., 2014) like cocoa. This is mainly because; the existing remote sensing based procedures, rely largely on remotely sensed data which have either; high spatial resolution but often with temporal resolution too low to derive crop phenology. Or with a high temporal frequency but low/medium spatial resolution (Durgun et al., 2016) unable to resolve crop fields.

In many tropical countries like Ghana, the hallmark of agricultural landscapes is extreme complex mixture of land covers (Estrada et al., 1994). In addition to relatively small crop fields (Ozdogan & Woodcock, 2006). In Ghana, average field size of the country's most important cash crop, cocoa is 2.5ha (MOFA, 2011). The crop is grown mostly under the shade of forest trees and at times mixed with food crops (Asare, 2005). Mapping the crop area in such a case using high spatial resolution procedure, requires images of medium to very high spatial resolution (50cm-30m pixel size) (Ozdogan & Woodcock, 2006). However, previous attempts at mapping cocoa using remote sensing utilized very high spatial resolution (50cm) data but reported difficulties in separating cocoa trees from that of natural forest due to observed similarities on the image (Boaitey, 2016; Tesema, 2015). This particular challenge affirms the unsuitability of mono-temporal, multispectral data only for crop classification (Conrad et al., 2010) despite the high spatial property.

Besides, high spatial resolution data often comes at a cost of temporal availability. This is mainly due to predetermined acquisition strategies and often clouds obstruction (Liu et al., 2014). Which then affects image availability. Also, this type of data is not available for all areas around the globe and often comes with relatively high cost in procurement.

An alternative then, is to use the hyper-temporal NDVI approach which uses low/medium spatial resolution data (often greater than 100m pixel size) with high temporal frequency. This type of data is able to retrieve crop phenological characteristics captured by temporal NDVI profiles and use to classify the crop. The data is also freely and continuously available online for all areas. In the case of cocoa, seasonal leaves flushing properties could be captured by the hyper-temporal NDVI datasets and used to distinguish from forest trees. However, the coarse spatial resolution of this type of data limits it's the ability to capture details at the 2.5ha field. The best approach therefore would be to combine the relative strengths of each method to improve its usage for mapping the crop.

To overcome the trade-off issues between the two procedures, this research attempts an improvement to mapping cocoa, by integrating hyper-temporal NDVI and high spatial resolution procedures to map the crop area(s) in Ghana.

1.4. Research Objective

The main objective of this research is to investigate the applicability of using hyper-temporal NDVI data from Moderate Resolution Imaging Spectroradiometer (250m-MODIS) with high spatial resolution data from Sentinel-2 (10-m) in a procedure to map cocoa area in Ghana. Specifically;

1. A. To identify and distinguish from hyper-temporal NDVI data, relevant ecological strata.
   B. Match these ecological strata statistically with cocoa area statistics as reported by the Ghana Cocoa Board (COCOBOD) to derive NDVI-based Cocoa Strata covering Ghana.

2. To further specify for the NDVI-based Cocoa Strata within Goaso, the exact cocoa cropping intensity through collected field data.

3. To stratify for Goaso high spatial resolution data from Sentinel-2 and specify the exact cocoa cropping intensity of these strata through collected field data.

4. To combine the approaches of 1,2 and 3 and derive an integrated product

5. To establish and compare the accuracies of;
   i. Hyper-temporal NDVI-based cocoa map of Goaso
   ii. High spatial resolution based cocoa map (baseline-method) of Goaso
   iii. Integrated product of Goaso

1.5. Research Questions

Which of the cocoa maps have the highest accuracy?

i. Hyper-temporal NDVI-based Cocoa map of Goaso?
ii. High spatial resolution based cocoa map (baseline-method) of Goaso?
iii. The integrated product of Goaso?

1.6. Research Hypothesis

The Adjusted R^2 of the cocoa map produced by the integrated approach explains 10% more variability of the crop area derived from actual cocoa fields polygons than the cocoa maps of the underlying approach used separately.

2. MATERIALS AND METHOD

2.1. Materials

2.1.1. Study Area

The study area comprised of Ghana cocoa districts with boundaries designated by the Ghana Cocoa Board (COCOBOD). Currently, 61 cocoa district's covering all of Ghana's major cocoa growing areas have been defined by the COCOBOD. However, the old district demarcations which consist of 41 cocoa districts was used in this research because of data availability.

For the hyper-temporal based cocoa area mapping, 38 out of the 41 cocoa districts in Ghana were used. The selection was based on the coverage extent of the MODIS tile H17v08 acquired and used for analysis. The cocoa districts are located in the south of Ghana and cuts across six administrative regions; Ashanti, Brong Ahafo, Central, Eastern, Western and Volta regions.

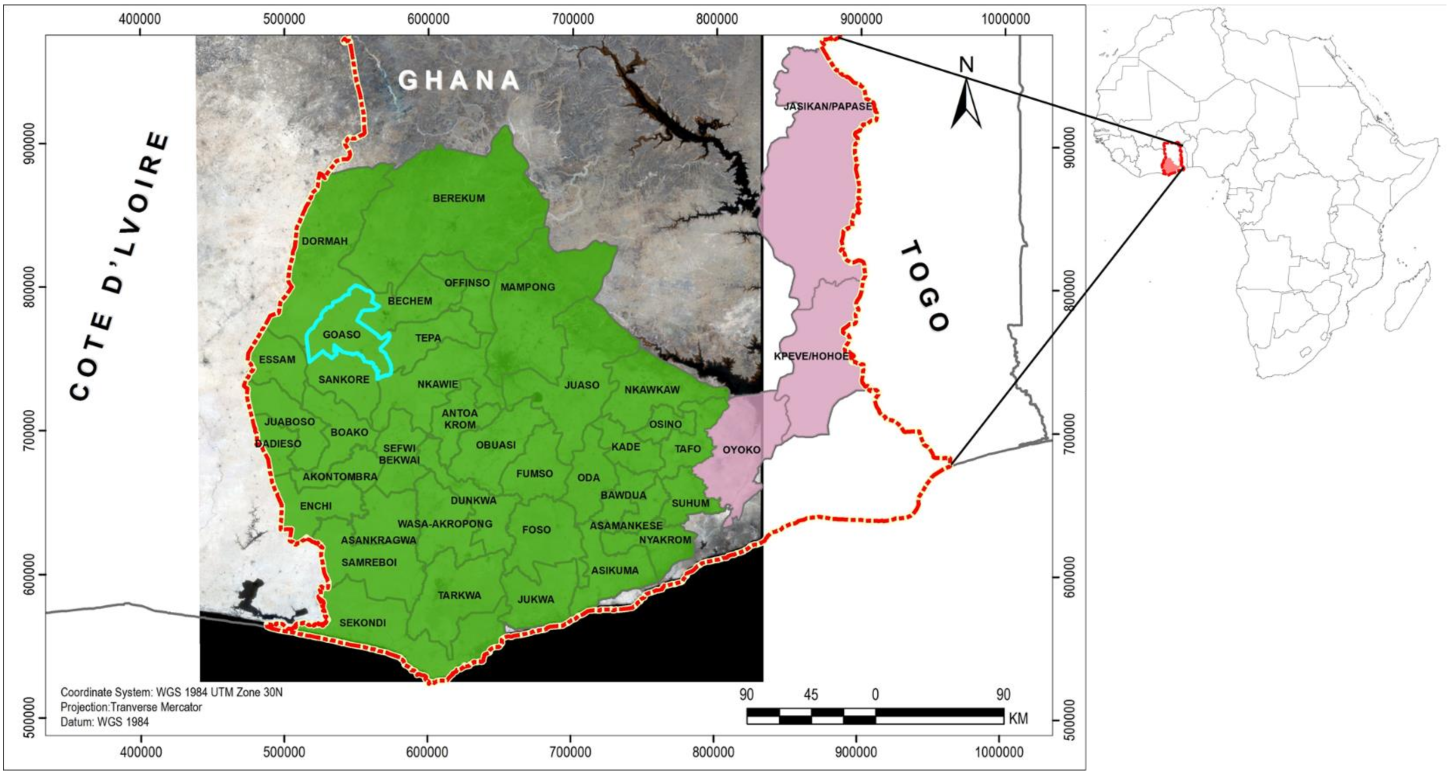


Figure 1: Map showing Cocoa Districts of Ghana with the MODIS tile h17v08 image in background and Goaso District highlighted as Area of Interest.

For the purpose of the high spatial and integrated approach cocoa area mapping, the Goaso district, located in the Brong Ahafo region of Ghana was selected out of the 38 districts. The coverage extent of the acquired Sentinel-2 image for the area, determined the area observed in the mapping process. Figure 1 shows a map of the study area and area of interest.

Climate & Soil Characteristics

The cocoa growing areas in Ghana cuts across three major agro-ecological zones including rain forest, deciduous forest and transitional zones all located within the southern belt of the country (Asante-Poku & Angelucci, 2013). These areas receive a mean annual rainfall of 2200mm in rain forest, 1500mm and 1300mm in the deciduous and transitional zones respectively. In these zones, rainfall distribution is bimodal giving a major and minor growing seasons (Ghana meteorological Services Department, Accra as cited in MOFA, 2011). Average temperature is 25^{0}C and Relative humidity is about 70-80% (Opoku-Ameyaw et al., 2010).

Major soil types found in the area are the Forest Ochrosol and Forest Ochrosol-Oxysol integrade (ME) great soil group of the Ghanaian soil classification system (Brammar, 1962 as cited in Asare, 2015) but generally classified as Acrisols in the FAO-UNESCO Revised Legend (FAO, 1988). These soil types are noted for their well-drained properties.

Cocoa Economy & Botany

Cocoa, *Theobroma cacao*, ranks third as most important beverage crop after tea and coffee globally (Lartey, 2013) and Ghana is the world second major producer of the crop. The total planted area of cocoa in the country is estimated at 1.6 million hectares with over 90% grown on small farms (Opoku-Ameyaw et al., 2010). The average field size of a cocoa farm in Ghana is estimated at 2.5 ha with a mean yield capacity of about 0.4MT/ha(MOFA, 2011; Opoku-Ameyaw et al., 2010).

The country's cocoa bean production statistics for past five years is presented in Figure 2 among other world producing countries of the crop.

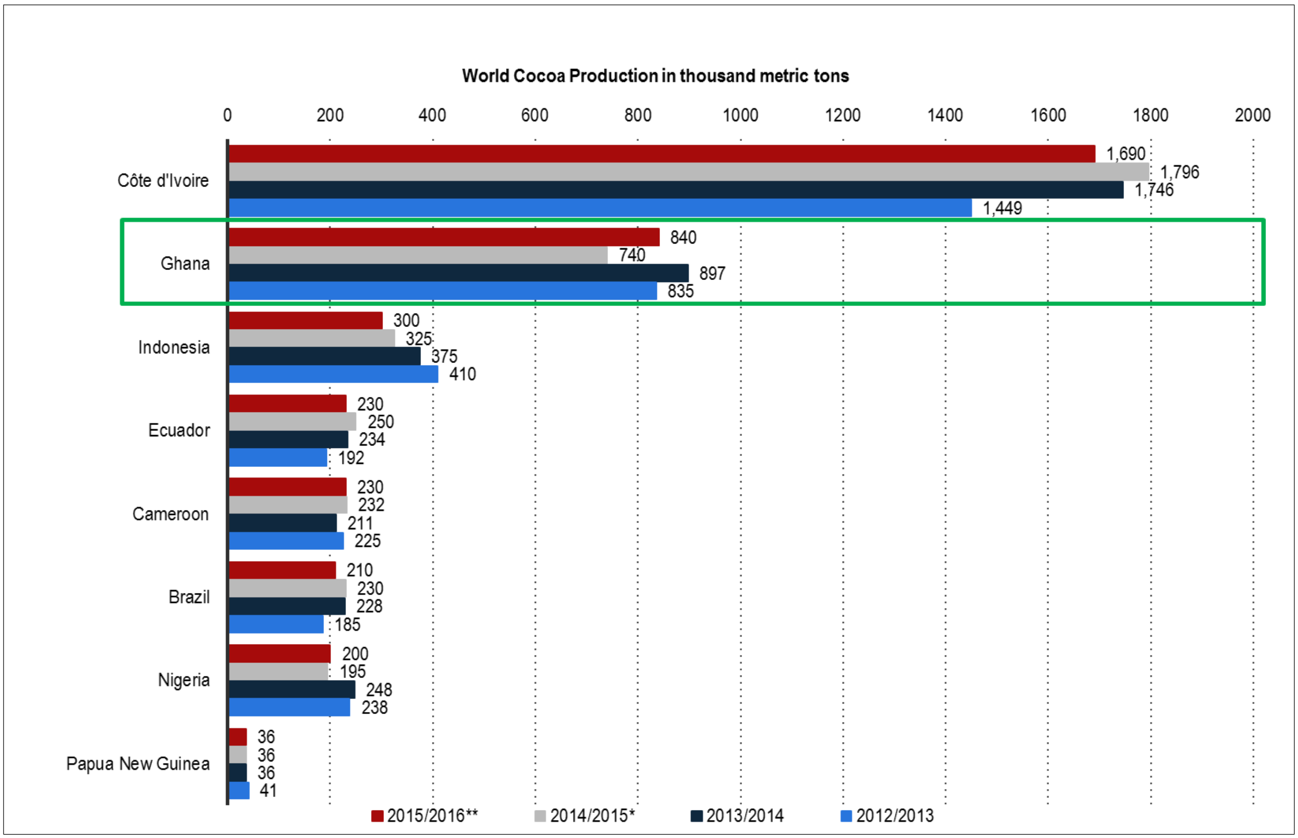


Figure 2: World cocoa bean production from 2012 to 2016 (in 1,000 metric tons). (ICCO, 2016)

Cocoa is grown in humid lowland tropics often on farms with a canopy of shade trees (Rice & Greenberg, 2000). In Ghana, it is cultivated traditionally under partially cleared forest with remaining trees often mixed with food crops to serve as shade to the cocoa trees (Anglaaere et al, 2011; Asare, 2005). Cocoa trees regardless of varieties grown require some initial shade for growth and are sensitive to drought (Carr & Lockwood, 2011).

The main varieties of cocoa trees grown in Ghana are, Amelonado (local name 'Tetteh Quashie') belonging to the Forastero group of cocoa, Amazonia and hybrid (local name 'akokora bedi') (Lartey, 2013; Opoku-Ameyaw et al., 2010; Asare, 2015). The Amelonado and Amazons, traditional types, requires shade throughout the crop growing and production cycle. The hybrid variety, which is relatively early-bearing, high-yielding and currently the main cocoa planting material are predominately sun-tolerant and so requires less to no shade (Acheampong et al., 2014).

The tree crop grows vertically upward (orthotropic growth) exhibiting cycles of leaf flushings (Vogel, 1975 as cited in Almeida & Valle, 2007). Opoku-Ameyaw et al. (2010) notes that the period between March-April and September-October are leaves flushing period in matured cocoa trees of West Africa. After achieving approximately 1 to 1.2m in height, it ceases orthotropic growth. The plant then forms about three to five branches that grow out at an angle as fan branches. These branches form crown of the cocoa tree (Cuatrecasas, 1964 as cited in Almeida & Valle, 2007).

2.1.2. Data Used

In this research data from primary and secondary sources were used. The primary data include, field data collected using Global Positioning System (GPS) and field observations. The secondary data include, Google Earth, MODIS NDVI data, Sentinel-2 Data, COCOBOD Statistics, Cocoa fields polygons, Ghana Cocoa districts and ancillary spatial data.

Field Data

Field data was collected between 21th September to 20th October in the Goaso cocoa district, Ghana. According to the research requirements, the main agricultural vegetation cover types, especially cocoa fields locations were collected using the stratified purposive sampling method. This sampling strategy was preferred to random sampling because the occurrence of cocoa fields in the study area are relatively uneven with fields size being main factor.

To get a representative sample, cocoa fields and other homogeneous cover types that are at least 1 hectare big (i.e. relatively the size of one Sentinel-2 pixel) in area size was considered. This was to allow comparison with information captured on a 10m*10m pixel of Sentinel-2 image.

By this sampling method, non-agricultural areas such as forest reserve and built up areas were identified and excluded by visual interpretation of Google Earth imagery from the area where samples were to be collected. The cocoa fields could not be distinguished from the Google earth images (see figure 3) received for the area. And so areas with possibility of finding major cocoa cover were located with the help of Cocoa district officials and local residents.

The selection of samples sites and determination of sample size was largely affected by the time available for the field data collection and proximity to roads for accessibility reasons. Samples were however taken across the study area as shown in figure 4.

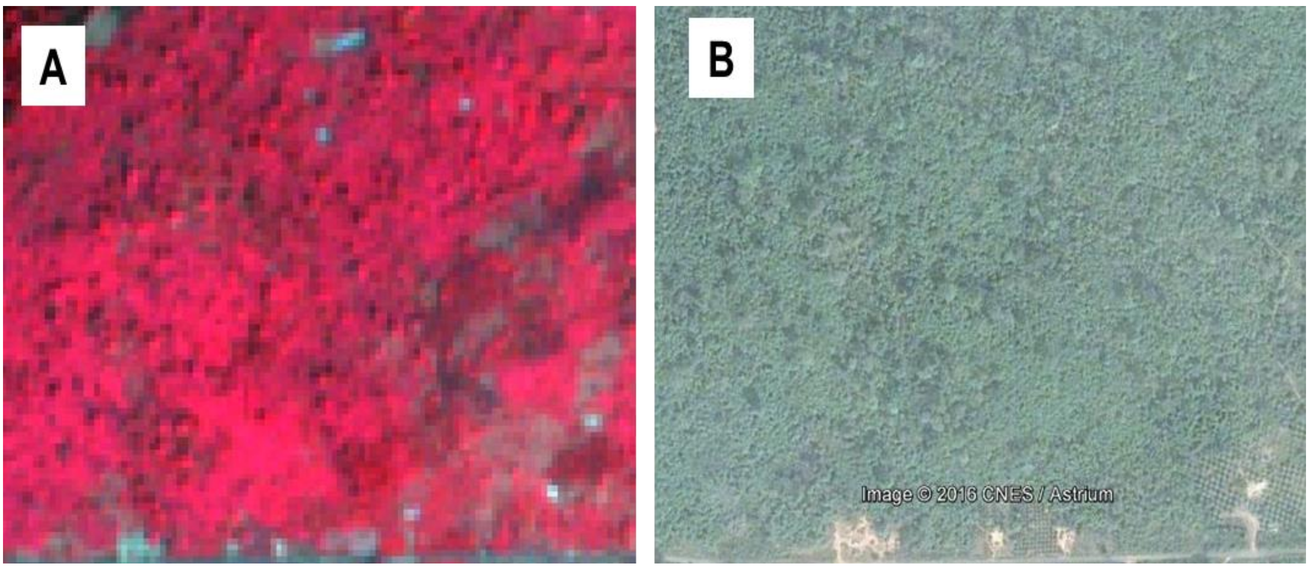


Figure 3: A cocoa field view from a Sentinel image (A) and Google earth (B) image.

In total 231 locations (sample size) were visited, out of which 127 were cocoa fields and 104 were non-cocoa cover types. GPS coordinates of the location were recorded for each site visited and 108 field photos geotagged to some of the locations visited taken. Figure 4 & 5 shows a spatial distribution of samples locations and some field photos respectively.

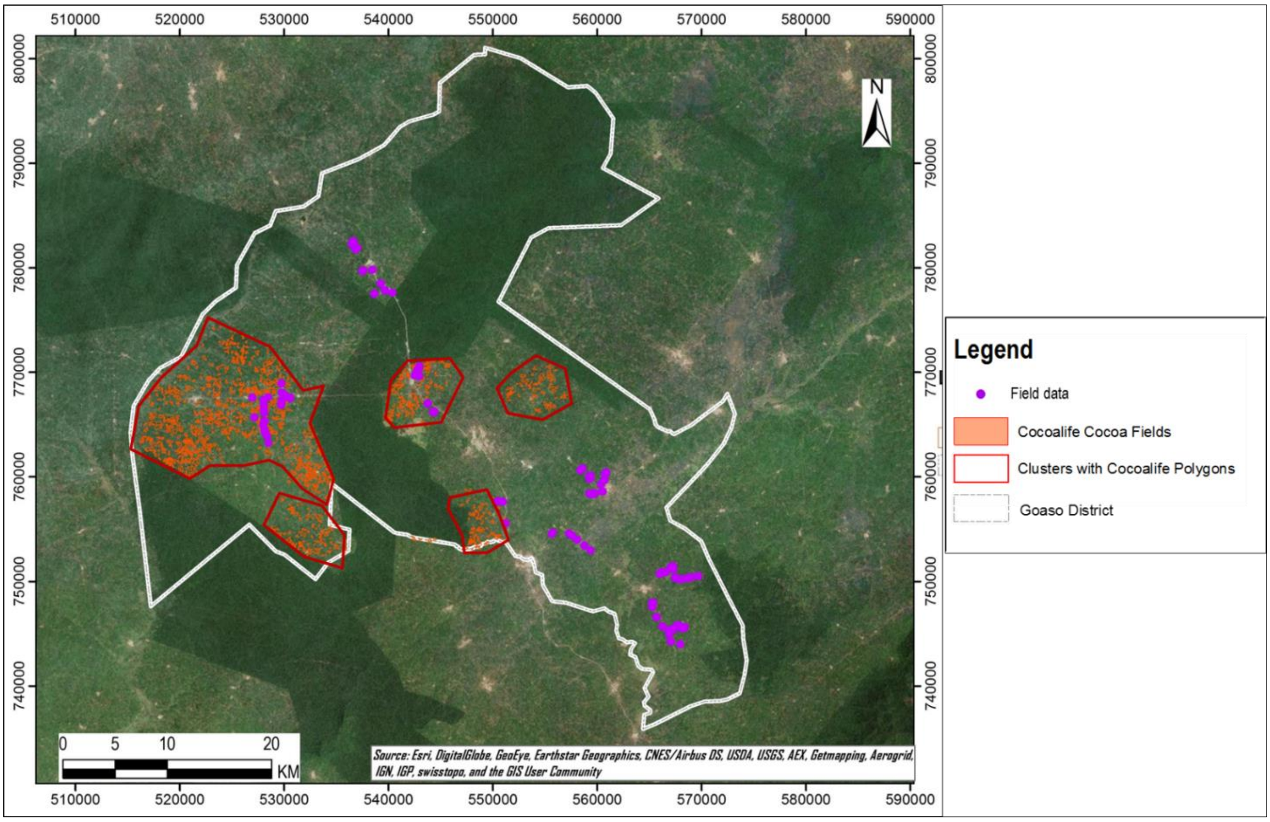


Figure 4: Spatial distribution of field data and CocoaLife cocoa polygons grouped into five clusters in Goaso

Figure 5: Pictures of some sample sites in Goaso, cocoa fields with shade trees (middle) & field boundary trees (right) (Photos taken during fieldwork).

