

# Master Thesis

---

## **Time Series Crop Monitoring of an Agricultural Scheme on a Plot Basis Using Complementary Remote Sensing**

Submitted by           Satarupa Chatterjee  
  
born on                 14.08.1988   in   India

Submitted for the academic degree of  
  
Master of Science (M.Sc.)

Date of Submission           08.09.2018

Supervisors                 Dr.rer.nat Nikolas Prechtel  
                                  Technical University of Dresden  
                                  Dr. Evangelia Drakou  
                                  University of Twente

## **Statement of Authorship**

Herewith I declare that I am the sole author of the thesis named

**“Time Series Crop Monitoring of an Agricultural Scheme on A Plot Basis Using Complementary Remote Sensing”**

which has been submitted to the study commission of geosciences today.

I have fully referenced the ideas and work of others, whether published or un-published. Literal or analogous citations are clearly marked as such.

Dresden 08/09/2018

Satarupa Chatterjee

## **Acknowledgement**

I would like to take this opportunity to thank my supervisor Dr. Nikolas Prechtel, for his continuous support and motivation. It wouldn't have been possible to complete this thesis on time without his continuous guidance and help.

I would also like to thank Ms. Juliane Cron for her support and motivation, and express my profound gratitude to my friends Ms. Karen Byrne, Ms. Natalie Barbosa and my mother for providing me with unfailing support and continuous encouragement throughout the phase of researching and writing this thesis. I am grateful for all your love and support.

Thank You.

# Contents

<b>LIST OF FIGURES.....</b>	<b>6</b>
<b>LIST OF TABLES.....</b>	<b>9</b>
<b>LIST OF ANNEXURE: .....</b>	<b>10</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>11</b>
<b>ABSTRACT .....</b>	<b>13</b>
<b>1. INTRODUCTION.....</b>	<b>14</b>
1.1 MOTIVATION.....	15
1.2 RESEARCH SIGNIFICANCE:.....	15
1.3 RESEARCH OBJECTIVES: .....	16
1.4 THESIS OUTLINE: .....	16
<b>2. BACKGROUND AND LITERATURE REVIEW .....</b>	<b>18</b>
CROP MONITORING USING REMOTE SENSING: .....	18
2.1 ELECTROMAGNETIC SPECTRUM AND VEGETATION: .....	18
2.2 VEGETATION INDICES:.....	21
2.2.1 NDVI ( <i>Normalised Difference Vegetation Index</i> ): .....	22
2.2.2 <i>Plant Phenology and NDVI</i> : .....	24
2.3 SENSOR FUSION: .....	26
2.3.1 <i>Fusion/Combining Methodologies</i> : .....	28
2.4 SOIL MOISTURE AND NDVI: .....	29
2.5 IMAGE PROCESSING: .....	30
2.5.1 MODIS:.....	30
2.5.2 PROBA-V:.....	31
2.6 INTERACTIVE VISUALISATION /ANALYTICAL TOOLS: .....	34
2.6.1 TIMESAT: .....	34
2.6.2 TimeStats: .....	34
2.6.3 PhenoSat: .....	34
2.6.4 SPIRITS: .....	35
<b>3. DATA AND STUDY AREA.....</b>	<b>36</b>
3.1 DATASETS:.....	36
3.1.1 MODIS Terra/ MODIS Aqua ( <i>collection 6</i> ): .....	36
3.1.2 MODIS Processing: .....	38
3.1.3 MODIS Land Surface Product Used ( <i>MOD09/ MYD09</i> ): .....	39
3.1.4 PROBA-V: ( <i>Wolters, 2018</i> ).....	42
3.1.5 PROBA-V Processing: <sup>13</sup> .....	44
3.1.6 PROBA-V Product Used :.....	46
3.1.7 SMAP: .....	48
3.1.8 SMAP Processing: .....	52
3.1.9 SMAP Data Used ( <i>R. Reichle 2012</i> ):.....	53
3.2 DATA AVAILABILITY:.....	55
3.3 STUDY AREA:.....	57



3.3.1	<i>Location</i> .....	57
3.3.2	<i>Climate and Soil:</i> .....	59
3.3.3	<i>Agriculture:</i> .....	60
<b>4.</b>	<b>METHODOLOGY .....</b>	<b>63</b>
4.1	DATA ACQUISITION: .....	63
4.1.1	<i>Image Download:</i> .....	63
4.1.2	<i>Data Format Conversion:</i> .....	64
4.1.3	<i>Reprojection:</i> .....	65
4.1.4	<i>Clipping:</i> .....	65
4.1.5	<i>Resampling:</i> .....	66
4.1.6	<i>Vector Data Preparation:</i> .....	66
4.2	DATA POST-PROCESSING: .....	67
4.2.1	<i>Cloud Masking:</i> .....	67
4.3	FINAL WORKFLOW: .....	72
<b>5.</b>	<b>RESULTS AND OBSERVATIONS.....</b>	<b>76</b>
5.1	MEAN MONTHLY NDVI: .....	76
5.2	DAILY NDVI:.....	78
5.3	NDVI ANOMALY: .....	79
5.4	INTERACTIVE VISUALISATION: .....	85
5.4.1	<i>Plot Based NDVI:</i> .....	87
5.4.2	<i>Region Based NDVI:</i> .....	88
5.5	PLOT BASED ANALYSIS: .....	89
5.5.1	<i>Multiplot Comparison-</i> .....	89
5.5.2	<i>Comparative Change detection:</i> .....	91
5.5.3	<i>Growth Period:</i> .....	93
5.6	SOIL MOISTURE AND NDVI: .....	94
5.7	TOOLS FOR PROCESSING: .....	100
<b>6.</b>	<b>CONCEPTUAL TOOL:.....</b>	<b>101</b>
6.1	TAB-1: MASKING .....	101
6.2	TAB-2: PROCESSING.....	103
6.3	TAB-3: PARAMETER.....	104
<b>7.</b>	<b>CONCLUSION .....</b>	<b>106</b>
	<b>REFERENCES .....</b>	<b>109</b>
	<b>TABLE.....</b>	<b>113</b>
	<b>ANNEXURE .....</b>	<b>118</b>

## List of Figures

FIGURE 2.1: REFLECTED SOLAR RADIATION RECORDED BY THE SENSOR. ....	18
FIGURE 2.2: SPECTRAL REFLECTANCE OF SOIL, VEGETATION AND WATER. (SOURCE: LILLESAND & KIEFER) .....	19
FIGURE 2.3. VEGETATION SPECTRAL REFLECTANCE IN DETAIL. (SOURCE: HTTP://WWW.MARKELOWITZ.COM).....	20
FIGURE 2.4. CELLULAR LEAF STRUCTURE AND ITS INTERACTION WITH ELECTROMAGNETIC ENERGY. (SOURCE: JEFF CARNS, NASA SCIENCE).....	20
FIGURE 2.5: NIR AND R INTERACTION AND NDVI CALCULATION.(SOURCE: NASA).....	23
FIGURE 2.6: START OF SEASON.....	24
FIGURE 2.7: END OF SEASON .....	24
FIGURE 2.8: TIME OF MAXIMUM OF NDVI PEAK. ....	25
FIGURE 2.9: GROWTH SEASON / DURATION. ....	25
FIGURE 2.10: AMPLITUDE .....	25
FIGURE 2.11: NDVI INTEGRAL .....	26
FIGURE 2.12 : GEOMETRIC APPROACH. (SOURCE: WOLTERS 2018) .....	32
FIGURE 2.13: RADIOMETRIC APPROACH FOR CLOUD MASK CREATION IN PROBA-V. (SOURCE: WOLTERS 2018) .....	33
FIGURE 3.1: GLOBAL LAND SURFACE REFLECTANCE IN TRUE COLOUR (MODIS-TERRA) ACQUIRED ON 02.09.2018. ....	37
FIGURE 3.2: GLOBAL LAND SURFACE REFLECTANCE IN TRUE COLOUR (MODIS-AQUA) ACQUIRED ON 02.09.2018.....	38
FIGURE 3.3: THE MODIS SINUSOIDAL GRID CONSISTS OF 460 NON-OVERLAPPING TILES WHICH MEASURE APPROXIMATELY 10° X 10°.(SOURCE: STRABALA) .....	39
FIGURE 3.4: MOD09GQ BAND 2 (NIR) GRAY-SCALE SURFACE REFLECTANCE PRODUCT. THE DATA WAS COLLECTED ON JUNE 2, 2017 OVER SUDAN, ERITREA AND ETHIOPIA.....	40
FIGURE 3.5:MOD09GQ BAND (2,1,1) FALSE COLOUR SURFACE REFLECTANCE PRODUCT. THE DATA WAS COLLECTED ON JUNE 2, 2017 OVER SUDAN, ERITREA AND ETHIOPIA.....	40
FIGURE 3.6: MYD09GA RGB BAND (1,4,3) TRUE COLOUR SURFACE REFLECTANCE PRODUCT. THE DATA WAS COLLECTED ON JUNE 1, 2016 OVER SUDAN, ERITREA AND ETHIOPIA. ....	41
FIGURE 3.7: PROBA-V SPACECRAFT IN ORBIT.( SOURCE: ESA) .....	42
FIGURE 3.8: PROBA-V INSTRUMENT LAYOUT.(SOURCE: WOLTERS, 2018) .....	43
FIGURE 3.9: PROBA-V PROCESSING CHAIN FLOWCHART (SOURCE:WOLTERS 2018).....	45
FIGURE 3.10: THE BRIGHTER WHITE AREAS INDICATE OVERLAPPING OBSERVATIONS.(SOURCE: WOLTERS 2018) .....	46
FIGURE 3.11: PROBA-V COLOUR COMPOSITE(1,2,3), RED(BAND 1), BLUE (BAND 2), NIR (BAND 3) AS SEEN ON 20/06/2017. ....	46
FIGURE 3.12: S1 TOC 330M BAND (1,2,3) COLOUR COMPOSITE OF SURFACE REFLECTANCE PRODUCT. THE DATA COLLECTED ON 27/08/2017 OVER SUDAN. ....	47
FIGURE 3.13: S1 TOC 330M BAND (2,1,3) FALSE COLOUR COMPOSITE OF SURFACE REFLECTANCE PRODUCT. THE DATA COLLECTED ON 27/08/2017 OVER SUDAN. ....	47
FIGURE 3.14: S1 TOC 330M BAND (1) RED BAND SURFACE REFLECTANCE PRODUCT. THE DATA COLLECTED ON 27/08/2017 OVER SUDAN.....	48
FIGURE 3.15: SMAP CONCEPT FIGURE. (SOURCE: NASA).....	49
FIGURE 3.16: SMAP IMAGE SWATH OF RADAR AND RADIOMETER.( SOURCE: JPL/ NASA) .....	51
FIGURE 3.17: GLOBAL MAP OF SMAP RADIOMETER. (SOURCE: NASA/JPL-CALTECH/GSFC).....	51

FIGURE 3.18: SMAP DATA PROCESSING AND DATA FLOW. (SOURCE: ENTEKABI 2014) .....	52
FIGURE 3.19: DATASET OF SMAP L3_SM_AP. ....	53
FIGURE 3.20 : SHOWING AVAILABLE SCENES FROM 2014-2015 FOR MODIS (AQUA/TERRA) AND PROBA-V. ....	55
FIGURE 3.21: TOTAL IMAGE AVAILABILITY. ....	56
FIGURE 3.22: CHARTS SHOWING THE YEARLY DATA AVAILABILITY, ALONG WITH THE TOTAL DATA AVAILABLE. ....	56
FIGURE 3.23: STUDY AREA. AL QADARIF SUDAN. (BASE MAP: GOOGLE HYBRID).....	57
FIGURE 3.24: DETAILED SCHEME MAP (SOURCE: DR. NIKOLAS PRECHTEL, TUD) .....	58
FIGURE 3.25: AFRICAN SAHEL EXTENT. ....	59
FIGURE 3.26: SORGHUM PRODUCTION/YIELD OF SORGHUM IN SUDAN BETWEEN 014-2016. (SOURCE: FAOSTAT). ....	62
FIGURE 4.1: RASTER IMAGE PREPARATION FLOWCHART. ....	63
FIGURE 4.2: UPLOAD SHAPEFILE TO SELECT TILE FOR DOWNLOAD. ....	63
FIGURE 4.3: SELECTING AND CUSTOMISING PROBA-V DATA.....	64
FIGURE 4.4: SELECTING AND CUSTOMISING SMAP DATA. ....	64
FIGURE 4.5: MODIS ACTUAL TILE AT 5.5KM (MODIS TERRA- 03/06/2014) .....	65
FIGURE 4.6: CROPPED DOWN TO STUDY AREA.....	66
FIGURE 4.7: VECTOR DATA CREATION FLOWCHART. ....	66
FIGURE 4.8 : MAIN PLOT WITH PARTS OF SUB-PLOTS DIGITISED FROM YEARLY IMAGES. ....	67
FIGURE 4.9: CLOUD MASKING FLOWCHART. ....	67
FIGURE 4.10: METHOD TO READ BIT-CODE.....	68
FIGURE 4.11: MODIS TERRA IMAGE TAKEN ON 01/07/2014. (A) ACTUAL IMAGE WITH CLOUD (B) CLOUD MASKED IMAGE.....	69
FIGURE 4.12: QA SDS WHICH IS USED AS CLOUD MASK FOR ABOVE IMAGE TAKEN ON 01/06/2014 ...	69
FIGURE 4.13: PROBA-V CLOUD MASK FOR 01/07/2014. (A) THE ACTUAL CLOUD MASK PROVIDED IN THE DATA. (B) CLOUD MASK DILATED BY 1 (C) CLOUD MASK DILATED BY 2. ....	70
FIGURE 4.14: PROBA-V MULTISPECTRAL BANDS SPLITTED INTO 4 PANCHROMATIC BAND. ....	71
FIGURE 4.15: STEP-BY- STEP PROBA-V CLOUD MASKING IMAGE IS FROM 10/06/2017. (A) ACTUAL CLOUD MASK (B) ACTUAL IMAGE (C) FIRST MASKED IMAGE FROM USING ACTUAL MASK (D) NEW MASK CREATED BY RADIOMETRIC THRESHOLDING. (E) FINAL MASKED OUTPUT. ....	71
FIGURE 4.16: FINAL DATA FUSION FLOWCHART. ....	72
FIGURE 4.17: (A) VIEWING ZENITH ANGLE (VZA) OF MODIS, (B) VIEWING ZENITH ANGLE (VZA) OF PROBA-V. ....	73
FIGURE 4.18: FINAL FUSED IMAGE CREATION. (A) PROBA-V IMAGE FOR 01/06/2014 (B) MODIS IMAGE FOR 01/06/2014 (C) FINAL FUSED IMAGE. ....	75
FIGURE 5.1: MEAN MONTHLY NDVI IN 2014 .....	76
FIGURE 5.2: MEAN MONTHLY NDVI IN 2015. ....	76
FIGURE 5.3: MEAN MONTHLY NDVI IN 2016. ....	77
FIGURE 5.4: MEAN MONTHLY NDVI IN 2017 .....	77
FIGURE 5.5: DAILY NDVI VALUE FOR THE YEAR 2014 FROM JUNE - DECEMBER.....	78
FIGURE 5.6: NDVI ANOMALY IN 2014 REFERRED TO (A) 2015, (B) 2016, (C) 2017.....	80
FIGURE 5.7: NDVI ANOMALY IN 2015 REFERRED TO (A) 2014, (B) 2016, (C) 2017.....	81
FIGURE 5.8: NDVI ANOMALY IN 2016 REFERRED TO (A) 2014, (B) 2015, (C) 2017.....	82
FIGURE 5.9: NDVI ANOMALY IN 2017 REFERRED TO (A) 2014, (B) 2015, (C) 2016.....	83
FIGURE 5.10: PLOT-75 YEARLY CHANGE COMPARISON.....	85
FIGURE 5.11: AREA UNDER CURVE. ....	86
FIGURE 5.12: NDVI VALUES PLOT BASIS FOR THE YEAR 2014.....	87

FIGURE 5.13: NDVI VALUES REGION BASED FOR THE YEAR 2014 .....	88
FIGURE 5.14: MULTILOT NDVI-2014 .....	89
FIGURE 5.15: MULTILOT NDVI-2015 .....	89
FIGURE 5.16: MULTILOT NDVI-2016 .....	90
FIGURE 5.17: MULTILOT NDVI-2017 .....	90
FIGURE 5.18: COMPARISON OF PLOT-75 AND PLOT-146. ....	91
FIGURE 5.19: COMPARISON OF PLOT-266 AND PLOT-335. ....	92
FIGURE 5.20: GROWTH SEASON OF A SINGLE PLOT. ....	93
FIGURE 5.21: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-75 FOR 2015. (A) CCF GRAPH, (B) LAG 0-LAG 6 COMPARED TO NDVI, (C) LAG 0 IN RELATION TO NDVI, (D) LAG 4 IN RELATION TO NDVI. ....	95
FIGURE 5.22: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-75 FOR 2016. (A) CCF GRAPH, (B) LAG 0-LAG 6 COMPARED TO NDVI, (C) LAG 0 IN RELATION TO NDVI, (D) LAG 5 IN RELATION TO NDVI. ....	96
FIGURE 5.23: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-75 FOR 2017. (A) CCF GRAPH, (B) LAG 0-LAG 6 COMPARED TO NDVI, (C) LAG 0 IN RELATION TO NDVI, (D) LAG 6 IN RELATION TO NDVI. ....	97
FIGURE 5.24: EFFECT OF SOIL MOISTURE AT TIME (T) ON CROP GROWTH. ....	99
FIGURE 6.1: MOCK VISUALISATION OF MAKING TOOL (TAB-1). ....	101
FIGURE 6.2: WORKFLOW OF TAB-1. ....	102
FIGURE 6.3: MOCK VISUALISATION OF PROCESSING TOOL (TAB-2). ....	103
FIGURE 6.4: WORKFLOW OF TAB-2. ....	103
FIGURE 6.5: MOCK VISUALISATION OF PARAMETER TOOL (TAB-2) .....	104
FIGURE 6.6: WORKFLOW OF TAB-3 .....	105

## List of Tables

TABLE 1: PEAK NDVI VALUES WITH DATE. ....	91
TABLE 2: PEAK NDVI VALUES WITH DATE. ....	92
TABLE 3: KENDALL CORRELATION VALUES PLOT-75 (2015).....	95
TABLE 4: KENDALL CORRELATION VALUES PLOT-75 (2016).....	96
TABLE 5: KENDALL CORRELATION VALUES PLOT-75 (2017).....	97
TABLE 6: PROBA-V CLOUD MASK (STATE MAP) .....	113
TABLE 7: MODIS (AQUA/TERRA) SPECTRAL BANDS. ....	114
TABLE 8: MODIS QUALITY ASSESSMENT PARAMETERS FOR THE QA SDS LAYER. ....	116

## List of Annexure:

ANNEXURE 1: ANOMALY BETWEEN FINAL FUSED IMAGE AND MODIS.....	118
ANNEXURE 2: ANOMALY BETWEEN FINAL FUSED IMAGE AND PROBA-V .....	118
ANNEXURE 3 : DAILY NDVI VALUE- 2015 FROM JUNE - DECEMBER.....	119
ANNEXURE 4: DAILY NDVI VALUE- 2016 FROM JUNE - DECEMBER.....	120
ANNEXURE 5: DAILY NDVI VALUE- 2017 FROM JUNE-DECEMBER .....	121
ANNEXURE 6: NDVI VALUES PLOT BASIS FOR THE YEAR 2015.....	122
ANNEXURE 7: NDVI VALUES PLOT BASIS FOR THE YEAR 2016.....	122
ANNEXURE 8: NDVI VALUES PLOT BASIS FOR THE YEAR 2017.....	123
ANNEXURE 9: DIFFERENT CURVE FITTING METHODS FOR NDVI .....	123
ANNEXURE 10: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-146 FOR 2015.(A) CCF GRAPH, (B) LAG0-LAG6 COMPARED TO NDVI, (C) LAG0 IN RELATION TO NDVI, (D) LAG4 IN RELATION TO NDVI .....	124
ANNEXURE 11: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-146 FOR 2016. (A) CCF GRAPH, (B) LAG0-LAG6 COMPARED TO NDVI, (C) LAG0 IN RELATION TO NDVI, (D) LAG4 IN RELATION TO NDVI .....	124
ANNEXURE 12: SOIL MOISTURE LAGS COMPARED TO NDVI VALUE OF PLOT-146 FOR 2017. (A) CCF GRAPH, (B) LAG0-LAG6 COMPARED TO NDVI, (C) LAG0 IN RELATION TO NDVI, (D) LAG6 IN RELATION TO NDVI .....	125

## **List of Abbreviations**

- AOI- Area of Interest
- BRDF - Bidirectional Reflectance Errors
- BDF- Bayesian Data fusion
- CCF- Cross Correlation/Coefficient Function.
- DVI- Difference Vegetation Index
- ESA - European Space Agency
- EASE 2 – Equal Area Scalable Earth
- ERS – European Remote Sensing
- EOS – End of Season / Earth Observing System
- ECS - EOS Core System
- EVI - Enhanced Vegetation Index
- FAPAR - Fraction of Absorbed Photosynthetically Active Radiation.
- HDF – Hierarchical Data Format
- IGFOV- Instantaneous geometric field of view.
- LAI - Leaf Area Index
- MODIS - Moderate Resolution Imaging Spectroradiometer
- MVC- Maximum Value Compositing
- MASAVI - Modified Soil Adjusted Vegetation Index
- NASA - National Aeronautics and Space Administration.
- NIR - Near InfraRed
- NDVI- Normalised Difference Vegetation Index
- OSAVI - Optimized Soil Adjusted Vegetation Index
- PROBA-V - Project for On-Board Autonomy Vegetation
- QC- Quality check
- QA - Quality Assurance
- ROI- Region Of Interest.
- SAA- Solar Azimuth Angle
- SAVI - Soil Adjusted Vegetation Index
- SAR – Synthetic Aperture Radar
- SDS- Science data sets
- SMAP – Soil Moisture Active Passive
- SNR - Signal to Noise Ratio

- SOS - Start of Season
- STVI - Stress related Vegetation Index
- SWIR- Short Wave InfraRed
- SZA – Solar Zenith Angle
- USGS - United States Geological Survey
- VAA - Viewing Azimuth Angle
- VNIR - Very Near InfraRed
- VZA – Viewing Zenith Angle
- VITO – Vision on Technology
- WDRVI - Wide Dynamic Range Vegetation Index



## **Abstract**

Crop Monitoring plays an important role in estimating the crop yield of an area resulting to the food security especially for less developed countries where domestic food sources plays a key role. Mapping and monitoring of crop growth is crucial for identifying the crop health and there potential yield. This thesis investigates a daily time series of data, analyses plot-based crop growth by processing vegetation signal from sowing to harvest and investigates the effect of soil moisture on crop growth by establishing a relationship between the soil moisture and NDVI. The main objective of this thesis is to harmonise and fuse two complementary sensor data to produce a plot-based time series of vegetation development and extrapolate vegetation signal as a function of soil moisture. The goals where achieved by, fusing MODIS and PROBA-V after smoothening based on a conditional criteria and the fused data is further statistically analysed to establish a relationship with soil moisture data obtained from SMAP. The results distinctively showed the crop growth for each plot with phenological information along with the lagged effect of soil moisture on the crop growth.

**Keywords: Crop Monitoring, Crop Growth, Crop Phenology, NDVI Anomaly, Time-series, NDVI and Soil Moisture, MODIS, PROBA-V, SMAP**

## 1. Introduction

According to FAO. 2002. (*The State of Food Insecurity in the World 2001*. Rome ) “Food security is a situation that exists when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life”, but achieving an optimum Food Security has always been a challenge, especially for less developed countries with high dependency on domestic food sources.

About 4.6 million people are food insecure in the Sudan. Above average harvests and good pastures by the end of 2014 helped to improve food security in most of the country, but got affected intensively in 2015 because of drought. Sudan continues to face steep rise in staple food prices, floods destroying field crops, pest infestations, continued blocking of internal and transboundary transhumance routes, localized disease outbreaks, as well as environmental factors such as desert encroachment and soil degradation.<sup>1</sup>

An estimated 80% of the country’s rural population relies on agriculture-based production for their food and income. In some areas, however, production is severely constrained by persistent insecurity, a lack of quality agricultural inputs, unpredictable rainfall, restricted livestock movement and over-stretched natural resources. Persistent insecurity and continued constraints on resources mean farmers are unable to fully recover and are always on the precipice of food insecurity.

Remote sensing-based crop monitoring has proven to be one of the most efficient ways of estimating and forecasting crop yield over time, it assists in crop harvest prediction and monitoring crop growth based on a time lapse convenient and faster, aiding authorities to inform the farmers and formulate precautionary measures.

Crop monitoring helps in analysing and studying the current crop conditions, aiding the observation of crop growth, along with identifying problems negatively affecting the growth, and with the current volatile climatic conditions, it is of utter importance to have a sustainable, frequent, updated and accurate crop yield estimation and forecast for food security.

---

<sup>1</sup> <http://www.fao.org/emergencies/countries/detail/en/c/148725>

## 1.1 Motivation

This thesis is conducted as part of a joint project between TU Dresden and Gadarif University (Sudan) which aims to develop a stable recording and visualization tool of crop development in a rain-fed mechanized agricultural scheme (Gadambalia), for the study area Gadarif, and can be further used for predicting crop yields prior to harvest, and can contribute to a spatially fine-grained estimation of productivity,

The study area has a very uncertain and erratic rainfall conditions which effects the crop growth leading to an indefinite crop yield and as the area produces about one third of the countries sorghum and half of sesame, monitoring the crop growth and changing patterns over years is highly significant, along with identifying the factors affecting the crop yield.

Thus the thesis is going to investigate a daily time series of data, analysing the crop growth by processing vegetation signal received from two complementary sensor (MODIS and PROBA-V), and creating a homogenous and continuous database of the plant from sowing to harvest. Moreover the thesis will investigate the effect of certain climatic parameters on crop growth by establishing a relationship between the soil moisture (selected as a parameter) and vegetation index identifying the effect of soil moisture on crop growth.

## 1.2 Research Significance:

The approaches adapted in the thesis methodology for crop growth monitoring and establishment of a relationship between soil moisture and vegetation indices are consequential.

The focus of the thesis is to create a dataset of daily vegetation indices for the analysis of the crop growth, by using complementary sensor data, from MODIS and PROBA-V. The approach is of adapting a daily dataset instead of composites, as the composite datasets are usually generated based on the MVC (Maximum value compositing) process, which minimizes the artefacts like cloud cover or BDRF effects but simultaneously introduces degradation of temporal accuracy, the maximum value of the composited period (ranging from 8 to 16 days) is assigned to a single image rather than the date when this value was captured. As concluded by (Zhao, Yang et al. 2012), a daily based study is better than composites, especially when dealing with plant growth.

The observations are plot-based, since there are different crop types grown in different plots and requires a detailed plot and subplot based analysis, unlike most of the similar researches which focuses on a large homogenous field of similar crops.

The methodology of fusing two sensor data ( PROBA-V and MODIS) attempts to create, optimise and analyse the time series while at the same time preserving the reflectance values of both the sensors, negating the effect of resolution degradation. Also the fusion of these two sensors have not been adapted yet.

The second adapted approach is to establish a relationship between soil moisture and plant growth by incorporating the MODIS and PROBA-V fused data with SMAP (Soil Moisture Active Passive) data. SMAP provides, highest resolution soil moisture information compared to the other available remote sensing based soil moisture systems, as availability of researches related to use of SMAP with MODIS and PROBA-V data to establish a relationship is limited, the observations can be fine-tuned for further research works. This research is processed based on free and open source software and programming language.

### **1.3 Research Objectives:**

The primary focus of this project is:

- Efficient harmonisation and fusion of NDVI from MODIS and PROBA-V in order to optimise spatial coverage, by a mutual signal transformation. Resulting to fused time series of growing seasons and increase accuracy of the output.
- To produce a plot-based time series of vegetation development from complementary sensor data sources.
- Analyse to what degree existing soil moisture recordings influence the vegetation signal at various crop development stages, mainly by using MODIS data with SMAP, through statistical correlation.
- Extrapolate vegetation signal as a function of soil moisture at a specific development stage.
- Discuss visualisation techniques synoptically displaying the source of information and the uncertainty (statistical errors) involved.

The research questions answered, for achieving the above mentioned objectives are discussed in the later chapters.

### **1.4 Thesis Outline:**

Chapter 2, will provide background information and literature review of existing research and methods along with an insight for the adapted methodologies and techniques. Chapter 3 describes the data used, specifications of the satellite imagery used and the study area.

Chapter 4 outlines the methodologies used, providing a detailed description for the pre-processing and post-processing of data. Chapter 5, will focus on the results achieved with discussion and analysis of the information generated. Chapter 6, provides a conceptual framework to the processing and visualisation tool that will provide a conclusion to the findings. Chapter 7, gives a final conclusion to the observations made in the thesis.

## 2. Background and Literature Review

This chapter focuses on the researches and studies conducted in crop monitoring, and the different methodologies available.

### Crop monitoring using remote sensing:

Monitoring and assessment of crop is critical in ensuring a good agricultural productivity. This continuous and intensive process requires remote sensing imageries. Crops generally do not grow evenly across the fields because of crop rotation, multiple cropping, or differences in soil nutrients, soil moisture or other factors, which effects crop yield from one plot to another within the same area. Remote sensing allows the farmer to identify areas within a field which are affected and adapt specific mitigating measures.

Remote sensing based monitoring has many advantages above regular field observations, such as the presence of Very Near Infrared (VNIR) that can see beyond the visible wavelengths and is highly sensitive to crop growth and changes. It also gives a complete overview of the area unlike field survey methods, thus having a temporal advantage over in-situ methodologies. Regular images obtained throughout the season from sowing to harvest provides more detailed information to the farmer.

### 2.1 Electromagnetic spectrum and vegetation:

Remote sensing has stages of data acquisition and processing. As observed in (Figure 2.1) solar radiation interacts with various objects depending on the type of object on which the light is incident the spectrum is reflected, refracted or absorbed. The reflected radiation is recorded by the satellite sensor, and further processed and sent to the ground station.

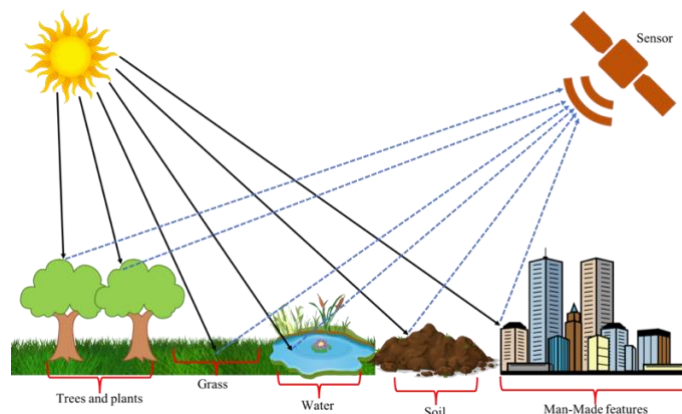


Figure 2.1: Reflected solar radiation recorded by the sensor.

The physical characteristics of the material cause the electromagnetic energy to be reflected, refracted, or absorbed in a way that is unique to each material, and these interactions are measured in discrete section of the electromagnetic spectrum that provides the spectral reflectance of the specific object. (Figure 2.2).