Also showing in figure 4 are CocoaLife cocoa polygons grouped into five clusters based on proximity. This data is collected independently by the Centre for Remote Sensing and Geographic Information Systems (CERSGIS). More details of this particular data is presented in the section under cocoa fields polygons data

MODIS NDVI Data

MODIS is a sensor aboard the Terra and Aqua satellites launched by National Aeronautics and Space Administration (NASA) in 1999 (NASA, n.d.). Data from this sensor is supplied in different products format distributed through platforms like the NASA Land Process Distributed Active Archive Center (LP DAAC) facility via http://reverb.echo.nasa.gov/ , from where the NDVI data product for this research was retrieved.

The NDVI data is a sub dataset of the MODIS Terra MOD13Q1 and Aqua MYD13Q1 products acquired as a time series from 24, December 2010 to 7 January, 2016. This five-year period of data was taken for this research to ensure continuous and long-term identification and mapping of the different land cover types based on their unique temporal NDVI profiles. A total of 231 images were available within the specified time. The products were available as 16-day L3 Global 250-m SIN Grid (Sinusoidal) tiles (tile h17v08 was used) and corrected for effects of atmospheric gases and aerosols by data provider (Didan, 2015a, 2015b). That is each pixel of the NDVI composite image is the best available pixel values from all the acquisitions from the 16-day period, with low view angle and minimum amount of clouds, haze or cloud shadow and aerosol loading (Solano et al., 2010). The NDVI sub dataset was retrieved and used for analysis.

Sentinel 2 Data

The most recent and less cloud contaminated Sentinel 2 image from European Space Agency (ESA) over the area of interest, Goaso cocoa district was obtained from https://scihub.copernicus.eu/ website. The available image dated 26, January, 2016, a Level-1C, Top of Atmosphere (TOA) reflectance product was downloaded and used in this research. This product is corrected for radiometric, geometric errors and orthorectified to highly accurate geolocated products by the data provider (ESA, 2015). Therefore, no further radiometric or geometric pre-processing was applied; it is however noted that clouds are visible in the study site scene (Figure 1, Goaso Scene).

COCOBOD Statistics

Cocoa area statistics by district was collected from the office of Cocoa Health and Extension Division (CHED) of the Ghana Cocoa Board (COCOBOD) in a tabular format (see Appendix 1). This office unit is in charge of conducting land surveys and compiling cocoa area statistics for the country. The statistics received was available as cumulative area survey results. The surveys are still ongoing for some of cocoa districts and as such incomplete. The percentage of land surveyed within each district and the estimated percentage of cocoa area are shown in figure 6 and 7 respectively.

Since statistics were incomplete for some districts, extrapolations were made based on the area surveyed and relative cocoa area for such districts. Extrapolated cocoa area within the district was calculated as;

Equation 1: Total Cocoa area in District (Ha) = Cocoa area in Surveyed (%) * District Area (ha)

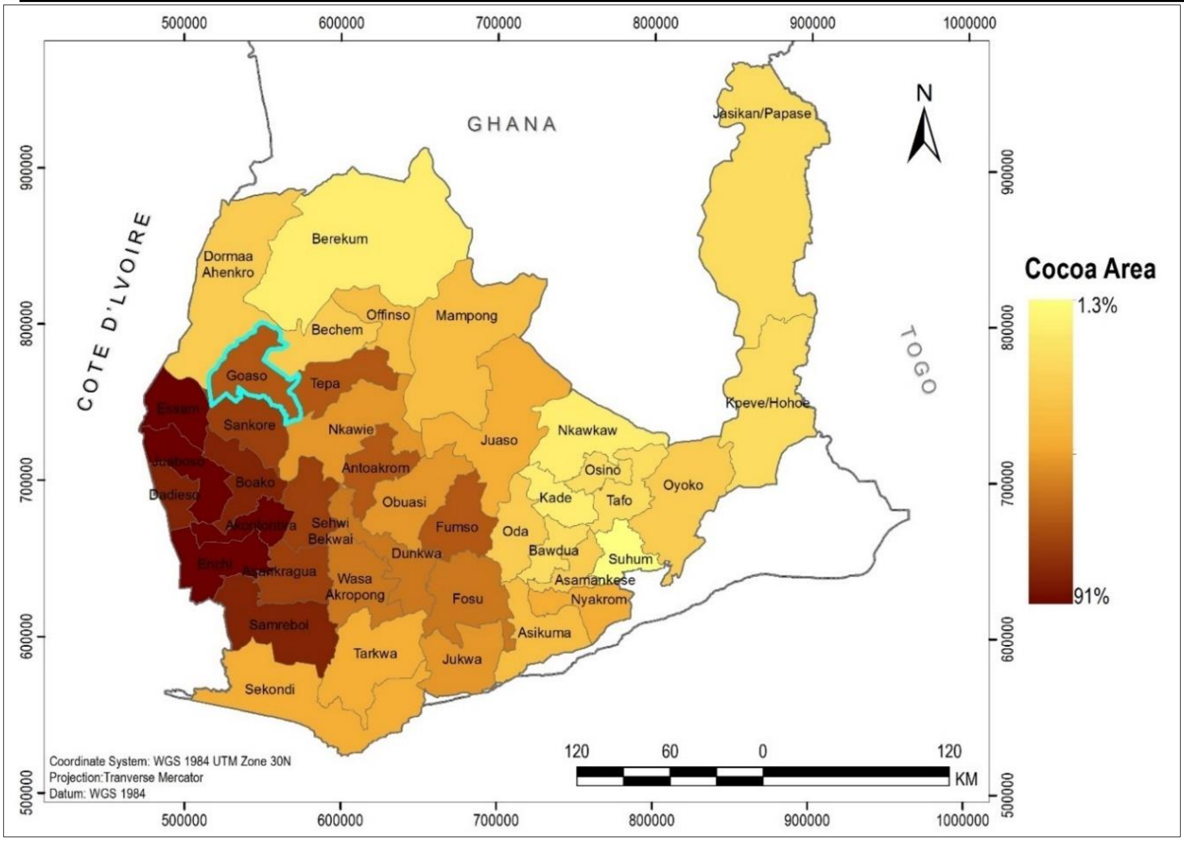


Figure 6: Area Extent Surveyed by COCOBOD to determine their relative cocoa areas (2016)

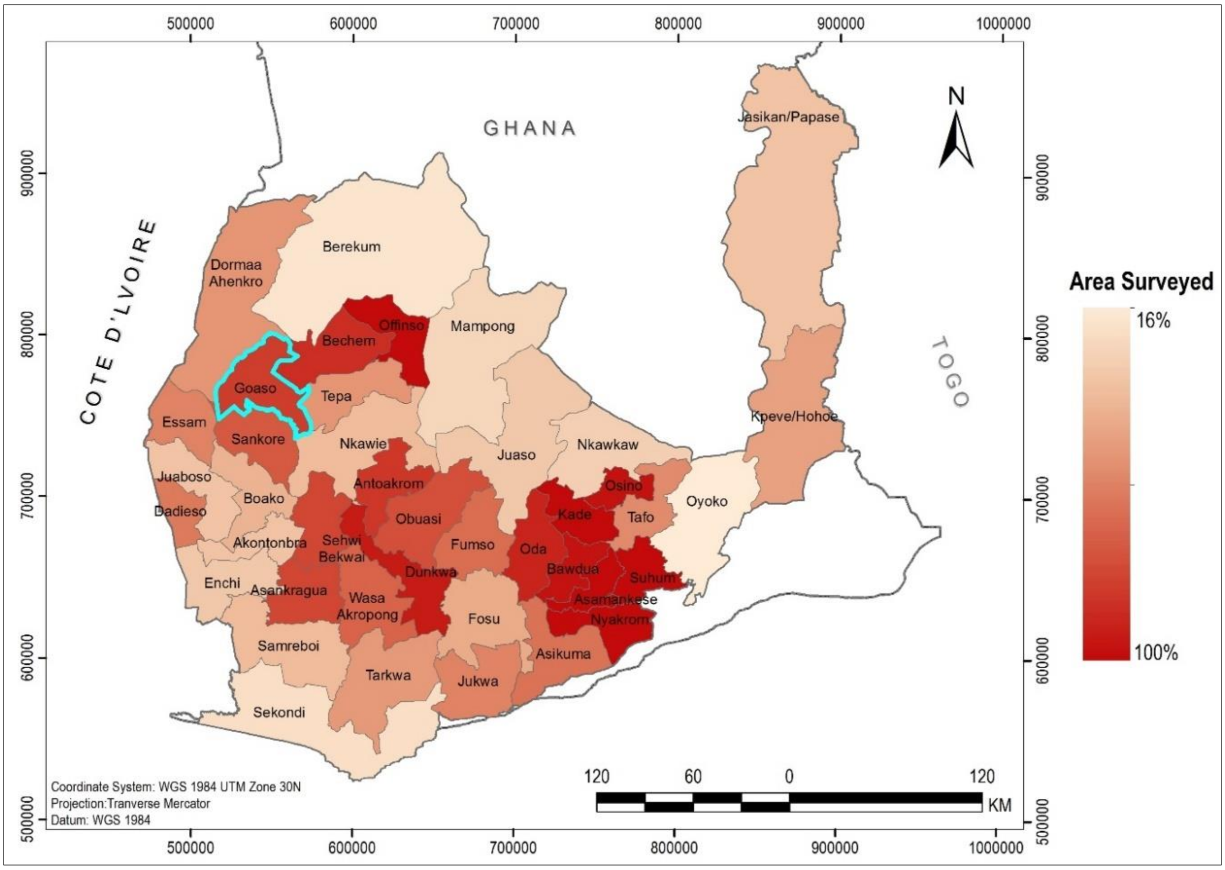


Figure 7: Estimated Percentage of Cocoa Area Within The COCOBOD Surveyed Districts.

Appendix 1 shows the COCOBOD statistics table with extrapolated cocoa figures for some districts.

Cocoa Fields Polygons Data

The Goaso cocoa fields data was obtained from Centre for Remote Sensing and Geographic Information Systems (CERSGIS) in Accra, with approval from the COCOALIFE organization. This field data was collected from March 2016 to July 2016 by CERSGIS for COCOALIFE https://www.cocoalife.org/the-program .

The database consisted of actual cocoa fields locations which were taken by recording the coordinates of individual cocoa farm boundaries using a GPS with precision of 3m. A total of 1,773 individual cocoa fields were available within the area of interest with area size ranging between 0.02 to 32 ha. The majority of the fields, 91%, were below 5 ha. The data do not cover the whole study area but only for some specific areas of interest to the COCOALIFE.
Figure 4 shows the clustered locations of the cocoa fields polygons in Goaso. This data is used for accuracy assessment for the final cocoa maps. It is however noted that the area assessed only includes the areas within the five clusters containing the cocoa fields delineated on the map. Within the clusters, areas within cocoa fields polygons are the confirmed cocoa areas, outside the boundaries are unconfirmed areas.

Ghana Cocoa Districts

The Ghana cocoa district shapefile was collected from CHED. The 41 districts boundary data was used in this study.

Ancillary Spatial Data

Other ancillary data sets were also collected and used in this research. They include forest/Game reserves, roads and rivers shapefiles retrieved from the OpenStreetMap (OSM) online database for Ghana https://download.geofabrik.de/africa/ghana.html.

2.1.3. Software Used

Software such as El-Shayal Smart GIS, MAPC2MAP, Locus Map were used for pre-field work data preparation, on site navigation and geotagging photos during the field work. ERDAS IMAGINE 2016, ENVI 5.3, Notepad++, TIMESAT, ILWIS 3.7, ESA SNAP, ArcGIS 10.4.1, were used for image and vector data preparation, processing and map composition. Microsoft office (Word, Powerpoint Excel, Visio) and SPSS were used for statistical analysis, flowchart design and report writing.

2.2. Method

The methodological steps followed in mapping cocoa area; hyper-temporal, high spatial resolution and the integrated approach are presented in flow chart shown in figure 8.

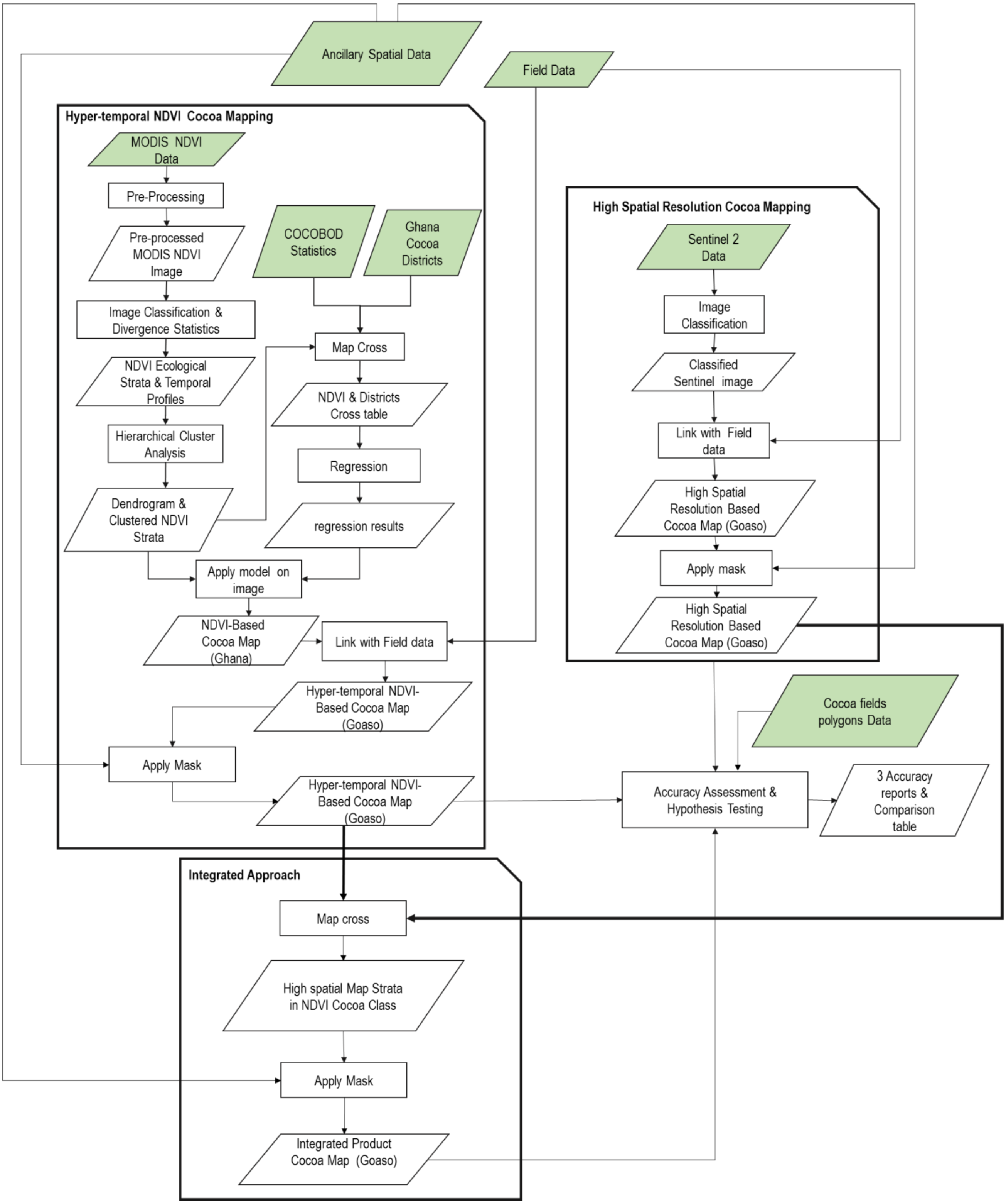


Figure 8: Flowchart of Method

2.2.1. Hyper-Temporal NDVI Cocoa Mapping

This section presents the methodological steps followed in mapping cocoa area using hyper-temporal NDVI Approach

Pre-Processing

In this research, the NDVI data from the MODIS Terra and Aqua satellites were combined. This was done to improve temporal frequency (Terra 16-day period starting Day 001, Aqua 16-day period starting Day 009) and also reduce significantly the potential noise of cloud contamination over the study area, southern Ghana, a region known for its persistent cloud cover (Ali et al., 2013). This gave also an added advantage of reduced cloud contamination since it is based on two images a day instead of one acquired at different times of the day (Terra 10:30 am and Aqua 01:30 pm local)(Ali et al., 2013). The combination was feasible since both share similar spatial, spectral and radiometric characteristics with a proven strong correlation in NDVI values (R2=0.97, RMSE=0.04) (Gallo et al., 2005). Figure 9 shows the flow of the pre-processing procedure followed to prepare data for analysis.

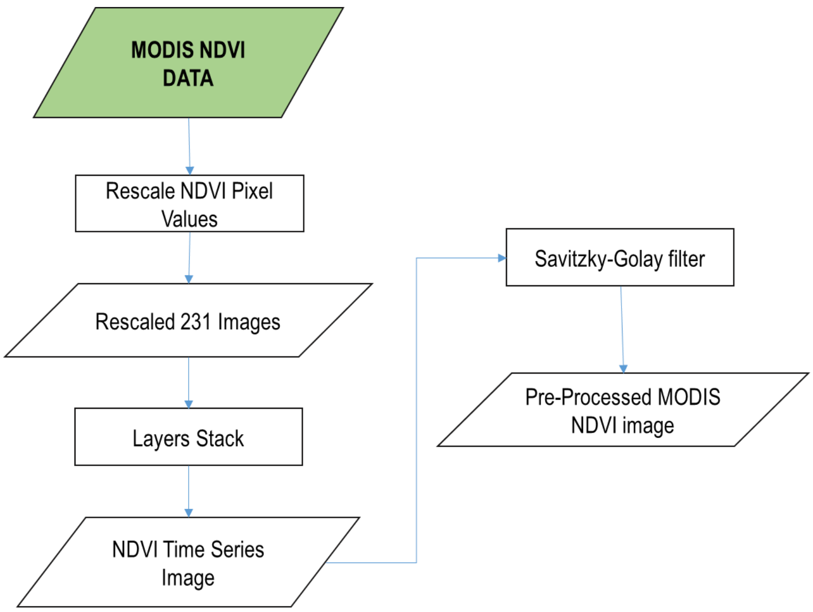


Figure 9: Pre-processing Expanded Flowchart

The MODIS NDVI data is a normalized transform of the NIR to red reflectance ratio, ϱNIR/ϱRED, and is designed to standardize Vegetation Index values to between -1 and +1 (Solano et al, 2010). The raw NDVI pixel values ranged -2000 to 10000 with -3000 flagged as nodata and so pixels values were rescaled to range 0-255 DN (where 0 is -1 and 255, +1) using the formulae;

Equation 2: Rescaled NDVI =original NDVI*0.02125+42.5+0.5 (de Bie, 2016)

The rescaled 231 images were stacked into a single NDVI time series following a sequential order using the image dates (See figure 10).

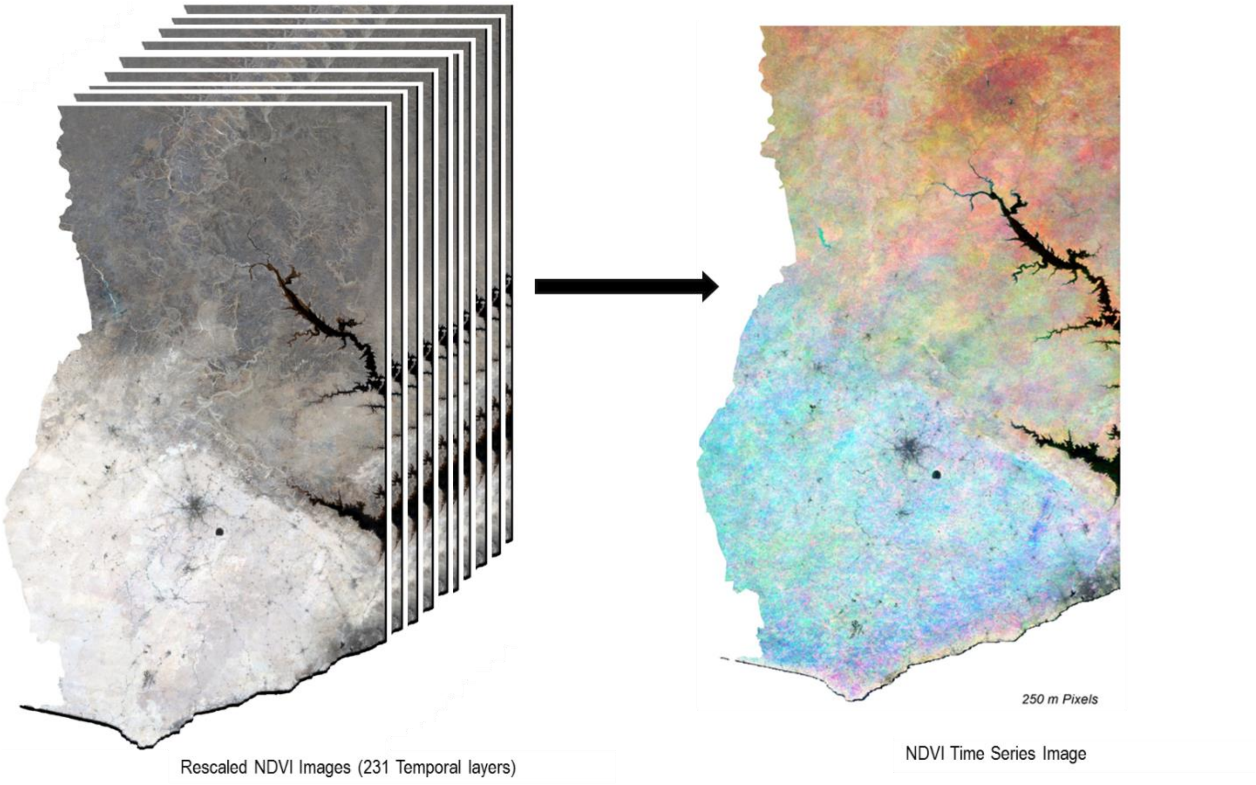


Figure 10: Rescaled 231 MODIS NDVI Images Stacked into an NDVI Time Series image

The adaptive Savitzky-Golay procedure in TIMESAT was then applied to clean and smoothen the NDVI temporal profiles of time series image as shown in figure 11.

The Savitzky-Golay allows to retain the upper envelope of NDVI time series which reduce data noise and account for any data gaps (Ali et al., 2013). This method is found useful for noisy and non-uniform NDVI time series datasets (Beltran-Abaunza, 2009; Jönsson & Eklundh, 2004) collected over areas with mostly persistent cloud cover . Final pre-processed MODIS NDVI image of this procedure was used for further analysis.

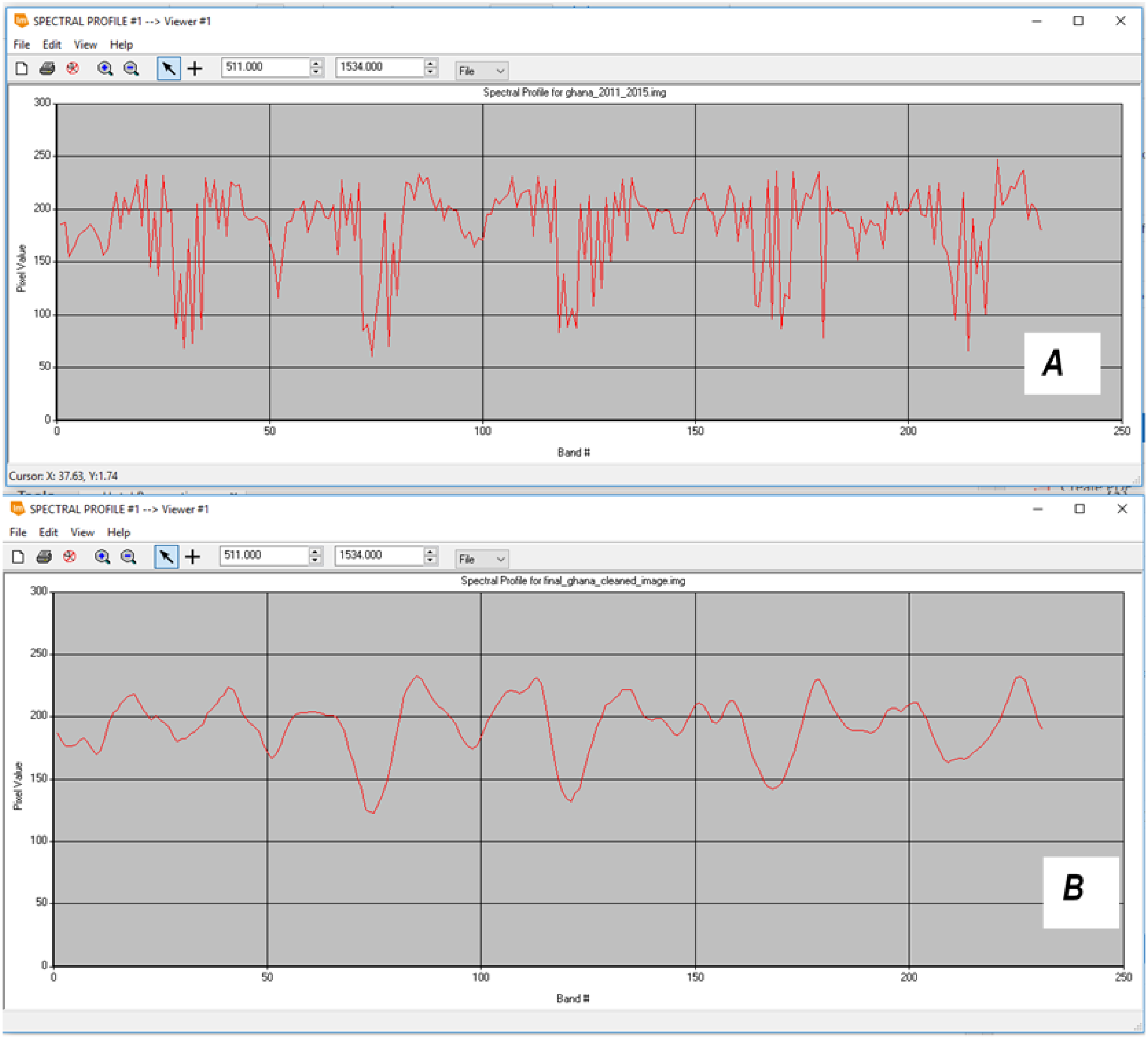


Figure 11: Spectral profile behaviour of a pixel in NDVI Time Series (A) Without applying Savitzky-Golay filter (B) With Savitzky-Golay filter applied (Screen dumps from Erdas)

Image classification & Divergence Statistics

To identify and distinguish relevant NDVI ecological strata from the pre-processed NDVI image, an approach developed by de Bie et al. (2008) was followed. First, the NDVI image was clustered using the Iterative Self-Organizing Data Analysis Technique (ISODATA) unsupervised classification method. This was done in Erdas Imagine software in runs from 10 to 100 classes (see figure 12). The maximum number of iterations which performs an entire image classification (self-organizing) per run was set to 50 and the convergence threshold set to 1.0. The unsupervised classification was done to capture the range of variability in phenology over the image across the study area.

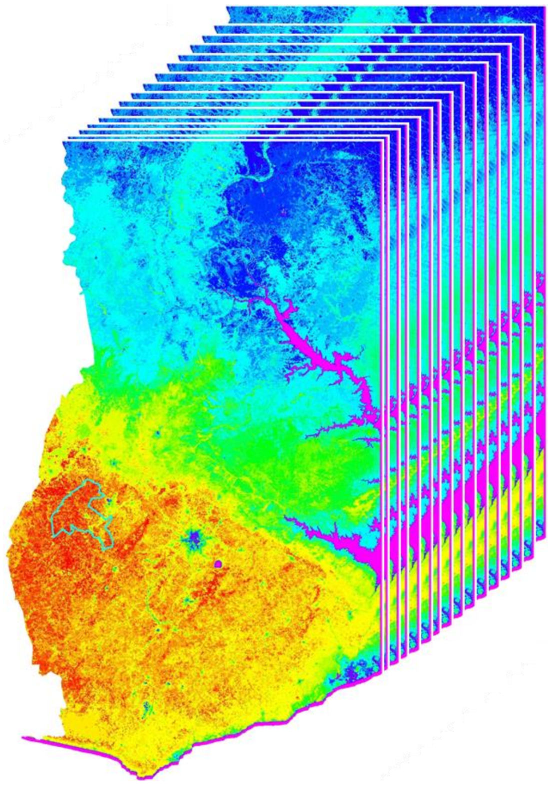


Figure 12: ISODATA runs 10-100 classes (sample classifications)

Second, to find the relevant number of NDVI classes to use for analysis, the various classification runs were examined on basis of separability using the divergence statistical distance measure (ERDAS, 2003). The divergence statistic's minimum (which denotes similarity between two classes) and average (which denotes similarity amongst all classes) separability values were obtained for all the classification runs and plotted to observe the common peak. To select the best classification from plot, the common peak where both separability values are high with the least number of NDVI classes is considered the optimum.

Finally, the selected classified image was converted to shapefile and re-projected to Universal Transverse Mercator (WGS_1984_UTM_Zone_30N) projection in Erdas then exported to ArcGIS. This was done to avoid issues of spatial shifts of pixels when using Sinusoidal (original projection of MODIS data) projection image data in ArcGIS (de Bie, 2016).

The re-projected shapefile was used for further spatial analysis.

Hierarchical Cluster Analysis

The NDVI temporal profiles of the of the selected classification were re-group using the Hierarchical Cluster Analysis (HCA). The HCA method identifies structures within the NDVI temporal profiles using the mean values extracted from signature file to establish a more natural grouping of the classes (Gauch Jr & Whittaker, 1981).

To do this, the mean NDVI values were extracted for each 16-day composite period from the signature file of the selected classified image. The 16-day mean NDVI values were group into monthly averages for the 5-year period in excel. The results were exported to SPSS for the Hierarchical Cluster Analysis (HCA). The HCA technique uses nearest neighbor method with cosine distance as similarity measure (Gauch Jr & Whittaker, 1981). The output gave a dendrogram which groups the NDVI classes that exhibit similar temporal behavior into clusters

Dendrogram results were used to produce an NDVI cluster map of the Ghana to allow visual assessment of spatial pattern.

Map Cross & Regression Analysis

To derive the NDVI-based cocoa strata, the clustered NDVI classes from the Hierarchical Clustered Analysis are correlated with cocoa area statistics from COCOBOD using stepwise multiple linear regression. This method was used because it is proved that regression models fit crop area data, with more than 70% of variability explained (GONZÁLEZ-Alonso & Cuevas, 1993; Maselli & Rembold, 2001). It is however on the assumption that, variables (e.g. NDVI classes) and crop area statistics are independent of each other and crop area of a particular crop is represented by a linear combination of multiple predictors.

The clustered NDVI classes were intersected with Ghana cocoa district polygons. The area in hectare(s) was calculated for each NDVI cluster in every district and lined to the COCOBOD statistics by district in a matrix. This matrix was exported to SPSS for the stepwise multiple regression analysis.

In the regression, the area covered by each NDVI cluster in the districts (38 districts) was used as explanatory (independent) variables and cocoa area statistics by district as dependent. No constant was used and coefficients were constrained between 0.0 and 1.0., because the cropped area can neither be in negative nor more than 100% of the district area (Khan et al., 2010). The resulting coefficients from the model were used to identify the NDVI-based cocoa strata.

The results of the regression analysis are summarized in the form of; selected NDVI cluster(s) with associated coefficient, Adjusted R^2, and significance level. The derived coefficient represented the percentage under cocoa crop area in the NDVI cluster. This was used to determine the cropping intensity of cocoa within the NDVI cocoa cluster. Results of the procedure was applied to generate cocoa map based on selected NDVI Cluster which shows the broad land cover types were cocoa is grown.

In the next step, the NDVI Classes of only the selected NDVI cluster(s) were linked again to the COCOBOD statistics using the same regression procedure. This was done to derive coefficients that was applied to only the NDVI cocoa cluster to select the final NDVI-based cocoa classes. The results of this procedure was used to finally select NDVI cocoa classes with corresponding temporal profiles. The temporal profiles of the selected NDVI-based cocoa classes were further examined visually using scatter plots and temporal NDVI profiles to see relationship with leaf flushing dates noted by Opoku-Ameyaw et al (2010).

NDVI-based Cocoa Map Link with Field Data

To specify the exact cocoa mapping intensity within the districts, Goaso district was selected and the clustered NDVI classes of that area crossed with field data to generate a matrix. This was done to improve the reliability of the hyper-temporal NDVI-based Cocoa map since reported crop area statistics where incomplete so coefficients generated from regression were less reliable.

To generate the matrix, the clustered NDVI classes were intersected with the field data. In the matrix generated, the percentage of the total samples for a particular cover type (e.g. cocoa or non-cocoa) within the NDVI Class was calculated. This calculated percentage was used to estimate the cover intensity within the NDVI class.

Final results of this procedure was used to create the final legend and hyper-temporal NDVI-based cocoa map of Goaso. The accuracy of this cocoa map is assessed using actual cocoa fields polygons received from CERSGIS, Accra (Ghana) with approval of COCOALIFE as discussed in section 2.2.4.

2.2.2. High Spatial Resolution Cocoa Mapping

The methodological steps followed in mapping cocoa area using the high spatial resolution procedure is presented in this section. This procedure for mapping cocoa area was the baseline method for mapping cocoa.

Image Classification

To identify cocoa land cover, a single-date Sentinel-2 image obtained for the Goaso district was clustered using the ISODATA unsupervised classification algorithm. This classification method was preferred to using supervised classification mainly because cocoa is grown within an area with a mixture of cover types; forest trees, food crops which function as shade in most cases. Also, because of the spectral reflectance similarity between cocoa trees and that of natural forest it is difficult to visually without prior knowledge of the area. Therefore, unsupervised classification method which allows what can be differentiated based on the image's properties without influence was advantageous. The map classes were then identified with ground verified data as next step.

The Sentinel image in this research was classified into 20 map classes to allow best discrimination between the cover types present in the image. To maintain the spatial resolution quality of the image, only the 10-m multispectral bands (4 bands) were used for the classification.

Classified Sentinel Image Link with Field Data

To identify and estimate intensity of cocoa cover within the 20 map classes of sentinel, the classified image was linked to the field data for legend construction. Following a similar procedure used in the hyper-temporal NDVI method, the classified image was intersected with field data to derive a matrix. In the matrix, the percentage of the total samples for a particular cover type within the class was calculated. This calculated percentage was used to estimate the cover intensity within the map class.

The classes with associated cocoa cover intensities from the matrix was used to produce the final cocoa map of the high spatial resolution procedure. The accuracy of this map was assessed using actual cocoa fields polygons as discussed in section 2.2.4.

2.2.3. Integrated Approach

This section integrates the hyper-temporal NDVI and high spatial resolution cocoa mapping methods into a top-to-down mapping procedure for cocoa in Goaso district. To derive the integrated product, hyper-temporal NDVI Based Cocoa map and the high spatial resolution based cocoa map were combined and cropping intensities estimated using simple arithmetic.

To do this, three main steps were implemented. First cocoa intensities were estimated for integrated approach using the formulae.

Equation 3: $$\mathrm{Y} = (\alpha * \beta)$$

Where Y= Integrated cocoa fractions (%)
α = Hyper-temporal NDVI cocoa fractions (%)
β = High spatial resolution cocoa fractions (%)

However, the results of this step do not show actual cocoa fractions in the map since it was based only the fractions. Therefore, to get the actual intensities two more steps were implemented using the area of cocoa estimated as a factor. Have estimated the integrated cocoa fractions in addition to the two approaches, the total cocoa area in Goaso was calculated for all three using equation 4.

Equation 4: $$\chi = \omega * \tau$$

Where χ = Cocoa area
ω = Mapped cocoa fraction (All cocoa fractions in α, β, Y)
τ = Total area

Finally, the average cocoa area estimated by the hyper-temporal NDVI and high spatial resolution approaches are divided by integrated cocoa area to derive a coefficient (See equation 5). The coefficient was used to multiply the values of the integrated cocoa fractions to get the actual cocoa area fractions. Any cocoa fractions values were constrained between 0 to 1.

Equation 5: $IG = (\alpha_i + \beta_i)/Y_i$

Where IG= Integrated Coefficient

α_i = Integrated cocoa area (ha)

β_i = High spatial resolution cocoa area (ha)

Y_i = Integrated cocoa area (ha)

The final estimated cocoa fractions result of was used to produce a final cocoa map for the integrated approach. The accuracy of the integrated approach cocoa map was also assessed using actual cocoa fields polygons as discussed in section 2.2.4.

2.2.4. Accuracy Assessment & Hypothesis Testing

The accuracies of cocoa maps from the three procedures were assessed using CocoaLife cocoa polygons in simple linear regression analysis where an Adjusted R^2 was obtained. The independently collected cocoa field polygons data included a total of 1,773 fields with an average field size of 2.2 ha. Cocoa fields with a minimum field size of 2 ha were only considered and selected for analysis. A total of 700 cocoa fields with the defined area minimum of 2 ha and were thus used for the accuracy assessment. The cocoa field polygons were examined with field geotagged photos to confirm at least that cocoa is grown within the polygons.

In the accuracy assessment process, the five clusters cocoa polygons (see figure 4) were intersected with cocoa maps to derive the mapped area within the clusters. The total area mapped within the actual cocoa polygons were taken as the confirmed cocoa areas and those outside unconfirmed. The area of cocoa fraction strata of the confirmed and those of the unconfirmed within the clusters were calculated. The results for the confirmed areas were correlated with the actual cocoa polygons in a simple linear regression. This was done for all three cocoa maps. The results of the accuracy assessment for each of the cocoa map is summarized in an Adjusted R^2 which was used as the measure of accuracy.

A final comparison table showing accuracy results of the three mapping approaches was created and based on that, the hypothesis was tested.

3. RESULTS

The results presented in this section cover the methodological steps carried out for each of the three crop area mapping procedures; hyper-temporal NDVI, high spatial resolution and the integrated approach for cocoa in Ghana.

3.1. Hyper-Temporal NDVI Cocoa Mapping

Image classification & Divergence Statistics

The results of the divergence statistics analysis are presented in the plot of figure 14. From the plot, both average and minimum separability were found to be relatively high at point, 35. However, the point 86 with a relatively high average separability value and good minimum separability value chosen. This was because, using 36 classes would be too general especially at the Goaso district level analysis.

The selected 86 class map with temporal profiles are presented in figure 15.

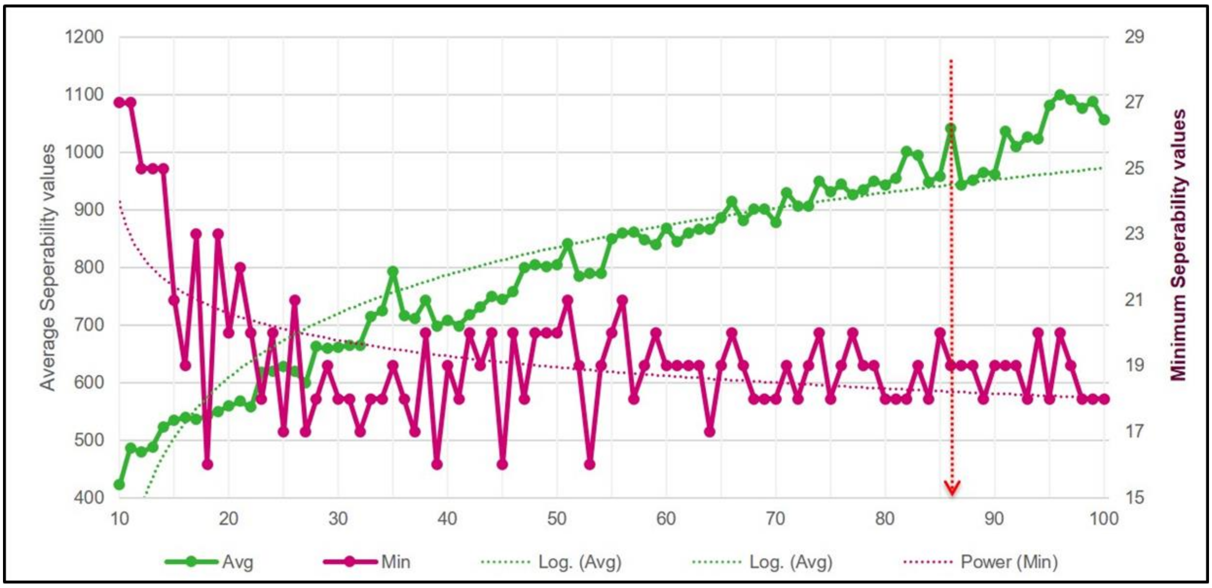


Figure 13: Divergence Statistics for Best NDVI Class Map Selection

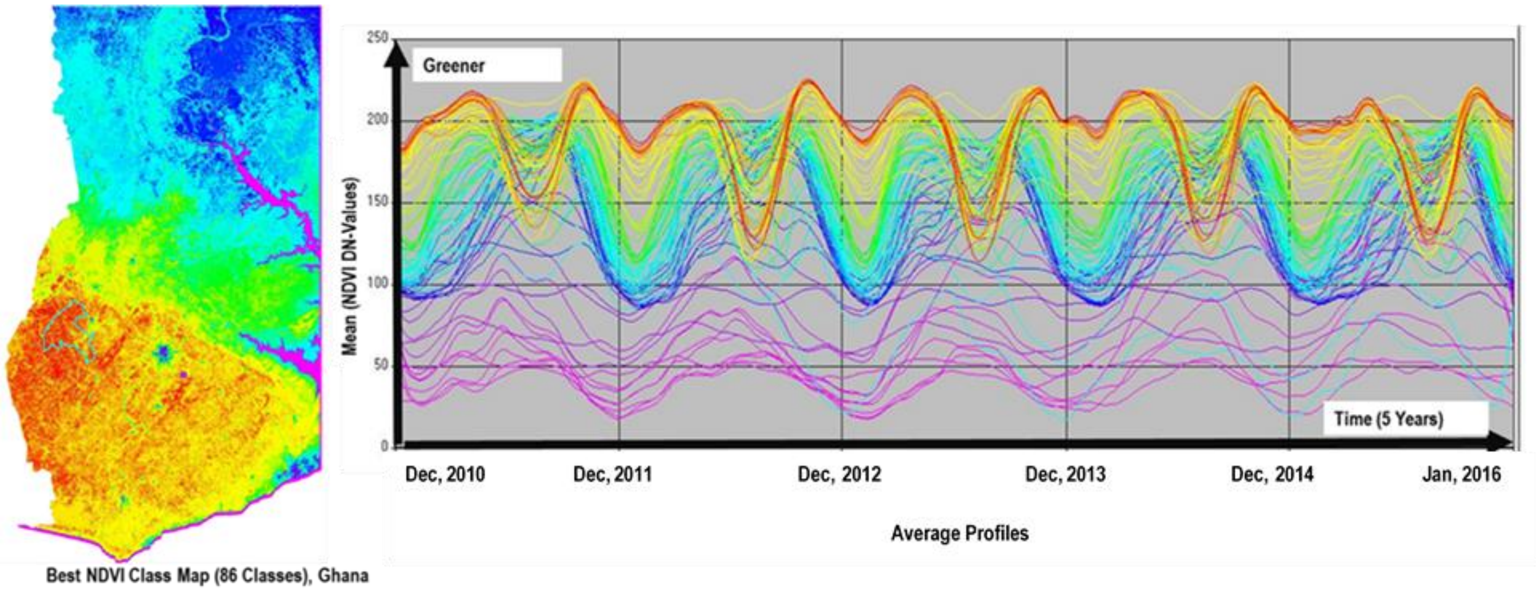


Figure 14: 86 Class NDVI Map and Temporal Profiles

From figure 15, the average profiles show the seasonal characteristics of the different land cover types present in the observed area in Ghana for the five-year period. The profiles suggest that there exist inter and intra annual variation in NDVI values in the area. Low NDVI DN (pixel) values indicates less vegetation presence and high NDVI values signifies the presence of very dense vegetation coveer. The associated map clearly shows the spatial stratification in the area.