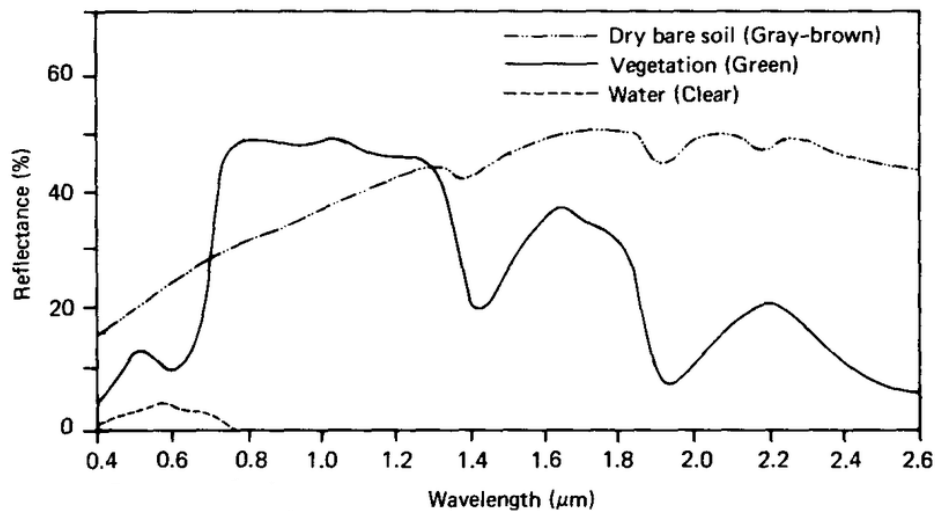


Figure 2.2: Spectral reflectance of Soil, Vegetation and Water. (Source: Lillesand & Kiefer)

Vegetation has an unique interaction with solar radiation compared to other natural materials. The vegetation spectrum typically absorbs in the Red and Blue wavelengths, reflects in the Green wavelength, and strongly reflects in the Near Infrared (NIR) wavelength and shows strong absorption feature in wavelength where atmospheric water is present. Different plant materials, water content, pigment, carbon content, nitrogen content, and other properties cause further variation across the spectrum<sup>2</sup>. As seen (Figure 2.3) Vegetation that is stressed shows higher reflectivity in SWIR portion of the spectrum and healthy vegetation is detected in NIR.

<sup>2</sup> <https://www.harrisgeospatial.com/Learn/Whitepapers>

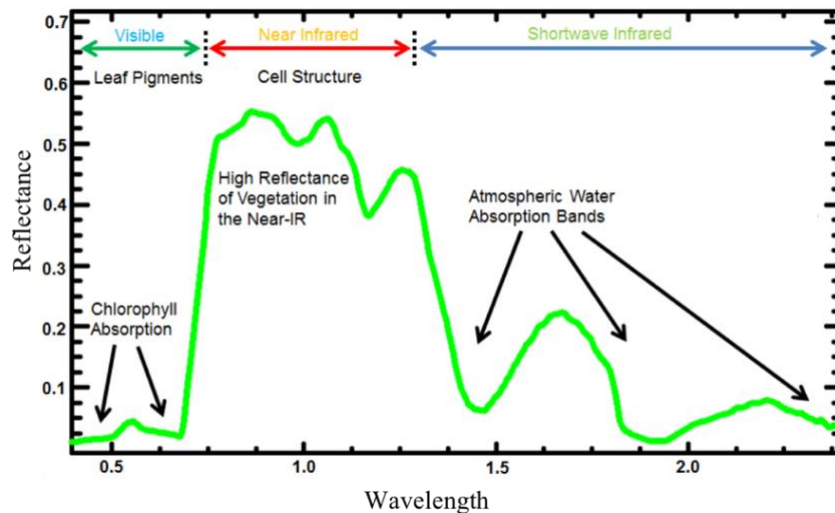


Figure 2.3. Vegetation Spectral reflectance in detail. (Source: <http://www.markelowitz.com>)

As seen in Figure 2.4 most visible light is absorbed, while almost half of the near infrared energy is reflected. This is because under the top surface of the leaf (epidermis) there are primarily two layers of cells the top one is the palisade parenchyma and consists of elongated cells tightly arranged in a vertical manner in this layer resides most of the chlorophyll, a protein that is responsible for capturing the solar energy and power photosynthesis. The lower level is the spongy parenchyma, consisting of irregularly shaped cells with a lot of air spaces between them in order to allow the circulation of gases. In addition to chlorophyll, the palisade parenchyma contains other pigments, such as carotenoids, anthocyanins and others, because of these pigments, most of the visible electromagnetic energy is absorbed, especially in the blue and red region.

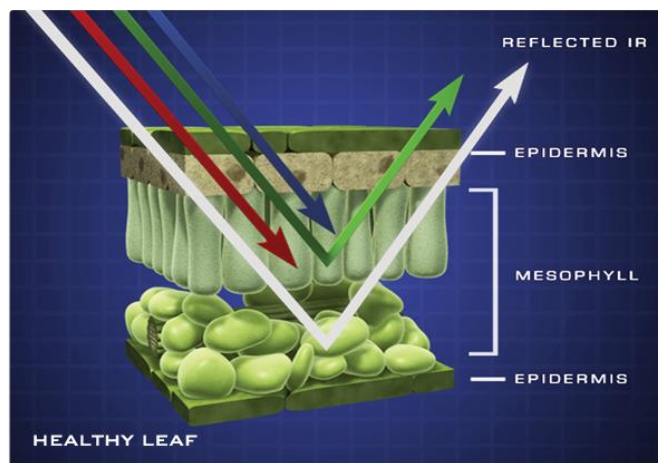


Figure 2.4. Cellular leaf structure and its interaction with electromagnetic energy. (Source: Jeff Carns, NASA Science).



Whereas, near infrared (NIR) energy is not affected by these pigments and almost completely penetrates the palisade parenchyma when it reaches the spongy parenchyma the presence of air spaces causes the refraction of the NIR energy in various directions, this results in approximately half the energy exiting the leaf from the lower epidermis and the other half from the top epidermis, towards the sky.<sup>3</sup>

## 2.2 Vegetation Indices:

Vegetation indices in remote sensing based crop monitoring forms the basis for scientific approach towards phenology detection, type classification or crop growth. As mentioned by (Zeng et al., 2016) monitoring crop phenology provides an essential information for crop management as well as for understanding regional to global scale vegetation dynamics, they developed a hybrid technique to detect the critical vegetative stages and reproductive stages from MODIS 250-m Wide Dynamic Range Vegetation Index (WDRVI) time-series data. But the usage of this methodology is more distinctive in case of moderate-dense vegetation with higher LAI, it has not yet proved to be much effective in case of plants with low biomass or LAI.

Another widely used vegetation index is LAI (Leaf Area Index). It is used to predict photosynthetic primary production, evapotranspiration and as a reference tool for crop growth. It can be determined directly by taking a statistically significant sample of foliage from a plant canopy measuring the leaf area per sample plot and dividing it by the plot land surface area<sup>4</sup>. But this index is limited to be used for higher coverage areas, or regions with vegetation having broader and denser canopies or bigger leaf surface area.

EVI (Enhanced Vegetation Index) is also largely used as a vegetation index, according to (Sakamoto, Yokozawa et al. (2005) EVI is more practical than NDVI when humidity is high and thus is appropriate for crops that grows in high humidity and water content like rice. EVI can be linearly correlated with the leaf area index but has high sensitivity in conditions of higher biomass only. According to (Matsushita, Yang et al. 2007) their results indicated that the soil adjustment factor in EVI makes it more sensitive to topographic conditions than in NDVI. Based on their results, it is recommended that the topographic effect should be removed in the reflectance data before the EVI is calculated, but when these indices are used in the area

---

<sup>3</sup> <http://www.seos-project.eu/modules/agriculture/agriculture-c01-s01.html>

<sup>4</sup> [https://en.wikipedia.org/wiki/Leaf\\_area\\_index](https://en.wikipedia.org/wiki/Leaf_area_index)

of rough terrain, where the topographic effect on the vegetation indices having only a band ratio format (e.g., the NDVI) it can usually be ignored

The research by (Raghavendra and Aslam 2017), compared the ability of different vegetation indices to support Leaf Area Index (LAI) at different growing stages of rice crops, and concluded that LAI indicates density of vegetation whereas other indices used like, Normalized Difference vegetation Index (NDVI), Difference Vegetation Index (DVI), Enhanced Vegetation Index 1 and 2 (EVI, EVI2), Soil Adjusted Vegetation Index (SAVI), Optimized Soil Adjusted Vegetation Index (OSAVI), Modified Soil Adjusted Vegetation Index 2 (MASAVI2) and Stress related Vegetation Index (STVI) mainly concentrate on vegetation conditions, with DVI, NDVI and STVI showing more relation to the crop growth compared to the other parameters. It was observed by (Wardlow, Egbert et al. 2007), that between the EVI and NDVI profiles that were, the NDVI maintained higher values than the EVI specifically through the senescence and/or peak phase for all crops and especially for Sorghum.

According to (Fensholt, Sandholt et al. 2004) when estimating the productivity of terrestrial ecosystems from satellite images, canopy spectral measures of vegetation intensity in relative units must be converted into quantitative biophysical variables. This conversion has been performed using models of varying complexity which have analysed that a Top-down statistical model provides an empirical relationship between variables like the NDVI and yield. Also the ongoing research has sought to establish more physically based relationships between vegetation development and NDVI through the use of parametric models. Based on satellite data the relationship between NDVI and FAPAR has been found to be linear or approximately linear for green vegetation.

Along with the various stand-alone methodologies adapted for crop monitoring, there are also complimentary parameters like soil moisture, precipitation or temperature that are analysed to support the vegetation indices. Among most of the vegetation indices used NDVI seemed to be most sensitive to crop growth from growing to senescence and also sensitive to soil moisture fluctuations in relation to the growth. After weighing out the pros and cons of applicability for the study area NDVI is considered to be aptly suitable.

### **2.2.1 NDVI (Normalised Difference Vegetation Index):**

Normalized Difference Vegetation Index (NDVI) quantifies vegetation by measuring the difference between Near Infrared (which the vegetation strongly reflects) and Red light (which the vegetation absorbs). NDVI always ranges from -1 to +1. But there is no distinct boundaries

for objects that the value represents, for example a cloud shadow is represented as more than -1, water is represented as -1, 0 represents no vegetation but at the same time it can be an urban area or other feature without vegetation cover on it. However, no green leaves gives a value zero and +1, 0.8 to 0.9 indicates the highest possible density of green leaves, with shrubs and grasslands or senescing crops with NDVI values approximately 0.2 to 0.5.

NDVI is one of the most dynamic and certain index to provide plant health and growth information. For healthy vegetation with high chlorophyll content it reflects more near-infrared (NIR) and Green light compared to other wavelengths and absorbs more Red and Blue light, whereas unhealthy vegetation will reflect less NIR. Similarly when the plant reaches Greenup peak it will have a higher reflectance of NIR than when it's growing or is in the senescence stage providing distinct signal values to identify crop growth cycle.

The formula used for measuring NDVI value is : **NDVI= (NIR – R) / (NIR +R)**

Where, NIR is Near Infrared and R is Red light from the visible spectrum.

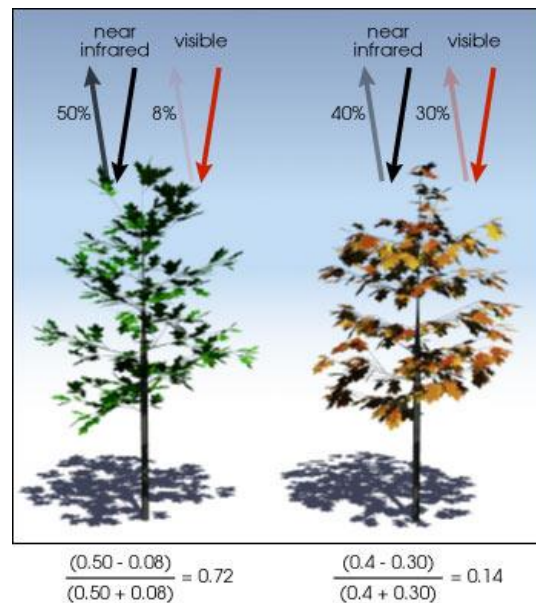


Figure 2.5: NIR and R interaction and NDVI calculation.(Source: NASA)

The NDVI has become a standard tool used by scientists to visualize vegetation cover, plant growth and biomass production from multispectral satellite data. Many NDVI products have been derived over the past years from different sensors and can be used for comparative studies or to support extensive time series analyses. <sup>5</sup>

<sup>5</sup> <http://www.fe-lexikon.info/lexikon-n.htm#ndvi>

## 2.2.2 Plant Phenology and NDVI:

<sup>6</sup>Plotting time-series NDVI data produces a temporal curve that summarizes the various stages that green vegetation undergoes during a complete growing season. Such curves can be analysed to extract key phenological variables, or metrics, about a particular season, such as the start of the growing season (SOS), peak of the season (POS), and end of the season (EOS). These characteristics may not necessarily correspond directly to conventional, ground-based phenological events, but do provide indications of ecosystem dynamics. The phenological metrics that can be extracted are:

Start of Season (SOS) - Beginning of measurable photosynthesis in the vegetation canopy. It is identified as the Day of a year from which a consistent upward trend in time-series NDVI is observed.

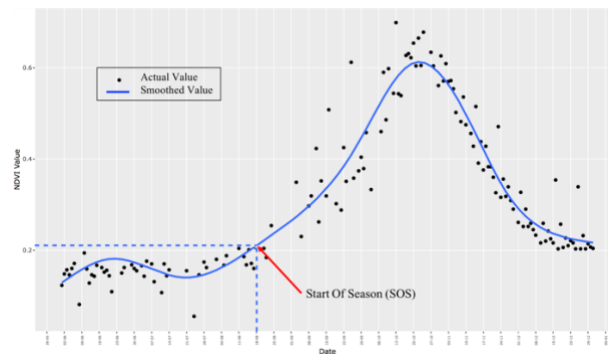


Figure 2.6: Start of Season

End of Season (EOS) – Identified as the Day of a year from which a consistent downward trend in time-series NDVI is observed. It is the end of measurable photosynthesis in the vegetation canopy.

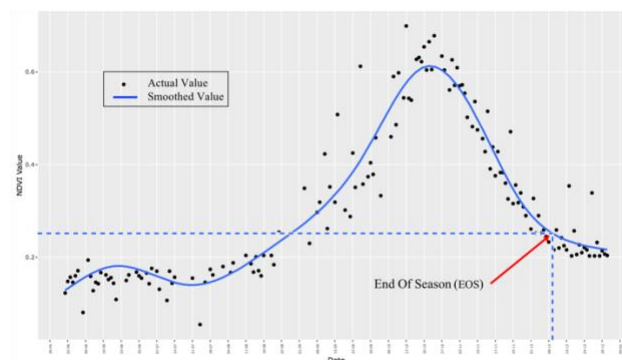


Figure 2.7: End of Season

<sup>6</sup> [https://phenology.cr.usgs.gov/methods\\_metrics.php](https://phenology.cr.usgs.gov/methods_metrics.php)

Time of Maximum (Peak) – The Day of a year corresponding to the maximum NDVI in an annual time-series, it is the time of maximum photosynthesis in the canopy.

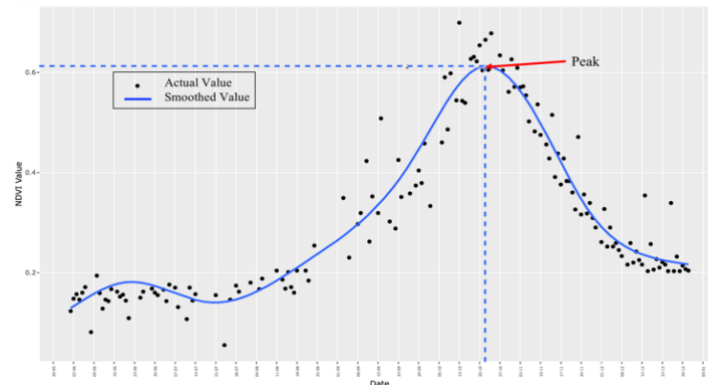


Figure 2.8: Time of maximum of NDVI peak.

Duration – It is the length of photosynthetic activity in the crop i.e., the growing season, calculated as the number of days from the SOS and EOS.

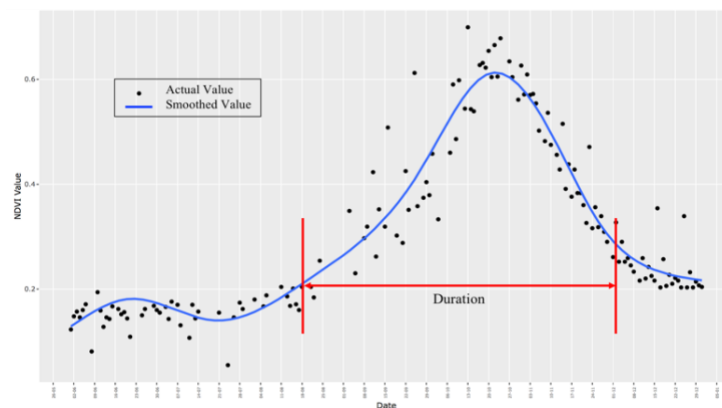


Figure 2.9: Growth Season / Duration.

Amplitude – It is calculated as the difference between Peak and SOS, referring to the maximum increase in canopy photosynthetic activity above the baseline.

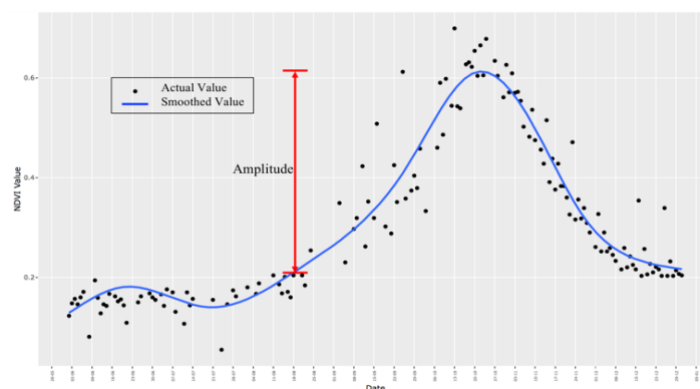
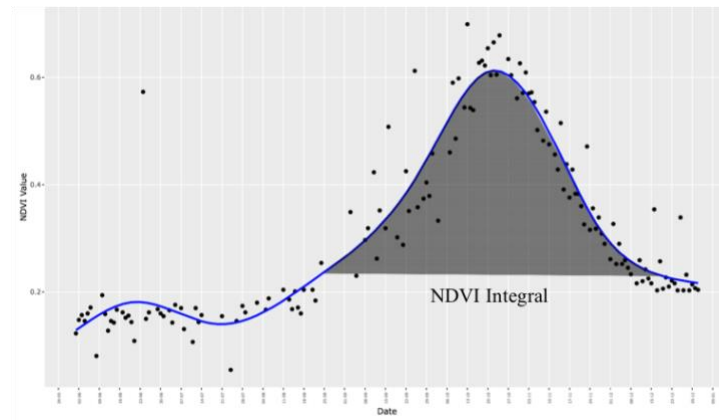


Figure 2.10: Amplitude

Time Integrated NDVI ( NDVI Integral) – It is the daily (interpolated) integration of NDVI above the baseline for the entire duration of the growing season, referring to the Canopy photosynthetic activity across the entire growing season.



*Figure 2.11: NDVI Integral*

## 2.3 Sensor Fusion:

Most of the crop monitoring researches use a single sensor for detection but the artefacts in the image remain to be a challenge in single sensor based studies. Also due to existing cloud cover in many images the cloud removal creates image gaps, leading to use of composite images or statistical processing of the image which compromises the variances sometimes. As an alternative to statistical processing or compositing, multiple sensor fusion or combination can be adapted to minimise the image gaps with an attempt to have maximum coverage.

As one of the aims is to create a daily time-series instead of using composites, MODIS and PROBA-V were selected as they have daily temporal coverage and also have close spatial resolutions. As pointed out by (Inglada, Hagolle et al. 2011) the resolution ratio between the images to be fused need to be not too far apart and that the spectral bands had to be very similar between images. Also the spatial resolution (250 m) and temporal (daily) coverage of MODIS data offers potential for retrieval of crop biophysical parameters and improved accuracy in crop yield assessment. Similarly PROBA-V was specifically launched to obtain vegetation information (Manakos, Gitas et al. 2012).

Although Landsat TM data would be more suitable in areas where the field sizes are small, according to (Doraiswamy 2004) the temporal frequency and cloud cover limits the retrieval of crop parameters that are changing during the season.

The research conducted by (Meroni, Fasbender et al. 2016) with the aim to investigate the capacity of PROBA-V sensor in delivering more spatially detailed and accurate forest monitoring and detecting changes in the forest cover caused by sudden events, such as forest fires and storm damage. They supplemented the PROBA-V data with MODIS data for certain levels where simulated PROBA-V data could not be used and compared the SPOT data to that of PROBA-V and MODIS. The conclusion was that PROBA-v and MODIS have higher accuracy in classification.

Also an extensive study was conducted by JRC (Commission) 2015) in order to evaluate the quality and usability for operational crop monitoring of the new PROBA-V satellite NDVI data and compared it to other moderate to low spatial resolution instruments like eMODIS. The analysis showed that in case of disagreement, PROBA-V indicates a positive anomaly than MODIS by a mean bias gain of 0.03.

An alternative approach for crop monitoring is the use of SAR data. (Toan, Ribbes et al. 1997) concluded that the approach of synthesising of ERS-1 with available onsite data with experimental results from fields at a tropical site and a temperate site has shown that the radar backscattering coefficients of flooded rice fields have a characteristic of increasing temporal behaviour. And the variations of the radar backscattering coefficients were expressed as a function of the crop biomass. The study concluded that a future incorporation of ERS and RADARSAT can be used for quantitative validation of the crop growth and condition. Also the technique of cross-validation of theoretical and actual values are easier when scattering is happening between 2 different mediums, like in this case the rice fields are flooded and thus have scattering from plant and water at the same time

But as the growth of the plantation in our study area has to be identified without any prior field spectral information and the area is mostly dry with very limited rainfall along with low crop biomass, incorporation of SAR is not a suitable approach.

The possibility of combining hyperspectral images with multispectral images were considered by (Adam, Mutanga et al. 2010), though the research objectives were oriented towards mapping the wetland vegetation, they concluded that the use of images with different spectral combinations has disadvantages in case of identifying different species as the biophysical and biochemical characteristics are highly correlated, and due to varying spectral responses it become difficult to be used over smaller area.

### 2.3.1 Fusion/Combining Methodologies:

In order to combine or fuse multiple sensor images for NDVI time series, there are some commonly accepted and used techniques.(Inglada, Hagolle et al. 2011)

Resampling, the simplest approach for fusion of time series by resampling the low resolution image onto the high resolution image, based on different resampling techniques like Nearest-neighbour interpolation, Bilinear and bicubic algorithms, or Lanczos resampling.

Savitzsky-Golay interpolation, this is a particular kind of interpolation which is usually applied to NDVI time profile, as NDVI time profiles are affected by the presence of clouds and the result is an underestimation of the vegetation index. Savitzsky and Golay (1964) proposed a simplified least squares- fit convolution for smoothing and computing derivatives of a set of consecutive values (a spectrum). The convolution can be understood as a weighted moving average filter with weighting given as a polynomial of a certain degree. The weight coefficients when applied to a signal, perform a polynomial least-squares fit within the filter window. This polynomial is designed to preserve higher moments within the data and to reduce the bias introduced by the filter. This filter can be applied to any consecutive data when the points of the data are at a fixed and uniform interval (Chen, Jönsson et al. 2004). The filter is more appropriate for composite data as they are based on MVC these images have a temporal resolution of 8 or 16 days which doesn't consider daily variations. The main disadvantage of this filter is that it requires an amount of experimentation to find the appropriate filter values required to best filter a specific signal.

Whitaker Smoother, is a modification of Savitzsky-Golay. It assumes an unknown piecewise polynomial smooth curve and puts a penalty on the integral of the squared second derivative. A banded system of linear equations results which can be solved efficiently with software for sparse matrices (Eilers 2003). Thus it is automatic and doesn't need training set, but similarly it is also more appropriate for composite images and can't efficiently handle missing data.

Most of the algorithms need the panchromatic band to be consistent with the spectral range covered by the multispectral channels, the Bayesian Data Fusion (BDF) method developed by (Fasbender, Radoux et al. 2008) estimates the statistical link between the images in order to find the most probable value for a pixel that can be given to another. It can also be successfully incorporated in any combination of panchromatic bands and is not affected by missing values.



## 2.4 Soil Moisture and NDVI:

While monitoring and analysing crop growth it is important to consider climatic factors, soil type etc, which effects the growth of the plant. As suggested by (Kang 2003) the phenological stages can be detected more precisely through a combination of climatic data. Thus in the thesis soil moisture is chosen as a parameter to understand and establish a relationship between soil moisture and plant growth.

According to (Doraiswamy and Hodges 1979) severe water stress delays development of crop plants capable of suspending growth, and kills plants. For sorghum, a 10-day wilting period before floral initiation was reflected in a 10-day delay in flowering. Stress after floral initiation of 14, 21, and 28 days delayed flowering by 10, 24, and 30 days. Water stress of sorghum just before floral initiation reduced the number of leaves initiated, suggesting a time lag between floral induction and floral initiation. Excess water reduces soil temperature variability and also reduces soil oxygen supply and so may delay development or injure the plant.

According to (Zhang and Zhou 2016) soil moisture is an important variable in land surface system research. It links the energy and water exchange between the atmosphere and land surface and indicates moisture conditions in soil surface. It is also used to indicate drought in vegetated environments thus helping in identifying the crop growth pattern because water content is an essential factor for vegetation growth. They concluded that the combination of optical and thermal and microwave remote sensing may have broad application prospects. As microwave remote sensing can obtain data all-time and all-weather it can provide great help for soil moisture products over long time series especially in humid tropical regions.

(Chen, de Jeu et al. 2014) found strong positive relations particularly between soil moisture and NDVI. They observed that soil moisture leads to certain behaviour of vegetation and confirmed the possibility to use satellite soil moisture data to predict vegetation dynamics which were checked based on anomalies. In the temporal domain this relation became stronger and both soil moisture and NDVI exhibit coherent trend changes during the whole study period along with a typical lag character between the soil moisture and NDVI.

The NDVI and soil moisture relationship is sometimes considered to be subjected to hydrological phases of ‘recovery’ stage when the rainfall has been less over a time period and ‘stable’ stage when the rainfall has been consistent over a time period. According to (Ahmed, Else et al. 2017) soil moisture is more tightly coupled with dynamics of plant photosynthesis

and respiration than rainfall since a variable amount of precipitation will be lost through interception and runoff. Shedding light on changing lagged relationships between NDVI and soil moisture could potentially provide useful information for predicting future responses to climate change as well as for developing adaptation strategies. There research concluded that the patterns of change in NDVI and soil moisture are similar for the recovery and stable phase unlike rainfall which varies.

(T. J. Farrar 1994) compared the effect of soil type on the relationship between NDVI and soil moisture. They concluded that soil moisture cannot account for variations among the soil types in rain use efficiency and the relationship between soil moisture and NDVI are dynamic to the soil type. But as our study area has the similar soil type throughout rain use efficiency will not have an effect.

According to (Wang, Xie et al. 2007) vegetation may develop quasi-equilibrium within a local climate condition leading to similar root-zone soil moisture in the climate gradient zone. The long-term mean root-zone soil moisture in such a system follows a general seasonality forced by long-term mean local climate and vegetation life cycles which is predictable. However, varying climatic conditions at various temporal scales result in temporal deviation of soil moisture from its long-term mean conditions. This soil moisture deviation from the mean condition affects vegetation and cause a change in vegetation characteristics from the mean condition. As well as in a short time frame (hours) the NDVI may decrease due to sudden soil moisture increase since increasing top-layer soil moisture would result in a larger decrease of near-infrared reflectance compared to the red reflectance of vegetation. However in a longer time frame it is expected that NDVI increases as soil moisture increases over the growing season.

## **2.5 Image Processing:**

To get optimised information post-processing of images is important and as cloud cover is a persistent problem in satellite images, removal of the clouds in post processing plays a vital role data quality. The following section will discuss the cloud masking technology adapted by the two sensors used in this research.

### **2.5.1 MODIS:**

Most of the MODIS composite data (8 and 16) at Level-3 are cloud free as they are composited of images from mainly cloud free days and is mostly clear, but in daily data a

systematic cloud removal is required. The existing MODIS State Quality Assurance SDS in the MOD09GA is used for identifying and removing the cloud cover more details about the SDS can be found in Table 8.

The MODIS cloud mask is a science data product that is produced regularly as an Earth Observing System (EOS) standard product. It's main purpose is to identify scenes where land, ocean and atmosphere products should be retrieved based upon the amount of obstruction of the surface due to clouds and thick aerosol. Like all official EOS data products the MODIS cloud mask is created in Hierarchical Data Format (HDF) format and consists of nine Scientific Data Set (SDS) objects.<sup>7</sup>

As the SDS layer consists of both the QA reflectance and QA state layer an important distinction exists between them that requires clarification. The reflectance band QA provides information that qualifies the quality of atmospheric correction at the pixel level, in contrast the data state QA provides information about each pixel's state, i.e., those characteristics that are band or resolution independent. The State QA SDSs is tapped for cloud-specific information because the reflectance band quality SDSs in the MOD09GQ products carry a parameter called Cloud State that is not populated since the V3 MODIS collection and therefore is not a reliable source of information.

The MODIS SDS quality product is assured by two methodologies of processing them. Firstly product quality on a field-of-view basis in this approach the Quality Assurance SDS contains information about the quality of the MODIS cloud mask based upon the number of individual spectral tests that are actually executed for the given field-of-view. The logic being that more tests will yield a better result, this implies that processing paths such as polar night and desert which have less tests associated with them will potentially be of lesser quality than other processing paths such as ocean day or land day paths. Secondly, product quality on a granule basis in which information on the quality of the MODIS cloud mask is provided on a per granule basis through EOS Core System (ECS) Inventory Metadata as values in the global attribute Core Metadata.

### **2.5.2 PROBA-V:**

Clouds obstruct land surface parameter retrieval in satellite observations therefore a proper cloud screening is pivotal in the pre-processing for the various value-added products.

---

<sup>7</sup> <http://cimss.ssec.wisc.edu/modis/CMUSERSGUIDE.PDF>

In (section 3.1.4) the method of data acquisition by the PROBA-V satellite is discussed and the scan difference in the NIR band and SWIR band can be understood.

According to (Wolters 2018) the Collection 0 algorithm is based on the use of static thresholds applied to the BLUE and SWIR spectral bands, false cloud detection over bright surfaces such as deserts and salt lakes and flagging of thick ice clouds as ‘snow/ice’ were among the key problems of the operational cloud screening method. To overcome these limitations a new algorithm as a hybrid between the radiometric approach (Ackerman, Holz et al. 2008) and a geometric approach (Simpson, McIntire et al. 2000) was developed and implemented for the PROBA-V reprocessing.

The geometric part of the cloud shadow detection algorithm has a cloud pixel located at position  $p$ , with the actual cloud being at height  $h$  from the tangential plane, i.e., the intersection of the sun beam and the line of sight from the satellite to the cloud pixel. The cloud shadow can then be found as the intersection of the sun beam and the tangential plane at the center. Solar zenith and azimuth angles are assumed to be equal in the cloud and cloud shadow pixels. (see Figure 2.12)

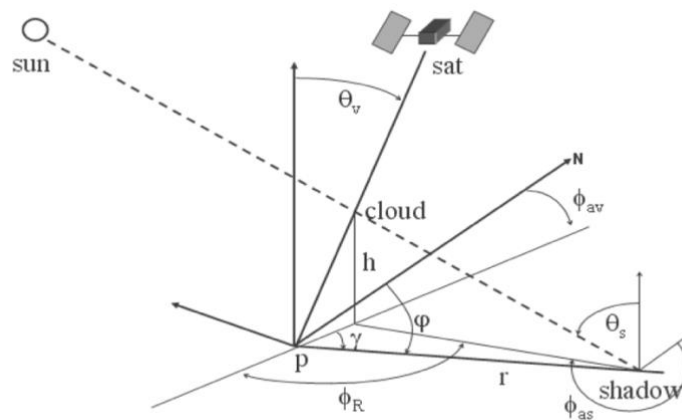


Figure 2.12 : Geometric Approach. (Source: Wolters 2018)

Cloud heights are estimated using the gradient in NIR reflectance along the projected path from a cloud to its shadow in the image. In case of a cloud shadow the NIR reflectance will decrease towards a minimum from cloud to shadow edge. If this change is above a threshold of 20% a shadow edge is detected. From the locations of the cloud and shadow edge the cloud height can subsequently be calculated.