Hierarchical Cluster Analysis

The results of the re-grouping of the NDVI Classes from the hierarchical cluster analysis gave the dendrogram presented in figure 15. The dendrogram show the relationship among the NDVI classes at different level of similarity on which basis clusters are made. A total of 15 clusters were identified based on the lowest cosine distance between the NDVI Class profiles. The 15 clusters from the dendrogram are presented in the HCA cluster map (figure 16) created for visual inspection of spatial pattern.

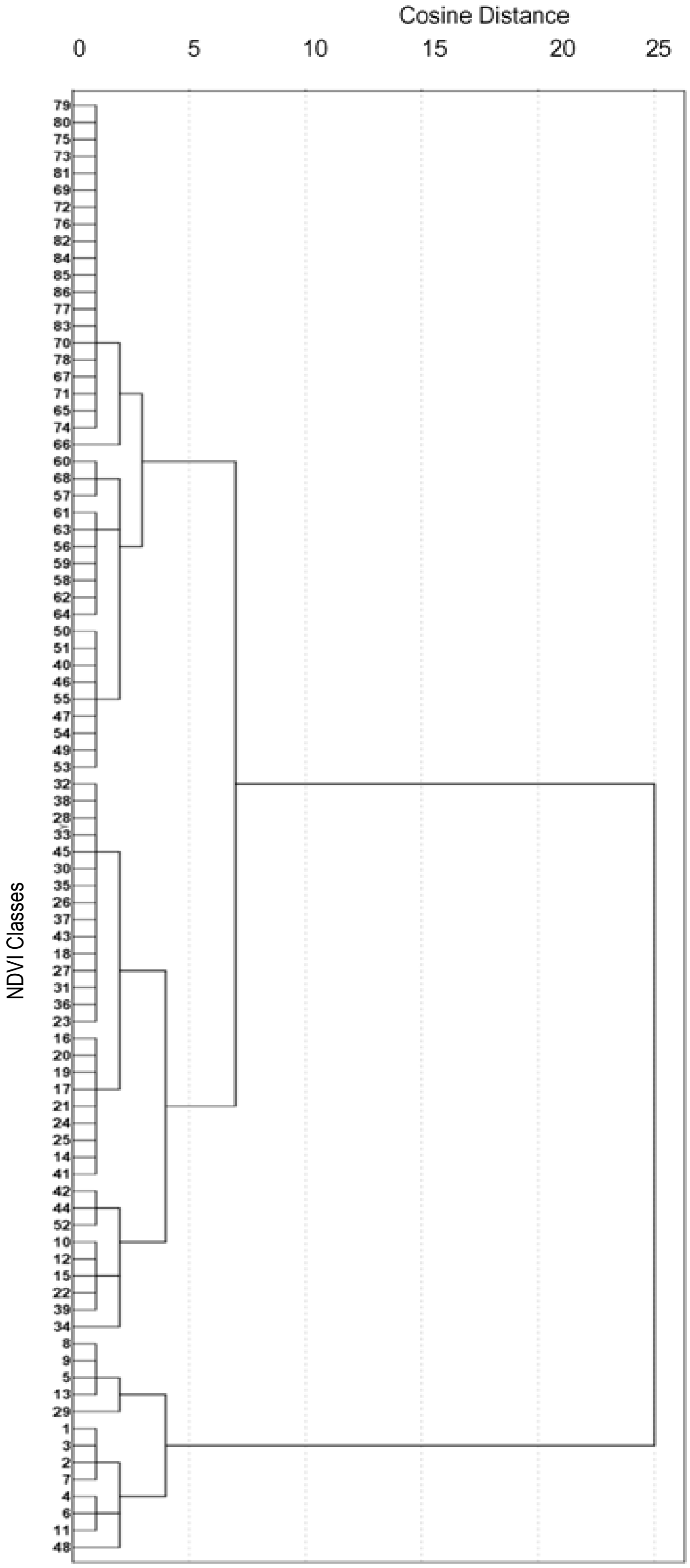


Figure 15 A dendrogram showing the re-grouping of the 86 NDVI Class Profiles into Clusters

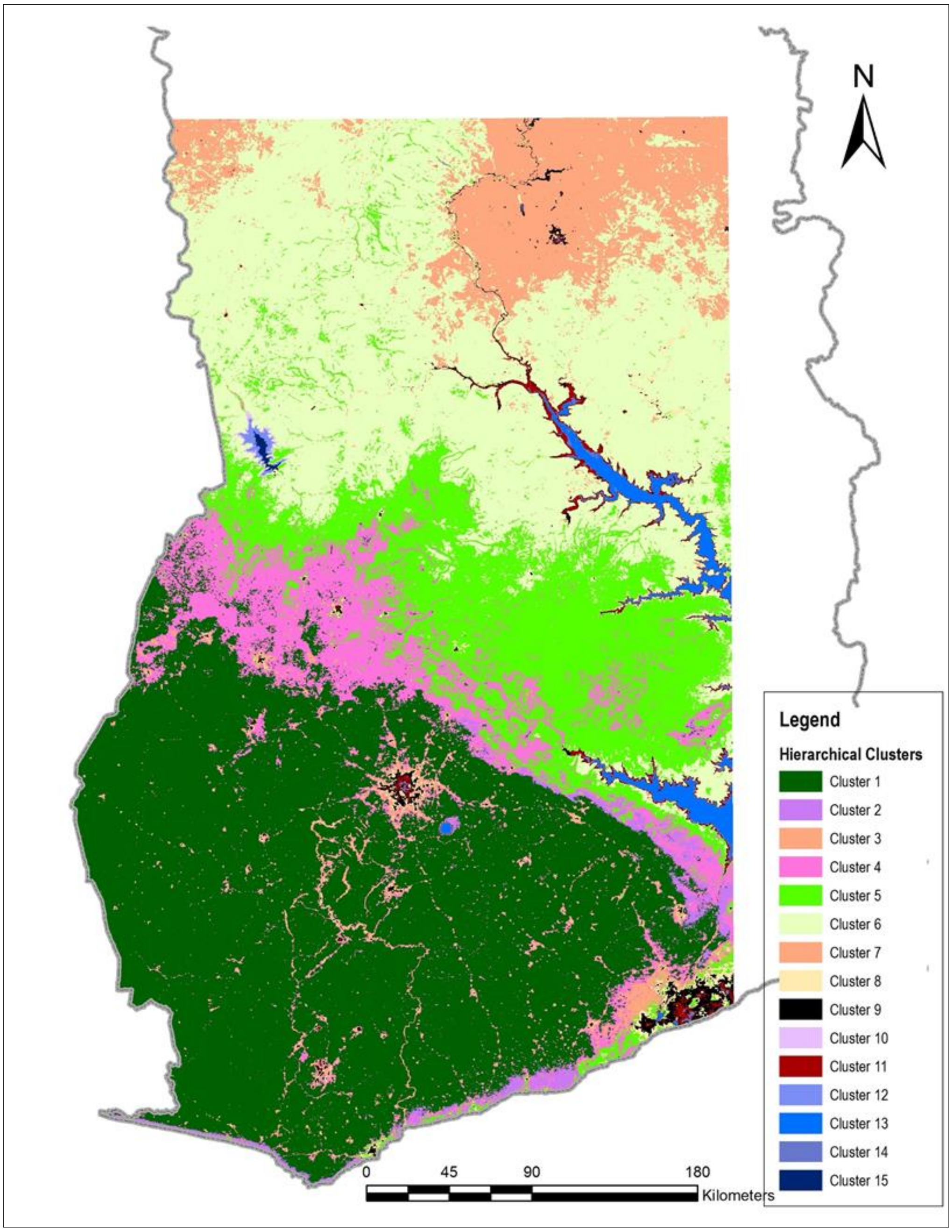
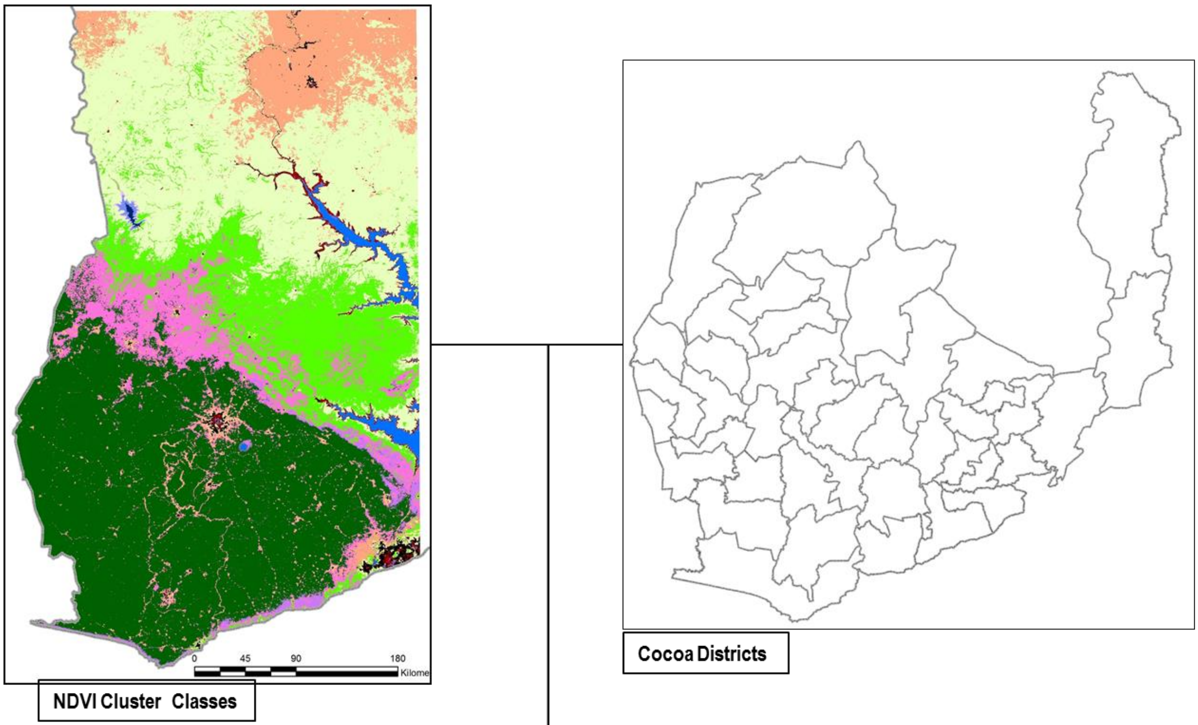


Figure 16: Clustered NDVI Strata, Ghana

Map Cross & Regression Analysis

The results of the combination between the clustered NDVI strata, COCOBOD statistics and the 38 districts are in the form of matrix part of which is presented in figure 17.



Cocoa Districts	COCOBOD Cocoa Area (ha)	NDVI Cluster 1 (ha)	NDVI Cluster 2 (ha)	NDVI Cluster 3 (ha)	NDVI Cluster 4 (ha)	---	etc
AKONTOMBRA	91,920	98,161	0	786	982	---	---
ANTOA KROM	59,971	109,892	42	13,564	10,812	---	---
ASAMANKESE	7,459	74,877	21	1,148	1,634	---	---
ASANKRAGWA	91,120	145,976	72	2,113	2,792	---	---
ASIKUMA	25,897	147,724	13,369	5,046	19,276	---	---
BAWDUA	3,786	82,143	52	3,537	3,705	---	---
BECHEM	19,455	121,218	16	6,059	51,565	---	---
BEREKUM	26,414	78,730	2,302	11,479	461,844	---	---
BOAKO	127,582	150,417	0	1,550	1,409	---	---
DADIESO	65,066	80,742	0	512	313	---	---
DORMAH	42,199	272,389	105	10,681	173,792	---	---
DUNKWA	65,767	154,764	230	6,594	7,023	---	---
ENCHI	151,922	175,008	16	750	982	---	---
ESSAM	122,086	137,053	0	1,136	554	---	---
FOSO	87,018	222,584	249	4,600	6,795	---	---
FUMSO	84,029	162,432	133	1,820	3,779	---	---
GOASO	91,412	176,963	26	1,501	4,607	---	---
JUABOSO	146,252	156,518	0	1,082	725	---	---
JUASO	105,105	234,239	9,913	14,244	89,574	---	---
---	---	---	---	---	---	---	---
etc	---	---	---	---	---	---	---

Figure 17: Clustered NDVI Strata, Ghana Cocoa Districts and COCOBOD Statistics Map Cross

The NDVI and Districts cross table in figure 17 show the total area covered by each NDVI Cluster in hectares within the various districts lined with COCOBOD cocoa area. These area values by NDVI clusters and reported cocoa statistics were entered into a multiple linear regression. The results are presented in table 1.

Table 1: Results of Stepwise Multiple Linear Regression for NDVI Clusters and Cocoa Area Statistics

NDVI Clusters	Co-efficient	Adjusted R^2	Sig. (%)
NDVI Cluster 1	0.427	0.743	0

This result shows that Cluster 1 which contains 20 NDVI classes (Appendix 2) was the most significant cluster, with 74% of the variability in cropped area by districts explained by the regression model. This interpretation was based on the adjusted R^2 value obtained. The derived coefficient of 0.427 indicates the intensity of the cocoa per hectare.

This coefficient value was used to produce cocoa map shown in figure 18.

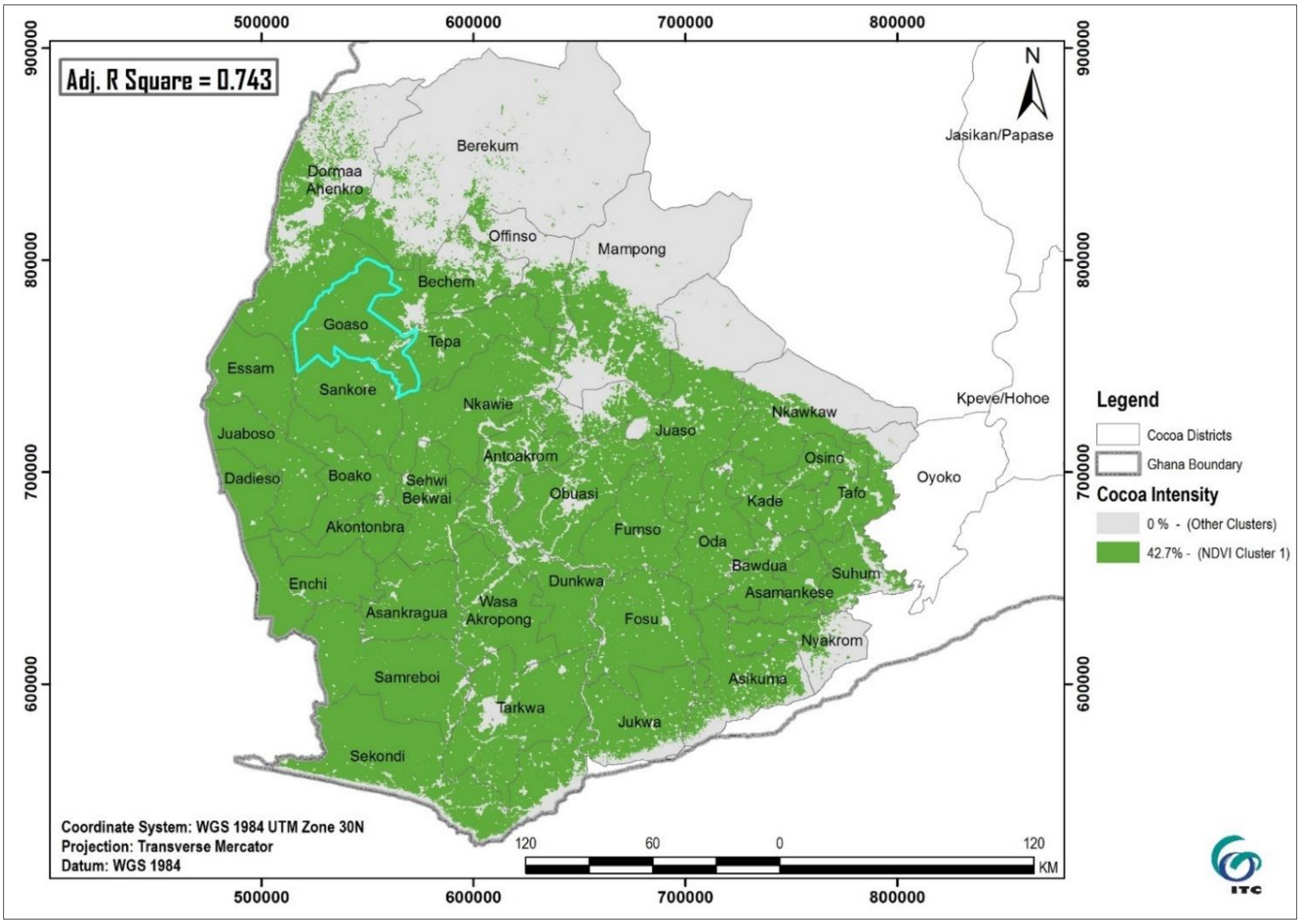


Figure 18: NDVI-Based Cocoa Map, Ghana

From the figure 18, it is obvious that the whole area discriminated as cocoa based on the selected NDVI cluster poorly reflects cocoa cropped areas. At this stage, only the broad land cover classes containing the crop are discriminated.

And so, repeated regression analysis for only classes of the selected NDVI cluster gave the results presented in table 2. The results of the multiple linear regression gave an Adjusted R^2 value of 0.907 indicating a very strong correlation with reported statistics. Results suggest that NDVI Class 85, 81 and 78 are the cocoa classes. However, all the coefficient which were expected to range from 0 to 1 were all more than 1. The selected classes were further examined using scatter plots and temporal NDVI plot shown in figure 19 and 20 respectively.

Table 2: Results of Stepwise Multiple Linear Regression for Significant Cluster NDVI Classes & Cocoa Area Statistics

Cluster NDVI Class	Coefficient	Sig. (%)
NDVI Class 85	4.286	0.00
NDVI Class 81	1.943	0.00
NDVI Class 78	1.662	0.00

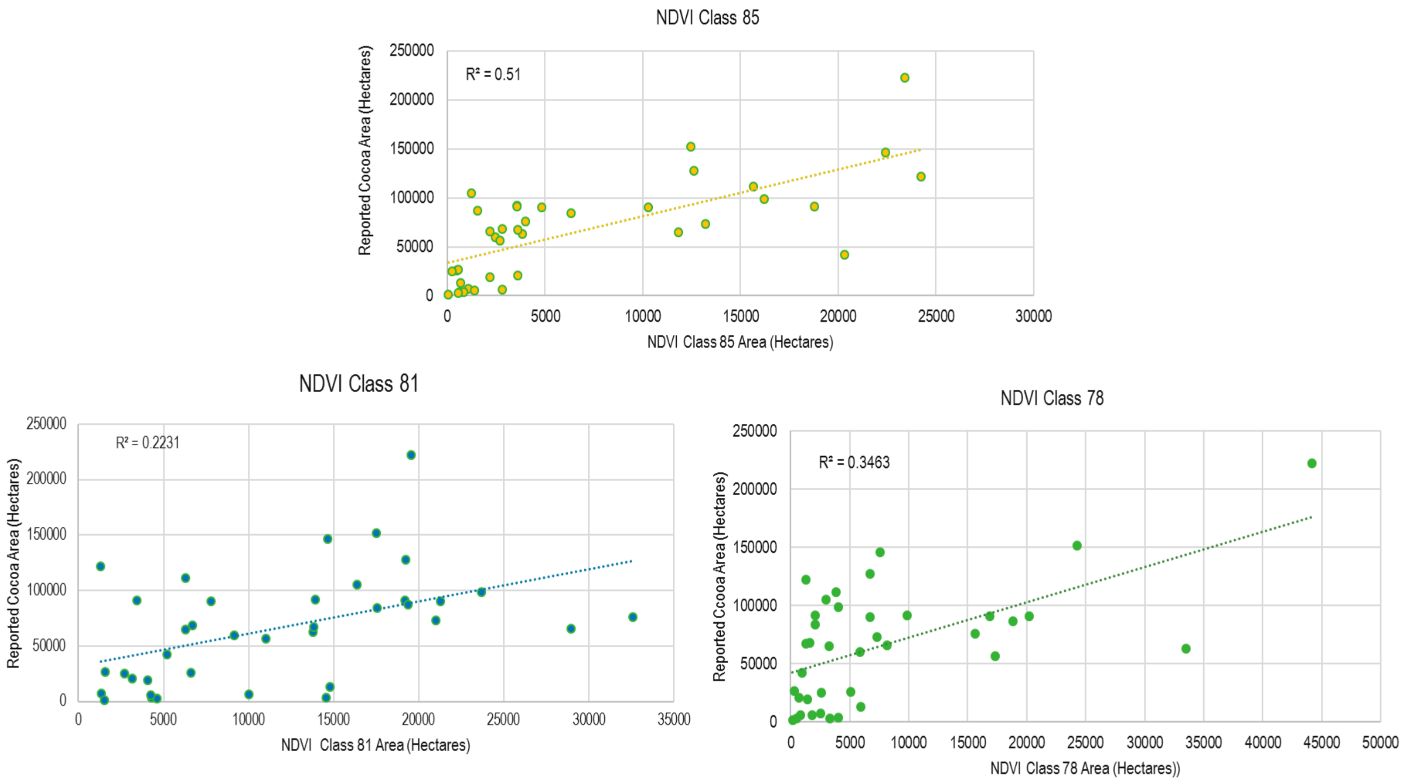


Figure 19: Scatter Plots showing correlations between Selected NDVI-Based Cocoa classes area and Reported Cocoa Area by COCOBOD in the 38 Cocoa Districts.

From the scatter plots, the strength of the relationship between the individual NDVI cocoa classes with COCOBOD statistics was low to moderate with R^2 ranging between 0.22-0.51.

The temporal profiles of the selected classes seem to share more similar temporal properties based on the five-year mean curves which shows that all classes have relatively high NDVI DN values. High NDVI DN-values signify presence of dense vegetation canopy (greenness) based on the curves, drops around August for all three profiles. Green up periods are seen to start around March – April and September up till November. This confirms the leaves flushing periods noted by Opoku-Ameyaw et al. (2010) in matured cocoa trees.

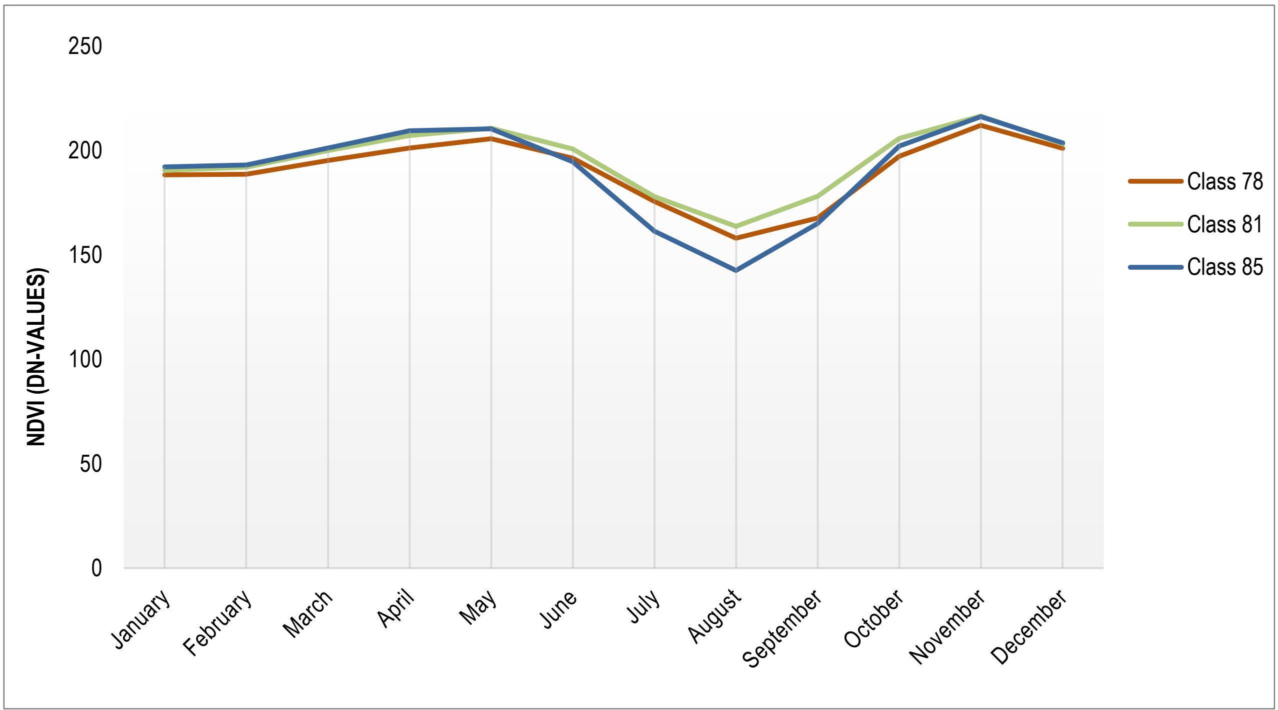


Figure 20: Five-Year Temporal NDVI Profiles of Selected Cocoa NDVI Classes

Although the selected NDVI classes regression results show significant result, the derived coefficient for each class (table 2) poorly reflects cocoa cropped area by district. The value of the coefficient indicates that for every 1 hectare of district area, cocoa presence is estimated to be 4 times higher than the hectare, in the case of NDVI Class 85 for instance. This was a possible statistical error resulting possibly from the incomplete COCOBOD statistics used.

Thus the results of this procedure could not be used to make a realistic cocoa map for the study area.

Therefore, further analysis used the results of the NDVI clusters used to produce the NDVI-based cocoa map in figure 18. Since the selected NDVI cluster 1 were not able to differentiate forest reserves, they were excluded by masking. Also, all further analysis excludes forest reserves by masking.

NDVI-based Cocoa Map Link with Field Data

The cross between NDVI Clusters and field data (see appendix 2) gave the matrix presented in table 3. The table show details for NDVI Classes for Cluster 1 since it was the most significant cluster from the regression analysis. All other clusters were grouped into one. The estimated intensities made based on the number of cocoa and non-cocoa field data NDVI strata was used to make the final legend for the hyper-temporal cocoa map.

It is however noted that, NDVI Classes of Cluster 1 with less than 10 field samples were combined and one estimate given for that group of classes.

The final cocoa map created for this method is presented in figure 21.

Table 3: NDVI-Clusters and Field Data Cross Table

NDVI Cluster 1 Classes	**Cocoa Intensity**	**Non-Cocoa Intensity**	**Samples**
Class 65	10%	90%	10
Class 75	80%	20%	30
Class 77	23%	77%	13
Class 80	49%	51%	43
Class 82	59%	41%	34
Class 84	50%	50%	14
Class 85	56%	44%	16
Class 86	92%	8%	25
Class 67, 70, 71, 73, 74, 76, 78, 81 & 83	55%	45%	29
Class 69, 72 & 79	Not surveyed	Not Surveyed	0
Other Clusters	17%	83%	17

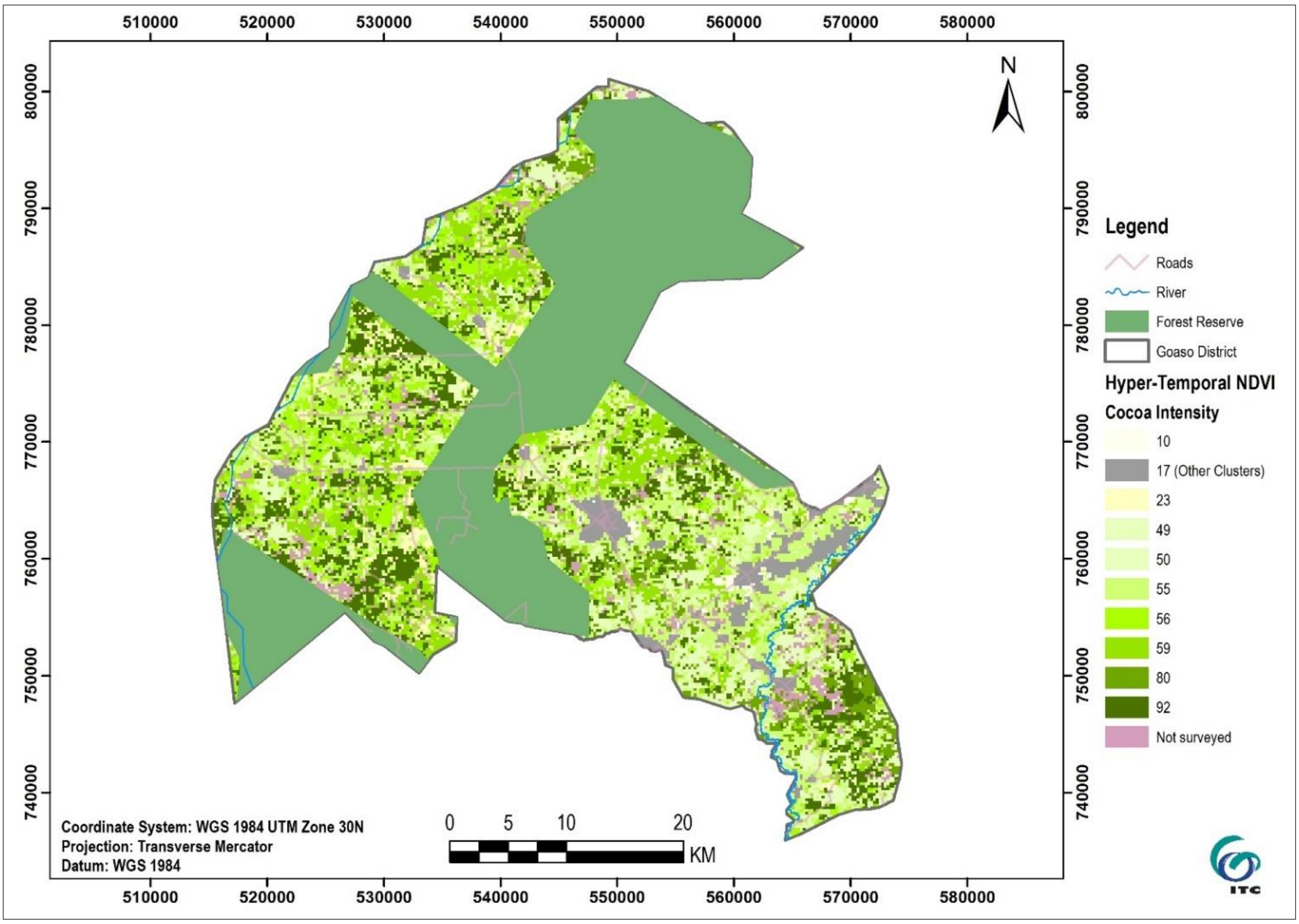


Figure 21: Hyper-Temporal NDVI Based Cocoa Map for Goaso Cocoa District

Accuracy Assessment of Hyper-Temporal Based Cocoa Map

The results of the correlation analysis performed for the hyper-temporal based cocoa map and actual cocoa polygons is presented in figure 23. Table 4 shows details of the various NDVI strata intensities for cocoa and their corresponding area in the five clusters (see figure 4).

From the correlation analysis, Adjusted R^2 value of 0.2506 was obtained, indicating a low relationship with about 25% of the variability in the cocoa fields polygons explained by the cocoa map stratification.

Figure 22 shows a part of the actual cocoa polygons overlaid with some geotagged photos taken during the field work in Goaso.

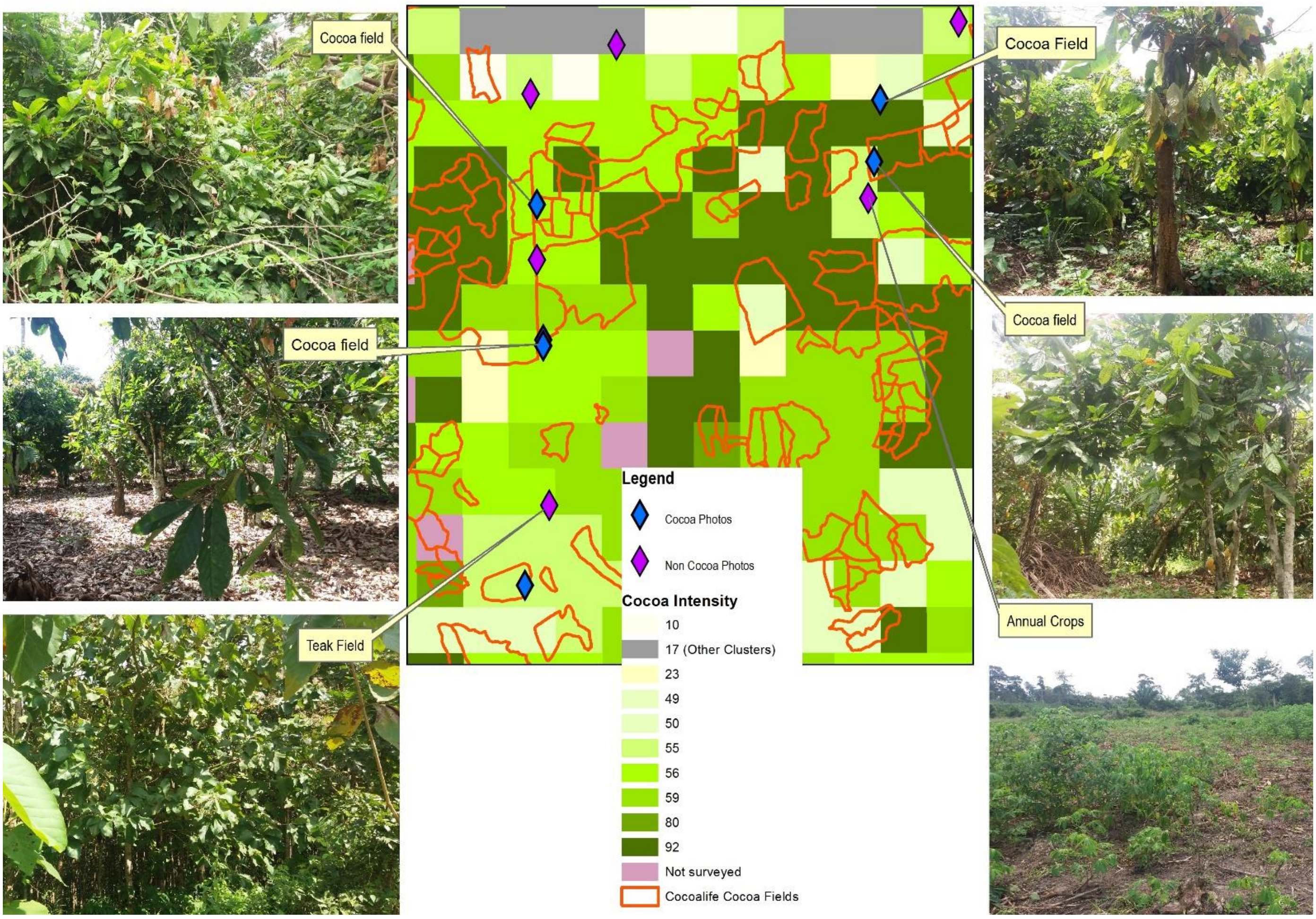


Figure 22: Hyper-Temporal Method Cocoa Map with Cocoa fields polygons and Geotagged Photos

Table 4: Validation of Hyper-Temporal Method Cocoa Map, Goaso

NDVI-Based Cocoa Fraction	Sum of 5 Clusters Area	CocoaLife Confirmed Area		Unconfirmed Area	
(%)	(Ha)	(Ha)	Percentage	(Ha)	Percentage
10	306	16	5%	290	95%
17	458	7	2%	451	98%
23	2279	154	7%	2125	93%
49	1605	179	11%	1426	89%
50	2575	214	8%	2361	92%
55	5227	469	7%	6150	93%
56	5096	649	13%	4448	87%
59	6217	620	10%	5597	90%
80	753	42	6%	710	94%
92	6439	696	11%	5743	89%

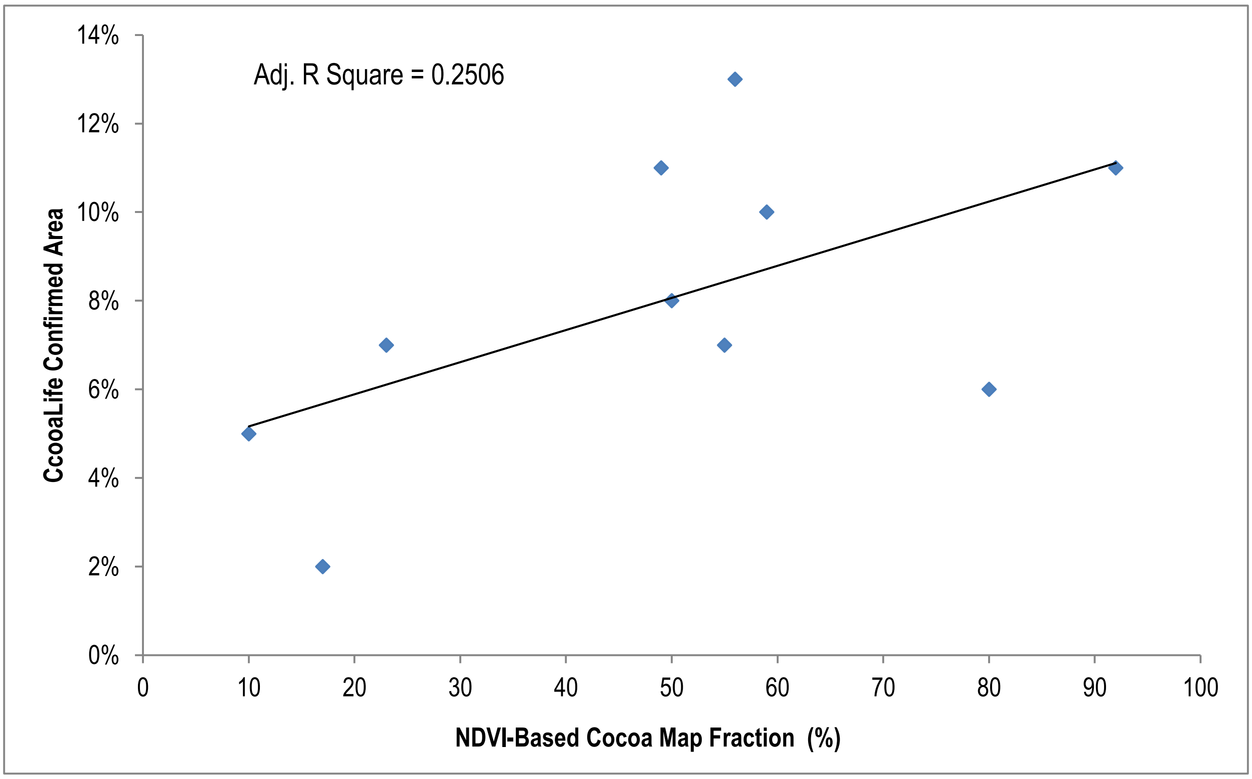


Figure 23: Accuracy results of Hyper-temporal NDVI Based Cocoa Map

3.2. High Spatial Resolution Cocoa Mapping Results

Image Classification

The results of the ISODATA unsupervised classification of the Sentinel-2 image (Image taken in January, 2016) into 20 map classes is shown in figure 24. A visual inspection of the 20 map classes shows a clear class discrimination of the different cover types present in the image.

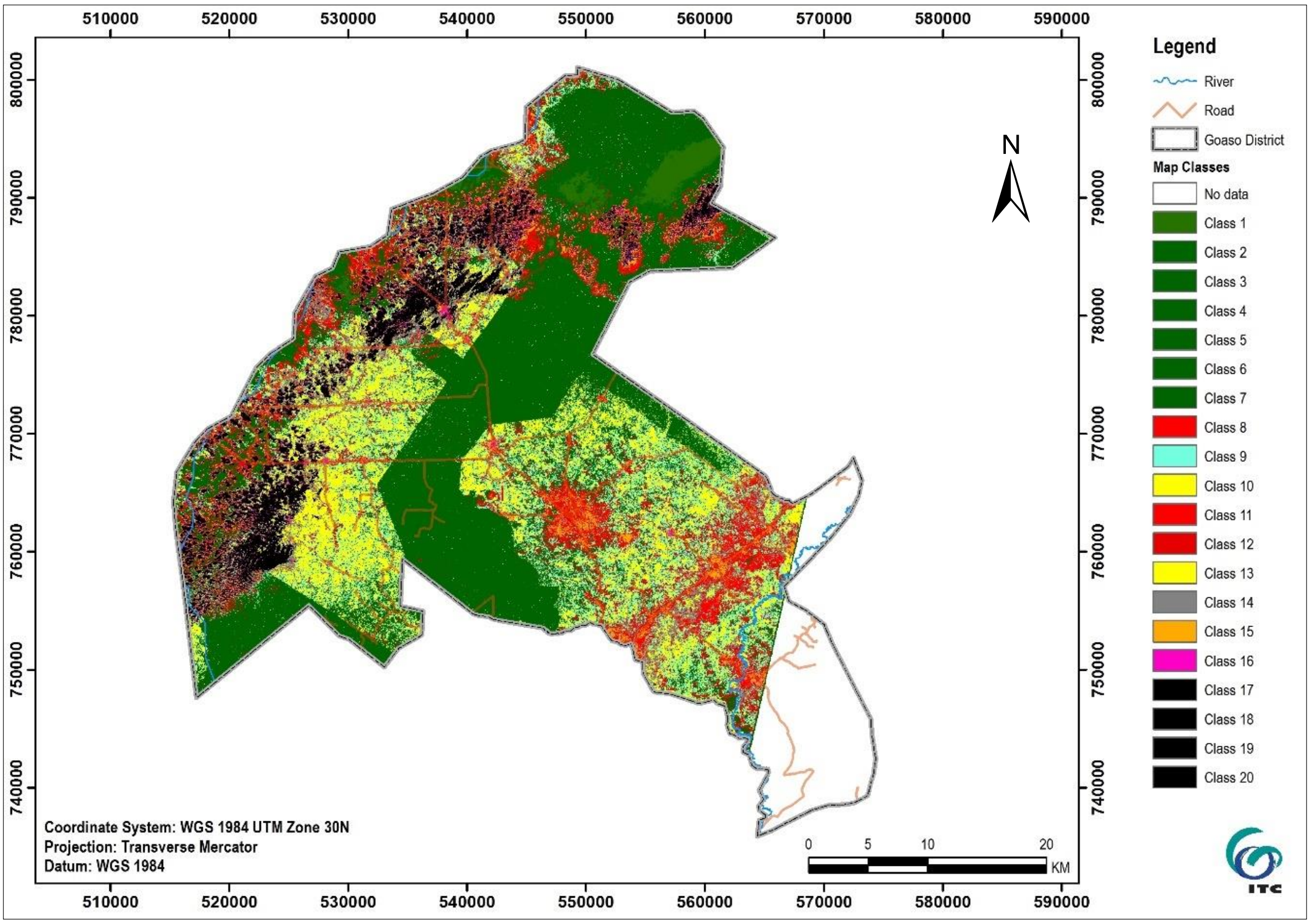


Figure 24: Classified Sentinel 2 image of Goaso Cocoa District.