The radiometric approach involves, using the BLUE and SWIR band reflectance. Two separate cloud masks are created  $3 \times 3$ -pixel search mask is applied to determine the matching

SWIR pixel for the BLUE band cloud mask, and the matching BLUE pixel for the SWIR band cloud mask, the final cloud mask is a merge of these two masking results. For the cloud detection based on the BLUE band reflectance, it is checked whether the observed value exceeds the BLUE band reflectance threshold of 0.2465. In addition it is checked whether the maximum SWIR reflectance value in a 3×3 pixel box above the BLUE pixel in the image exceeds the SWIR band threshold, if both conditions are satisfied, the BLUE pixel is classified as cloudy, this is only applied on the 300 m cloud mask.(see Figure 2.13)

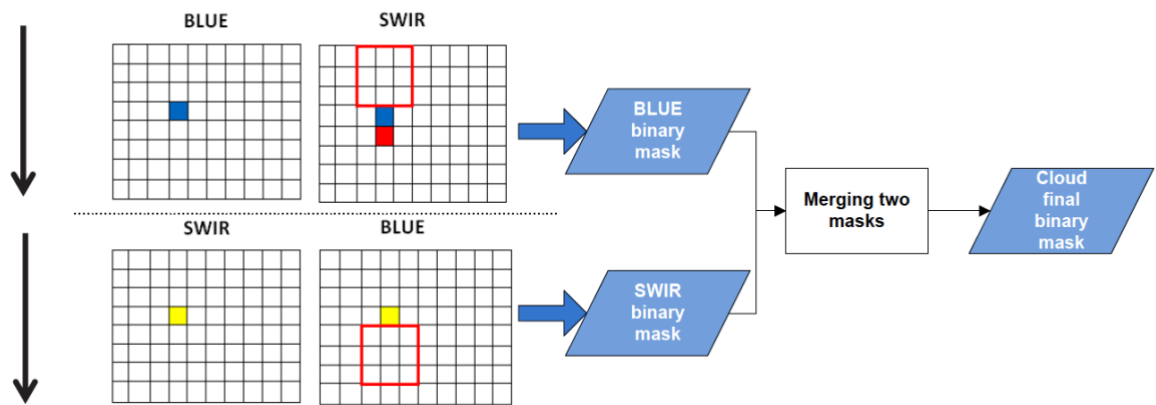


Figure 2.13: Radiometric Approach for cloud mask creation in Proba-V. (Source: Wolters 2018)

For creating the final State Map (Fensholt, Sandholt et al.) layer each PROBA-V pixel is assigned to a land cover class based on monthly images generated by ESA's Land Cover Climate Change Initiative-1 (CCI) of which the classes 'land', 'water', 'snow/ice', and 'unknown' are used for further processing. Subsequently, for each land cover class background surface reflectance for the BLUE spectral band are generated. Based on a monthly clear-sky climatology obtained from Medium Resolution Imaging Spectrometer (MERIS) 0.413 and 0.443  $\mu\text{m}$  observations over the period 2002 – 2012. In case of missing data (e.g. over areas in the winter season) coarse-resolution (5 km) broad-band (0.3 – 0.7  $\mu\text{m}$ ) ESA's GlobAlbedo<sup>8</sup> surface reflectance data are used. In the final step, a set of decision rules are defined, consisting of threshold tests (on band reflectance's, reflectance ratios or amplitude differences) and similarity checks. The similarity checks are performed using the Spectral Angular Difference (SAD) which measures the cosine of the angle between two vectors.

<sup>8</sup> European Space Agency (ESA), *GlobAlbedo Project*, <http://www.globalbedo.org>

## **2.6 Interactive Visualisation /Analytical Tools:**

<sup>9</sup>As one of the objectives is to create an interactive tool for processing and visualisation for analysis based on Free and Open Source, an overview about the currently existing software's and tools are given below .

### **2.6.1 TIMESAT:**

TIMESAT is a software package for time-series processing and assessment of vegetation dynamics, The package consists of routines developed in MATLAB and Fortran for Windows and tested also under Linux. The objective behind the package is to represent the data of seasonal vegetation signal in a meaningful way, and to identify the underlying vegetation. A number of processing steps are taken to transform the noisy signals into smooth seasonal curves, including fitting asymmetric Gaussian or double logistic functions, or smoothing the data using a modified Savitzky-Golay filter. The methods incorporate qualitative information on cloud contamination from ancillary datasets. The resulting smooth curves are used for extracting seasonal parameters related to the growing seasons. It has been used in a large number of applied studies for phenology parameter extraction, data smoothing, and general data quality improvement.

### **2.6.2 TimeStats:**

Time-Stats is a free tool for the analysis of multi-temporal equidistant georeferenced remote sensing data archives, such as MODIS, AVHRR, MERIS and SPOT-Vegetation. Key features include parametric and nonparametric methods for trend detection, generalized-least square regression, distributed lag models, cross spectra analysis, windowed trend and frequency analysis, continuous wavelet transform, empirical mode decomposition and extraction of phenological indexes (peaking times and magnitudes). Time-Stats is programmed in the Interactive Data Language (IDL) and freely distributed with the IDL virtual machine. Generated raster output files are saved in the standard ENVI format with appropriate header files and are portable to common geospatial satellite imaging processing software packages.

### **2.6.3 PhenoSat:**

PhenoSat is a software tool that allows to extract satellite VI metrics related to vegetation phenology in a simple and easy way. As input it takes the yearly vegetation index images to

---

<sup>9</sup> <http://unigis.sbg.ac.at/files/Mastertheses/Full/102723.pdf>

process, and outputs the data processing steps and the phenological information. It is developed based on MATLAB, and uses 6 smoothing methods to fit the data: Savitzky-Golay (SG), Piecewise Logistic (PL), Fourier Series (FS), Gaussian Model (GM), Polynomial Curve Fitting (PCF) and Cubic Smoothing Splines (CSS). The PhenoSat outputs are two excel files with data and phenological information. In case of using the images as input, three phenological maps (start of season, maximum vegetation development and end of season) are also created.

#### **2.6.4 SPIRITS:**

The Software for the Processing and Interpretation of Remotely sensed Image Time Series (SPIRITS) is an integrated and flexible free software environment for analysing satellite derived image time series in crop and vegetation monitoring developed by JRC Europe. The software was originally developed as a toolbox for crop monitoring with remote sensing, but has evolved into an independent tool for the processing and analysis of time series of raster data. It can be used to perform and to automatize many spatial and temporal processing steps on time series and to extract spatially aggregated statistics. Vegetation indices and their anomalies can be rapidly mapped and statistics can be plotted and interpreted in seasonal graphs to be shared with analysts and decision makers. It has been developed to answer the specific needs of the agriculture monitoring community, with the objective to support the whole chain of image time series processing steps and the production of synoptic visual analysis outputs in a user-friendly and flexible manner. Since its first release in 2012 the software has evolved into an advanced and comprehensive time series processing tool. SPIRITS is written in java programming language.

As discussed above most of the existing softwares are well designed and integrates many features, especially SPIRITS the newer version of the software plans to incorporate PostgreSQL as a connection for database. But none of the softwares incorporates the feature of fusioning data neither it consider the incorporation of additional parameter likes precipitation, soil moisture temperature to compliment the vegetation indices calculations.

### 3. Data and Study Area

This chapter will discuss the data, its acquisition and availability, along with an overview of the study area with the climatic and agricultural patterns.

#### 3.1 Datasets:

##### 3.1.1 MODIS Terra/ MODIS Aqua (collection 6):

Moderate Resolution Imaging Spectroradiometer (MODIS) provides high radiometric sensitivity (12 bit) in 36 bands of the electromagnetic spectrum of wavelengths ranging from 0.4  $\mu\text{m}$  to 14.4  $\mu\text{m}$ . Two bands (NIR and IR) are imaged at a resolution of 250 m at nadir with five bands at 500 m and the remaining 29 bands at 1 km. A  $\pm 55^\circ$  scanning pattern at the EOS orbit of 705 km achieves a 2,330-km swath and provides global coverage every one to two days<sup>10</sup>. The Scan Mirror Assembly (Whisk-broom) uses a continuously rotating double-sided scan mirror to scan  $\pm 55^\circ$  and is driven by a motor encoder built to operate at 100 percent duty cycle throughout the 6-year instrument design life.

MODIS is a key instrument aboard the Terra satellite (EOS AM-1) launched in 1999 and Aqua satellite (EOS PM-1) launched in 2002. It was built by Santa Barbara Remote Sensing to NASA specifications<sup>11</sup> and was designed with the ability of monitoring atmospheric, land, and ocean imaging in a single instrument. The sensor is ideal for monitoring large-scale changes in the biosphere it measures the photosynthetic activity of land and marine plants to yield better estimates in plant productivity, and helps in gauging the plants health.

##### Specifications:

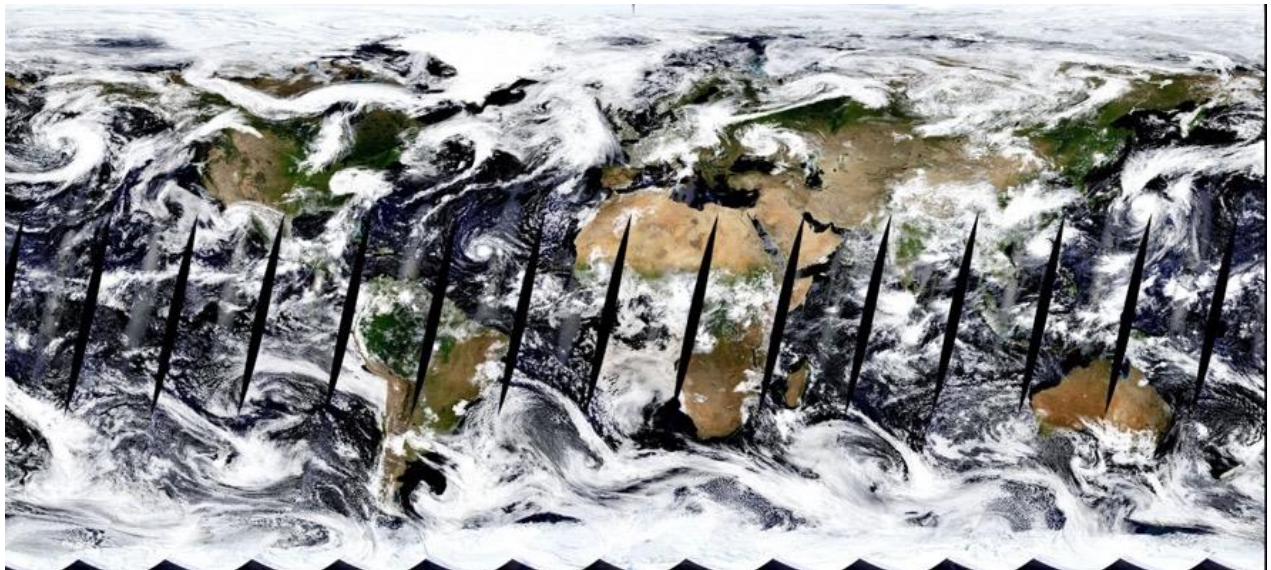
Orbit	705 Km, sun - synchronous, near – polar, circular.  10:30 a.m. descending node for Terra. (Figure 3.1)  1:30 p.m. ascending node for Aqua.(Figure 3.2)
Temporal Resolution	Daily.

<sup>10</sup> <https://modis.gsfc.nasa.gov/about/design.php>

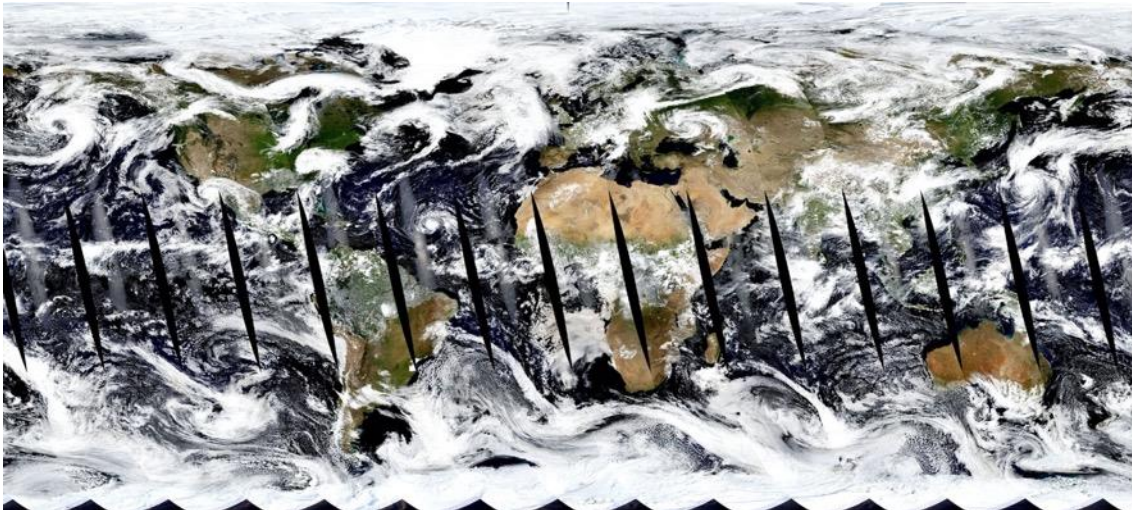
<sup>11</sup> <https://modis.gsfc.nasa.gov/about/components.php>



Spatial Resolution	250 m (bands 1-2), 500 m (bands 3-7), 1000 m (bands 8-36). (see: Table 7)
Scan Rate	20.3 rpm, cross-track (Whisk-broom).
Swath Dimension	2330 km (cross track) by 10 km (along track at nadir).
Telescope	17.78 cm diameter off-axis, a-focal (collimated), with intermediate field stop.
Size	1.0 x 1.6 x 1.0 m
Weight	228.7 kg
Power	162.5 W (single orbit average)
Data Rate	10.6 Mbps (peak daytime); 6.1 Mbps (orbital average)



*Figure 3.1: Global Land Surface Reflectance in true colour (MODIS-Terra) acquired on 02.09.2018..*



*Figure 3.2: Global Land Surface Reflectance in true colour (MODIS-Aqua) acquired on 02.09.2018.*

### 3.1.2 MODIS Processing:<sup>12</sup>

MODIS has five distinct levels of processing. Level 0 data is raw satellite feeds. Level 1 data is radiometrically calibrated. Level 2 data is level 1 data that is atmospherically corrected to yield a surface reflectance product. Level 3 data is level 2 data that has been gridded into a map projection and has been temporally composited or averaged. Level 4 data are products that have been put through additional processing. All data up to and including level 2 are delivered in an un-gridded orbital swath format with each swath typically cut into small segments. Data at level 3 and higher are geolocated into a specific map projection typically with the geolocated products in a set of non-overlapping tiles.

The level 2 MOD09 output includes all corrected bands and band quality data, as well as aerosol retrieval data and data for assessing the quality of the aerosol retrieval algorithm. It consists of 1km State QA SDS (Quality Assurance Scientific Data Sets) with band quality SDSs containing information about the quality of the atmospheric correction of each pixel. and the 500m State QA SDS containing information about the pixel's state.(see Table 8)

The data is run for each tile in the MODIS sinusoidal grid for each data (see Figure 3.3). The number of observations at each pixel is determined not only by the number of orbits at that location but also by the spread of observational coverage of off-nadir pixels.

<sup>12</sup> MODIS Surface Reflectance User's Guide, collection 6, <http://modis-sr.ltdri.org>.

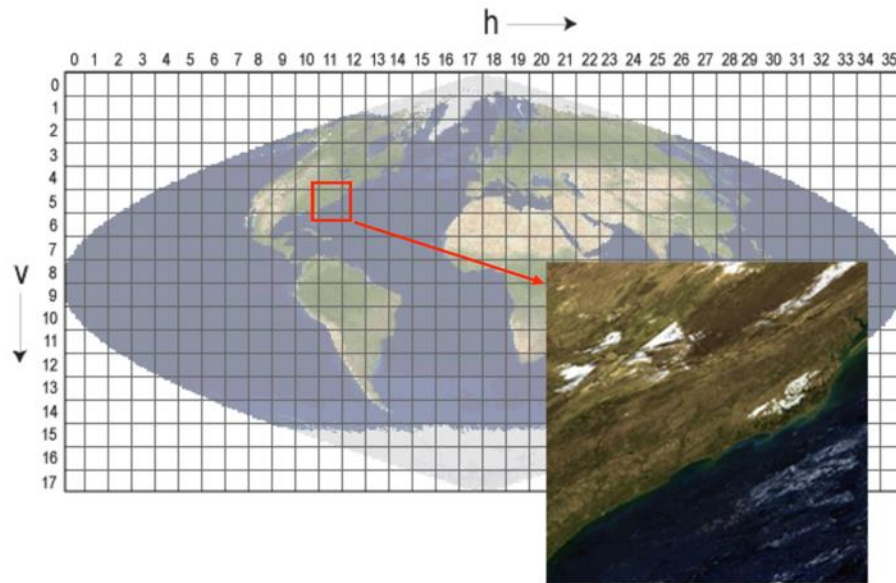


Figure 3.3: The MODIS sinusoidal grid consists of 460 non-overlapping tiles which measure approximately  $10^{\circ} \times 10^{\circ}$ . (Source: Strabala)

### 3.1.3 MODIS Land Surface Product Used (MOD09/ MYD09):

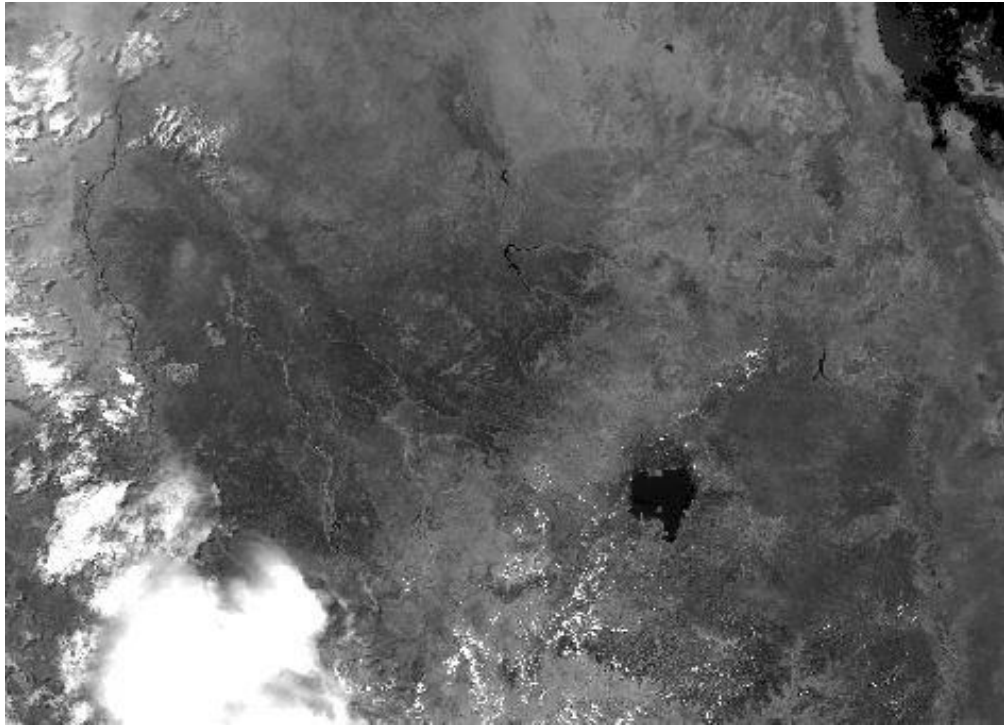
#### MOD09GQ/ MYD09GQ: (Vermote 2015)

MODIS Terra/Aqua Surface Reflectance Daily L2G Global 250. It provides an estimate of the surface spectral reflectance of Terra MODIS 250 m bands 1-2 corrected for atmospheric conditions like gasses, aerosols, and Rayleigh scattering. Along with the 250 m bands are the QC 250 m layer and five observation layers. The SDS layers in this dataset which were considered are:

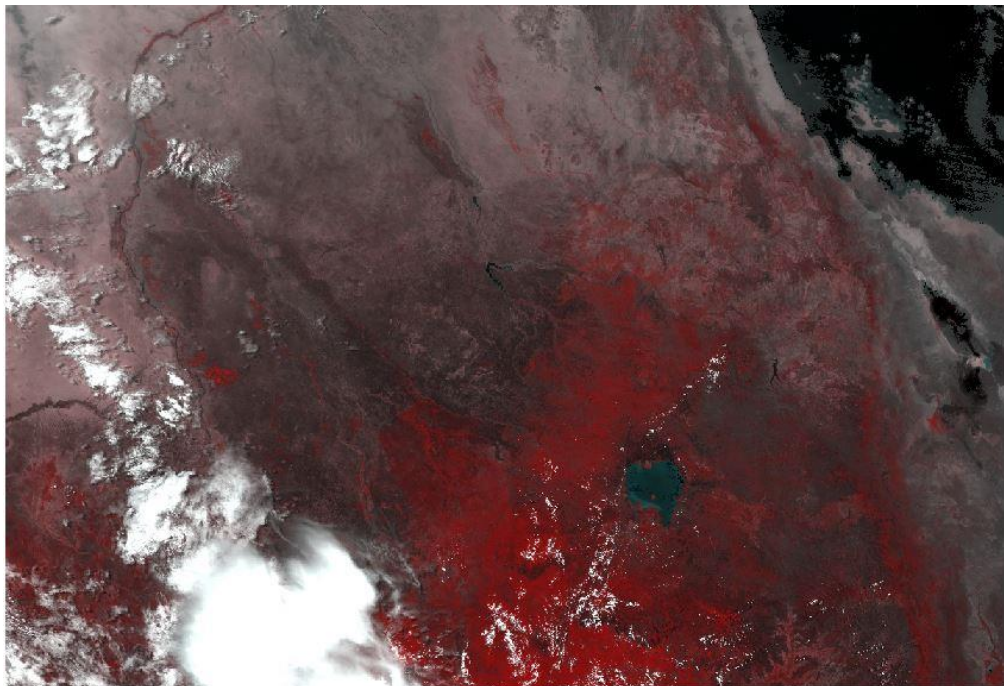
- sur\_refl\_b01\_1: 250m Surface Reflectance Band 1 (620-670 nm), (see Figure 3.4).
- sur\_refl\_b02\_1: 250m Surface Reflectance Band 2 (841-876 nm).
- QC\_250m\_1: 250m Reflectance Band Quality. (Instead state QA from 09GA is used).
- obscov\_1: Observation Coverage (percentage of the grid cell area covered by the observation).
- num\_observations: number of observations within a pixel.

This product is intended to be used in conjunction with the quality and viewing geometry information of the 500 m product (MOD09GA/ MYD09GA), as the QC SDS within this layer is not updated from collection 3 onwards of MODIS data product.





*Figure 3.4: MOD09GQ Band 2 (NIR) gray-scale surface reflectance product. The data was collected on June 2, 2017 over Sudan, Eritrea and Ethiopia.*



*Figure 3.5: MOD09GQ Band (2,1,1) False colour surface reflectance product. The data was collected on June 2, 2017 over Sudan, Eritrea and Ethiopia.*

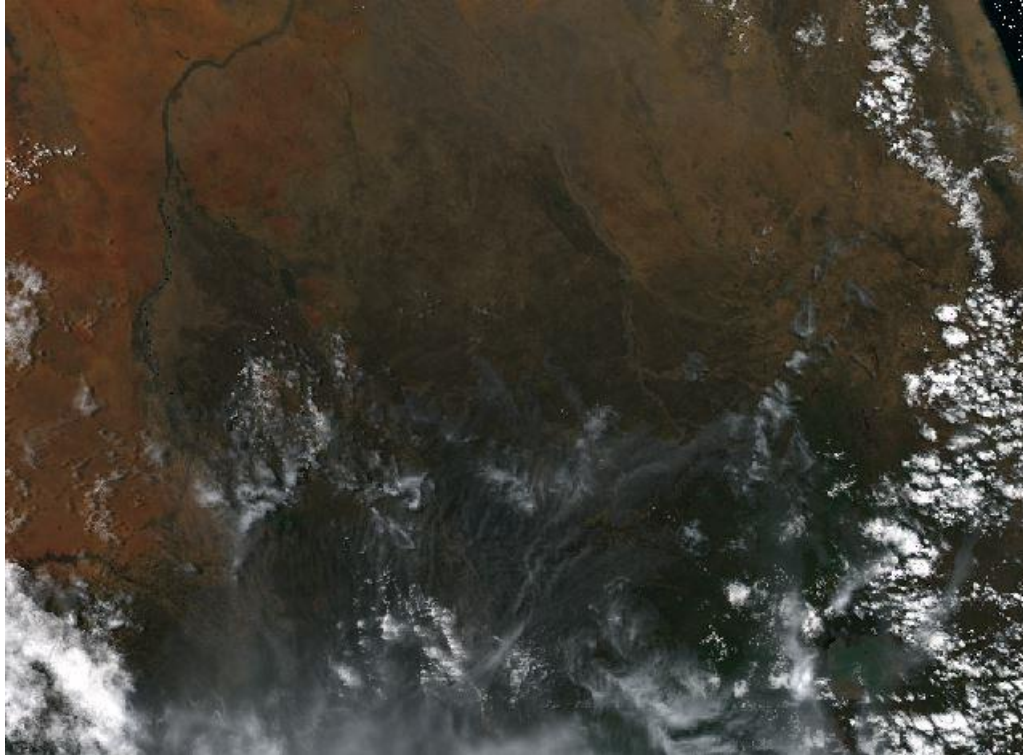
#### MOD09GA/ MYD09GA:(Vermote 2015)

MODIS Terra/Aqua Surface Reflectance Daily L2G Global 500 m and 1 km. It provides an estimate of the surface spectral reflectance of Terra MODIS Bands 1 through 7 and these

are also similarly corrected as MOD09GQ/ MYD09GQ. Provided along with the 500 m reflectance and four observation bands are a set of nine 1 km observation bands. The SDS layers in this dataset which were considered are:

- sur\_refl\_b01\_1: 500m Surface Reflectance Band 1 (620-670 nm)
- sur\_refl\_b02\_1: 500m Surface Reflectance Band 2 (841-876 nm).
- sur\_refl\_b03\_1: 500m Surface Reflectance Band 3 (459-479 nm).
- sur\_refl\_b04\_1: 500m Surface Reflectance Band 4 (545-565 nm).
- sur\_refl\_b05\_1: 500m Surface Reflectance Band 5 (1230-1250 nm).
- sur\_refl\_b06\_1: 500m Surface Reflectance Band 6 (1628-1652 nm).
- sur\_refl\_b07\_1: 500m Surface Reflectance Band 7 (2105-2155 nm).
- state\_1km\_1: Reflectance Data State, used as quality layer (see Table 8)).
- obscov\_500m\_1: Observation coverage.
- num\_observations\_500m: Number of observations within a pixel.

The reflectance layers from the MOD09GA are used as the source data for many of the MODIS land products. The quality SDS (state\_1km\_1) in this dataset is Big-Endian bit coded, and supplements the pixel quality or state information in the image.



*Figure 3.6: MYD09GA RGB Band (1,4,3) True colour surface reflectance product. The data was collected on June 1, 2016 over Sudan, Eritrea and Ethiopia.*

### 3.1.3.1 MODIS Naming Convention:

MOD09GQ.A2017152.h21v07.006.2017154030207.hdf

MYD09GA.A2014153.h21v07.006.2015287030226.hdf

Product Name: <u>MOD09GQ/</u> <u>MYD09GA</u>	MOD- Terra MODIS MYD- Aqua MODIS  09GQ / 09GA- Surface Reflectance product
<u>A2017152</u>	Julian date of acquisition (A-YYYYDDD)
<u>h21v07</u>	Tile identifier. h- horizontal axis(XX), v- vertical axis(YY)
<u>006</u>	Collection Version
2017154030207	Julian Date of production.(YYYYDDDDHHMMSS)
.hdf	Data format. (HDF EOS)

### 3.1.4 PROBA-V: (Wolters, 2018)

The PROBA-V (Project for On-Board Autonomy–Vegetation), is the fourth satellite in the European Space Agency's PROBA series. It was launched on 6<sup>th</sup> May 2013 and has been developed and built by QinetiQ Space N.V and subcontractors for the Directorate of Technology (DTEC) of ESA.



Figure 3.7: PROBA-V spacecraft in orbit.( Source: ESA)



PROBA-V was aimed to bridge the gap in space-borne vegetation measurements between SPOT-VGT and the Sentinel-3 satellite, with an objective to ensure continuity with the SPOT-VGT mission's heritage. The PROBA-V mission had a designed life of 2.5 years, but the platform performance was well within requirements and in May 2015 it was decided to extend the mission with another 2.5 years until May 2018. Further extension through October 2019 was decided on in May 2017

It has a Field Of View of  $102^\circ$ , with a swath width of 2285 km. The swath width ensures a daily global coverage of the land masses in the latitudes  $35^\circ$  and  $75^\circ$  North and in the latitudes between  $35^\circ$  and  $56^\circ$  South, with a 90% daily coverage of equatorial zones and 100% two-daily imaging, during day time, of the land masses in the latitudes between  $35^\circ$  North and  $35^\circ$  South.<sup>13</sup>

It is a multispectral across-track scanner (push-broom) spectrometer with 4 spectral bands; Blue (at  $0.463 \mu\text{m}$ ), Red (at  $0.655 \mu\text{m}$ ), NIR (at  $0.837 \mu\text{m}$ ), and SWIR (at  $1.603 \mu\text{m}$ ). The VNIR detector consists of three lines of 5200 pixels, and SWIR detector is a linear array composed of three staggered detectors of 1024 pixels (see Figure 3.8). The optical design consists of three cameras. Each camera has two focal planes, one for the short wave infrared (SWIR) and one for the visible and near-infrared (VNIR) bands. It has been designed such that the NIR band observes the Earth first, followed by the RED, BLUE, and SWIR bands.

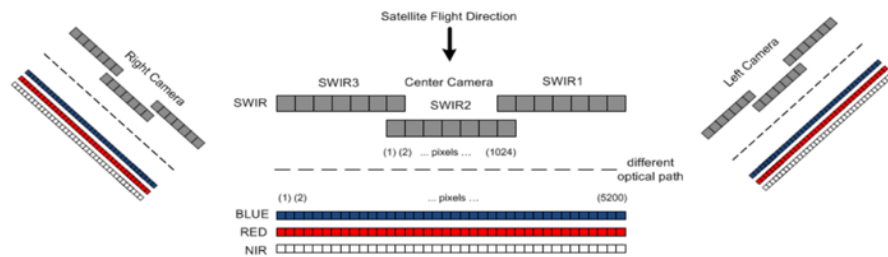


Figure 3.8: PROBA-V Instrument Layout.(Source: Wolters, 2018)

### Specifications:<sup>13</sup>

Orbit	819 - 827 Km, sun - synchronous.
LTDN (Local Time on Descending Node)	10:30 hours with a drift limited between 10:30 and 11:30 AM during the mission lifetime, because the satellite has no

<sup>13</sup> <https://earth.esa.int/web/eoportal/satellite-missions/p/proba-v>

	onboard propellant and the overpass time will continue to decrease as a result of increasing atmospheric drag.
Inclination	98.7°
Temporal Resolution	Daily near-global coverage (90%).
Spatial Resolution	330m, 100m and 1km.  VNIR: 100 m at nadir, 360 m at edge of swath.  SWIR : 200 m at nadir, 600 m at edge of swath.
Spectral Bands	B1: Blue- 0.415-0.500 $\mu\text{m}$  B2: Red- 0.580-0.770 $\mu\text{m}$  B3: NIR- 0.730-0.960 $\mu\text{m}$  B4: SWIR- 1.480-1.760 $\mu\text{m}$
Swath Dimension	2285 km.
Telescope	3 identical TMA telescopes mounted on an optical bench together.
Size	200 mm x 812 mm x 350 mm
Weight	35 kg
Power	30 W
Data Rate	7.15 Mbit/s (after compression).
IGFOV	96.9 m for VNIR and 193.8 m for SWIR

### 3.1.5 PROBA-V Processing:<sup>13</sup>

PROBA-V has data processing on 4 levels. Level 0, it ingests raw data and produces instrument packets, sorted by time, duplicated packets removed and wrong packets taken out. Level 1 data is used for geometric and radiometric processing. Level 1A produces files containing instrument data annotated with instrument and platform housekeeping as well as ancillary (position, velocity, pointing) information. Whereas Level 1B consists of associated



geometry and geometry derived information at pixel level. Level 1C consists of radiometrically corrected L1B at pixel level.