Classified Sentinel Image Link with Field Data

The results of the cross between classified sentinel image and the field data is presented in table 5. Based on the estimated intensities of cocoa and non-cocoa within map classes in the table, results show no single map class has a 100% cocoa cover. This result suggests then that cocoa is not a separate cover type on its own rather a mixture. Appendix 3 shows the details of the non-cocoa cover types that are within the same classes with the cocoa.

The final cocoa map presented in figure 25, therefore shows the cocoa intensities only in the Goaso area.

Table 5: Sentinel map classes and field data cross table

Map Classes	Cocoa Intensity	Non-Cocoa Intensity	Samples
Class 7	14%	86%	14
Class 9	30%	70%	10
Class 10	80%	20%	15
Class 11	23%	77%	31
Class 13	80%	20%	15
Class 14	60%	40%	35
Class 15	17%	83%	12
Classes 6, 8 & 16	20%	80%	10
Class 1-5 & 12	Not surveyed	Not surveyed	0
Class 17-20	Clouds/Shadows	Clouds/Shadows	14

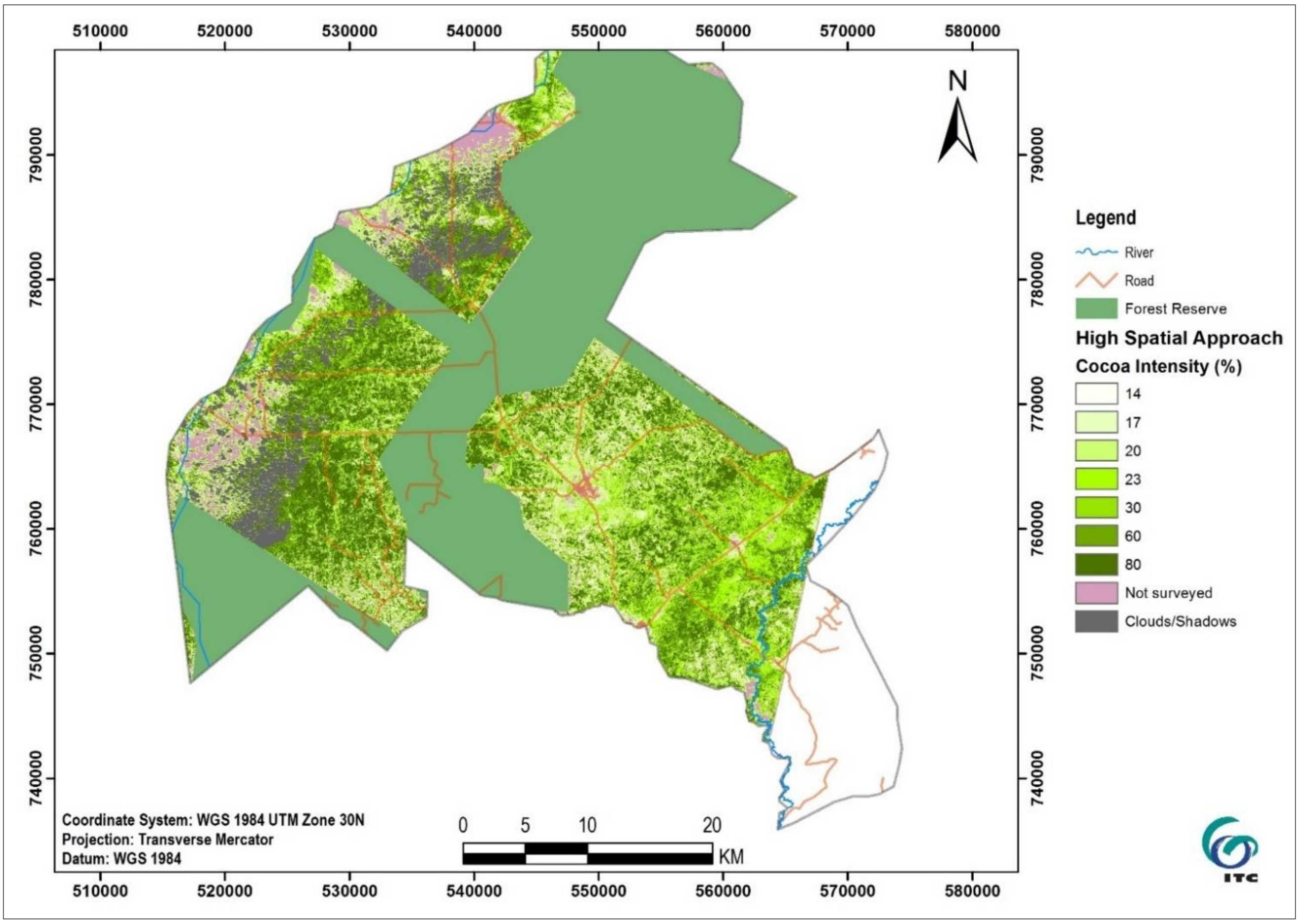


Figure 25: High Spatial Resolution Based Cocoa Map, Goaso

Accuracy Assessment of High Spatial Resolution Based Cocoa Map

The results of the correlation analysis performed for the high spatial resolution cocoa map and actual cocoa polygons is presented in figure 27. Table 6 shows details of the various cocoa intensities and their corresponding area in the five clusters.

From the correlation analysis, an Adjusted R^2 value of 0.684 was obtained, indicating that the 68% of the variability in the cocoa fields polygons is explained by the cocoa map stratification. This shows a strong correlation between cocoa area estimated and cocoa field polygon area.

Figure 26 shows a few of the cocoa fields overlaid with geotagged photos.

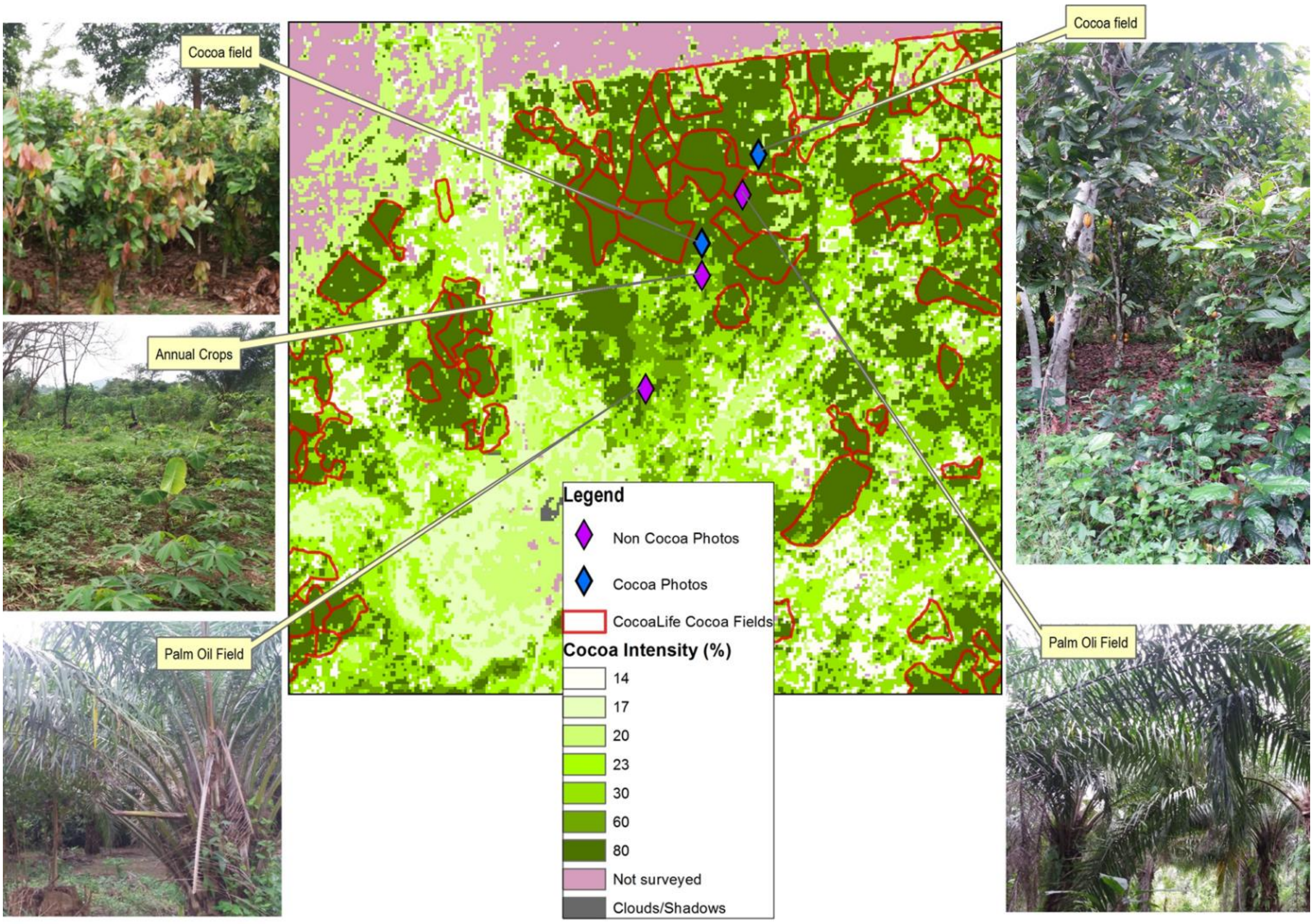


Figure 26: High Spatial Resolution Based Cocoa Map with Cocoa fields polygons & Geotagged Photos

Table 6: Validation of High Spatial Resolution Method Cocoa Map, Goaso

High Spatial Resolution Map Cocoa Fraction	**Sum of 5 Clusters Area**	**CocoaLife Confirmed Area**		**Unconfirmed Area**	
(%)	(Ha)	(Ha)	(%)	(Ha)	(%)
14	3,206	151	5%	3,055	95%
17	743	54	7%	688	93%
20	3,726	247	7%	3,479	93%
23	1,606	68	4%	1,538	96%
30	5,344	387	7%	4,956	93%
60	2,053	163	8%	1,890	92%
80	8,529	1,061	12%	7,468	88%

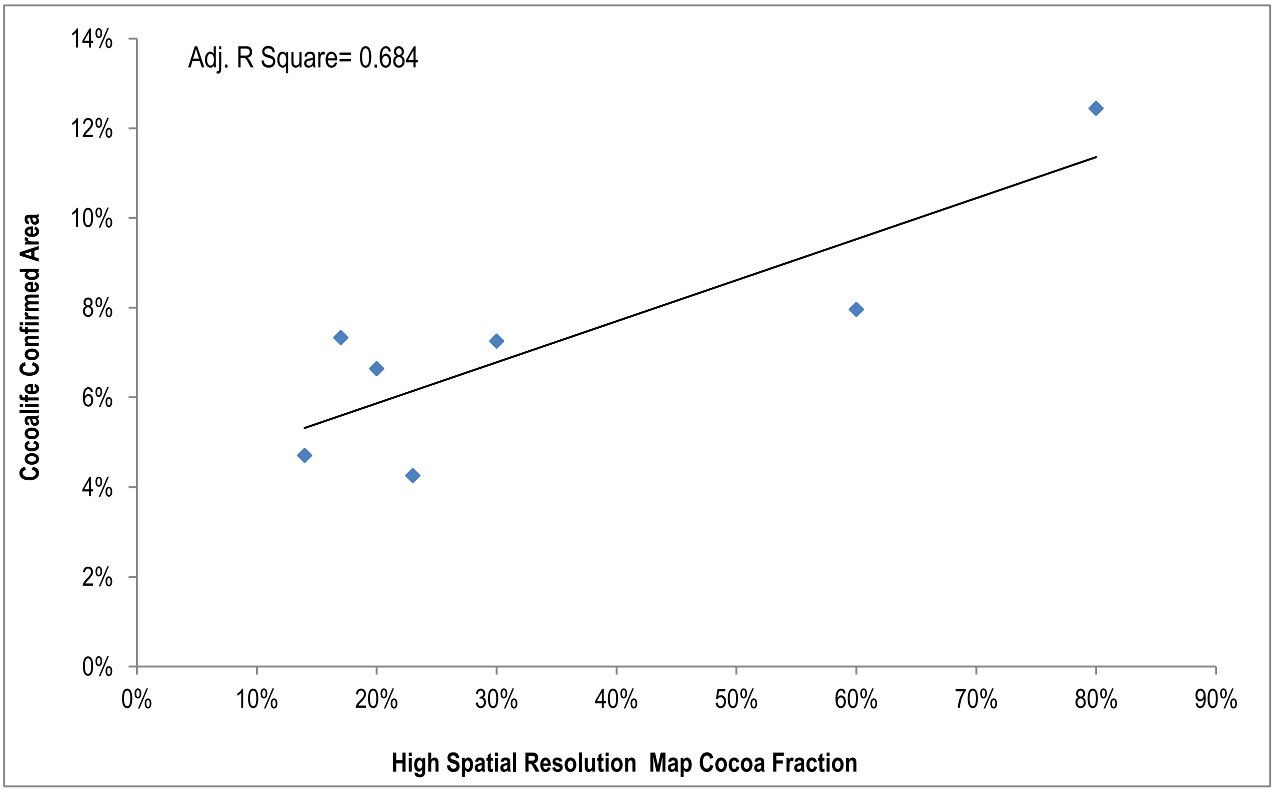


Figure 27: Accuracy results of High Spatial Resolution Based Cocoa Map

3.3. Integrated Appraoch Cocoa Mapping

For the first step of calculating cocoa intensities for the combined hyper-temporal based cocoa map and high spatial resolution based cocoa are, detailed results are attached in appendix 4. Based on the generated cocoa intensities the total cocoa area estimated was 23,573 hectares for Goaso. This value compared to those of the underlying approaches as shown in table 7, is under estimated.

Next step taken by dividing integrated estimate of 23,573 by the average cocoa area of the two approaches gave a coefficient of 2.084. This coefficient was multiplied with generated cocoa fractions to derive the actual cocoa fractions for the integrated product (see appendix 7 for matrix).

The final results of this approach is used to make the final cocoa map presented in figure 28.

Table 7: Summary of integrated map analysis

Cocoa Mapping Method	**Total Area Observed (Goaso)**	**Estimated Area of Cocoa**	
		Ha	**Percent**
Hyper-Temporal NDVI Based Cocoa Map	100, 385	58,743	59%
High Spatial Resolution Cocoa Map	100, 386	39,511	40%
Integrated Approach Cocoa map	100, 387	45,675	47%

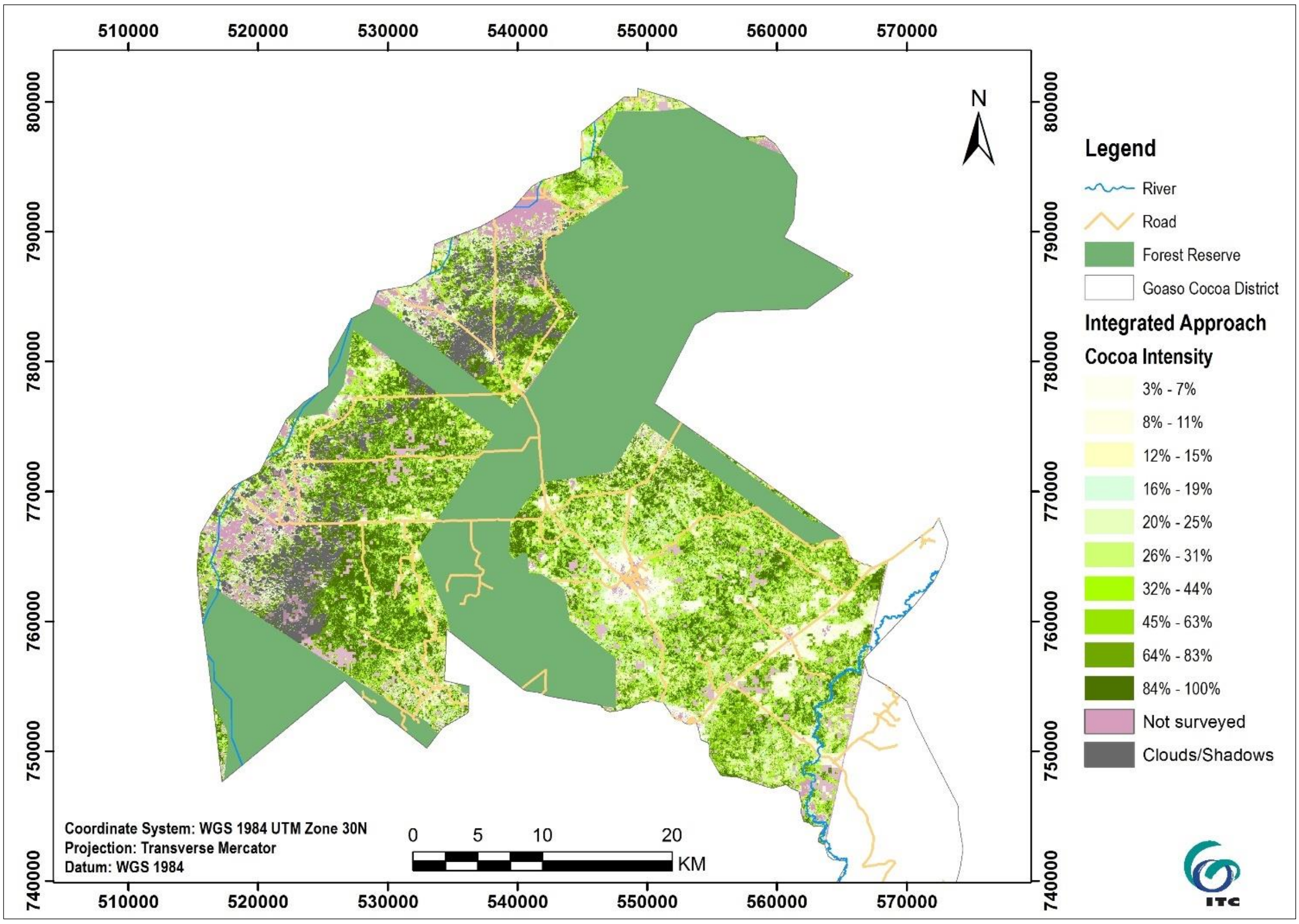


Figure 28 Integrated Approach Cocoa Map

Accuracy Assessment of Integrated Approach Cocoa Map

Integrated cocoa map and actual cocoa fields correlation results are presented in figure 29. Table 8 shows the details of the figures correlated.

From the correlation analysis, an Adjusted R^2 value of 0.3758 was obtained. This shows that the integrated shows improvement relative to hyper-temporal approach alone but poor when compared with the results to the high spatial resolution approach.