The Level 1C data is used as an input for Level 2 data which performs mapping and SWIR mosaicking, snow/ ice detection, cloud and cloud shadow detection and atmospheric correction. The compositing into synthesis images is performed at Level 3, the aim is to optimally combine multiple observations into a single and cloud-free synthesis image producing the final user accessible data.(see Figure 3.9).

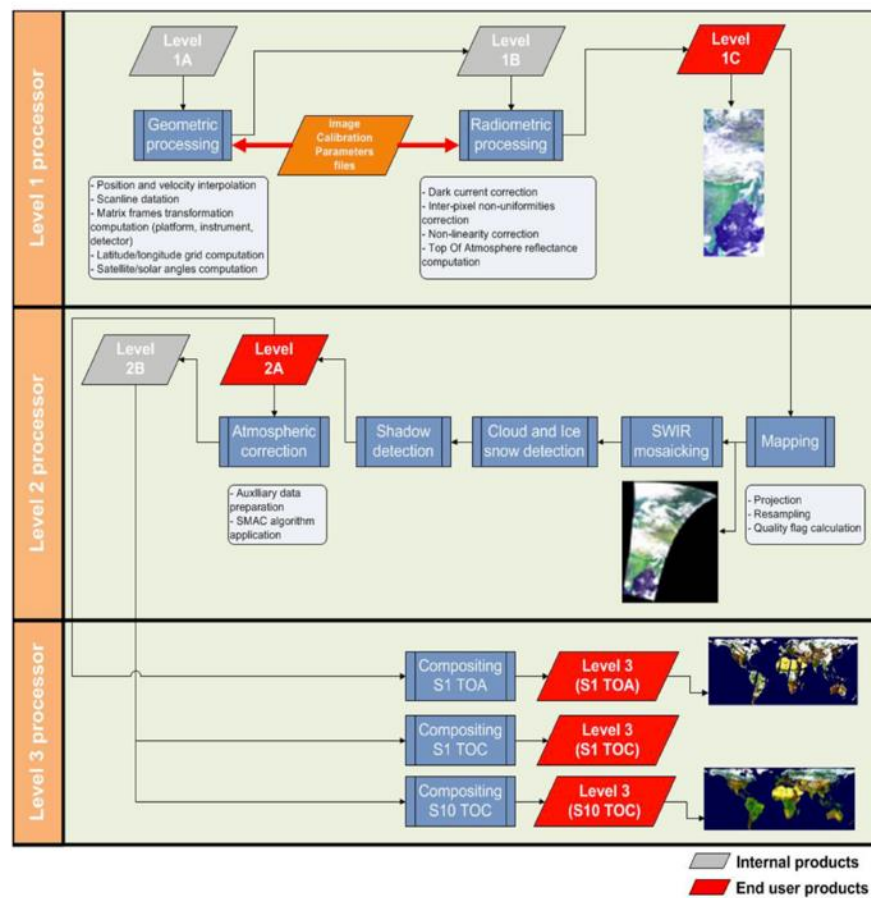


Figure 3.9: PROBA-V processing chain flowchart (Source: Wolters 2018).

The compositing rules for the 300 m and 100 m syntheses are: observations covered by all spectral bands are preferred over observations covered by only a few spectral bands. Observations with a good pixel quality indicator for all bands are preferred over observations of less quality. Cloud-free observations are preferred over ice/snow observations, which in turn are preferred over cloudy observations. The final PROBA-V data products are projected in a standard WGS-84 projection similar as for the SPOT-VGT products.

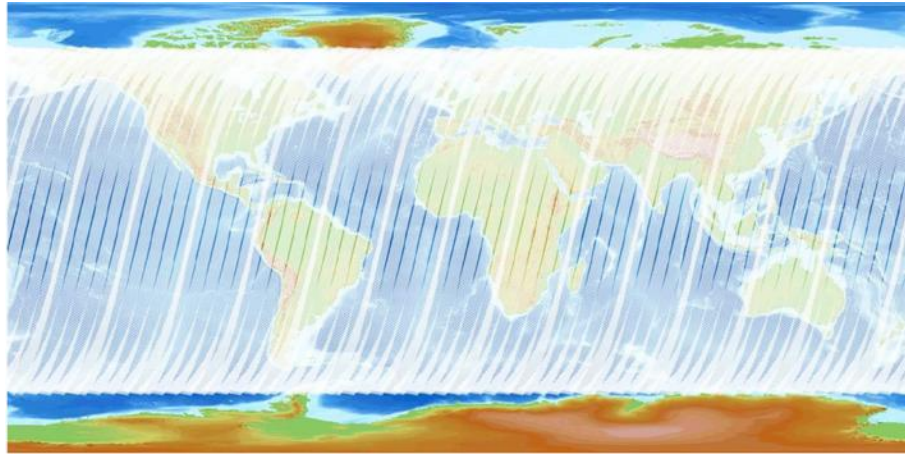


Figure 3.10: The brighter white areas indicate overlapping observations. (Source: Wolters 2018)

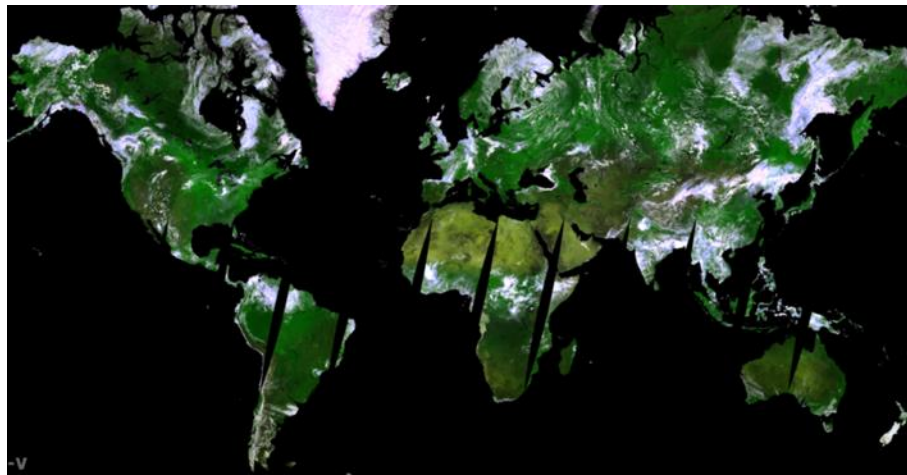
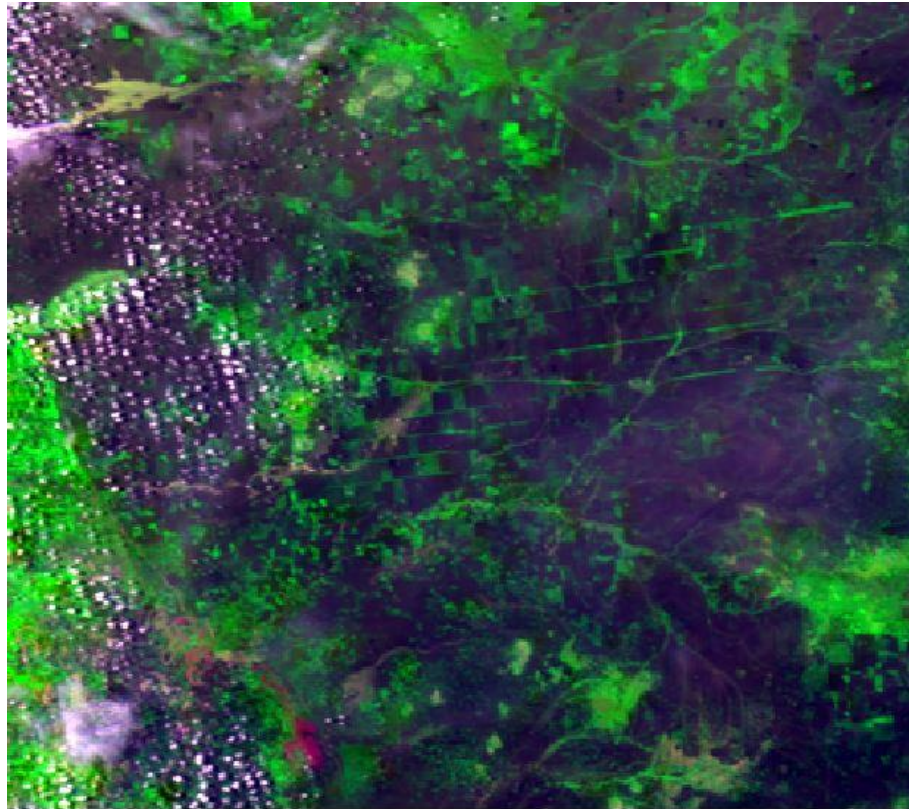


Figure 3.11: PROBA-V Colour composite(1,2,3), Red(Band 1), BLUE (Band 2), NIR (Band 3) as seen on 20/06/2017.

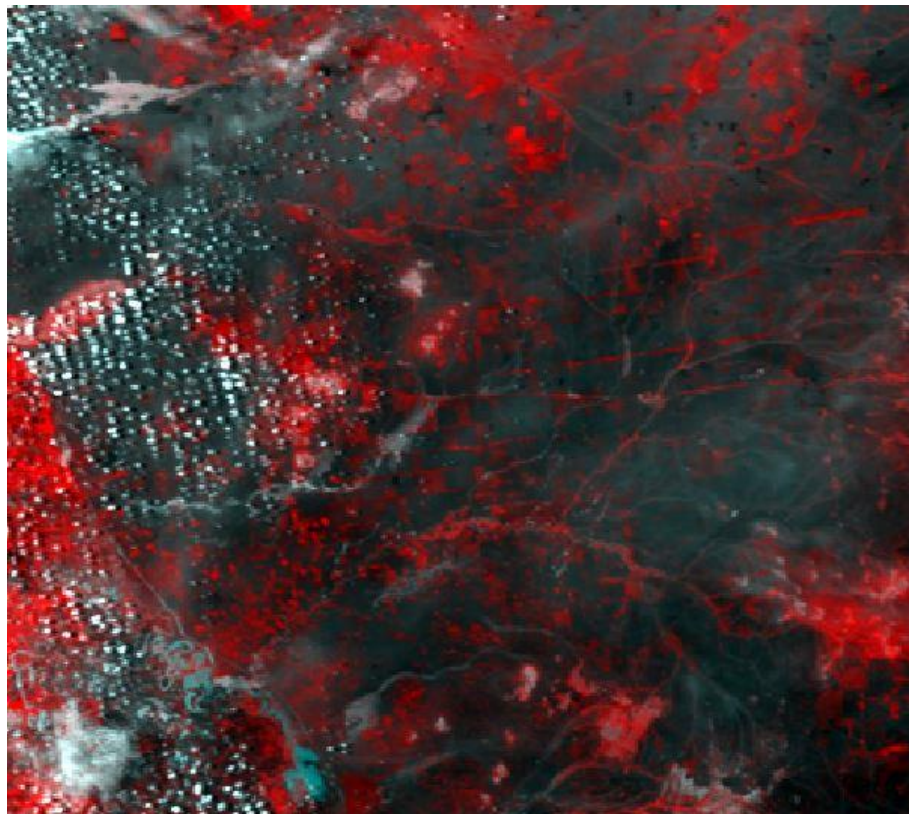
### 3.1.6 PROBA-V Product Used :

#### S1-TOC (Top of Canopy) 330m:

This is Level 3 ground reflectance values at a spatial resolution of 330m, with best-quality and cloud-free data selected for the daily composite time period. It has 12 SDS layer; Geometry layer: Sun Azimuth Angle (SZA), Viewing Azimuth Angle (VAA) SWIR, Viewing Azimuth Angle (VAA) VNIR, Solar Zenith Angle (SZA) SWIR, Solar Zenith Angle (SZA) VNIR, Viewing Zenith Angle (VZA) SWIR, Viewing Zenith Angle (VZA) VNIR. Radiometry Layer: Band 1 (Biro, Pradhan et al.), Band 2 (NIR), Band 3 (BLUE), Band 4 (SWIR). Quality Layer: Quality State Map.



*Figure 3.12: S1 TOC 330m Band (1,2,3) Colour composite of surface reflectance product. The data collected on 27/08/2017 over Sudan.*



*Figure 3.13: S1 TOC 330m Band (2,1,3) False Colour composite of surface reflectance product. The data collected on 27/08/2017 over Sudan.*



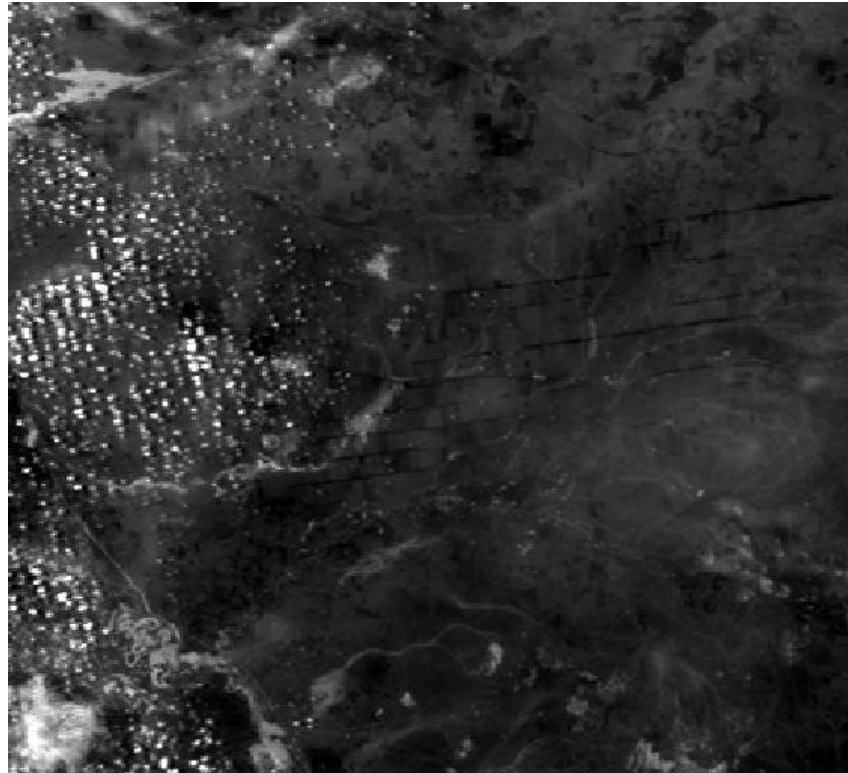


Figure 3.14: S1 TOC 330m Band (1) Red Band surface reflectance product. The data collected on 27/08/2017 over Sudan.

#### 3.1.6.1 PROBA-V Naming Convention:

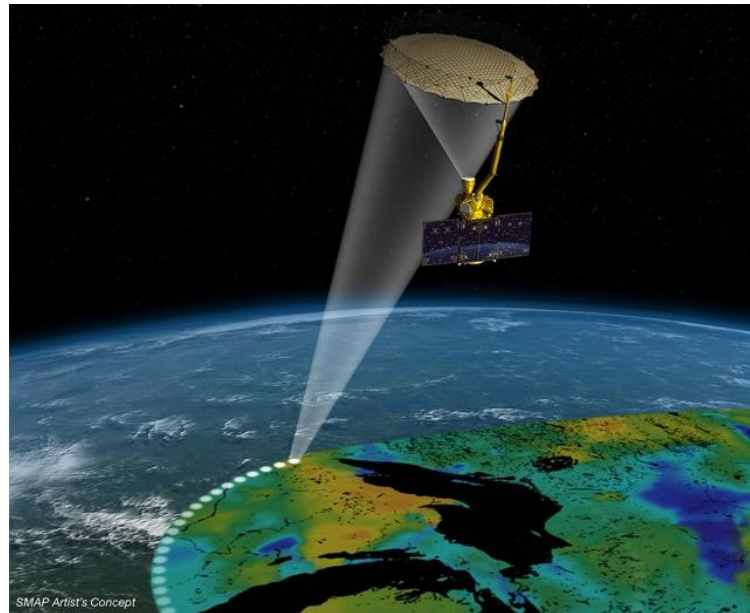
PROBAV\_S1\_TOC\_20140601\_333M\_V101.HDF5

<u>PROBAV</u>	Sensor name
S1_TOC	Composite S1 level 3 at Top of Canopy.
20140601	Date of acquisition (YYYYMMDD)
333M	Resolution
V101	Collection Version
.HDF5	Data format. (HDF EOS)

#### 3.1.7 SMAP:

Soil Moisture Active Passive (SMAP) is a remote-sensing observatory with two instruments a synthetic aperture radar (SAR) and a radiometer that maps soil moisture and determine the freeze or thaw state of the area being mapped. The SMAP instrument

incorporates an L-band radar and an L-band radiometer that share a single feedhorn and parabolic mesh reflector at an offset from nadir and rotates about the nadir axis at 14.6 rpm (nominal), providing a conically scanning antenna beam with a surface incidence angle of approximately  $40^\circ$  (R. Reichle 2012).



*Figure 3.15: SMAP concept figure. (Source: NASA)*

Soil moisture initially was measured at scales ranging from point scale (in situ) to satellite footprint scales of about 40 km at various temporal resolutions. Measurement networks of in-situ sensors such as USDA's Soil Climate Analysis Network (SCAN) or NOAA's Climate Reference Network (CRN) have potentially high soil moisture measurement accuracy but are spatially very sparse. Radar scatterometers have also been used to retrieve soil moisture like the European Remote Sensing Satellite (ERS) C-band scatterometer with resolution of 50 km has been used to retrieve surface soil moisture over sparsely vegetated regions with moderate accuracy.

But all these measurement technologies do not have high spatial and temporal resolution and lacks wide spatial coverage, optimal sensing depth, and desired accuracy in retrieved soil moisture over moderate vegetation conditions. Thus, to meet these needs, NASA's Soil Moisture Active Passive (SMAP) mission was launched which uses the L-band radar and L-band radiometer for concurrent, coincident measurements integrated as a single observation system. This combination takes advantage of the relative strengths of both active (radar) and passive (radiometer) microwave remote sensing for soil moisture mapping.

SMAP, was launched on January 31, 2015. It is developed by NASA's Jet Propulsion Laboratory, which build the spacecraft, the instrument (except for the radiometer), and the science processing system. NASA Goddard Space Flight Center provide the L-band radiometer and Level 4 science processing. The Canadian Space Agency (CSA) is also a mission partner to provide critical support to science and calibration/validation pre- and post-launch.

**Specifications:**

Orbit	685 Km, sun - synchronous, near – polar. Equator crossings 6 a.m. and 6 pm. local time.
Inclination	98.1°
Temporal Resolution	2 Daily global coverage.
Swath Dimension	1000 km (see Figure 3.16)
Size	1.5 x .9 x .9 m
Weight	944 kg
Power	1450 Watts
Radar: L-Band	Frequency: 1.2 to 1.3 GHz. Polarizations: VV, HH, HV. Relative accuracy (3 km grid): 1 dB (HH and VV), 1.5 dB (HV). Data acquisition: High-resolution (SAR) data acquired over land, low-resolution data acquired globally.
Radiometer: L-Band	Frequency: 1.41 GHz. Polarizations: H, V, 3rd & 4th Stokes. Relative accuracy (30 km grid): 1.3 K. Data collection: High-rate (sub-band) data acquired over land, low-rate data acquired globally.

Antenna	Shared, 6 m diameter.
---------	-----------------------

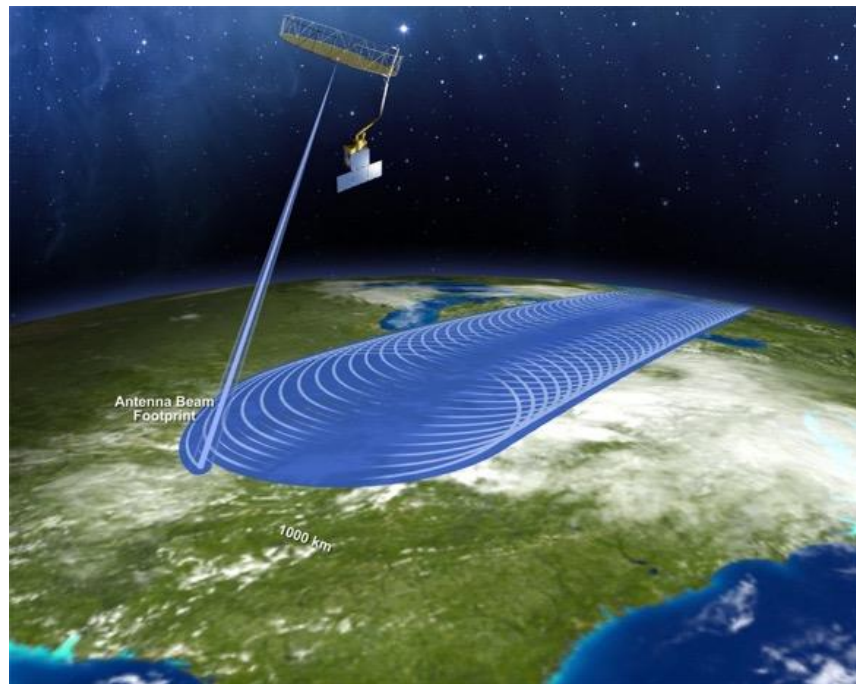


Figure 3.16: SMAP image swath of radar and radiometer.( Source: JPL/ NASA)

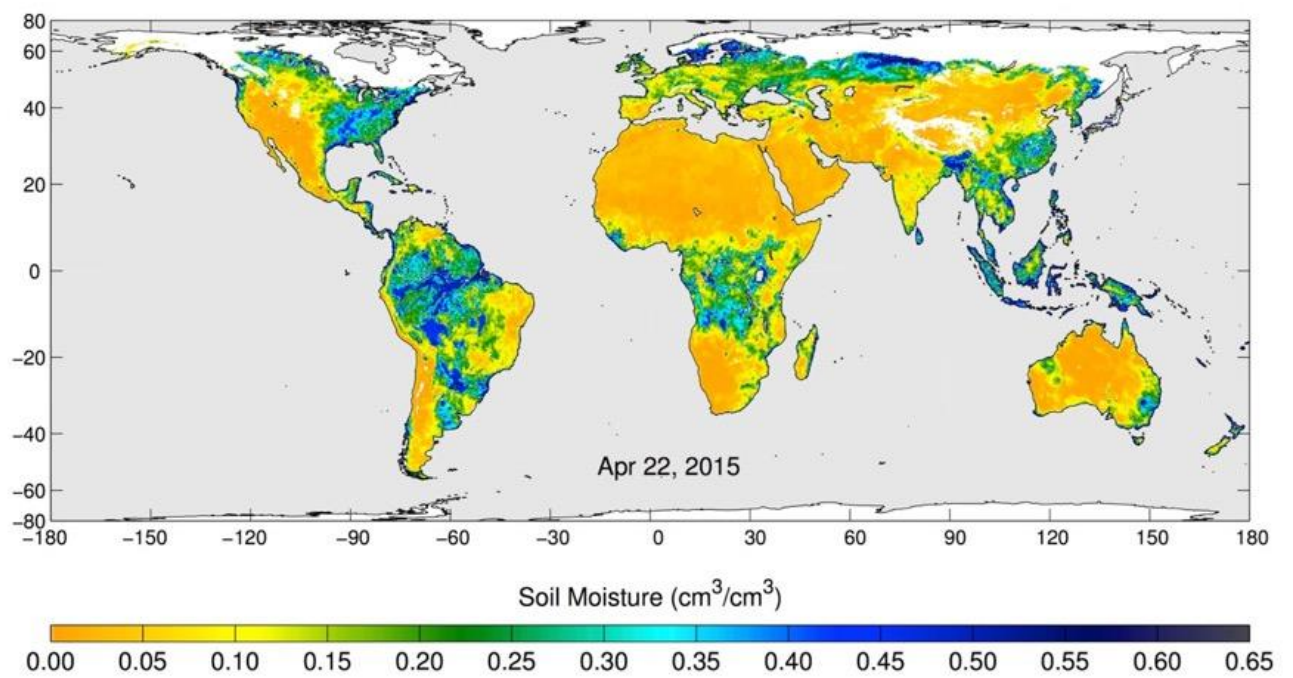


Figure 3.17: Global map of SMAP radiometer. (Source: NASA/JPL-Caltech/GSFC)

### 3.1.8 SMAP Processing:<sup>14</sup>

SMAP has a 5 level processing like most of the other satellites. Level 0 is the reconstructed, unprocessed instrument data at original resolution with all communication artefacts removed and time ordered. Level 1 consisting of Level 1A which is level 0 data time referenced and annotated with ancillary information. Level 1B and 1C data products are calibrated and geolocated instrument measurements of surface radar backscatter cross-section and brightness temperatures. Level 2 products are geophysical retrievals of soil moisture on a fixed Earth grid based on Level 1 products and ancillary information. The Level 2 products are output on a half-orbit basis. Level 3 products are daily composites of Level 2 surface soil moisture and freeze/thaw state data. Level 4 products are model-derived value-added data products of surface and root zone soil moisture and carbon net ecosystem exchange that support key SMAP applications.

In total, the SMAP mission generates 15 distributable data products representing four levels of data processing. Level 1 and Level 2 appear in granules that are based on half orbit granules of the SMAP satellite. Level 3 products are daily global composites of the Level 2 geophysical retrievals for an entire UTC day. Level 4 products contain output from geophysical models utilizing SMAP data. The SMAP data are projected on EASE 2 grid equal area projection.

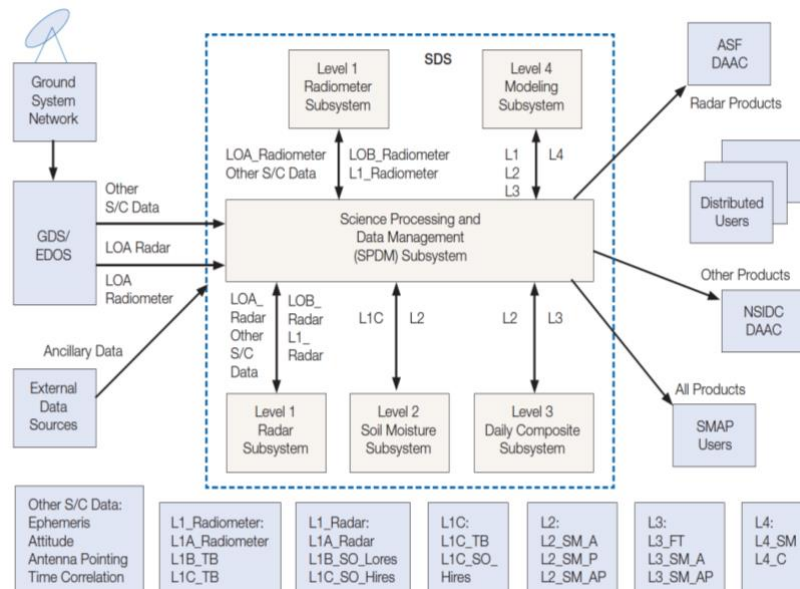


Figure 3.18: SMAP Data processing and data flow. (Source: Entekabi 2014)

<sup>14</sup> <https://smap.jpl.nasa.gov/data/>



### 3.1.9 SMAP Data Used (R. Reichle 2012):

#### L3\_SM\_AP (Enhanced):

The SMAP L3\_SM\_AP (Level 3 Soil Moisture Active Passive) product is a daily global composite of the SMAP L2\_SM\_AP product. It represents gridded data of SMAP radiometer based soil moisture retrieval, ancillary data, and quality-assessment flags on the global Surface soil moisture (0-5 cm) in  $\text{cm}^3/\text{cm}^3$  derived from brightness temperature and sigma nought measurements. The output is on a Earth-fixed, global, cylindrical 9 km Equal-Area Scalable Earth Grid version 2.0 (Das 2015).

To generate the standard L3\_SM\_AP product the processing software ingests one day's worth of L2\_SM\_AP granules and create individual global composites as two-dimensional arrays for each output parameter defined in the L2\_SM\_AP product. Wherever data overlap occurs (typically at high latitudes), data whose acquisition time is closest to the 6:00 am local solar time is chosen (Das 2015).

The dataset consists of total 58 bands that includes the albedo, boresight incidence, EASE column index, ease row index, freeze thaw fraction, latitude, longitude, latitude centroid, longitude centroid, radar water body fraction, retrieval quality flag, roughness coefficient, soil moisture, soil moisture error etc., for both AM and PM dataset.

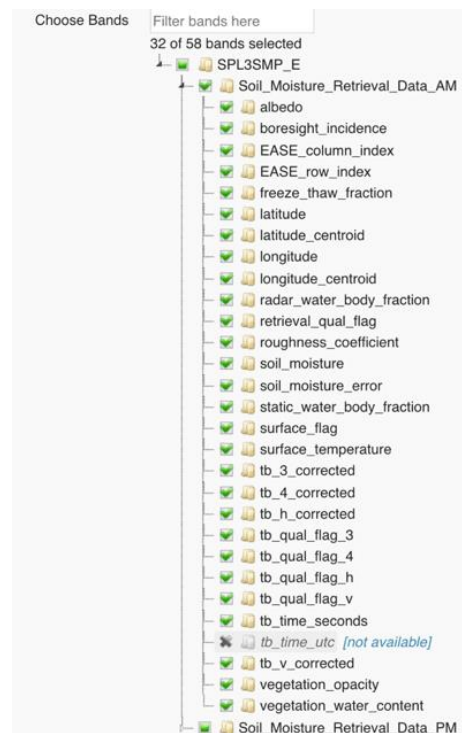


Figure 3.19: Dataset of SMAP L3\_SM\_AP.

### 3.1.9.1 SMAP naming convention:

SMAP\_L3\_SM\_P\_E\_20161207\_R14010\_001\_Soil\_Moisture\_Retrieval\_Data\_AM\_soil\_moisture.tiff

SMAP	Indicates Mission
L3_SM_P	Indicates Specific product: L4- Level 4, SM- Soil Moisture, P- Passive radiometer (Read note below).
E	Enhanced
20161207	Date (YYYYMMDD)
R14010	Composite release ID. R- Release, l- launch indicator (post-launch standard data), V(4)- 1 digit Major version Number, vvv(010)- 3 digit Minor version
001	Number of times the file was generated under the same version for a particular date/time interval (001: 1st time).
Soil_Moisture_Retrieval_Data_AM	File type in level 3, AM- time
Soil_moisture	Data type
.tiff	Data format. (geotiff, customized and downloaded)

#### **Note<sup>15</sup>:**

The SMAP dataset is using only passive radiometer data because NASA's SMAP observatory had determined that its radar, one of the satellite's two science instruments, can no longer return data. However, the mission continues to produce high-quality science measurements supporting SMAP's objectives with its passive radiometer instrument. On July 7, 2015, SMAP's radar stopped transmitting due to an anomaly involving the radar's high-power amplifier (HPA).

<sup>15</sup> <https://www.jpl.nasa.gov/news/news.php?feature=4710>

Initially SMAP's active radar and passive radiometer instruments were designed to complement each other and mitigate the limitations of each measurement alone. The radar enabled high-resolution measurements of up to 3km, but with lower accuracy for sensing surface soil moisture. In contrast, the microwave radiometer is more accurate in its measurements, but has lower resolution of about 40km. By combining the active and passive measurements, SMAP was designed to estimate soil moisture at a resolution of about 9km.