Table 8: Validation of Integrated Cocoa Map, Goaso

Integrated Cocoa Map Fraction	Sum of Area of Five Clusters	CocoaLife Confirmed Area		Unconfirmed Area	
	(Ha)	(Ha)	Percent	(Ha)	Percent
3%	35	0	1%	34	99%
4%	20	1	3%	20	97%
4%	41	1	3%	40	97%
5%	50	1	2%	49	98%
5%	8	0	2%	8	98%
6%	91	0	1%	90	99%
6%	51	2	5%	49	95%
7%	270	7	3%	263	97%
7%	123	2	1%	121	99%
8%	90	1	1%	89	99%
10%	233	9	4%	224	96%
11%	12	1	4%	11	96%
11%	162	3	2%	159	98%
13%	28	3	12%	24	88%
14%	177	15	8%	162	92%
14%	530	31	6%	499	94%
15%	326	12	4%	314	96%
16%	548	19	4%	529	96%
16%	376	20	5%	356	95%
17%	35	3	9%	32	91%
17%	696	36	5%	660	95%
17%	55	8	14%	47	86%
18%	55	4	8%	51	92%
19%	170	8	5%	162	95%
20%	112	12	11%	100	89%
20%	212	20	9%	193	91%
21%	251	16	6%	235	94%
21%	52	4	7%	48	93%
21%	20	1	3%	20	97%
23%	629	37	6%	592	94%
23%	646	43	7%	603	93%
23%	75	9	12%	66	88%
24%	139	5	4%	134	96%
25%	731	46	6%	684	94%
26%	389	9	2%	380	98%
27%	775	49	6%	726	94%
28%	219	8	4%	211	96%

28%	19	2	9%	17	91%
29%	110	5	5%	105	95%
31%	195	21	11%	173	89%
31%	525	33	6%	492	94%
33%	88	11	13%	77	87%
33%	65	4	7%	61	93%
34%	814	49	6%	765	94%
35%	661	44	7%	617	93%
37%	1,181	95	8%	1,086	92%
38%	1,534	136	9%	1,397	91%
44%	209	13	6%	196	94%
50%	132	5	4%	127	96%
58%	1,094	98	9%	995	91%
61%	101	15	15%	86	85%
63%	106	11	10%	95	90%
69%	429	25	6%	404	94%
70%	445	41	9%	404	91%
74%	328	20	6%	309	94%
82%	259	35	13%	224	87%
83%	634	88	14%	546	86%
92%	1,085	106	10%	978	90%
93%	1,326	174	13%	1,152	87%
98%	1,971	255	13%	1,716	87%
100%	2,601	322	12%	2,279	88%

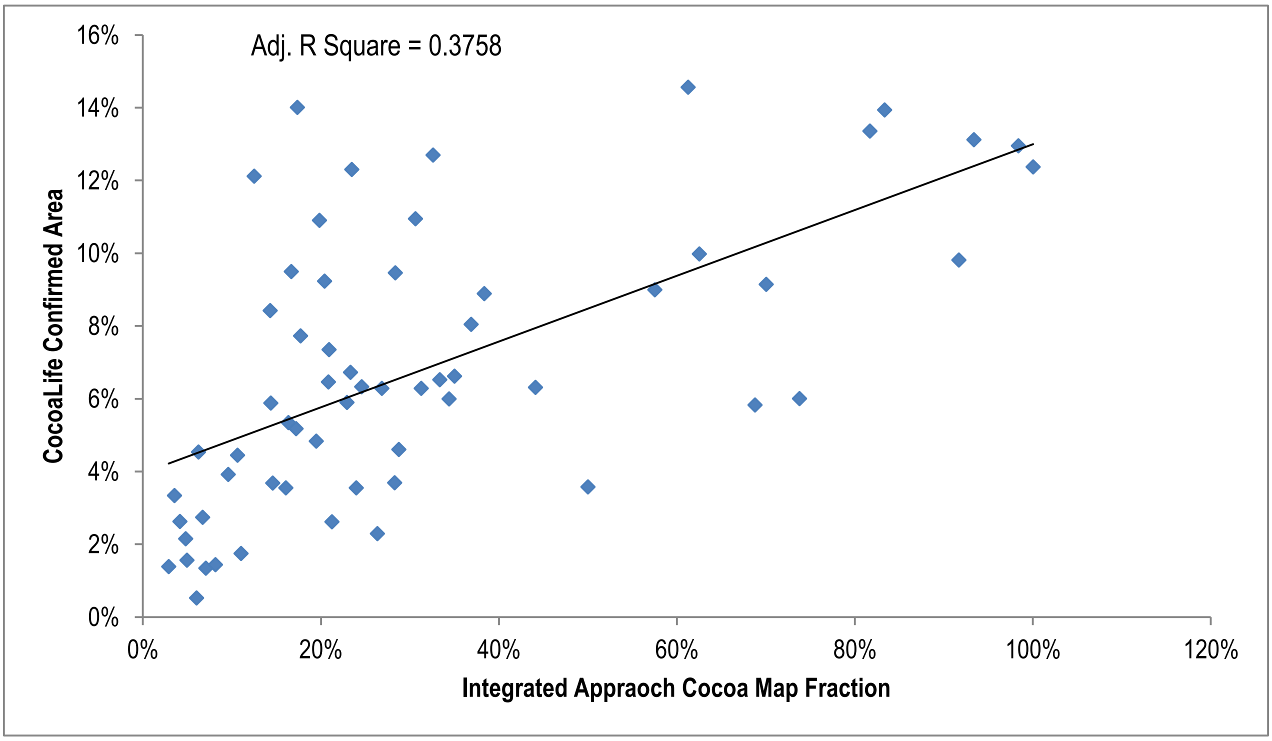


Figure 29: Accuracy assessment for Integrated Cocoa Map

3.4. Accuracy Comparison & Hypothesis Testing

Having assessed the accuracy of the three cocoa maps from each approach, this section presents an accuracy comparison and tests the hypothesis of this research. Table 9 presents the results of the comparison of the accuracy measured by the obtained Adjusted R^2 value. The value of the Adjusted R^2 shows the level by which each cocoa map stratification was able to explain the variability in the actual cocoa fields polygons and strength of relationship.

Table 9: Accuracy Comparison Table

Cocoa Mapping Method	Adjusted R^2
Hyper-Temporal NDVI Based Method	0.251
High Spatial Resolution Method	0.684
Integrated Approach	0.376

From table 9, of the three approaches, the hyper-temporal recorded the lowest accuracy with an Adjusted R^2 value of 25%. The high spatial resolution methods recorded the highest with an Adjusted R^2 value of 68% showing a relatively strong correlation with cocoa fields polygons. The integrated approach R^2 value was in between with 38%.

The results also indicate that the integrated approach is relatively better than only the hyper-temporal NDVI method used separately.

Based on these results, the following answers the research question and hypothesis of this study which states;

Research Question

Which of the cocoa maps have the highest accuracy?

i. Hyper-temporal NDVI-based Cocoa map of Goaso?
ii. High spatial resolution based cocoa map (baseline-method) of Goaso?
iii. The integrated product of Goaso?

The high spatial resolution cocoa map of Goaso have the highest accuracy based on the Adjusted R^2 obtained.

Research Hypothesis

The Adjusted R^2 of the cocoa map produced by the integrated approach explains 10% more variability of the crop area derived from actual cocoa fields polygons than the cocoa maps of the underlying approach used separately.

This hypothesis is rejected.

4. DISCUSSION

4.1. Mapping Cocoa By Hyper-Temporal NDVI Procedure

In this study, the hyper-temporal NDVI procedure successfully identified the broad land cover types containing cocoa on basis of clustered NDVI-classes for Ghana. However, distinguishing cocoa as a separate land cover performed poorly. This was evident in the cocoa map itself and the relatively low mapping accuracy of 25% of Adjusted R^2 recorded when correlated with actual cocoa polygons.

Regarding progress, the method facilitated mapping of cocoa across 38 cocoa districts in Ghana (figure 18), thus proving its utility for wide area mapping. This provides baseline information regarding where cocoa is grown within the districts and in what quantity evident in cocoa map of Goaso made using this approach (see figure 21). Such information can reduce the amount of time used in cocoa crop surveys by focusing attention on strata with the crop.

It is necessary to note that cocoa is a perennial tree crop, like rubber and coffee. Whilst a number studies have mapped perennial crops like rubber, coffee (Bispo et al., 2014; Dong et al., 2013; Fan et al., 2015) using hyper-temporal datasets, none has been done for cocoa. Discussions and comparisons of results for this method are therefore made in relation to findings for these crops with similar properties like cocoa.

The variability of NDVI profiles of the tree crops unlike annual food crops vary with age, as age strongly affects the tree canopy density. (Li & Fox, 2012). Young cocoa trees regardless of variety grown require shade for growth and so mostly cultivated alongside with trees often mixed with food crops, an example is shown in figure 30.

Figure 30: Young cocoa trees mix with shade trees and food crops. (Photo taken during fieldwork).

Even after they have matured, cocoa trees canopy covers are sometimes obstructed by canopy covers of shade trees i.e. viewing from space with optical images in the observed area (see figure 31).

Figure 31: Matured cocoa trees mix with shade trees. (Photo taken during fieldwork).

This shows cocoa trees grows in a complex landscape with a heterogeneous mix of land cover types which affects NDVI signal recording. It may be relatively easy to discriminate annual food crops by their temporal NDVI behaviour as demonstrated in studies by Kahubire (2002) using the same method and location. However, the ability of hyper-temporal NDVI to distinguish natural forest trees from tree crops like rubber with similar spectral confusion with forest was found to give low accuracy results (Dong et al., 2013).

In this study, the examined temporal profiles of the NDVI-based cocoa strata in figure 20 show to coincides with seasonal leaf flushing periods in matured cocoa trees noted by Opoku-Ameyaw et al (2010). This suggests that the young cocoa trees were not captured. A look at the selected NDVI Cluster cocoa map for Ghana (see figure 18) could not distinguish the forest reserves. This show that at these two scales cocoa trees could not be distinguished from natural forest trees.

A closer look at the behaviour of NDVI pixels values in an actual forest reserve and cocoa field in the study area is presented in figure 32 for visual examine phenological characteristics. It is evident from the profiles that are no clear differences to distinguish. Both seem to share relatively high NDVI pixel values indicating presence of dense vegetation cover. This is not surprising when looking at all photos (see figure 22, 26, 30 & 31) taken of the cocoa land cover. Also, the phenological characteristics observed over the five-year period seem to be same for the seven pixels taken at different locations within the land covers.

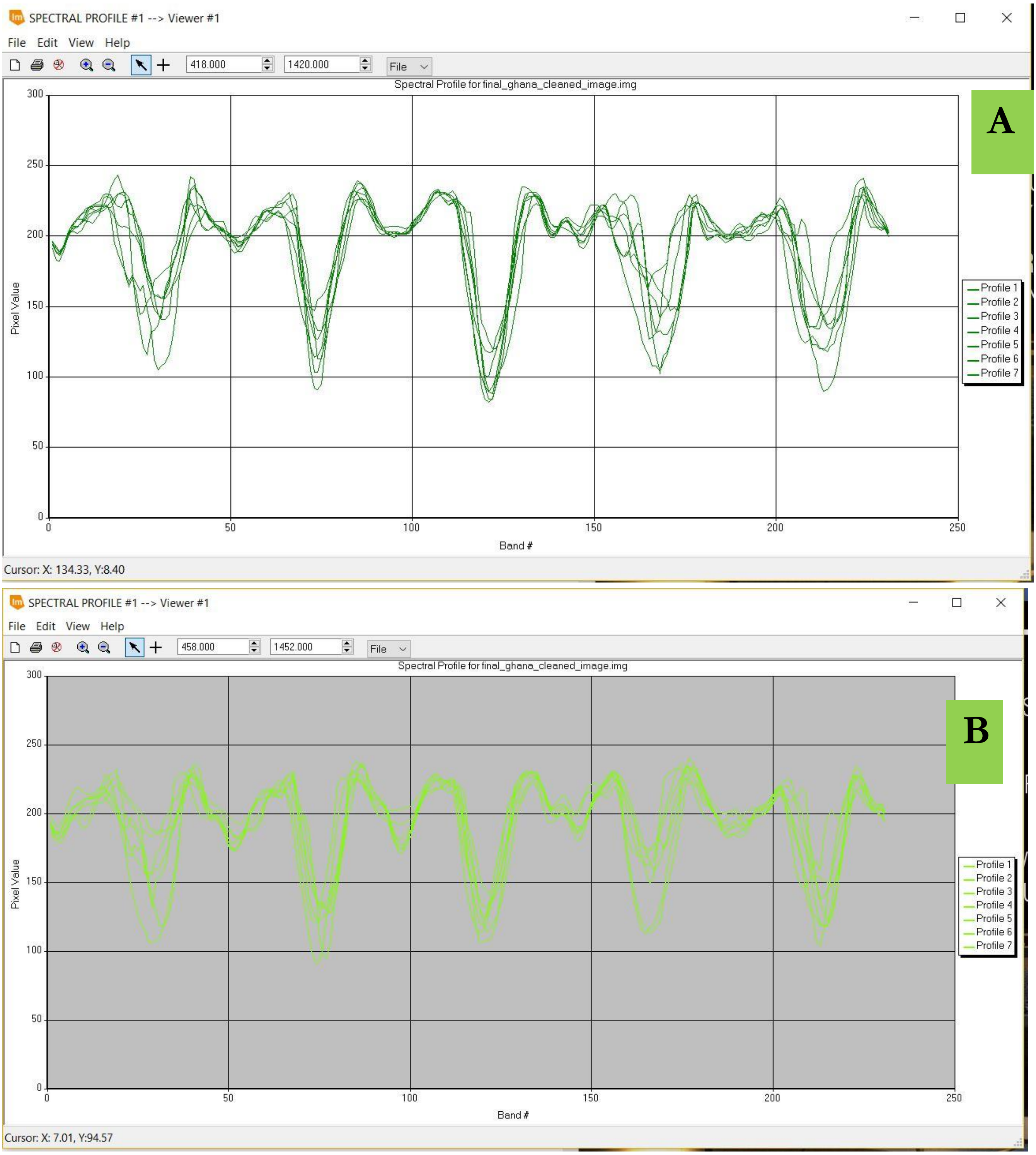


Figure 32: Spectral profile behaviour of a Pixel in NDVI Time Series (A) natural forest reserves (B) actual cocoa fields (Screen dumps from Erdas)

One reason for the low performance may be simply due to small field sizes and fragmented nature of the fields. As noted, the average field size of cocoa farms in Ghana is 2.5 ha (Opoku-Ameyaw et al., 2010). Compared to the 250-m pixel size (6.25 ha) of MODIS, it is far smaller and difficult to resolve fields at this resolution. In addition to the small field sizes, cocoa fields are established on fragmented pieces of land in Ghana (Acheampong et al., 2014). This implies that the resolution of MODIS data used is not appropriate to capture the cocoa fields of Ghana using this method separately.

Another reason also may be due to the use COCOBOD statistics to identify cocoa strata. That is, cocoa NDVI classes were identified through applying result of multiple linear regression obtain the cocoa strata Since statistics received were incomplete as surveys are still ongoing, extrapolations were made. The errors recorded in the coefficient, which was expected to range from 0 to 1 but gave otherwise higher values, indicates that the extrapolations made on the statistics could be wrong.

Nevertheless, the results of hyper-temporal NDVI method cocoa mapping provides basis for further studies.

4.2. High Spatial Resolution Based cocoa mapping

The high spatial resolution crop mapping procedure utilizing data from Sentinel-2 recorded the highest mapping accuracy of 68% in Adjusted R^2 value in this research Compared to the hyper-temporal NDVI mapping procedure, details at the field level were appropriately captured and mapped. However, cocoa land cover was not distinguished as a separate land cover type but a mix just like the hyper-temporal NDVI approach.

This high performance can be attributed mainly to the spatial resolution of 10-m (0.01 ha) which was able to capture the small field sizes (average size in Goaso is 2.2 ha) and also the isolated patches of the crop in the final cocoa map (see figure 26) more accurately. This confirms the findings of studies by Liu et al., (2014) and Conrad et al. (2014) among others that using high spatial resolution data to map crops with small field sizes and within heterogeneous landscapes gives a more accurate result than mapping low spatial resolution.

Regardless of this progress, the main issue still remains that this type of data may not be always available due to predetermined acquisition strategies and clouds obstruction (Liu et al., 2014) plus cost factor. Besides this mono-temporal crop classification is only valid for a short period of time and aspects of the crop phenology is not captured for further studies. Also, mapping the same crop over the same area using a another high resolution image taken at a different season may not give same results since crops reflectance properties varies with time (de Bie et al., 2008).

4.3. Integrated Approach

The integrated approach gave the second highest mapping accuracy in the procedure to map cocoa. But the performance was significantly lower than that of the cocoa map produced using high spatial resolution procedure. The greatest gain however is that, it gives a more reliable crop area discrimination since is based on two sensors and utilizes both time and space properties to map the crop. Another advantage is that, the crop map is fully reproducible at underlying resolutions (i.e. at 250-m and 10-m resolutions). That is, having simultaneous overview of cocoa in this case at both regional scale made possible by hyper-temporal NDVI resolution and the field scale with high spatial resolution.

This procedure maybe an easier and reliable option to viewing detailed crop composition within coarse resolution pixel footprint than subpixels fractions estimation using temporal unmixing approaches (Lobell & Asner, 2004; Ozdogan, 2010). Due to the low performance of hyper-temporal NDVI procedure to clearly capture cocoa, it affected the mapping accuracy of the integrated method. To improve the mapping accuracy of using this integrated approach requires that the ideal temporal NDVI profiles of a specific crop be clearly distinguished and an appropriate high spatial resolution image used to map details.

One benefit of the present approach is that, it saves the time in crop area survey in areas where the target crop may not be present and reduce the cost of purchasing high resolution images for wide area mapping. Since coarser resolution gives various strata with estimated intensities that allows decision to be made regarding which strata to survey and an appropriate high spatial image procured for just that area of interest.

The method is expected to work with any two sets of hyper-temporal NDVI and high spatial data to map any crop. A limitation however is the hardware limitations in the form of large computer memory and disk space requirements for storing and processing of hyper-temporal datasets. In addition, the combined shapefile gets slower in opening when the volume of data is high as a result area coverage.

.

5. CONCLUSION

Overall, the applicability of hyper-temporal NDVI data from MODIS with high spatial resolution crop used simultaneously (integrated) to map crops is feasible from the case of cocoa. It was encouraging to see that even though the hyper-temporal method could not perform very well to map the crop area of cocoa, it was able to identify at the principal land cover types containing the crop and cropping intensities estimated.

Using the underlying approaches of the integrated separately, the hyper-temporal NDVI procedure by itself was found to perform poorly in its ability map cocoa crop area in the study area. This showed in the overall mapping accuracy of the generated map with an Adjusted R^2 value of 25%. The high spatial resolution on the other hand was able to map out cocoa intensities more accurately with an Adjusted R^2 value of 68% when correlated with actual cocoa polygons in the observed area.

The integrated method, able to map cocoa area using the strengths of the two approaches; time and space with an Adjusted R^2 value of 38%. This gave a more reliable crop area information than just using procedures separately since is based on two sensors. Though not tested in other cases, the present method ability to utilize high resolution to review details within a low resolution pixel footprint provides a unique opportunity to solving mixed pixels' problems in hyper-temporal crop area mapping. Testing this ability and the utility of the method with a different set of hyper-temporal NDVI and high spatial resolution data or crop is recommended for research.

Timely, detailed and reliable crop area maps are relevant tools that can make a significant difference in the global efforts towards achieving food security and environmental sustainability. This research contributes to the efforts of utilizing the unique capabilities of remote sensing in providing such crop information needed.

LIST OF REFERENCES

Acheampong, E., Dawoe, K. E., Bosu, P. P., & Asante, A. W. (2014). *Moving Forward With Redd + in Ghana: a Closer Look At Shade Systems , Crown Cover , Carbon Stocks and Socio- Economic Dynamics of Smallholder Cocoa Agroforestry Systems*. Retrieved from http://theredddesk.org/sites/default/files/resources/pdf/snv_reap_ghana_lr_singlepages_1.pdf

Ali, A., de Bie, C. A. J. M., & Skidmore, A. K. (2013). Detecting long-duration cloud contamination in hyper-temporal NDVI imagery. *International Journal of Applied Earth Observation and Geoinformation*, *24*(1), 22–31. http://doi.org/10.1016/j.jag.2013.02.001

Ali, A., de Bie, C. A. J. M., Skidmore, A. K., Scarrott, R. G., Hamad, A., Venus, V., & Lymberakis, P. (2013). Mapping land cover gradients through analysis of hyper-temporal NDVI imagery. *International Journal of Applied Earth Observation and Geoinformation*, *23*(1), 301–312. http://doi.org/10.1016/j.jag.2012.10.001

Almeida, A.-A. F. de, & Valle, R. R. (2007). Ecophysiology of the cacao tree. *Brazilian Journal of Plant Physiology*, *19*(4), 425–448. http://doi.org/10.1590/S1677-04202007000400011

Anglaaere, L. C. N., Cobbina, J., Sinclair, F. L., & McDonald, M. A. (2011). The effect of land use systems on tree diversity: Farmer preference and species composition of cocoa-based agroecosystems in Ghana. *Agroforestry Systems*, *81*(3), 249–265. http://doi.org/10.1007/s10457-010-9366-z

Asante-Poku, A & Angelucci, F. (2013). Analysis of Incentives and Disincentives for Cocoa in Ghana. *Technical Notes Series*, (MAFAP, FAO, Rome.). Retrieved from http://www.fao.org/3/a-at551e.pdf

Asare, R. (2005). *Cocoa agroforests in West Africa: a look at activities on preferred trees in the farming systems*. *Forest & Landscape Working Papers no. 6*. Retrieved from http://orton.catie.ac.cr/repdoc/A3707i/A3707i.pdf

Asare, R. (2015). *The relationships between on-farm shade trees and cocoa yields in Ghana*. University of Copenhagen, Frederiksberg, Denmark.

Asilo, S., de Bie, K., Skidmore, A., Nelson, A., Barbieri, M., & Maunahan, A. (2014). Complementarity of two rice mapping approaches: Characterizing strata mapped by hypertemporal MODIS and rice paddy identification using multitemporal SAR. *Remote Sensing*, *6*(12), 12789–12814. http://doi.org/10.3390/rs61212789

Badhwar, G. D. (1984). Automatic corn soybean classification using Landsat. *Remote Sens. Environ.*, *14*(1–3), 31–37.