### 3.2 Data Availability:

Availability of MODIS and PROBA-V data are affected by various factors like scan angle as shown in Figure 3.1 the satellite creates a gap due to the scan angle and doesn't create tiles for certain days over specific regions. Also because of cloud cover most of the data were not available after post processing as the data considered is majorly covered by the monsoon season of the study area. The following charts shows the monthly availability of each sensor data:

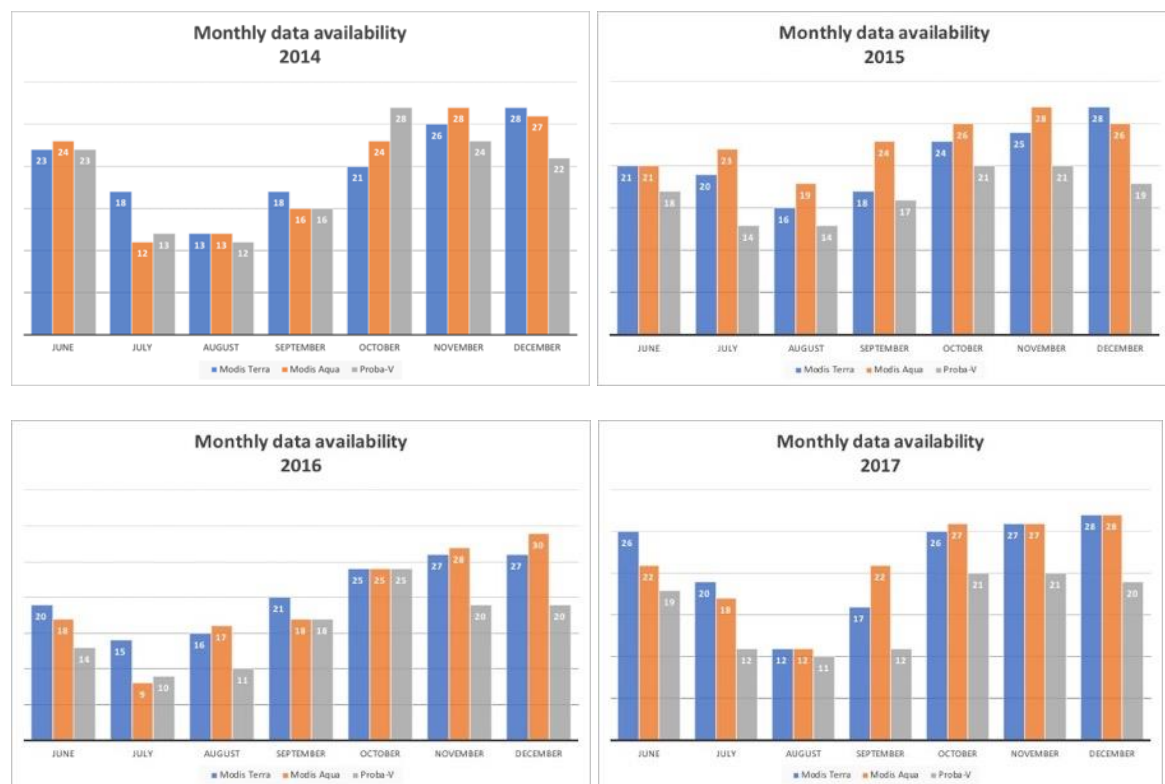


Figure 3.20 : Showing available scenes from 2014-2015 for MODIS (Aqua/Terra) and PROBA-V.

As observed in Figure 3.20 the drop in availability is mainly in July and August because of increased cloud cover. The following charts Figure 3.21 and Figure 3.22 provides an overall information about the data availability from 2014-2017:

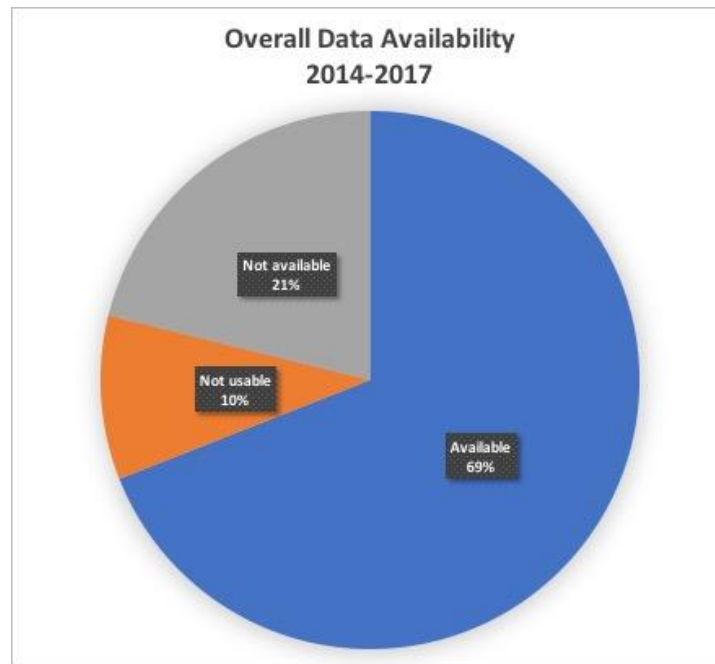


Figure 3.21: Total Image availability.



Figure 3.22: Charts showing the yearly data availability, along with the total data available.

### 3.3 Study Area:

#### 3.3.1 Location

The study area Gadambalia is located in the state of Gadarif near the town of Al Qadarif in south-eastern Sudan and lies approximately between the latitude  $13^{\circ} 15'N$  to  $14^{\circ} 15'N$  and longitude  $34^{\circ} 30'E$  to  $35^{\circ} 20'E$ . It is located at an average altitude of 608 m with a gentle slope of 0-1% the total area covered is about 2970 km<sup>2</sup>.

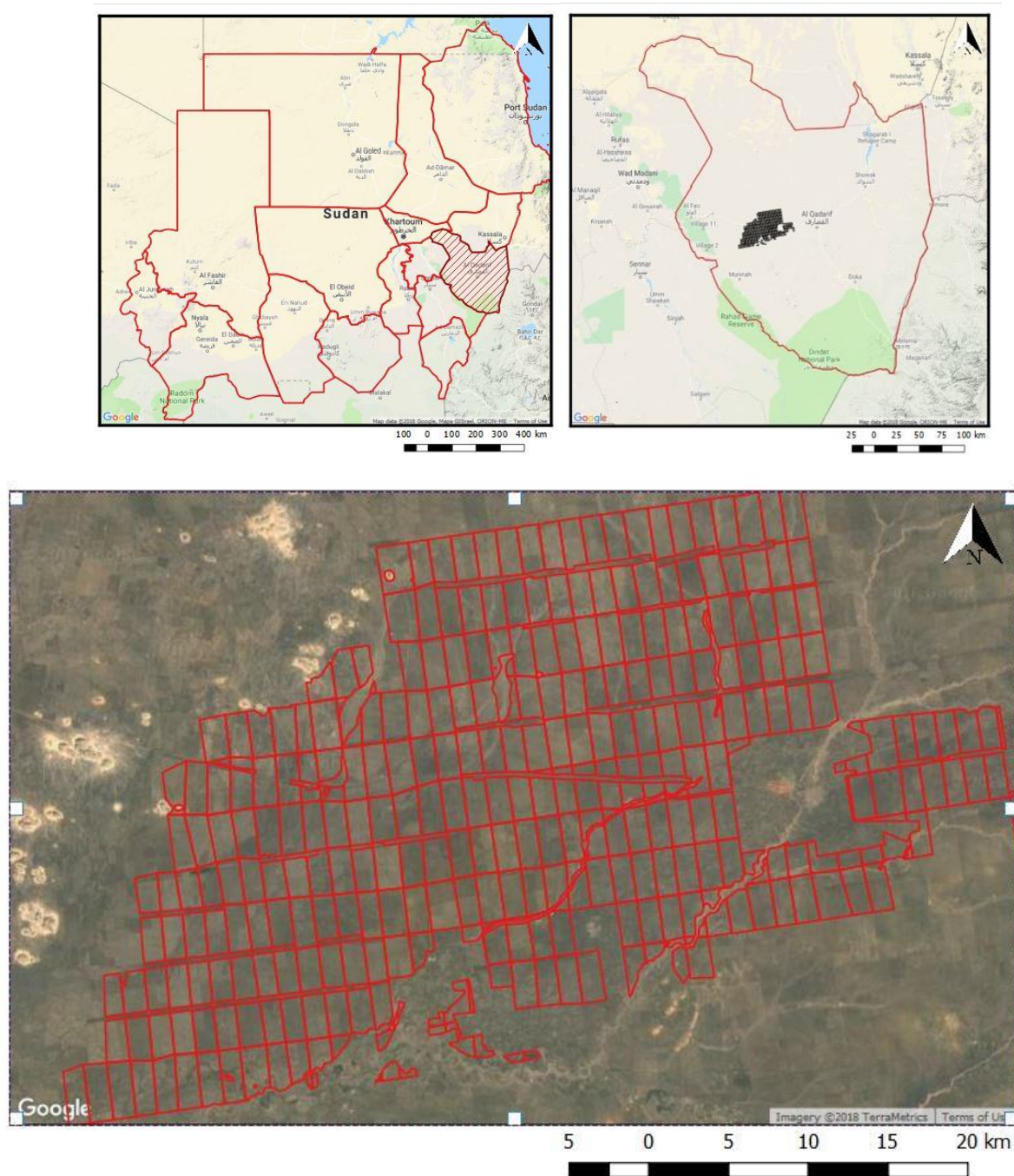


Figure 3.23: Study Area. Al Qadarif Sudan. (Base map: Google Hybrid)



# Mechanised Farming Scheme of Gadambalia

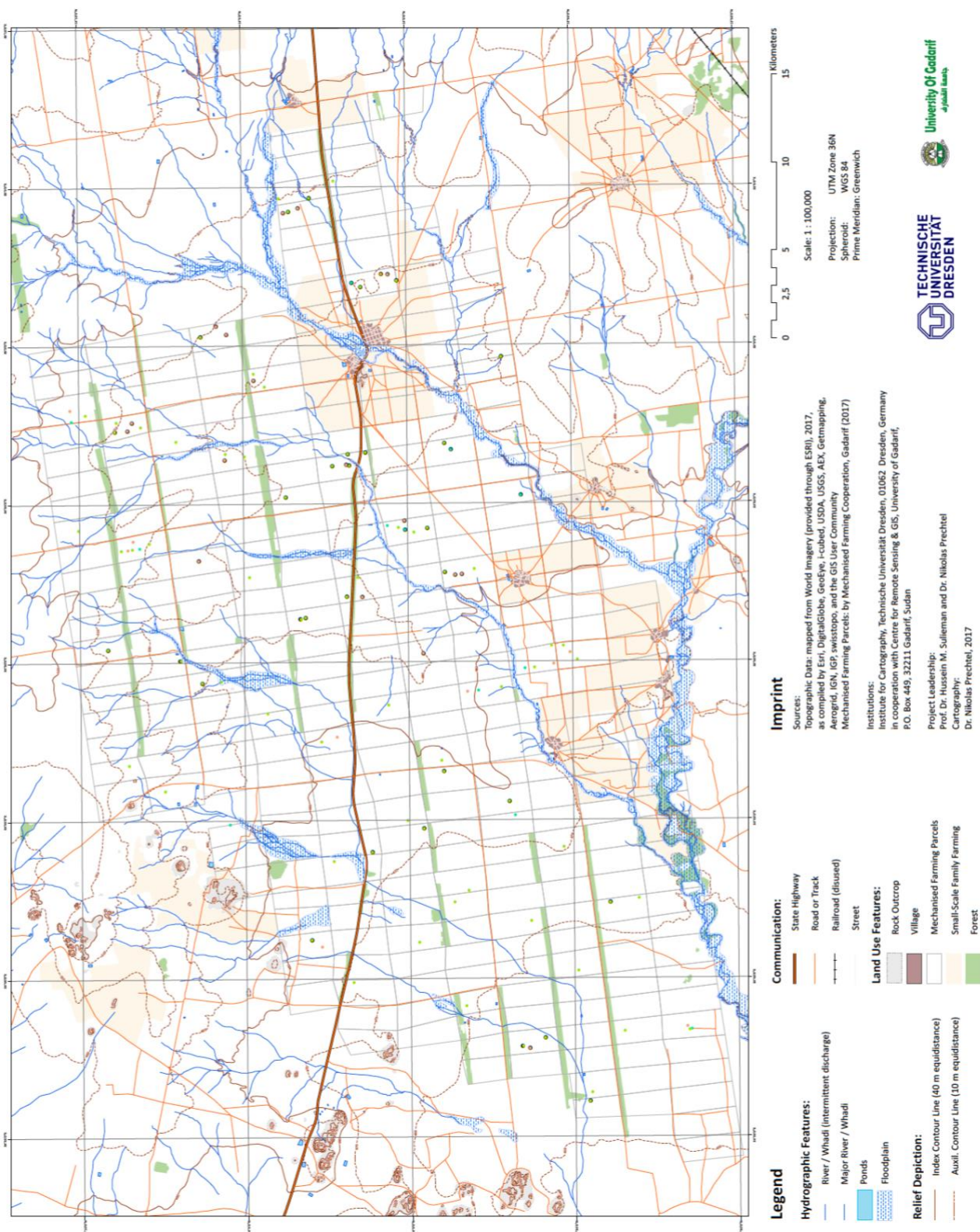


Figure 3.24: Detailed Scheme Map (Source: Dr. Nikolas Pechtel, TUD)

### 3.3.2 Climate and Soil:

The study area falls in the eco-climatic zone of the African Sahel (see Figure 3.25). The region lies between the Sahara in North and Savanna in south and the coast in west. It has semi-arid climatic conditions, with aridity index ranging from 0.2 to 0.4 (Biro, Pradhan et al. 2013), the rainfall is limited to only few months ranging from June-September peaking at July-August. The temperature ranges from 22° C in winter to a mean maximum of 37° C in summer while it may reach about 40°-42°C in April or May (Biro, Pradhan et al. 2012).

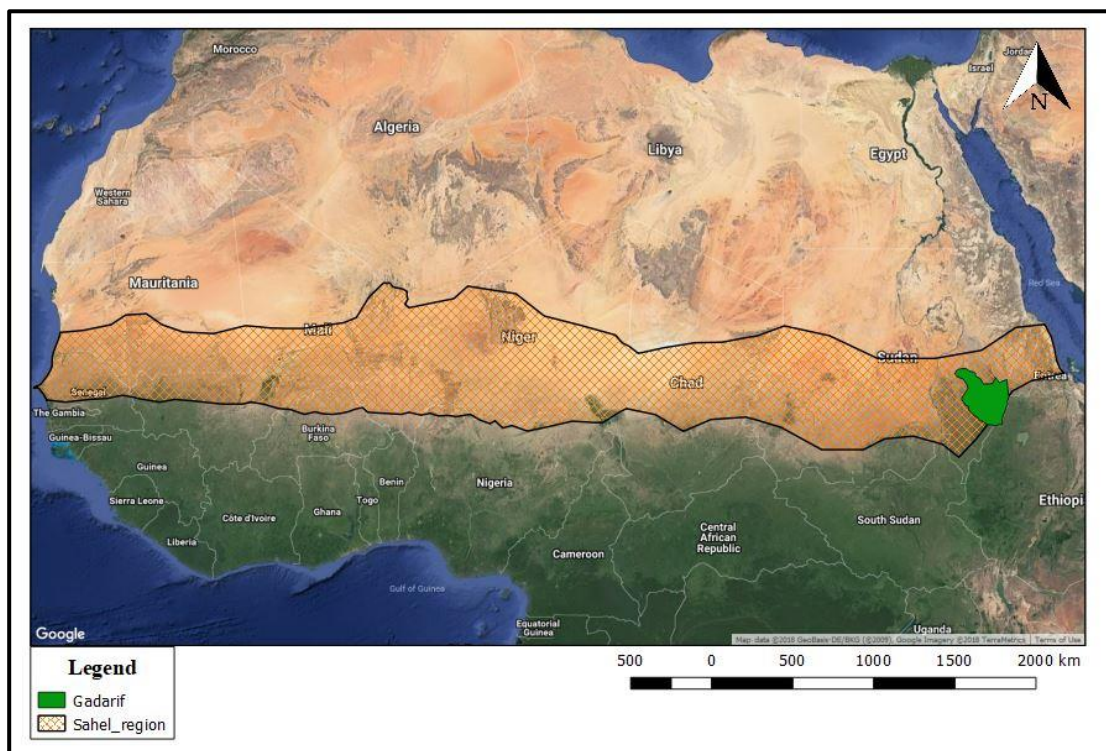


Figure 3.25: African Sahel extent.

Rainfall in the Gadarif region varies from 400 mm to over 700 mm with an annual average of 591 mm. The average monthly rainfall in June at the beginning of the rainfall season is around 66 mm at Gadambalia. At the peak of the rainfall season in August, rainfall is about 176 mm and ranges down to 72 mm in September towards the end of the rainy season. Although the total amount of rainfall during the rainy season seems sufficient the high temperatures of the rainy months substantially reduce the amount of effective rainfall (selwar 2014).

Another typical characteristic of the rainfall in this region according to (Hermance and Sulieman 2018) is most of the cumulative rainfall in this region is intense but localized and have convective storm activity associated with the seasonal northward migration of the tropical rain belt into this region. As most downpour from these convective storm have a lateral scale

a storm system passing through a region may result in two contrasting types of local extremes: one area may experience flash flooding, while an adjacent area may experience dry conditions. As the rainfall plays a dominant role in the weather conditions of this region because of its temperature extremes, any form of rainfall anomaly has a major effect.

The soil type in this region are heavy dark cracking clays. The clay content is very high, amounting 70 to 80%, (Biro, Pradhan et al. 2013), Montmorillonite is the common clay type in the area and its content varies between 47% and 75 %. The soil is moderately fertile, and has low content of organic matter and nitrogen (Biro, Pradhan et al. 2012)

### **3.3.3 Agriculture:**

The study area is a section of the Mechanised farming corporation namely Gadambalia, the mechanised farming was first introduced in 1940's in Gadarif by the British army and eventually became the largest rainfed mechanised farming area in Sudan. The main crops grown under this scheme are Sorghum and Sesame with some Millet, Cotton and Guar. It contributes to about one third of the national production of sorghum and about half of the sesame. According to FAOSTAT four countries, India, China, Myanmar, and Sudan, contribute >60% to the total world production of sesame seeds. Gadarif along with Kordofan region forms the back bone of the national economy as the Sesame and Sorghum are mostly exported<sup>16</sup>.

#### **3.3.3.1 Sesame:**

As the study area contributes to about half of the sesame production in the country it is important to know the growing conditions of this crop, so it can be related to the extremities of the weather conditions in the region. <sup>17</sup>Sesame varieties have adapted to many soil types and thrives best on well-drained, fertile soils of medium texture and neutral pH, however these have low tolerance for soils with high salt and water-logged conditions. Sesame is drought-tolerant, in part due to its extensive root system. However, it requires adequate moisture for germination and early growth.

While the crop survives drought as well as presence of excess water, the yields are significantly lower in either conditions. <sup>18</sup>Planting sesame is the most critical phase of its

---

<sup>16</sup> <https://reliefweb.int/sites/reliefweb.int/files/resources/Sudan%20Profiles%20Final%20en.pdf>

<sup>17</sup> <https://en.wikipedia.org/wiki/Sesame#Processing>

<sup>18</sup> [https://www.extension.iastate.edu/alternativeag/cropproduction/pdf/sesame\\_crop\\_guide.pdf](https://www.extension.iastate.edu/alternativeag/cropproduction/pdf/sesame_crop_guide.pdf)



management, successful establishment of sesame requires careful seedbed preparation and close attention to soil moisture. Sesame will not emerge from soils that are even slightly crusted and needs fairly warm soil temperatures of 21°C or more.

Ideally Sesame is sown during June-July, grows in August-October, and harvested during November-December.

### 3.3.3.2 Sorghum:

<sup>19</sup>Sorghum is mainly cultivated in drier areas, especially on low potential, shallow soils with high clay content. It usually grows poorly on sandy soils, except where a heavy textured subsoil is present. Sorghum is more tolerant of alkaline salts than other grain crops and can therefore be successfully cultivated on soils with a pH (KCl) between 5.5 and 8.5. Sorghum can better tolerate short periods of waterlogging. Soils with a clay percentage of between 10 and 30 % are optimal for sorghum production. The optimum growth requirements of sorghum plants, in order to exploit its inherit yield potential, are a deep well-drained fertile soil, a medium to good and fairly stable rainfall pattern during the growing season, temperate to warm weather (20–30°C) and a frost-free period.

Sorghum is able to tolerate drought better than most other grain crops. This can be attributed to an exceptionally well-developed and finely branched root system, which is very efficient in the absorption of water. It has a small leaf area per plant, which limits transpiration. Similar to sesame it requires better soil moisture during its initial growth period (seedling generation) but also fluctuating rainfall conditions of approximately 400mm during it's growing period. Sorghum is one of the best crops for producing a good yield in warm or hot, dry, non-irrigated conditions. It also grows well in hot humid conditions. Its chief disadvantage is that it depletes the soil nutrient supply, and the succeeding crop must be fairly heavily fertilized (Doraiswamy and Hodges 1979).

Ideally sorghum is sown during June-July, grows in August-November, and is harvested during December.

---

<sup>19</sup> [http://www.nda.agric.za/docs/Infopaks/FieldCrops\\_Sorghum.pdf](http://www.nda.agric.za/docs/Infopaks/FieldCrops_Sorghum.pdf)

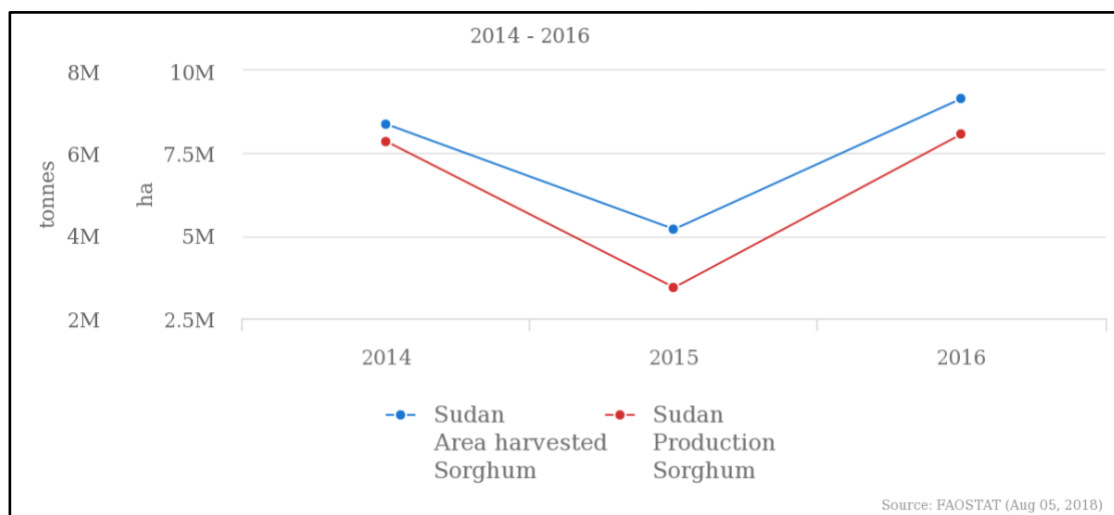


Figure 3.26: Sorghum Production/Yield of Sorghum in Sudan between 014-2016. (Source: FAOSTAT).

## 4. Methodology

This chapter focuses on the methods used and the workflow followed for execution of results. The workflow is mainly divided into three parts: 1) Data Acquisition 2) Data Post-Processing and 3) Main workflow.

### 4.1 Data Acquisition:

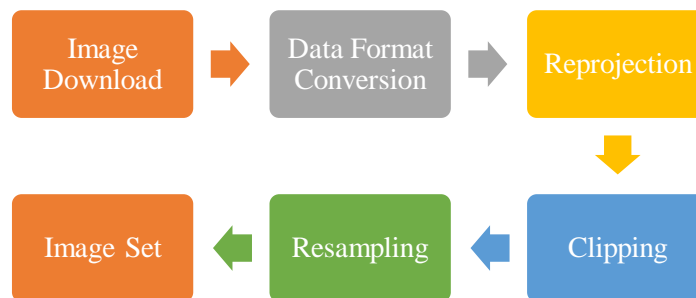


Figure 4.1: Raster Image Preparation Flowchart.

#### 4.1.1 Image Download:

Data products for MODIS Terra and MODIS Aqua are freely available on United States Geological Survey (USGS) Earth Explorer (<https://earthexplorer.usgs.gov>) registered users can select the place of interest by uploading shapefile (Figure 4.2) and further filter the dataset based on date and other required parameters, the data downloaded is in HDF format.

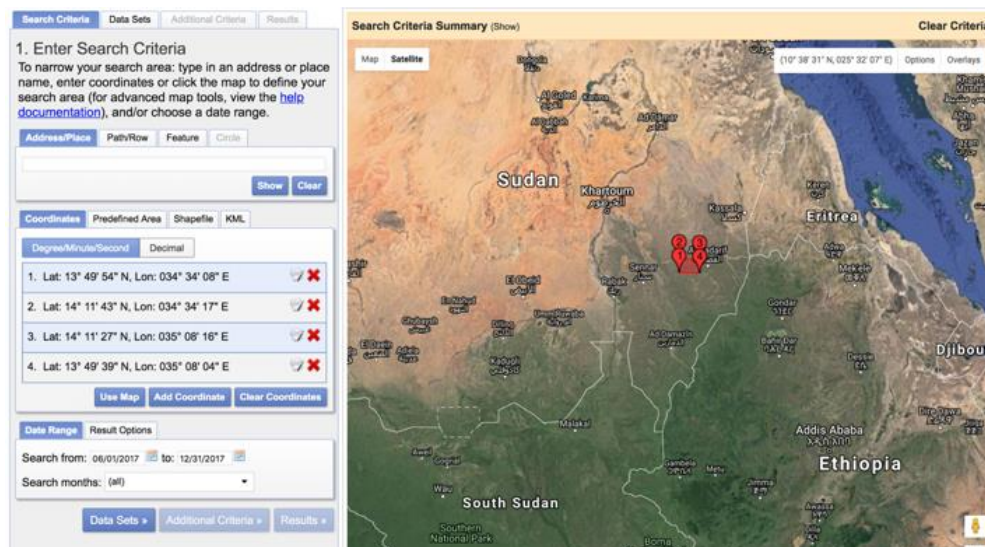


Figure 4.2: Upload shapefile to select tile for download.

PROBA-V data are freely available on Vision on Technology (VITO) product distribution portal (<http://www.vito-eodata.be/PDF/portal>) in collaboration with European

Space Agency (ESA) registered users can select the area of interest by creating an online ROI and further filter the data based on acquisition date. It allows the user to place a customized order by selecting the required projection system, file format (GeoTIFF, ASCII, HDF) and bounding box for the area.

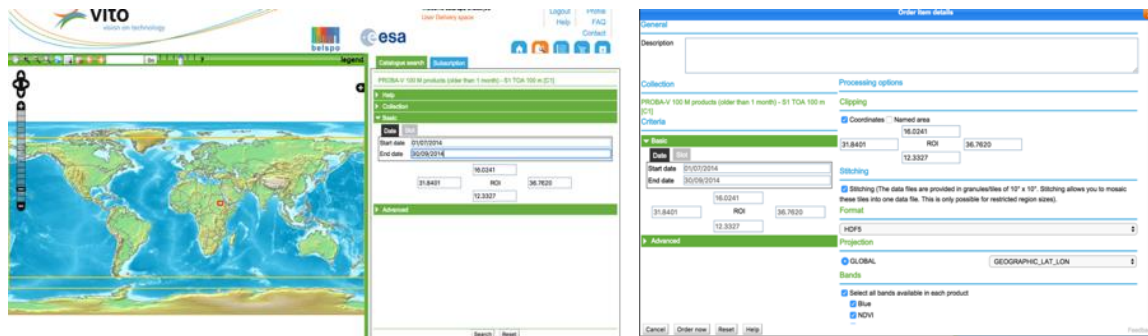


Figure 4.3: Selecting and customising PROBA-V data..

Products for SMAP (Soil Moisture Active Passive) are available on National Aeronautics and Space Application (NASA) US free access hub (<https://search.earthdata.nasa.gov>) it allows the user to select data similar to that of Vito hub above, additionally allows the user to select or deselect bands or data within the dataset which are not required by the user, thus reducing download size and increasing the speed.

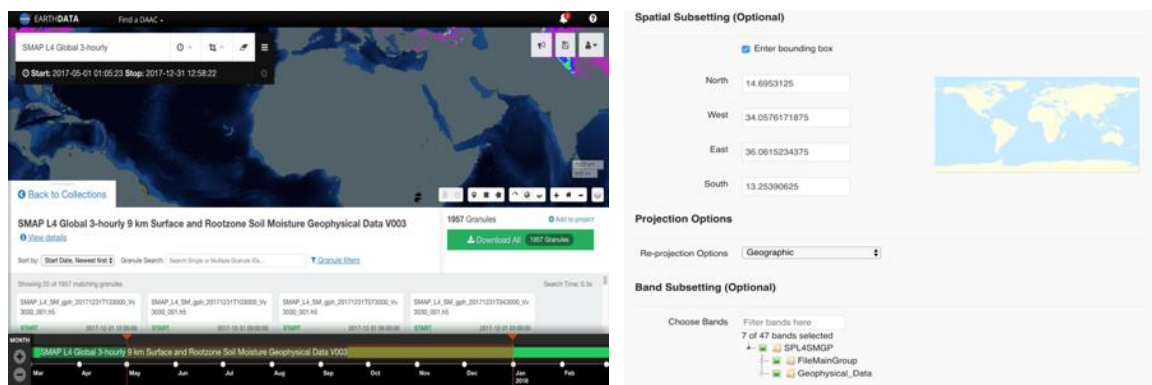


Figure 4.4: Selecting and customising SMAP data.

#### 4.1.2 Data Format Conversion:

For the MODIS data downloaded a python based batch processing conversion tool was developed to convert the data from HDF format to Tiff Format. For PROBA-V and SMAP customised orders in GeoTIFF format were placed for downloading.

### 4.1.3 Reprojection:

The MODIS image downloaded was in sinusoidal projection, PROBA-V in Geographic Lat/Long and SMAP was in EASE-GRID projection. To have all the images in the same projection, they were all re-projected to WGS-84/UTM Zone 36N using batch processing in QGIS.

### 4.1.4 Clipping:

The images were clipped down to the required chosen area, as discussed in the previous chapter. The images have a larger footprint like MODIS is created at 5.5km grid area for each tile generated over a region, thus the MODIS data had to be reduced to the study area. For PROBA-V it was customised to a smaller area instead of the entire world tile by assigning a bounding box before downloading, but it had to be further clipped down to the study area. SMAP tiles are similar to PROBA-V, thus they were customised with a bounding box and further clipped after download. The cropping was executed by batch processing extraction tool in QGIS.

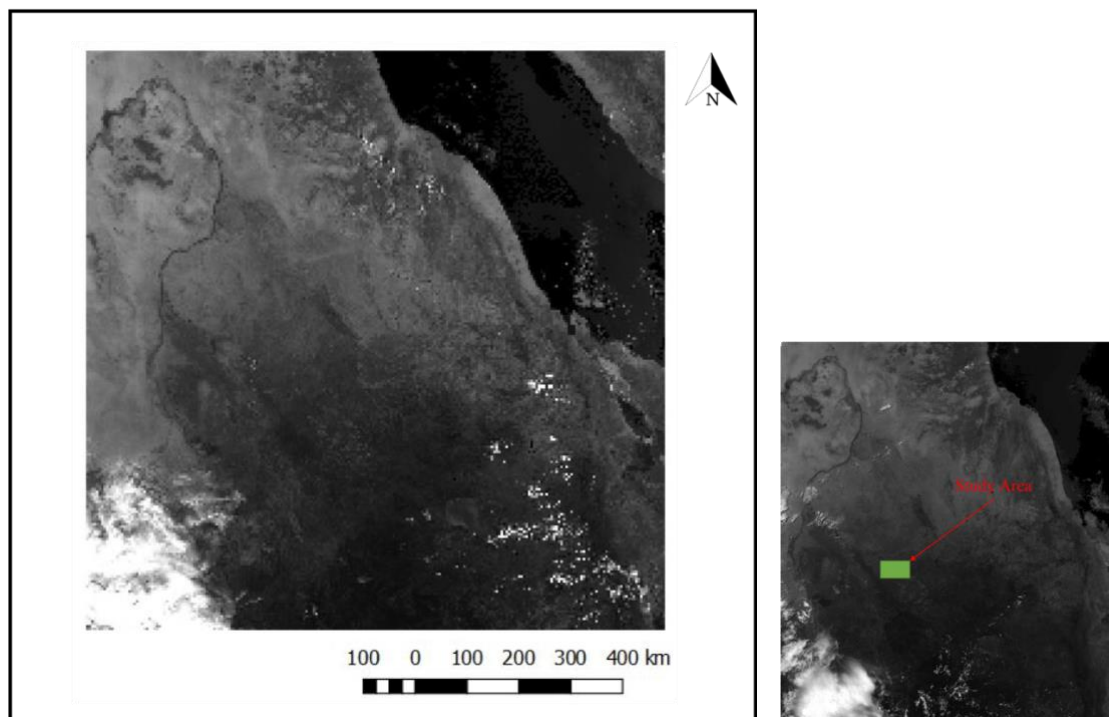


Figure 4.5: MODIS actual tile at 5.5km (MODIS Terra- 03/06/2014)

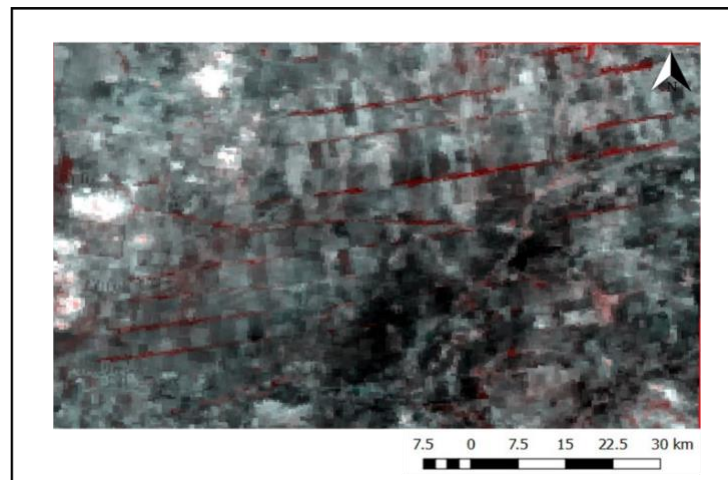


Figure 4.6: Cropped down to study area.

#### 4.1.5 Resampling:

The actual recorded resolution of MODIS is 231m and for PROBA-V at 350 m but disseminated at 250m and 300m respectively, thus to homogenise the spatial resolution of images it has been uniformly resampled to 250 m.

#### 4.1.6 Vector Data Preparation:



Figure 4.7: Vector data creation flowchart.

##### 4.1.6.1 Image Download:

High resolution images of Sentinel and Landsat were downloaded for 2014-2017 and for each month from June - December.

##### 4.1.6.2 Digitisation:

Sub-plots were identified from the high resolution image (sentinel) and were manually digitised for all the four years. This was done to identify the changes in cropping pattern within the main plots. Also the permanent vegetation cover changed over the years, which were identified and digitised.



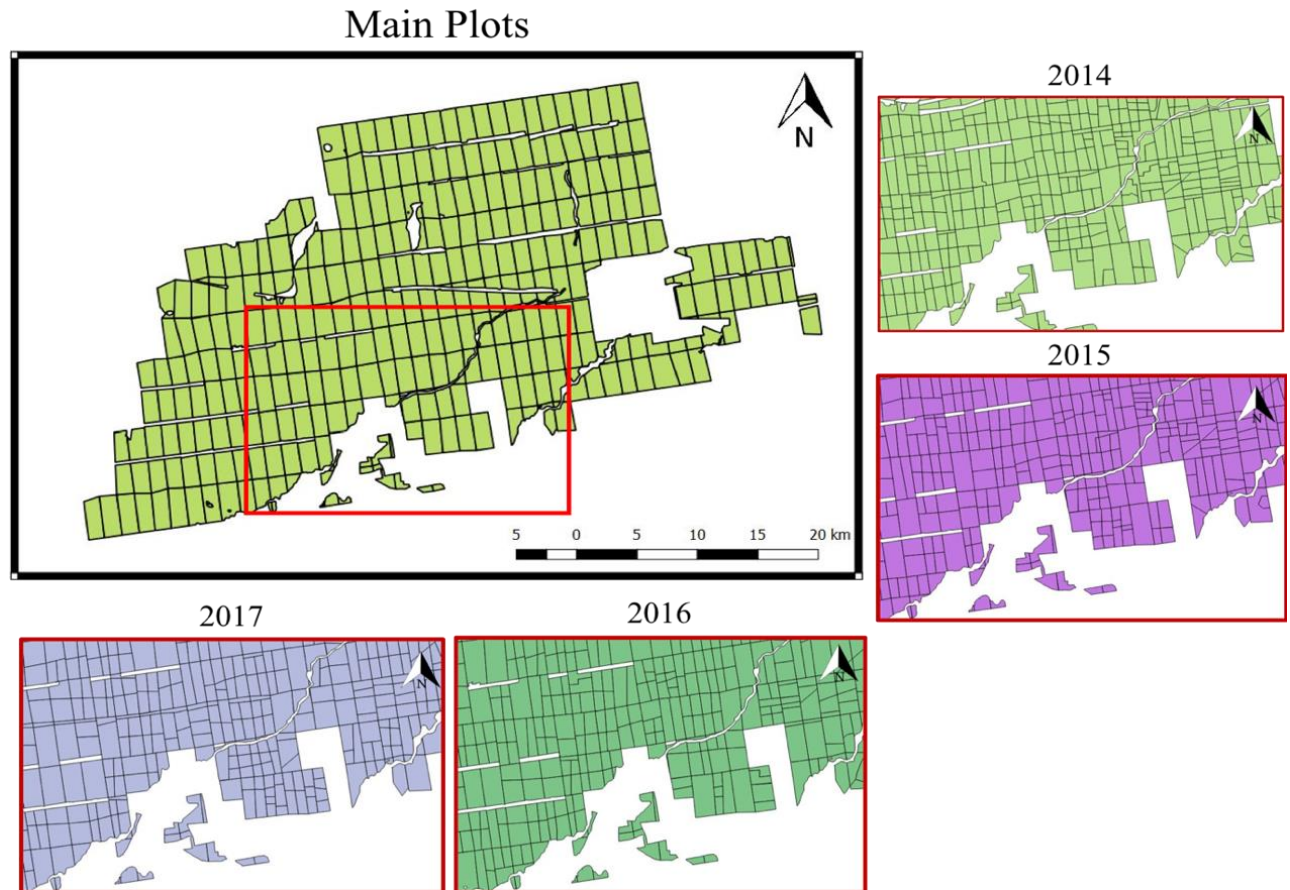


Figure 4.8 : Main plot with parts of sub-plots digitised from yearly images.

## 4.2 Data Post-Processing:

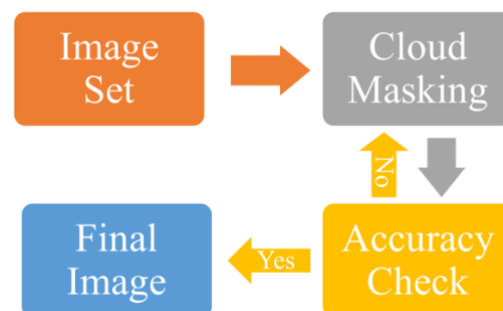


Figure 4.9: Cloud Masking Flowchart.

### 4.2.1 Cloud Masking:

#### 4.2.1.1 MODIS

Masking tool was created for batch processing based on Python, to read the layer bit by bit. Depending on the requirement and the quality of the image being masked the bits where

chosen to be used. For instance some images had Bit-Number 13 off as there were no effect on neighbouring pixel, while the rest had it on.

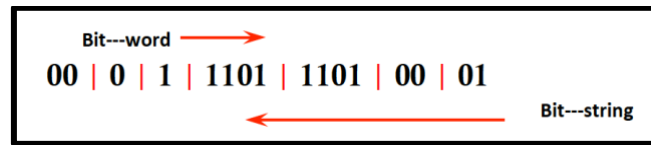


Figure 4.10: Method to read Bit-Code.

Decoding from binary example:

Bit value: 2197 16 bit binary value: 0|0|0|0|1|0|00|10|010|1|01

Interpretation based on Table 8:

Bit No.	Parameter	Used as filter in code	Bit word to be allowed (re table)	Bit word observed	Observed word interpretation
0-1	Cloud State	Yes	00/11	01	Cloudy
2	Cloud Shadow	Yes	0	1	Yes
3-5	Land/Water Flag	Yes	001	010	Land
6-7	Aerosol Quantity	Yes	01/00	10	Average
8-9	Cirrus Detected	Yes	00	00	None
10	Internal Cloud Algorithm Flag	Yes	0	0	No cloud
11	Internal Fire Algorithm Flag	Not applicable	Not used	0	Not used
12	MOD35 Snow/Ice Flag	Not applicable	Not used	0	Not used
13	Pixel is adjacent to cloud	Yes	0	0	No
14	Salt Pan	Not Applicable	-	0	No
15	Internal Snow Mask	Not Applicable	-	0	No



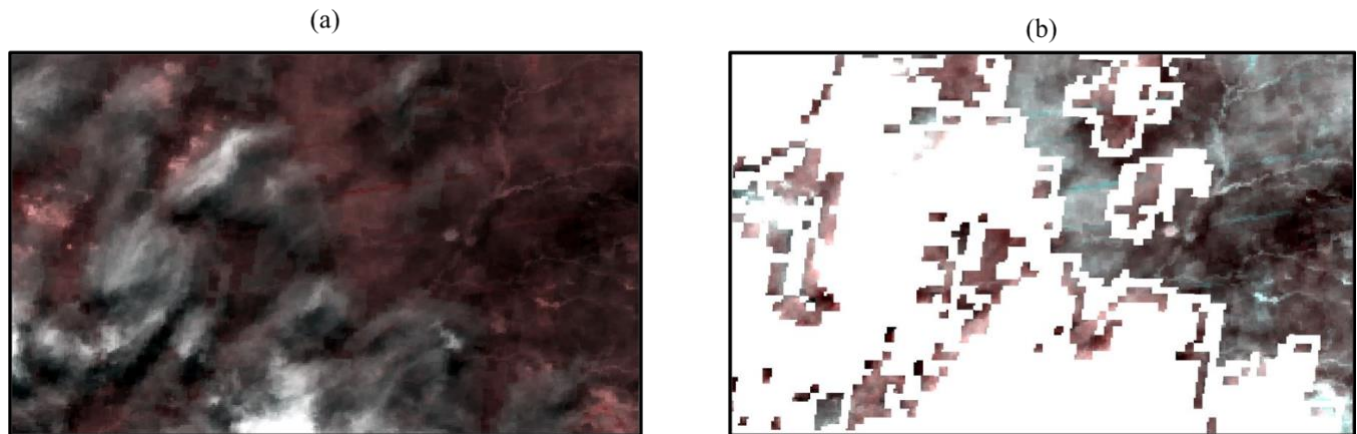


Figure 4.11: MODIS Terra image taken on 01/07/2014. (a) Actual image with cloud (b) Cloud masked image

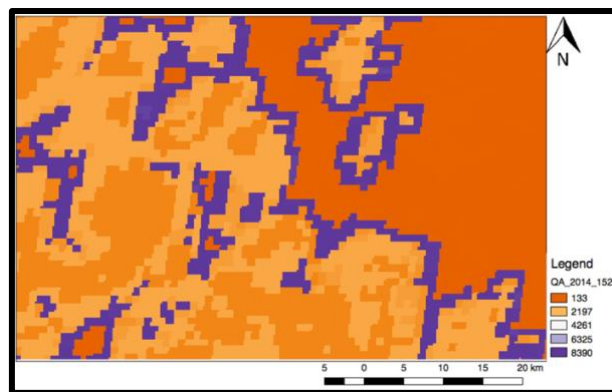


Figure 4.12: QA SDS which is used as Cloud mask for above image taken on 01/06/2014

In the Figure 4.11(a), the actual MODIS Terra image is shown after downloading with cloud covers on it, it is a RED and NIR band image with band combination of (2,1,1). Figure 4.11(b) shows the masked MODIS Terra image after it was processed following the bit parameters.

The values in the cloud mask represent the bit combinations for the pixel state in the image as seen in Figure 4.12, the legend shows the different bit combinations. It can be interpreted as:

133-	0	0	0	0	0	0	0	0	1	1	000	1	01
2917-	0	0	0	0	1	0	0	0	1	0	010	1	01
4261-	0	0	0	1	0	0	0	0	1	0	100	1	01
6325-	0	0	0	1	1	0	0	0	1	0	110	1	01
8390-	0	0	1	0	0	0	0	0	1	1	000	1	10

The masked image were evaluated based on correctly classified cloud pixel and shadow pixel with that of the actual image and have a kappa coefficient of 0.928, the agreement between the masked image and the quality image are strong with an overall accuracy of 93% .

#### 4.2.1.2 PROBA-V:(Wolters 2018)

Same tool mentioned above was used for masking PROBA-V with combination as mentioned in (Table 6), but an additional parameter of dilation (Figure 4.13) had to be added. As the PROBA-V mask had some indigenous issues discussed above (section 3.1.4) because of the band spacing, the layer was dilated up to 2 pixels in certain cases.

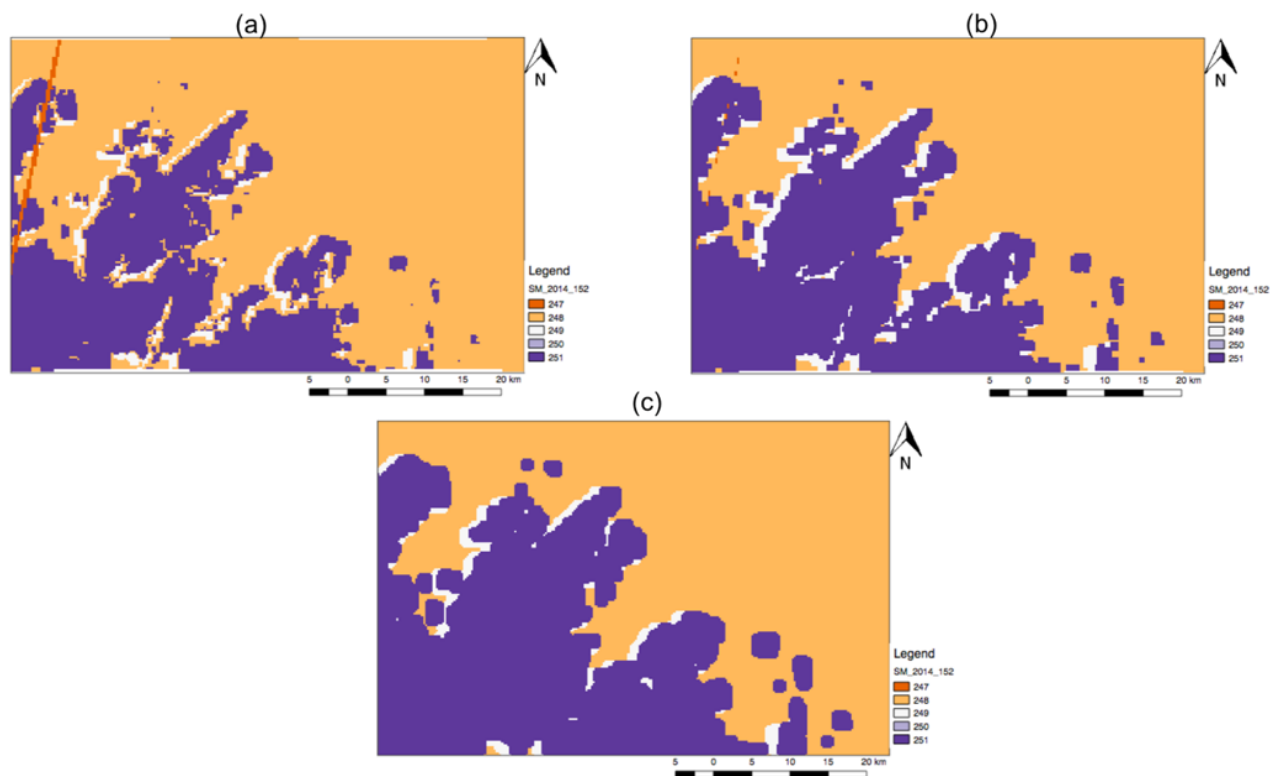


Figure 4.13: PROBA-V Cloud Mask for 01/07/2014. (a) The actual Cloud Mask provided in the data. (b) Cloud Mask dilated by 1 (c) Cloud mask dilated by 2.

In some images the thin layers of clouds are not removed by the mask, those images had to be manually masked by thresholding radiance values in the Blue band as cloud reflectance is the highest in Blue band (Figure 4.14).

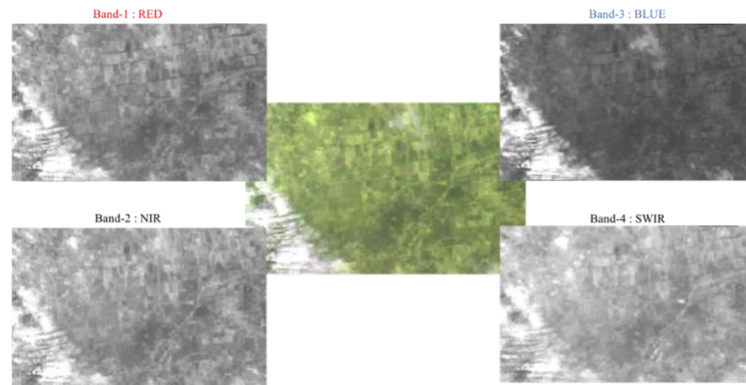


Figure 4.14: PROBA-V Multispectral bands splitted into 4 panchromatic band.

The manual cloud masking workflow is shown in Figure 4.15, the actual mask is used to remove the clouds from the image using the above mentioned code, but few clouds are not removed by mask as seen in (c), and thus based on the radiometric thresholding as discussed in (section 2.5.2) new mask (d) is created from SWIR and BLUE band and used as a mask on the first masked output image (c) resulting in the final output masked image (e).

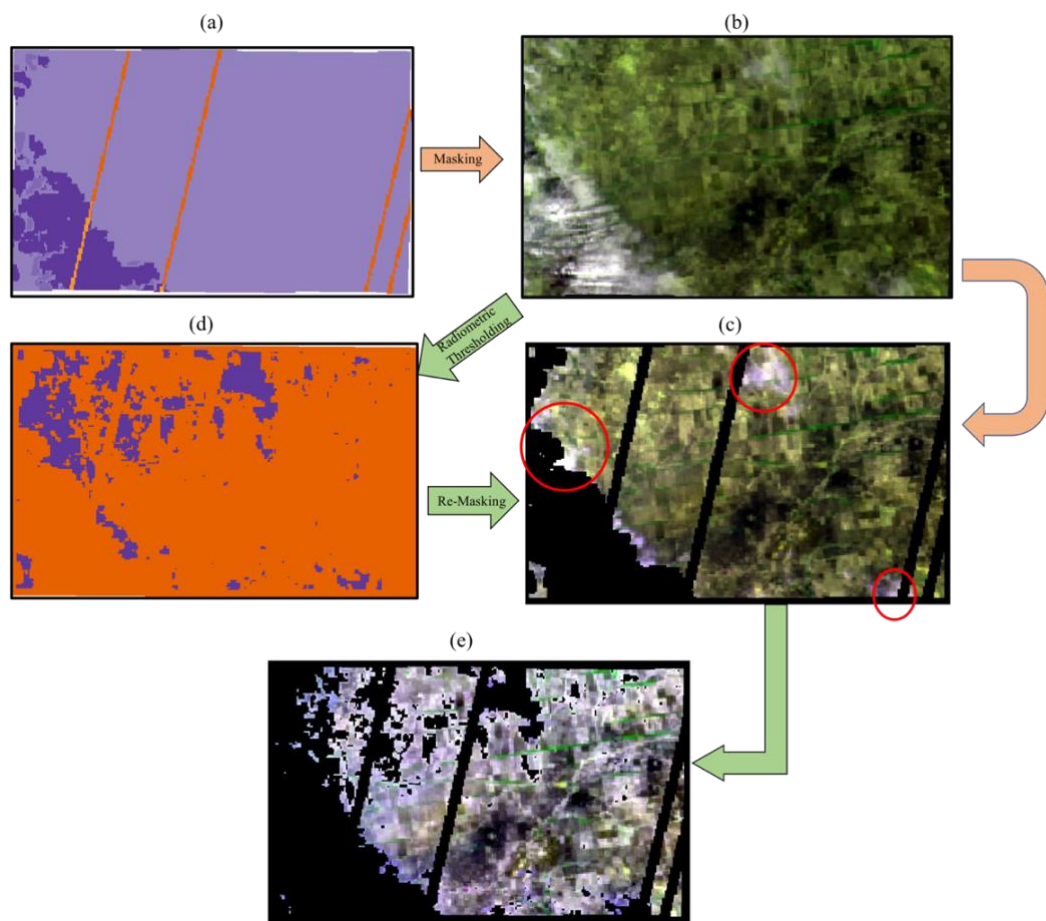


Figure 4.15: Step-by- Step PROBA-V Cloud Masking image is from 10/06/2017. (a)Actual Cloud Mask (b)Actual image (c) First masked image from using actual mask (d)New mask created by radiometric thresholding. (e)Final masked output.

The kappa coefficient of the final output image from PROBA-V was 0.789 which was calculated based on the classified cloud pixel, shadow pixel, and thin cloud pixel. It can be concluded the new pixels has an accuracy of 78% after classifying and removing cloud pixels.

### 4.3 Final Workflow:



Figure 4.16: Final data fusion Flowchart.

After post processing all the images, the next step was to fuse the MODIS and PROBA-V but before the fusion few consistency checks had to be done based on certain parameters of the dataset to ascertain their eligibility to be fused.

The data checks were conducted based on three parameters. Firstly the BRDF (Bidirectional Reflectance Errors), as they are already corrected in the MODIS level-3 and PROBA-V Level 3 data, BRDF will not cause considerable error. Secondly, cloud masking is already done in post-processing.

And lastly, the Viewing Zenith Angle (VZA) it is important as both the sensors have different scanning techniques and there is a difference in image registration. For example an object recorded at a certain angle by MODIS can be different from that recorded by PROBA-V, and realignment of either image to the angle of the other image is required but this realignment is required if the position of the object being studied is out of the required thresholds. The image set of PROBA-V used is scanned at the nadir and distortion increases with distance from the center and for MODIS the distortion increases when the angle is over  $30^\circ$  from center. The distortion for MODIS increases when  $VZA > 30$  (Duveiller, Lopez-Lozano et al. 2015) and for PROBA-V  $VZA > 36$ , and also the VZA has low impact on smaller study areas.

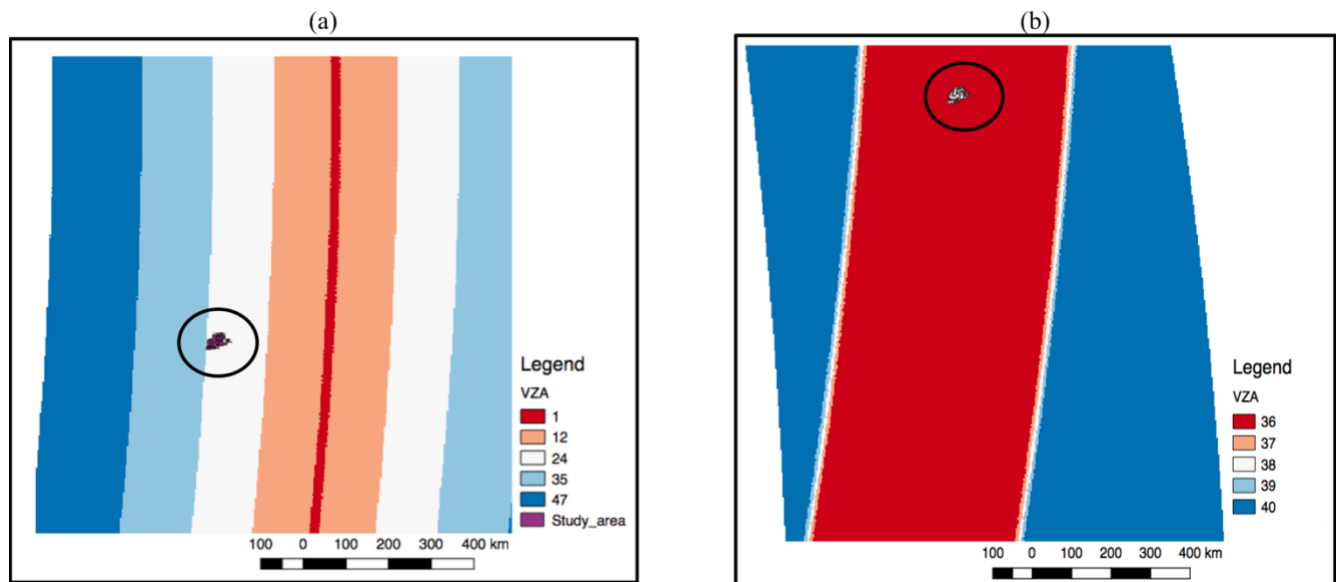


Figure 4.17: (a) Viewing Zenith Angle (VZA) of MODIS, (b) Viewing Zenith Angle (VZA) of PROBA-V.

Figure 4.17(a) shows the scanning swath of MODIS (whisk-broom scanner), with the center marked at  $1^\circ$  and angle increases from the center towards the sides. Figure 4.17(b) shows the scanning swath of PROBA-V (push-broom scanner) with the center up to  $30^\circ$ , and the distortion increases after that. Both the scan line swaths are obtained from the main tile time SDS.

The Viewing Zenith Angle was checked and as the study area falls within the acceptable range of  $VZA \leq 30$  for MODIS (see Figure 4.17(a)) and  $VZA \leq 36$  PROBA-V (see Figure 4.17(b)) realignment was not required for the images.

After consistency check of the image the next step was to calculate the NDVI values for each image, a tool was developed based on python to batch process the NDVI calculation for MODIS and PROBA-V respectively.

The NDVI images were pan sharpened based on Bayesian Fusion Method (section 2.3.1) with Weight-1 and S coefficient-2 in QGIS batch processing. A matching of the images were done between the two set of images so that the days missing doesn't cause an error while fusing. Finally the two images were merged based on the condition:

$$(MOD \geq 327676) * ((PV - 0.03) * (PV \geq 0 \ \&\& \ PV \leq 1)) + (MOD * (MOD \neq 327676) + (MOD > 0 \ \&\& \ MOD < 1))$$

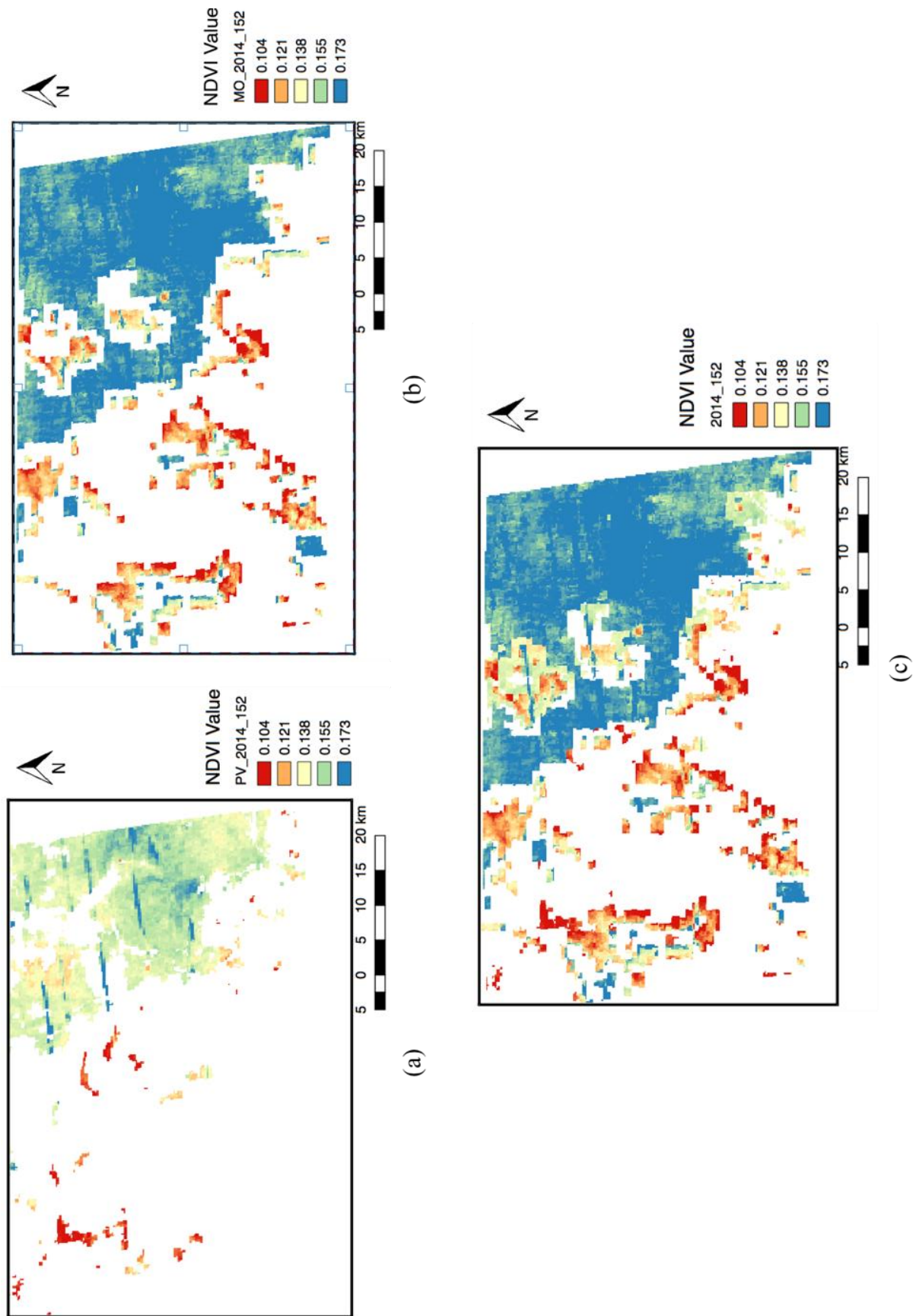
Where; MOD - Modis image, PV - PROBA-V image, 32767- Missing Data value. The condition specifies to replace MODIS null values with PROBA-V values ranging between 0 and 1 and only to keep MODIS values which falls within the NDVI range of 0-1, as to remove extreme or error values from the final product. Before substituting the value, the PV value is subtracted by 0.03 to reduce the gain in PV as suggested by (Commission) 2015) as they observed that PROBA-V has a positive mean-bias of 0.03 to MODIS. Thus to adjust the gain in PV values it is reduced by 0.03. The same was cross validated between sample images through SNR, and difference proved to be by 0.03.

The overall accuracy check on the final fused image had a R Square Error<sup>20</sup> of 0.3. Also an anomaly check was done on randomly chosen images to compare if there is any existing anomaly between the final fused image and the actual data set of PROBA-V and MODIS, the results are attached in the Annexure 1 and Annexure 2 it shows there is no anomaly between datasets and have a uniform relationship with anomaly 0 in all the images.

---

<sup>20</sup> R square or coefficient of determination shows percentage variation in y which is explained by all the x variables together. So, if R-square is 0.8, it means 80% of the variation in the output variable is explained by the input variables. So, lower the R squared, the less variation is explained by input variables.





1 Figure 4.18: Final Fused image creation. (a)PROBA-V image for 01/06/2014 (b) MODIS image for 01/06/2014 (c)Final fused image.