Beltran-Abaunza, J. . (2009). *Method Development to Process Hyper-Temporal Remote Sensing (RS) Images for Change Mapping*. University of Twente, Enschede, The Netherlands (Master's Thesis)). Retrieved from https://www.itc.nl/library/papers_2009/msc/gem/beltran-abaunza.pdf

Bispo, R. C., Lamparelli, R. A. C., & Rocha, J. V. (2014). Using fraction images derived from MODIS data for coffee crop mapping. *Engenharia Agrícola*, *34*(1), 102–111. http://doi.org/10.1590/S0100-69162014000100012

Boaitey, A. (2016). *Quantifying Spatial Indicators of Ecological Quality in a Cocoa Landscape in Goaso Forest.* University of Twente, Faculty of Geo-Information Science and Earth Observation. (Master's Thesis). Retrieved from http://www.itc.nl/library/papers_2016/msc/nrm/boaitey.pdf

Buechel, S. W., Philipson, W. R., & Philpot, W. D. (1989). The effects of a complex environment on crop separability with landsat TM. *Remote Sensing of Environment*, *27*(3), 261–271. http://doi.org/10.1016/0034-4257(89)90087-4

Carr, M. K. V., & Lockwood, G. (2011). The Water Relations and Irrigation Requirements of Cocoa (Theobroma Cacao L.): A Review. *Experimental Agriculture*, *47*(4), 653–676. http://doi.org/10.1017/S0014479711000421

Conrad, C., Dech, S., Dubovyk, O., Fritsch, S., Klein, D., Löw, F., … Zeidler, J. (2014). Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images. *Computers and Electronics in Agriculture*, *103*, 63–74. http://doi.org/10.1016/j.compag.2014.02.003

Conrad, C., Fritsch, S., Zeidler, J., Rücker, G., & Dech, S. (2010). Per-field irrigated crop classification in arid Central Asia using SPOT and ASTER data. *Remote Sensing*, *2*(4), 1035–1056. http://doi.org/10.3390/rs2041035

Crapper, P. F. (1980). Errors incurred in estimating an area of uniform land cover using Landsat. *Photogramm. Eng. Remote Sens.*, *46*(10), 1295–1301.

de Bie, C. A. . J. M., Nguyen, T. T. H., Ali, A., Scarrott, R., & Skidmore, A. K. (2012). LaHMa: a landscape heterogeneity mapping method using hyper-temporal datasets. *International Journal of Geographical*

Information Science, *26*(11), 2177–2192. http://doi.org/Doi 10.1080/13658816.2012.712126
de Bie, C. A. J. M. (2016). *Remote Sensing and Geographic Information Systems Analysis Methods to Support Food and Water Security Studies*. (2015-2017) (Class Lecture Notes). University of Twente, Enschede, The Netherlands.
de Bie, C. A. J. M., Khan, M. R., Toxopeus, A. G., Venus, V., Skidmore, A. K., & de Bie, C. A. (2008). Hypertemporal image analysis for crop mapping and change detection. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008*, *2*, 803–814. Retrieved from http://isprsserv.ifp.uni-stuttgart.de/proceedings/XXXVII/congress/7_pdf/5_WG-VII-5/13.pdf
Didan, K. (2015a). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. *NASA EOSDIS Land Processes DAAC*. http://doi.org/10.5067/MODIS/MOD13Q1.006
Didan, K. (2015b). MYD13Q1 MODIS/Aqua Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. *NASA EOSDIS Land Processes DAAC*. Retrieved from https://doi.org/10.5067/MODIS/MYD13Q1.006
Dong, J., Xiao, X., Chen, B., Torbick, N., Jin, C., Zhang, G., & Biradar, C. (2013). Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. *Remote Sensing of Environment*, *134*, 392–402. http://doi.org/10.1016/j.rse.2013.03.014
Duncan, J. M. A., Dash, J., & Atkinson, P. M. (2015). The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes. *Frontiers in Environmental Science*, *3*(August), 1–16. http://doi.org/10.3389/fenvs.2015.00056
Durgun, Y., Gobin, A., Van De Kerchove, R., & Tychon, B. (2016). Crop Area Mapping Using 100-m Proba-V Time Series. *Remote Sensing*, *8*(7), 585. http://doi.org/10.3390/rs8070585
ERDAS. (2003). *ERDAS Field Guide*. Retrieved from http://200.132.36.199/elodio/downloads/sr/Field_Guide_ERDAS_Imagine.pdf
ESA. (2015). SENTINEL-2 User Handbook, (1), 64. http://doi.org/GMES-S1OP-EOPG-TN-13-0001
Estrada, A., Coates-Estrada, R., & Meritt Jr., D. (1994). Non flying mammals and landscape changes in the tropical rain forest region of Los Tuxtlas, Mexico. *Ecography*.
Fan, H., Fu, X., Zhang, Z., & Wu, Q. (2015). Phenology-based vegetation index differencing for mapping of rubber plantations using landsat OLI data. *Remote Sensing*, *7*(5), 6041–6058. http://doi.org/10.3390/rs70506041
FAO. (1988). FAO/Unesco Soil Map of the World, Revised Legend, with corrections and updates. *World Soil Resources Report 60*. Retrieved from http://www.fao.org/fileadmin/user_upload/soils/docs/isricu_i9264_001.pdf
FAO. (2000). Guidelines for National FIVIMIS: Background and Principles. Retrieved from http://www.nfpcsp.org/agridrupal/sites/default/files/FIVIMS_National_Guidelines.pdf
Foerster, S., Kaden, K., Foerster, M., & Itzerott, S. (2012). Crop type mapping using spectral-temporal profiles and phenological information. *Computers and Electronics in Agriculture*, *89*, 30–40. http://doi.org/10.1016/j.compag.2012.07.015
Gallo, K., Ji, L., Reed, B., Eidenshink, J., & Dwyer, J. (2005). Multi-platform comparisons of MODIS and AVHRR normalized difference vegetation index data. *Remote Sensing of Environment*, *99*(3), 221–231. http://doi.org/10.1016/j.rse.2005.08.014
Gauch Jr, H. G., & Whittaker, R. H. (1981). Hierarchical classification of community data. *Journal of Ecology*, *69*, 537–557. Retrieved from http://www.science.fau.edu/biology/gawliklab/papers/GauchJrHGandRHWhittaker1981.pdf
GONZÁLEZ-Alonso, F., & Cuevas, J. M. (1993). Remote sensing and agricultural statistics: crop area estimation through regression estimators and confusion matrices. *International Journal of Remote Sensing*, *14*(6), 1215–1219. http://doi.org/10.1080/01431169308904405
Goward, S. N., & Huemmrich, K. F. (1992). Vegetation canopy PAR absorptance and the normalized difference vegetation index: An assessment using the SAIL model. *Remote Sensing of Environment*, *39*(2), 119–140. http://doi.org/10.1016/0034-4257(92)90131-3
Gumma, M. K., Nelson, A., Thenkabail, P. S., & Singh, A. N. (2011). Mapping rice areas of South Asia using MODIS multitemporal data. *Journal of Applied Remote Sensing*, *5*, 53547. http://doi.org/10.1117/1.3619838
Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213. http://doi.org/10.1016/S0034-4257(02)00096-2
ICCO. (2016). ICCO-Production of Cocoa Beans. *ICCO, Quarterly Bulletin of Cocoa Statistics*, *XLII*, *Coco*(3).

Retrieved from https://www.icco.org/statistics/production-and-grindings/production.html
Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., … Koetz, B. (2015). Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sensing*, *7*(9), 12356–12379. http://doi.org/10.3390/rs70912356
Jansen, L. J. ., & Di Gregorio, A. (2003). Land-use data collection using the "land cover classification system": results from a case study in Kenya. *Land Use Policy*, *20*(2), 131–148. http://doi.org/10.1016/S0264-8377(02)00081-9
Jönsson, P., & Eklundh, L. (2004). TIMESAT - A program for analyzing time-series of satellite sensor data. *Computers and Geosciences*, *30*(8), 833–845. http://doi.org/10.1016/j.cageo.2004.05.006
Kahubire, E. B. (2002). *SPATIAL DE-AGGREGATION OF CROP AREA STATISTICS USING REMOTE SENSING , GIS AND EXPERT KNOWLEDGE A Case Study of Ghana*. University of Twente, Faculty of Geo-Information Science and Earth Observation. (Master's Thesis).
Khan, M. R., de Bie, C. A. J. M., van Keulen, H., Smaling, E. M. A., & Real, R. (2010). Disaggregating and mapping crop statistics using hypertemporal remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *12*(1), 36–46. http://doi.org/10.1016/j.jag.2009.09.010
Lartey, L. L. (2013). *Mapping Cocoa Swollen Shoot Virus Disease Distribution in Western Region, Ghana*. Kwame Nkrumah University of Science and Technology (Master's Thesis). Retrieved from http://dspace.knust.edu.gh/bitstream/123456789/5364/1/Lilian Lucy Lartey.pdf
Leeuw, J. de, Georgiadou, Y., Kerle, N., Gier, A. de, Inoue, Y., Ferwerda, J., … Narantuya, D. (2010). The function of remote sensing in support of environmental policy. *Remote Sensing*, *2*(7), 1731–1750.
Li, Q., Wang, C., Zhang, B., & Lu, L. (2015). Object-Based Crop Classification with Landsat-MODIS Enhanced Time-Series Data. *Remote Sensing*, *7*(12), 16091–16107. http://doi.org/10.3390/rs71215820
Li, Z., & Fox, J. M. (2012). Mapping rubber tree growth in mainland Southeast Asia using time-series MODIS 250 m NDVI and statistical data. *Applied Geography*, *32*(2), 420–432. http://doi.org/10.1016/j.apgeog.2011.06.018
Liu, M. W., Ozdogan, M., & Zhu, X. (2014). Crop Type Classification by Simultaneous Use of Satellite Images of Different Resolutions. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(6), 3637–3649. http://doi.org/10.1109/TGRS.2013.2274431
Lobell, D. B., & Asner, G. P. (2004). Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of Environment*, *93*(3), 412–422. http://doi.org/10.1016/j.rse.2004.08.002
Maselli, F., & Rembold, F. (2001). Analysis of GAC NDVI Data for Cropland Identification and Yield Forecasting in Mediterranean African Countries. *Photogrammetric Engineering & Remote Sensing*, *67*(5), 593–602. Retrieved from https://pdfs.semanticscholar.org/1f7d/5eba601324ab23d94f9178ec36f0643761e1.pdf
MOFA. (2011). *Agriculture in Ghana. Facts and Figures (2010)*. Retrieved from http://mofa.gov.gh/site/wp-content/uploads/2011/10/AGRICULTURE-IN-GHANA-FF-2010.pdf
NASA. (n.d.). Moderate-resolution Imaging Spectroradiometer, 1–180. Retrieved from https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf
Nguyen, T. T. H., De Bie, C. A. J. M., Ali, A., Smaling, E. M. A., & Chu, T. H. (2012). Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam, through hyper-temporal SPOT NDVI image analysis. *International Journal of Remote Sensing*, *33*(2), 415–434. http://doi.org/10.1080/01431161.2010.532826
Opoku-Ameyaw, K., Baah, F., Gyedu-Akoto, E., Anchirina, V., Obiatey, H.K., Cudjoe, R., Acquaye, S., Opoku, S. Y. (2010). *Cocoa Manual: A source book for sustainable cocoa production*. Cocoa Research Institut of Ghana/Cadbury International Ghana. Retrieved from http://crig.org.gh/
Ozdogan, M. (2010). The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis. *Remote Sensing of Environment*, *114*(6), 1190–1204. http://doi.org/10.1016/j.rse.2010.01.006
Ozdogan, M., & Woodcock, C. E. (2006). Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment*, *103*(2), 203–217. http://doi.org/10.1016/j.rse.2006.04.004
Ozdogan, M., Yang, Y., Allez, G., & Cervantes, C. (2010). Remote sensing of irrigated agriculture: Opportunities and challenges. *Remote Sensing*. http://doi.org/10.3390/rs2092274
Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution*. http://doi.org/10.1016/j.tree.2005.05.011
Price, K.P., Egbert, S.L., Nellis, M.S., Nellis, Lee, R.Y., Boyce, R. (1994). Mapping land cover in a high

plains agro-ecosystem using a multidate Landsat thematic mapper modeling approach. *Trans. Kansas Acad. Sci*, *100*(1/2), 21–33.
Rice, R. A., & Greenberg, R. (2000). Cacao Cultivation and the Conservation of Biological Diversity. *AMBIO: A Journal of the Human Environment*, *29*(3), 167–173. http://doi.org/10.1579/0044-7447-29.3.167
Roumenina, E., Atzberger, C., Vassilev, V., Dimitrov, P., Kamenova, I., Banov, M., … Jelev, G. (2015). Single- and multi-date crop identification using PROBA-V 100 and 300 m S1 products on Zlatia Test Site, Bulgaria. *Remote Sensing*, *7*(10), 13843–13862. http://doi.org/10.3390/rs71013843
Sarkar, S., & Kafatos, M. (2004). Interannual variability of vegetation over the Indian sub-continent and its relation to the different meteorological parameters. *Remote Sensing of Environment*, *90*(2), 268–280. http://doi.org/10.1016/j.rse.2004.01.003
Singha, M., Wu, B., & Zhang, M. (2016). An object-based paddy rice classification using multi-spectral data and crop phenology in Assam, Northeast India. *Remote Sensing*, *8*(6). http://doi.org/10.3390/rs8060479
Smith, J. H., Ramey, D.B. (1982). A crop area estimator based on changes in the temporal profile of a vegetative index. *Proc. Survey Res. Methods Section*, 495–498.
Solano, R., Didan, K., Jacobson, A., & Huete, A. (2010). MODIS Vegetation Index User ' s Guide (MOD13 Series), *2010*(May), 1–42.
Tesema, M. A. (2015). *Estimating and Mapping Aboveground Carbon Stocks in Cocoa Plantation Using Worldview-2 Satellite Image and Lidar Data*. University of Twente Faculty of Geo-information Science and Earth Observation (ITC). Retrieved from https://ezproxy.utwente.nl:2315/library/2015/msc/nrm/tesema.pdf
Thomson, K. . (2003). World agriculture: towards 2015/2030: an FAO perspective. *Land Use Policy*, *20*(4), 375. http://doi.org/10.1016/S0264-8377(03)00047-4
Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. http://doi.org/10.1016/0034-4257(79)90013-0
Vaudour, E., Noirot-Cosson, P. E., & Membrive, O. (2015). Early-season mapping of crops and cultural operations using very high spatial resolution Pléiades images. *International Journal of Applied Earth Observation and Geoinformation*, *42*, 128–141. http://doi.org/10.1016/j.jag.2015.06.003
Wardlow, B. D., & Egbert, S. L. (2008). Large-area crop mapping using time-series MODIS 250??m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, *112*(3), 1096–1116. http://doi.org/10.1016/j.rse.2007.07.019

APPENDICES

Appendix 1: COCOBOD Cocoa Area Statistics with Extrapolated cocoa figures made for districts with incomplete survey results.

District	District Area (Ha)	Cumulative Area Surveyed (Ha)	District Area Surveyed	Area of Cocoa in Surveyed (Ha.)	Area of Cocoa in Surveyed (%)	Extrapolated (Ha)
Akontonbra	101,067	43,428	43%	39,498	91	91,920
Antoakrom	137,722	117,302	85%	51,079	44	59,971
Asamankese	78,604	78,604	100%	7,459	9	7,459
Asankragua	152,633	124,621	82%	74,398	60	91,120
Asikuma	190,640	138,975	73%	18,879	14	25,897
Bawdua	91,035	89,875	99%	3,738	4	3,786
Bechem	182,965	166,573	91%	17,712	11	19,455
Berekum	939,448	304,115	32%	8,551	3	26,414
Boako	155,265	87,537	56%	71,930	82	127,582
Dadieso	82,484	59,423	72%	46,875	79	65,066
Dormaa Ahenkro	486,006	315,612	65%	27,404	9	42,199
Dunkwa	173,282	161,091	93%	61,140	38	65,767
Enchi	178,563	74,286	42%	63,203	85	151,922
Essam	140,026	97,174	69%	84,724	87	122,086
Fosu	237,087	134,299	57%	49,292	37	87,018
Fumso	169,929	129,906	76%	64,238	49	84,029
Goaso	185,144	155,865	84%	76,956	49	91,412
Juaboso	160,087	69,561	43%	63,549	91	146,252
Juaso	452,366	189,019	42%	43,918	23	105,105
Jukwa	194,116	134,006	69%	38,931	29	56,394
Kade	108,440	108,440	100%	3,404	3	3,404
Mampong	551,716	201,393	37%	24,918	12	68,264
Nkawie	307,243	157,229	51%	50,574	32	98,828
Nkawkaw	257,206	103,019	40%	2,500	2	6,241
Nyakrom	128,729	128,729	100%	24,967	19	24,967
Obuasi	226,992	183,942	81%	59,281	32	73,155
Oda	150,693	138,169	92%	11,960	9	13,044
Offinso	141,389	141,389	100%	20,768	15	20,768
Osino	65,651	65,028	99%	2,939	5	2,967
Samreboi	279,018	142,455	51%	113,466	80	222,238
Sankore	164,727	131,698	80%	89,259	68	111,644
Sehwi Bekwai	150,097	122,551	82%	73,470	60	89,984
Sekondi	447,675	163,147	36%	33,045	20	90,676
Suhum	104,615	104,615	100%	1,393	1	1,393
Tafo	134,301	88,180	66%	3,739	4	5,695
Tarkwa	292,013	187,092	64%	40,342	22	62,966
Tepa	157,555	101,903	65%	43,429	43	67,146
Wasa Akropong	202,210	158,020	78%	59,585	38	76,247
***Jasikan/Papase	920,913	396,566	43%	15,744	4	36,560
***Kpeve/Hohoe	432,205	274,687	64%	12,617	5	19,853
***Oyoko	333,525	54,937	16%	4,163	8	25,271

*****Districts excluded from analysis (outside the coverage extent of the MODIS data)**

Appendix 2: NDVI Classes of Cluster 1 and Field Data

NDVI Clusters	**Area in Goaso (ha)**	**Samples**
Class 65	1710	10
Class 67	8376	6
Class 69	471	Not Surveyed
Class 70	1532	6
Class 71	448	1
Class 72	916	Not Surveyed
Class 73	2684	2
Class 74	176	1
Class 75	8897	30
Class 76	8620	3
Class 77	9853	13
Class 78	2120	7
Class 79	7694	Not Surveyed
Class 80	14039	43
Class 81	3489	2
Class 82	29778	34
Class 83	2524	1
Class 84	18894	14
Class 85	18959	16
Class 86	37507	25
Other Clusters	6455	17
Total		**231**

Appendix 3: Sentinel Map Classes and field data.

Map Classes	Goaso Area (ha)	Annual	Orange	Palm	Teak	Cocoa	Samples
Class 7	13,556	43%	36%	7%	0%	**14%**	14
Class 9	19,683	40%	0%	30%	0%	**30%**	10
Class 10	15,943	0%	0%	20%	0%	**80%**	15
Class 11	12,704	48%	0%	13%	16%	**23%**	31
Class 13	8,556	7%	7%	7%	0%	**80%**	15
Class 14	10,789	23%	0%	14%	3%	**60%**	35
Class 15	4,989	75%	0%	0%	8%	**17%**	12
Classes 6, 8 & 16	19,185	60%	0%	20%	0%	**20%**	10
Class 1-5 & 12	50,298	Not surveyed	Not surveyed	Not surveyed	Not surveyed	**Not surveyed**	0
Class 17-20	13,514	Clouds/Shadows	Clouds/Shadows	Clouds/Shadows	Clouds/Shadows	**Clouds/Shadows**	14
No Data							***75***
Grand Total	**169,216**						**231**

Appendix 4: Integrated cocoa fractions and area calculations. (Showing results for Equation 3 &4).

Integrated Cocoa Fractions	Total Area, Goaso (ha)	Cocoa Area (ha)
1.0%	133	1.33
2.0%	98	1.95
2.0%	228	4.56
2.0%	450	9.00
2.0%	71	1.43
3.0%	1,392	41.75
3.0%	209	6.28
3.0%	747	22.42
3.0%	945	28.36
4.0%	1,806	72.24
5.0%	1,011	50.53
5.0%	126	6.32
5.0%	618	30.92
6.0%	215	12.88
7.0%	1,228	85.94
7.0%	1,127	78.89
7.0%	1,438	100.69
8.0%	2,504	200.31
8.0%	1,151	92.06
8.0%	144	11.48
8.0%	2,268	181.48
8.0%	214	17.16
9.0%	374	33.64
9.0%	805	72.49
10.0%	360	35.98
10.0%	1,185	118.53
10.0%	2,096	209.55
10.0%	573	57.25
10.0%	477	47.70
11.0%	2,809	309.01
11.0%	2,789	306.79
11.0%	820	90.17
12.0%	1,102	132.26
12.0%	3,316	397.97
13.0%	3,724	484.11
13.0%	3,143	408.59
14.0%	1,248	174.68
14.0%	151	21.11
14.0%	569	79.69
15.0%	1,720	257.98
15.0%	2,040	306.05
16.0%	683	109.25
16.0%	735	117.57
17.0%	4,139	703.70
17.0%	1,740	295.85
18.0%	3,153	567.48
18.0%	6,147	1,106.41
21.0%	1,330	279.20
24.0%	1,105	265.09

28.0%	3,319	929.19
29.0%	736	213.49
30.0%	917	275.04
33.0%	2,701	891.26
34.0%	1,160	394.42
35.0%	1,439	503.75
39.0%	1,597	622.71
40.0%	2,390	956.18
44.0%	4,384	1,928.86
45.0%	2,930	1,318.66
47.0%	4,241	1,993.48
48.0%	323	154.96
55.0%	1,676	921.71
64.0%	1,029	658.71
74.0%	5,087	3,764.63
Total	100,385	23,573

Appendix 5: Hyper-temporal NDVI-Based Cocoa Area Calculation Using Cocoa fractions. (Showing results for Equation 4).

NDVI-Based Cocoa Fractions	**Total Area, Goaso (ha)**	**Area of Cocoa (ha)**
10	1,476	148
17	4,647	790
23	5,821	1,339
49	7,500	3,675
50	10,358	5,179
55	21,066	11,587
56	10,034	5,619
59	16,239	9,581
80	4,651	3,721
92	18,593	17,106
Total	100,385	58,743

Appendix 6: High spatial resolution cocoa area Calculation Using Cocoa fractions. (Showing results for Equation 4).

Sentinel Map Classes Cocoa Fraction	Total Area, Goaso (ha)	Area of Cocoa (ha)
14	12,780	1,789
17	4,820	819
20	18,375	3,675
23	12,139	2,792
30	18,678	5,604
60	10,213	6,128
80	23,380	18,704
Total	**100,385**	**39,511**

Appendix 7: Final Integrated cocoa fractions

Integrate cocoa fractions	Total Area, Goaso (ha)	Cocoa area (ha)
3%	133	4
4%	98	3
4%	228	9
5%	450	22
5%	71	4
6%	1,392	84
6%	209	13
7%	747	50
7%	945	67
8%	1,806	147
10%	1,011	97
11%	126	13
11%	618	68
13%	215	27
14%	1,228	176
14%	1,127	162
15%	1,438	210
16%	2,504	402
16%	1,151	188
17%	144	24
17%	2,268	390
17%	214	37
18%	374	66
19%	805	157
20%	360	71
20%	1,185	242
21%	2,096	437
21%	573	120
21%	477	101
23%	2,809	644
23%	2,789	651
23%	820	193
24%	1,102	264
25%	3,316	816
26%	3,724	982
27%	3,143	844
28%	1,248	353
28%	151	43
29%	569	164
31%	1,720	527
31%	2,040	638
33%	683	223
33%	735	245
34%	4,139	1423
35%	1,740	609
37%	3,153	1163
38%	6,147	2357
44%	1,330	586
50%	1,105	552
58%	3,319	1909

61%	736	451
63%	917	573
69%	2,701	1857
70%	1,160	812
74%	1,439	1062
82%	1,597	1304
83%	2,390	1993
92%	4,384	4020
93%	2,930	2736
98%	4,241	4172
100%	8,115	8115
Total	100,385	45,675