## 5. Results And Observations

### 5.1 Mean Monthly NDVI:

The following images shows the mean monthly NDVI for the study area over the period 2014-2017. According to FAO reports, there was less harvest in the year 2015 and 2017, as the country faced frequent droughts for these two years, and it can be observed from the maps as well that the NDVI values falls during the peak season of Aug-Oct in the year 2015 and 2017.

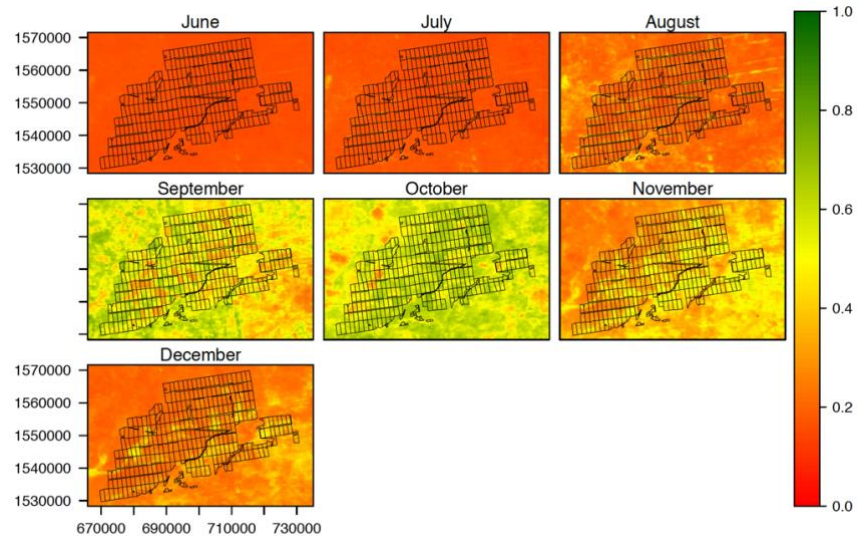


Figure 5.1: Mean Monthly NDVI in 2014

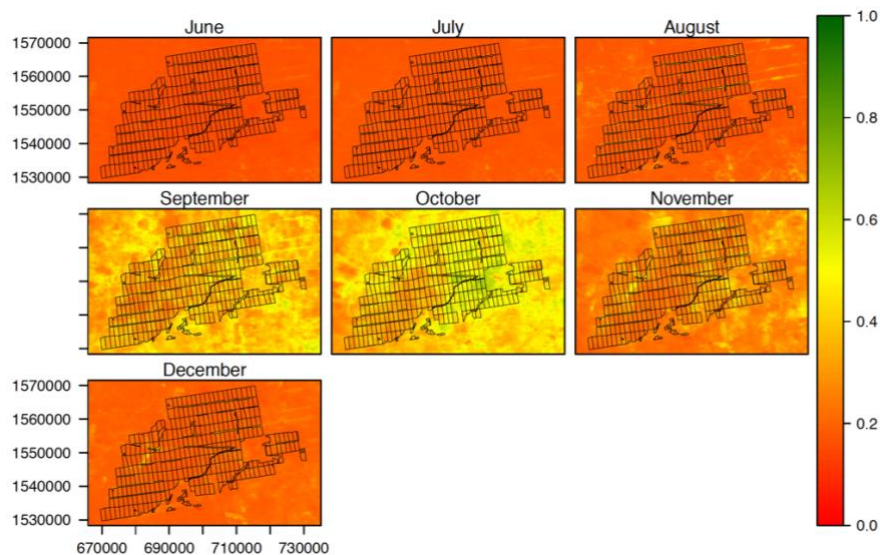


Figure 5.2: Mean Monthly NDVI in 2015.

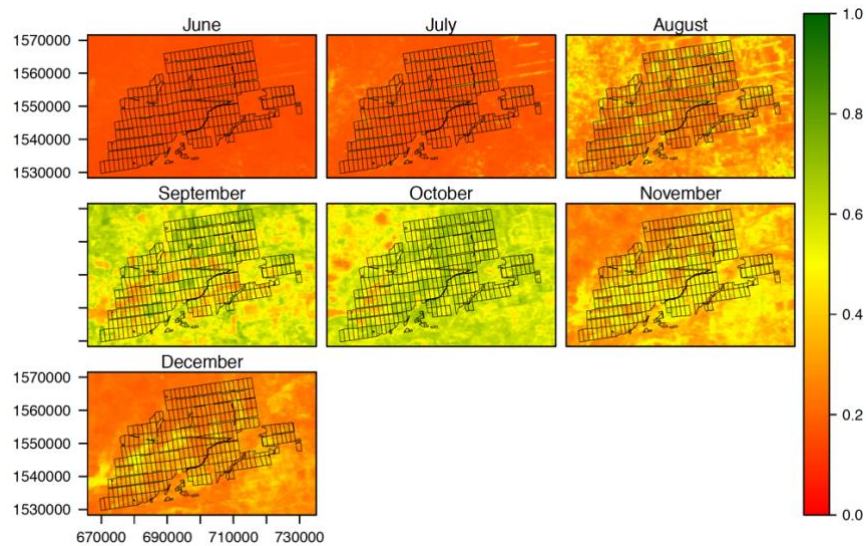


Figure 5.3: Mean Monthly NDVI in 2016.

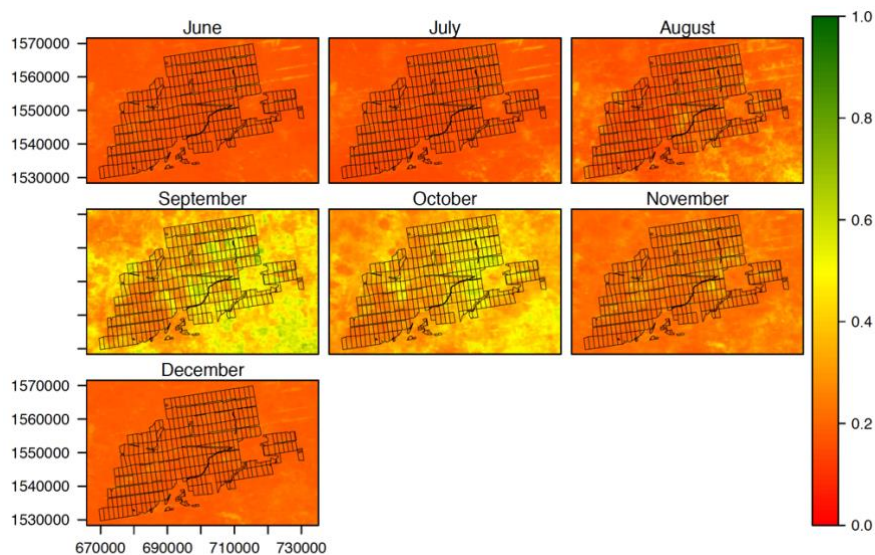


Figure 5.4: Mean Monthly NDVI in 2017

In Figure 5.1, the NDVI value for 2014 ranges from 0.4-0.6 in September and October which is the peak growth season for the crops. Whereas in Figure 5.2 (2015) the NDVI value barely crosses 0.4 except for few plots where the vegetation has a higher growth these sites which has a slightly higher green-up are the sites which has small moats or ponds in the field for harvesting rain water. Similarly in Figure 5.3, the NDVI values for 2016 are mostly around 0.6 with few plots ranging up to 0.7, but in 2017 as seen in Figure 5.4, the crops barely reaches the green-up, mostly throughout their growing season the NDVI ranges between 0.2-0.4, with barely any plots reaching a higher NDVI, the country as a whole faced severe drought this year.



## 5.2 Daily NDVI:

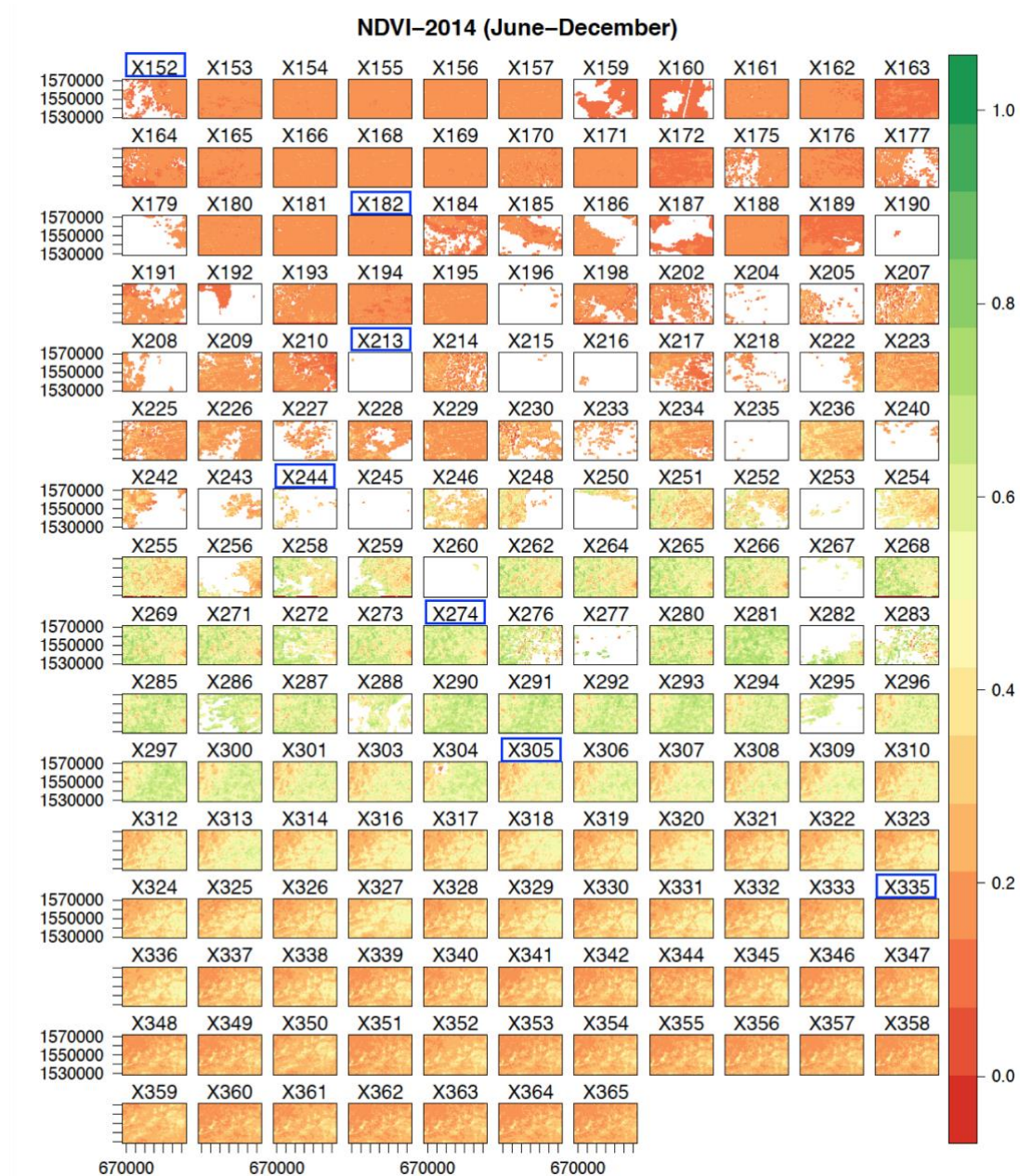


Figure 5.5: Daily NDVI value for the year 2014 from June - December.

In Figure 5.5, shows the daily NDVI value for the study area with first day of every month marked with blue box. The image name are based on their Julian day, the white space in the image have no data available after post-processing, and also the missing days can easily be identified from the image. The daily based observation gives a distinctive visual understanding of the onset of green-up from day 246 (03/10/2014) onwards, and the harvest starts from day 301 (28/11/2014). As there are two type of plantations of Sesame and Sorghum in the area the harvest and green-up is not same throughout the images, and thus a plot based

analysis is discussed later. In daily NDVI values for the year 2015 shown in Annexure 3 shows short season of growth ranging over few days, because of the drought this year the growth and productivity of the crops. For the year 2016 as observed in Annexure 4 the green-up started around Day 235 and the harvest started around Day 316 and the growth of crops were uniform over the area.

2017 was also declared as a drought year and the NDVI values in Annexure 8 justifies the same, the growth of the crops are hardly visible and most of the year is marked with very low NDVI values, this was mainly because of a prolonged El-Nino effect that had a cumulative effect on the overall crop production in the country.<sup>21</sup>

The productivity of the crop got largely affected by the rainfall and drought conditions, and thus soil moisture is considered as one of the parameters to study the plant growth later on this chapter. The El-Nino had a prolonged and large scale effect over 18 months throughout many countries in east Africa one of which was Sudan, the El-Nino effects started building-up on 2015 and the final effect was on 2017, at the same time 2016 showed NDVI values slightly improved indicating some crop growth which is presumably because of region specific rainfall.

### **5.3 NDVI Anomaly:**

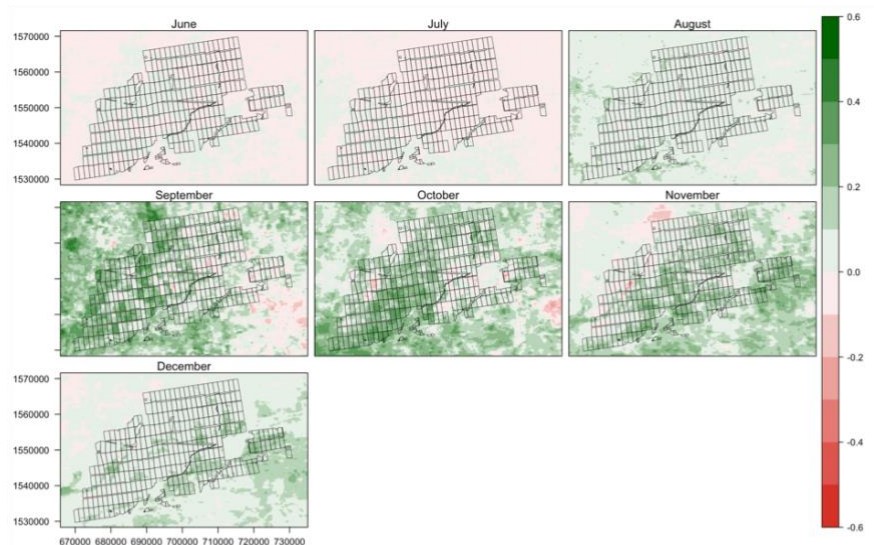
An NDVI anomaly is the difference between the average NDVI for a particular month of a given year and the average NDVI for the same month over a specified number of years. This approach can be used to characterize the health of vegetation for a particular month and year relative to another year, which is a good indicator of drought or declining vegetation health

To understand the difference in NDVI between the years anomaly maps are made which will represent the negative or positive relationship for NDVI values between different time series. The positive values represents positive anomaly i.e., the NDVI value was higher for that year compared to others, no change in values are 0 anomaly, and the negative values represents negative anomaly i.e., the NDVI for that year is less than the other years. From the anomaly map the crop growth behavior can be easily deduced. It is also important to maintain the reference to the types of crop grown in this region as discussed in (section 3.3.3) because the anomaly is different based on plots within the same time period due to varied crops. The anomaly can be classified as  $<-0.15$  = Large Decrease,  $-0.15:0$  = Small Decrease,  $0:0.15$  = No Change,  $0.15:0.25$  = Small Increase,  $>0.25$  = Large Increase.

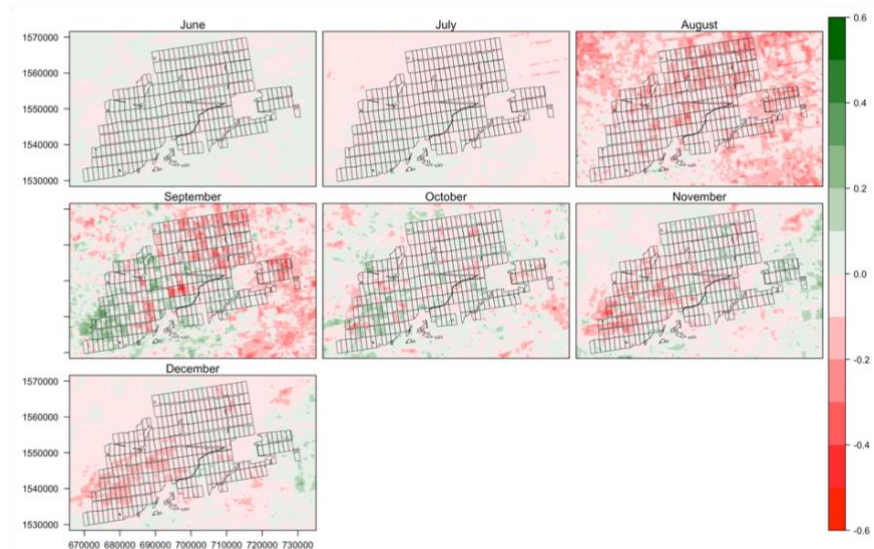
---

<sup>21</sup> <http://www.irinnews.org/feature/2017/03/17/drought-africa-2017>

(a) 2015



(b) 2016



(c) 2017

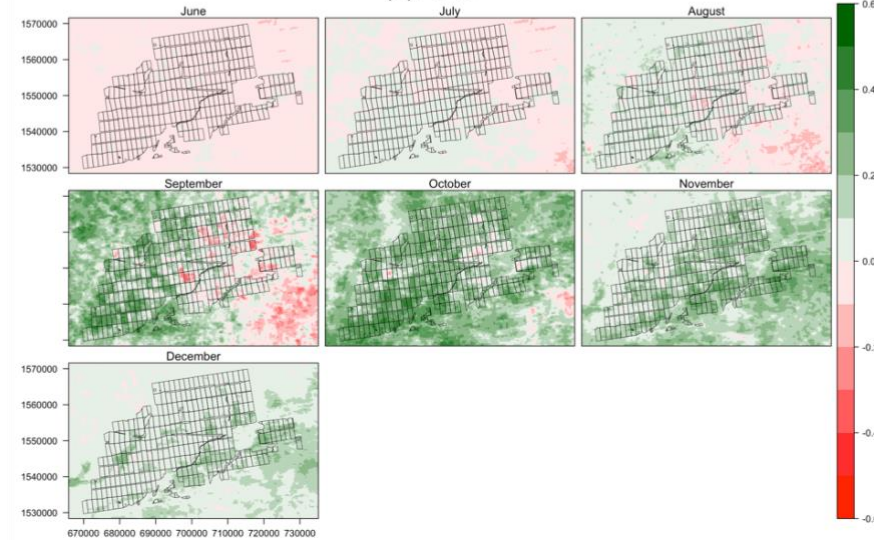
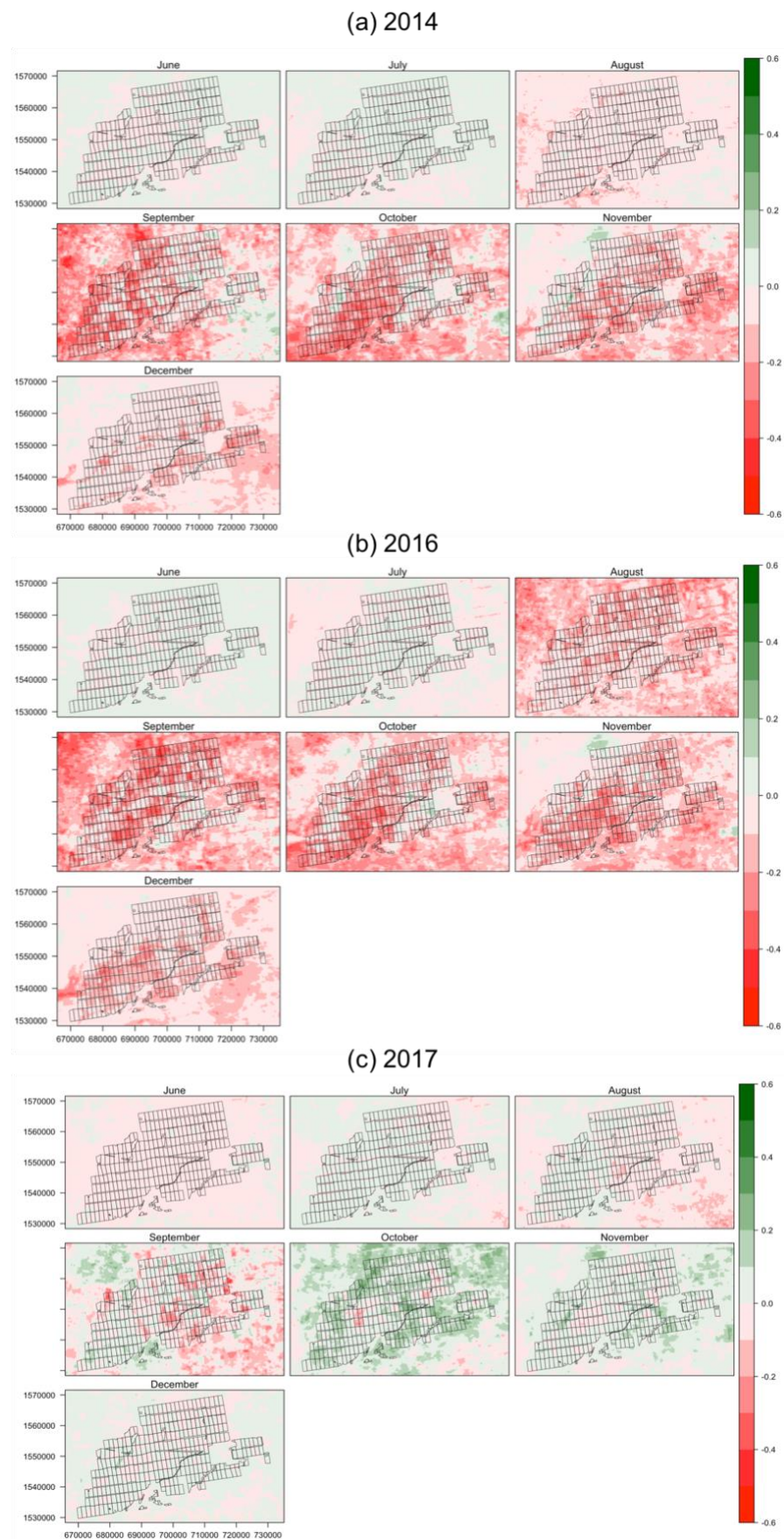
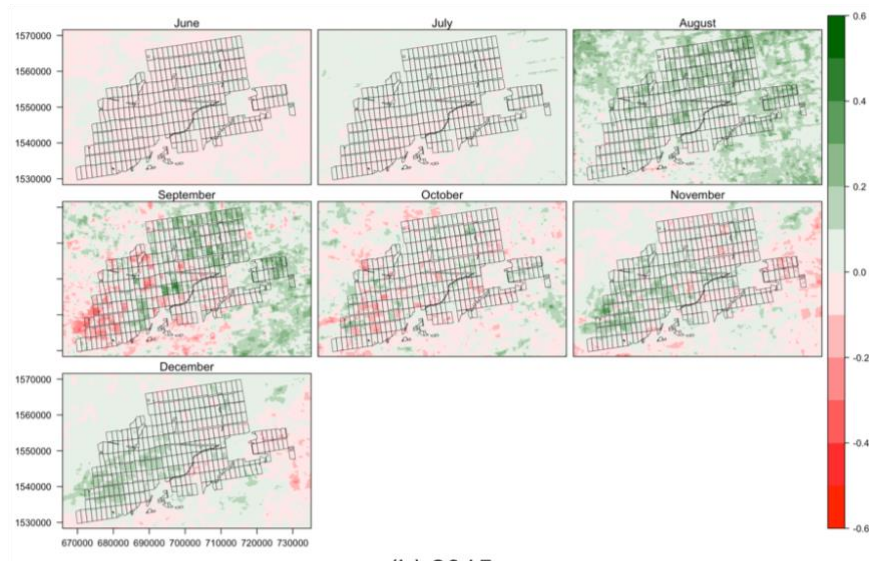


Figure 5.6: NDVI Anomaly in 2014 referred to (a) 2015, (b) 2016, (c) 2017.





(a) 2014



(b) 2015



(c) 2017



Figure 5.8: NDVI Anomaly in 2016 referred to (a) 2014, (b) 2015, (c) 2017





Figure 5.9: NDVI Anomaly in 2017 referred to (a) 2014, (b) 2015, (c) 2016

Figure 5.6 shows the anomaly of 2014 to the other 3 years of 2015, 2016, 2017 respectively. As 2014 was a normal growth year with steady crop growth, it has a large increase positive anomaly compared to the drought years in 2015 (Figure 5.6 (a)) and 2017 (Figure 5.6 (c)). With a stronger large increase and wide spread anomaly with 2017 as this was the worst drought affected year.

Also it is observed that the anomaly is higher specifically in the green-up seasons in September to October, since the NDVI reaches it's optimum peak during these seasons detection of the anomaly is easier. The observation between the two normal growth years of 2014 and 2016 (Figure 5.6 (b)), large decrease anomaly is observed mainly in the month of August and September and specific to certain plots, this can be attributed to the fact that the area observes crop rotation thus the crop in 2014 is different from the crop in 2016 which had different growing onset and so the anomaly.

Figure 5.7 shows the anomaly of 2015 to the years 2014, 2016 and 2017 respectively. As observed 2015 had a negative anomaly with the year 2014 (Figure 5.7 (a)) and 2016 (Figure 5.7 (b)), with large decrease anomaly in 2016 ranging from August to December covering more plots as compared to that of 2014 where it's observed predominantly between September to November. The anomaly between the two drought years of 2015 and 2017 shown in (Figure 5.7 (c)) helps to understand the intensity of the drought, as this shows a small increase positive NDVI for 2015 specifically in the peak growth month of October, concluding that the impact of drought was higher in 2017 than in 2015.

Figure 5.9 shows the anomaly of 2017 to the years 2014, 2015 and 2016 respectively. As 2017 was worst affected by the El-Nino effect and had a severe drought, it has a large decrease negative anomaly to all the other years. But it has the maximum anomaly to 2014 (Figure 5.9 (a)) whereas with 2016 (Figure 5.9 (c)) the effect is limited to the peak month of October and also the intensity seems to be less.

From the anomaly study 2014 can be considered as a reference year for normal growth of crops and all the other years has a dominant negative anomaly trend with 2014. A reference year for anomaly detection is usually the year with most consistent crop growth with an optimum NDVI and without any major climatic effect or changes it can change depending on the years considered for monitoring.

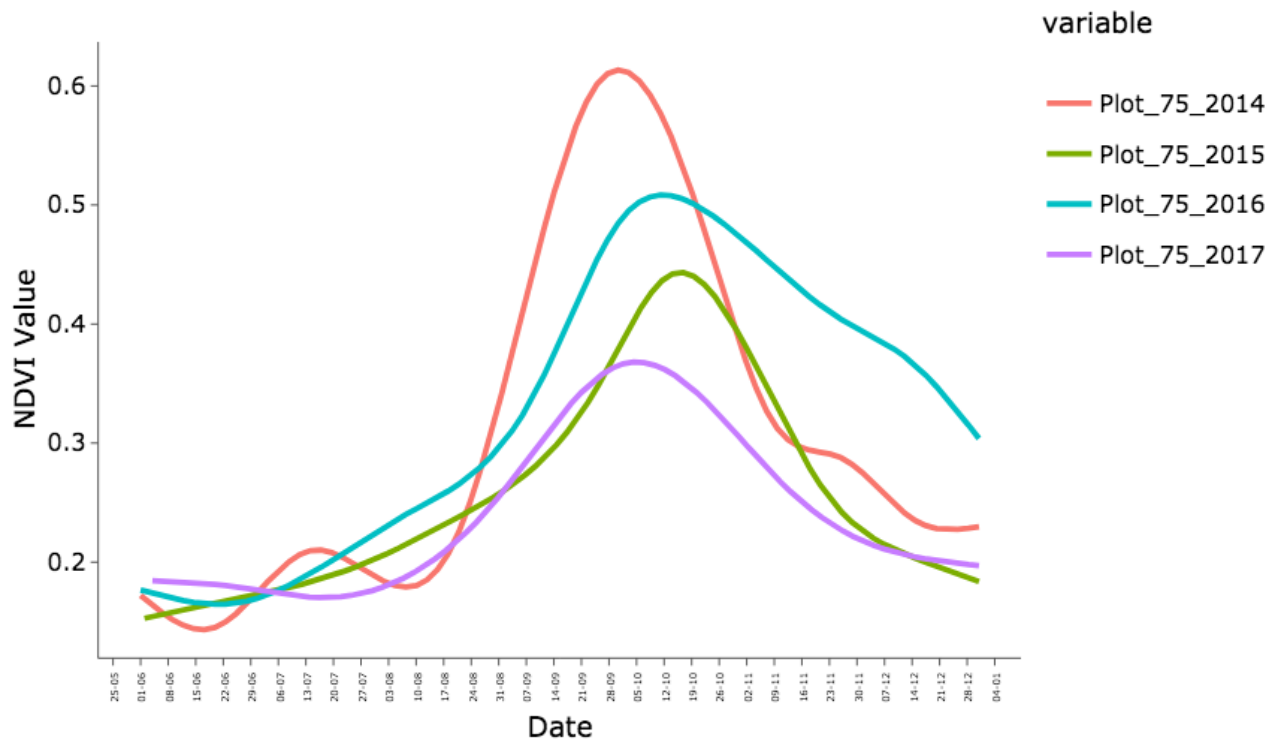


Figure 5.10: Plot-75 yearly change comparison.

Figure 5.10, shows the comparison of Plot-75 from 2014 – 2017, it can be seen from the NDVI curves that for the year 2015 and 2017 the values were lower compared to 2014 and 2016. Also the maximum NDVI was in 2014 having an uniform and steady growth along with timely harvest.

## 5.4 Interactive Visualisation:

The following results gives an interactive analysis of the plot based NDVI. For the interaction either a single plot can be selected or a region can be selected, and the output graph can be interacted with and is saved as a webpage.

The curve fitting technique used above is LOESS (LOcal polynomial regrESSion), which originated as generalisation of LOWESS (LOcally WEighted Scatter-plot Smoother). LOESS is a non-parametric method because the linearity assumptions of conventional regression methods have been relaxed, instead of estimating parameters like  $m$  and  $c$  in  $y = mx + c$ , a non-parametric regression focuses on the fitted curve. The fitted points are estimated with respect to the whole curve rather than a particular estimate<sup>22</sup>, so the overall uncertainty is measured as how well the estimated curve fits the population curve. It is called local regression because the

<sup>22</sup> [https://www.statsdirect.com/help/nonparametric\\_methods/loess.htm](https://www.statsdirect.com/help/nonparametric_methods/loess.htm)

fitting at say point  $x$  is weighted toward the data nearest to  $x$ , and since it is a local regression it can handle missing values more confidently without running into errors unlike piecewise fit or polynomial fit where the trade-off between bias and variance gets affected by missing values.

Also at USGS/EROS, temporal satellite data is smoothed using a weighted, least-squares linear regression approach that involves a moving temporal window to calculate a regression line. This family of lines is then averaged at each point, and interpolated between points, to provide a continuous, relatively smooth NDVI signal over time and the resulting relationship between raw and smoothed data is statistically based. A comparison of other curve fitting are added in the Annexure 9.

The AUC (Area Under Curve) represents the area between the graph of  $y = f(x)$  and the  $x$ -axis is given by the definite integral below. This formula gives a positive result for a graph above the  $x$ -axis, and a negative result for a graph below the  $x$ -axis. Thus the median is calculated from the plot summary which is assigned as the baseline for calculating AUC for the curve giving the NDVI integral of the plot.

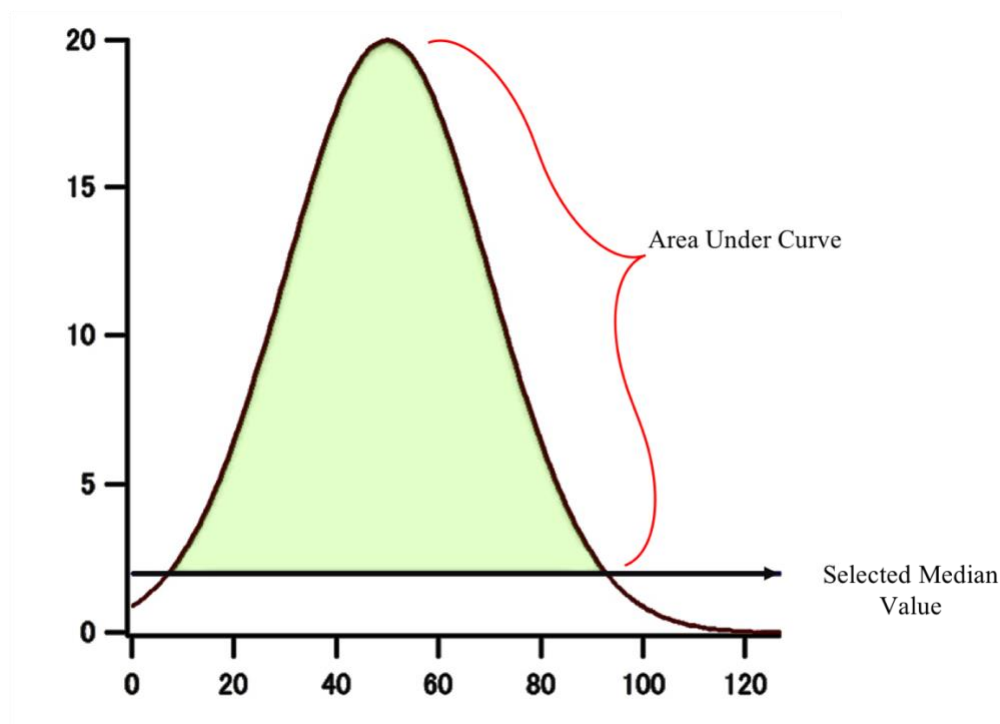


Figure 5.11: Area Under Curve.

### 5.4.1 Plot Based NDVI:

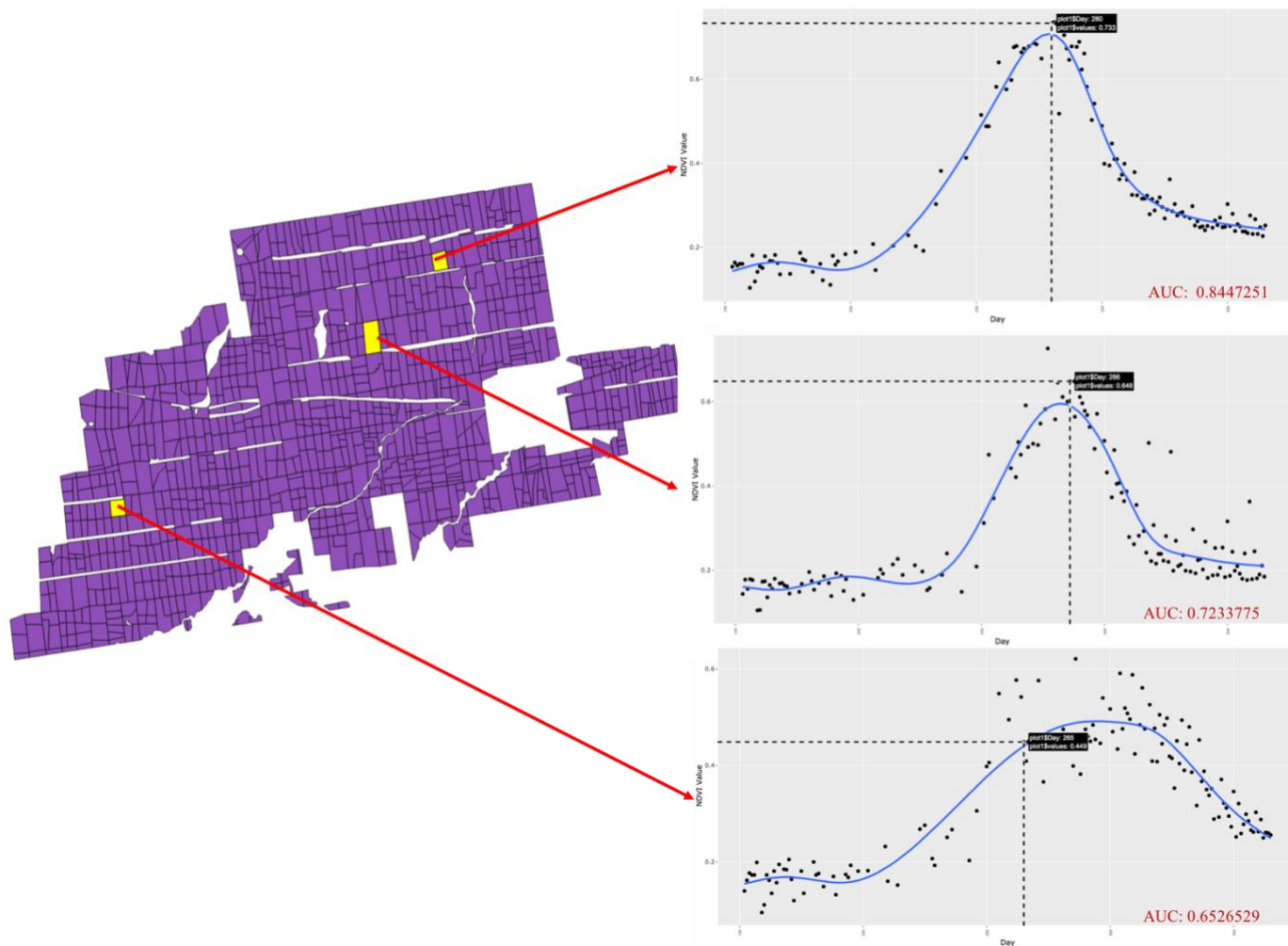


Figure 5.12: NDVI values plot basis for the year 2014.

Figure 5.12 shows the result of querying a plot for NDVI values, by clicking on the plot it generates a scatterplot with NDVI values from June to December, it is fitted with a LOESS curve and calculates the AUC ( Area Under the Curve) to get the NDVI integral value (0-1).

The sample NDVI graph for the plots in the year 2015, 2016, and 2017 are attached in the Annexure 6, Annexure 7, and Annexure 8 respectively.



### 5.4.2 Region Based NDVI:

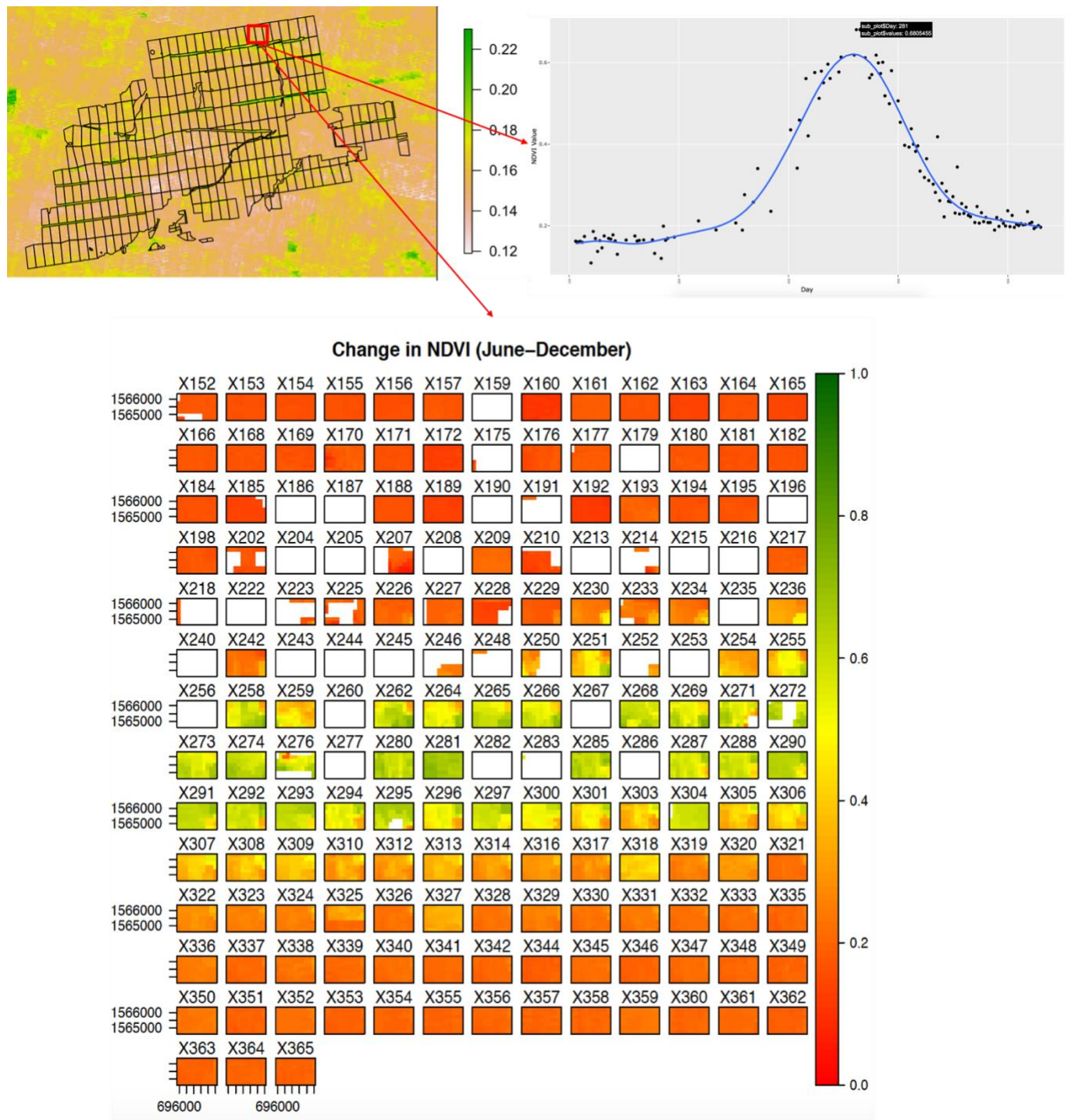


Figure 5.13: NDVI values region based for the year 2014.

Figure 5.13 shows a region based analysis for the year 2014, the user selects a region on the map and generates a NDVI graph for the selected region along with a daily basis NDVI map of the region.

## 5.5 Plot Based Analysis:

### 5.5.1 Multiplot Comparison-

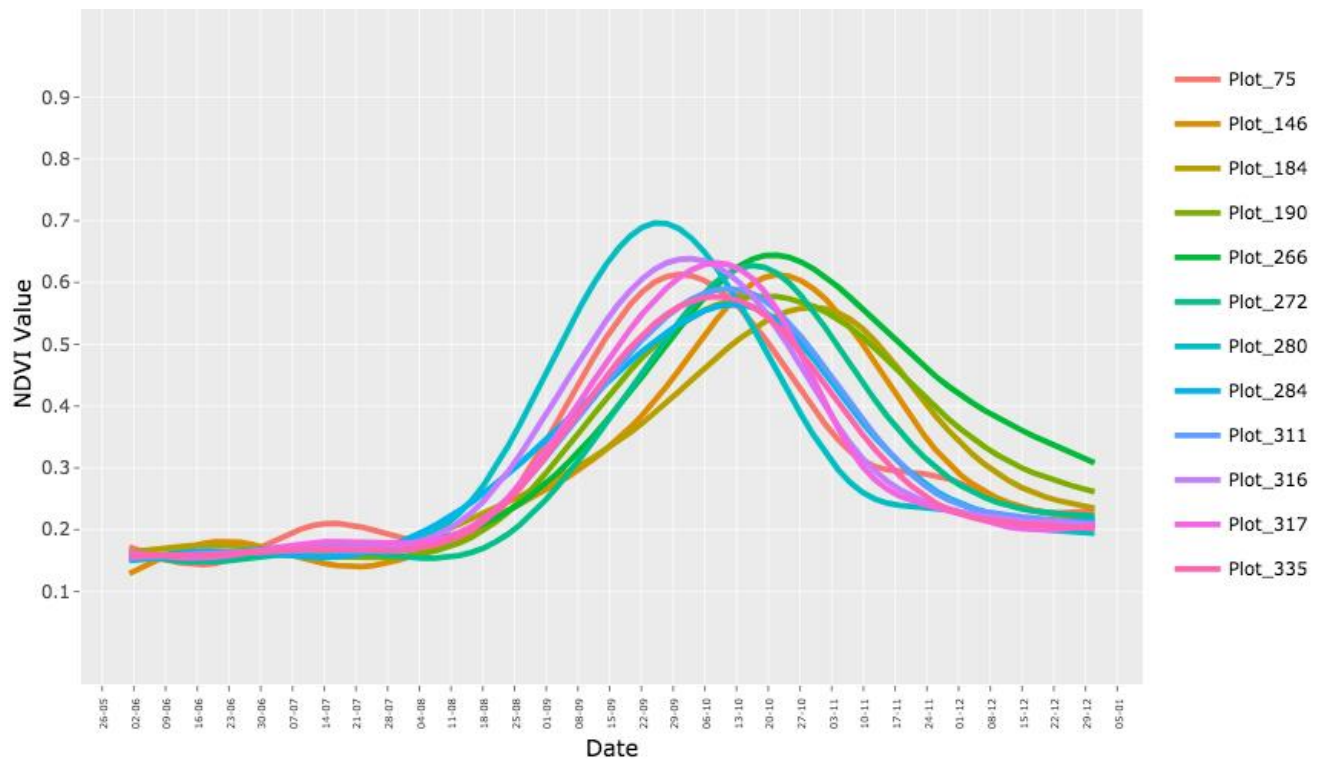


Figure 5.14: Multiplot NDVI-2014

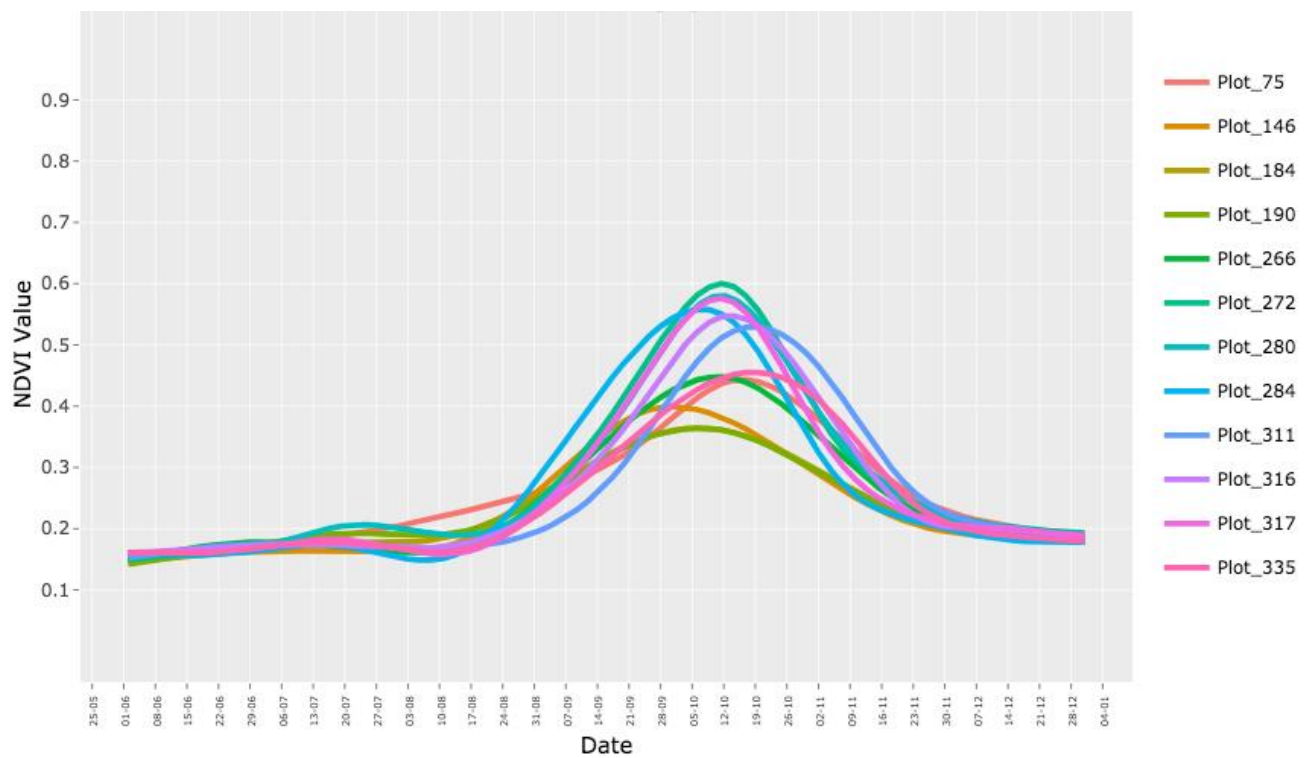


Figure 5.15: Multiplot NDVI-2015



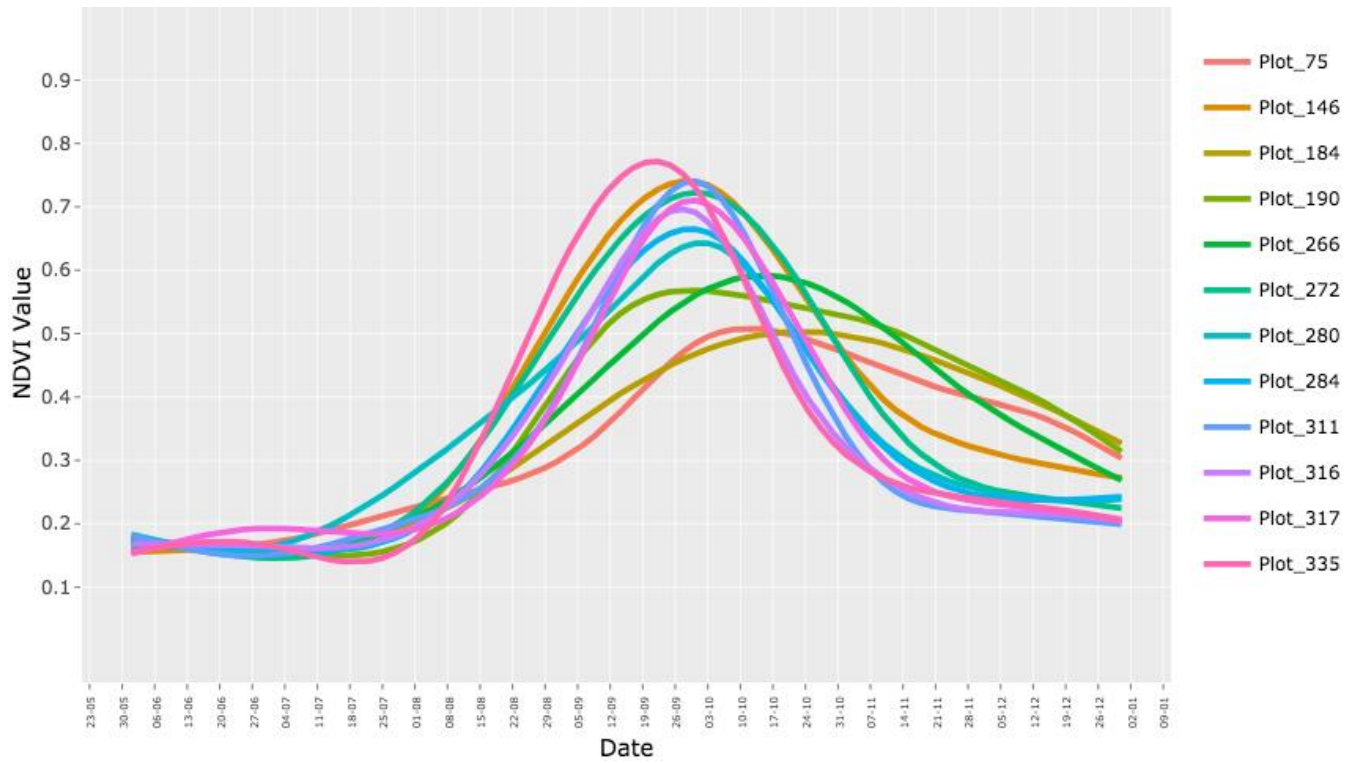


Figure 5.16: Multiplot NDVI-2016

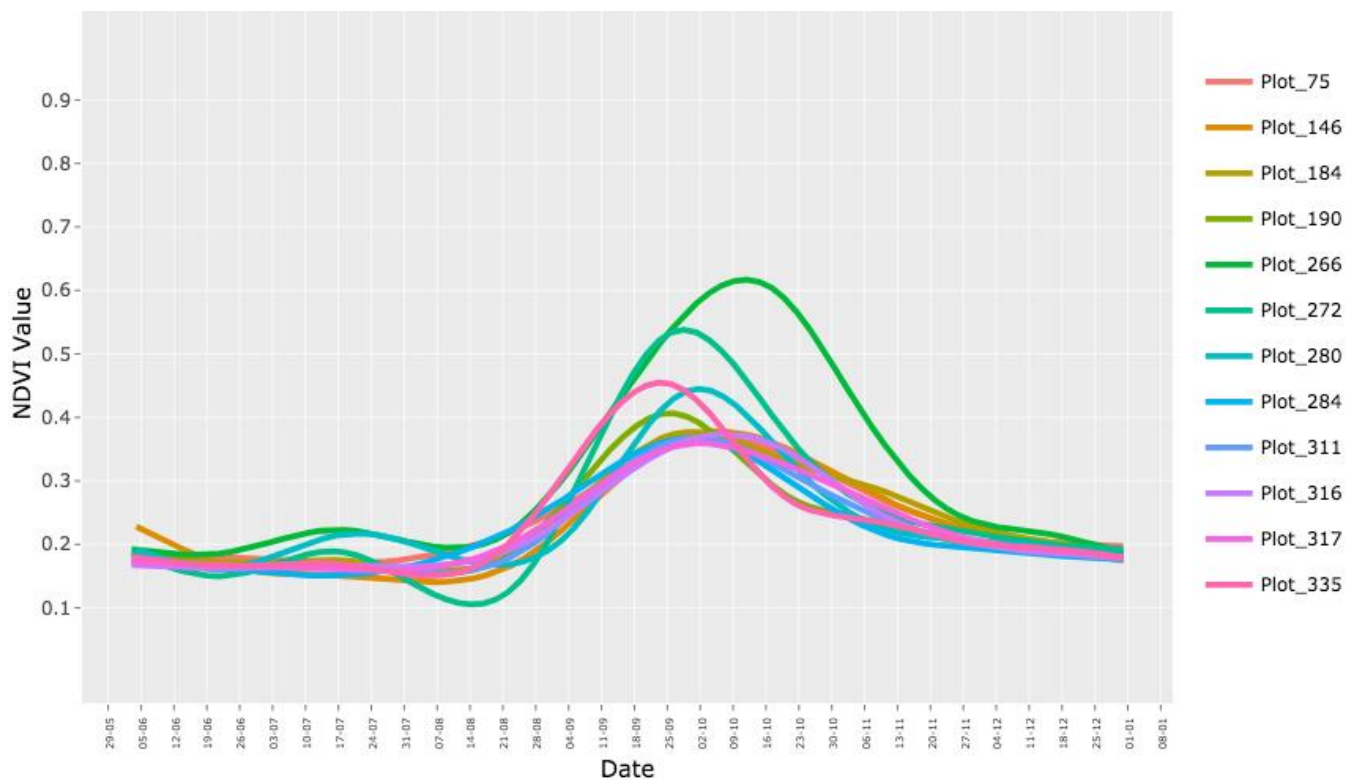


Figure 5.17: Multiplot NDVI-2017

Figure 5.14, Figure 5.15, Figure 5.16, and Figure 5.17 shows NDVI values of 12 randomly selected plot over the four years from 2014-2017 respectively. It is observed that the

same plot e.g. Plot-266 has different NDVI peak dates over the years. Also within the same year, plots has green-up peak at different times. This is because of crop rotation between years and different crop types sown in different plots within the same year.

The multiplot analysis for the plots would help to detect the changes of individual plot under same conditions and same crop type aiding the organisation to make more accurate yield predictions for a particular plot. Also this time series understanding of the plot would help to identify underperforming plots and also will give a better understanding of the crop type which is more sustainable to changing weather conditions.

### 5.5.2 Comparative Change detection:

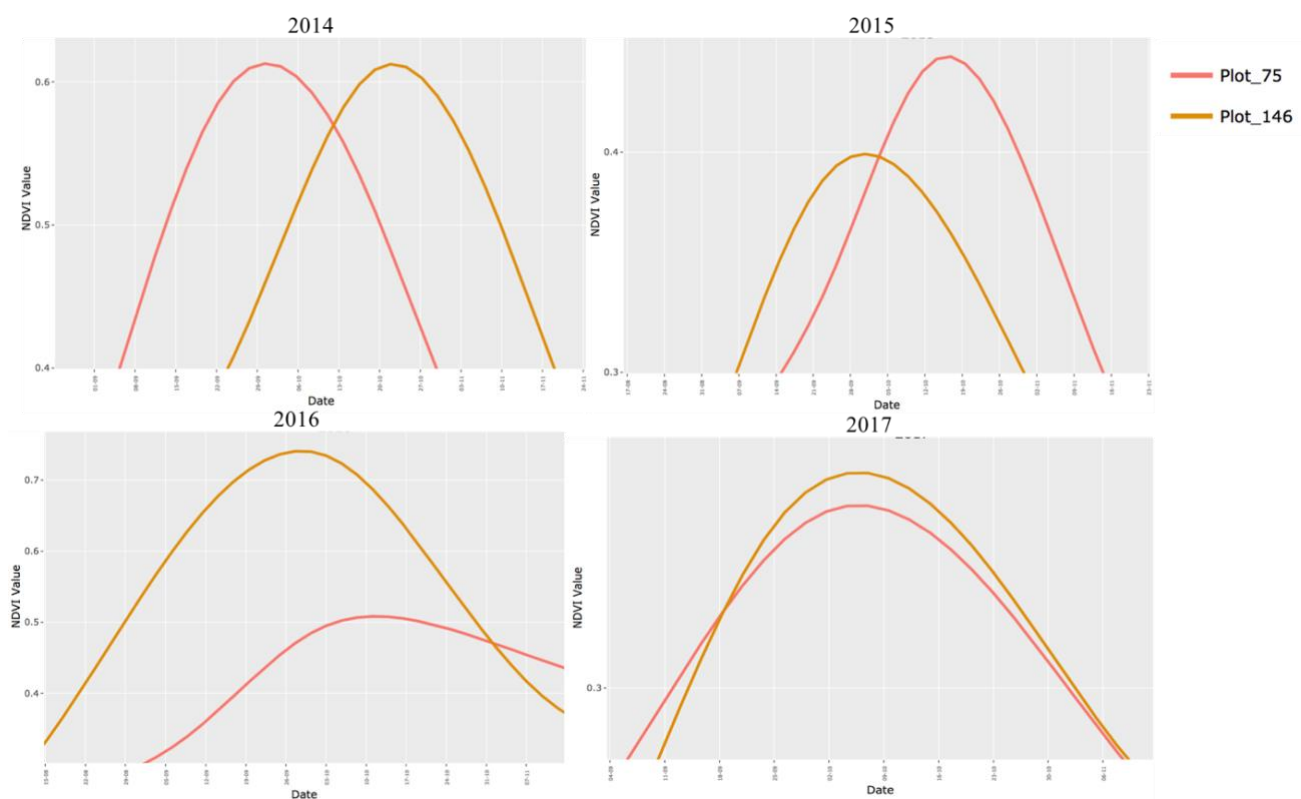


Figure 5.18: Comparison of Plot-75 and Plot-146.

	2014		2015		2016		2017	
Plot Number	75	146	75	146	75	146	75	146
Peak NDVI Value	0.612	0.612	0.443	0.399	0.508	0.740	0.366	0.378
Peak Date	30/09	22/10	17/10	01/10	11/10	28/09	07/10	07/10

Table 1: Peak NDVI values with date.

The Table 1 above for Figure 5.18, gives the maximum NDVI value reached by each plot during the four years. It is observed from the graph that Plot-75 and Plot-146 had crop rotation from 2014 -2016 but in 2017 they reached their maximum NDVI peak on the same day confirming the same crop type. Also the crop type can be assumed from the NDVI peaks based on the date they reached their maximum, like in 2014 Plot-75 had Sesame growing and Plot-146 has Sorghum .

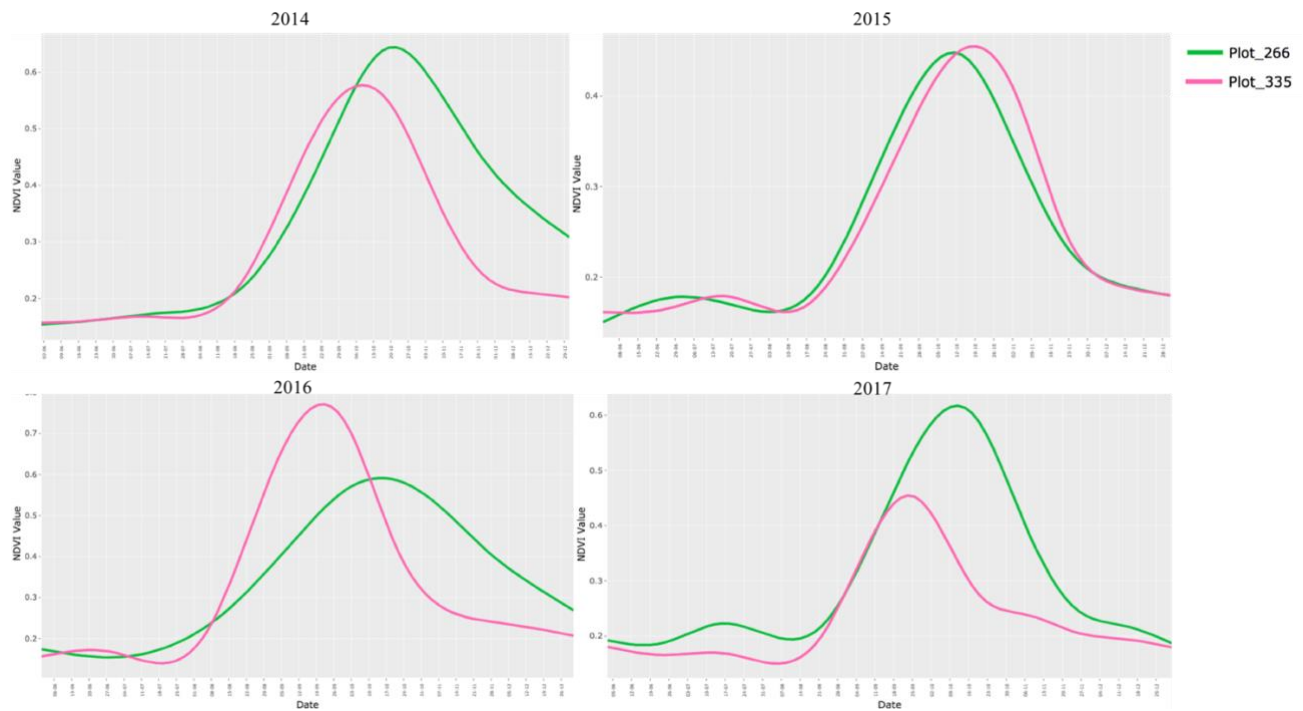


Figure 5.19: Comparison of Plot-266 and Plot-335.

	2014		2015		2016		2017	
Plot Number	266	335	266	335	266	335	266	335
Peak NDVI Value	0.577	0.644	0.448	0.455	0.591	0.771	0.616	0.454
Peak Date	08/10	22/10	12/10	20/10	14/10	29/09	12/10	23/09

Table 2: Peak NDVI values with date.

Table 2 gives the maximum NDVI value reached by each plot in Figure 5.19 during 2014-2017. The crop type in Plot-266 seems to be consistent without any rotation just variation in sowing time which resulted in different NDVI peak dates. But in Pot-335 the crop type changed between 2014-2015 and 2016-2017, which can be identified from the shift of NDVI peak dates in Table 2.

### 5.5.3 Growth Period:

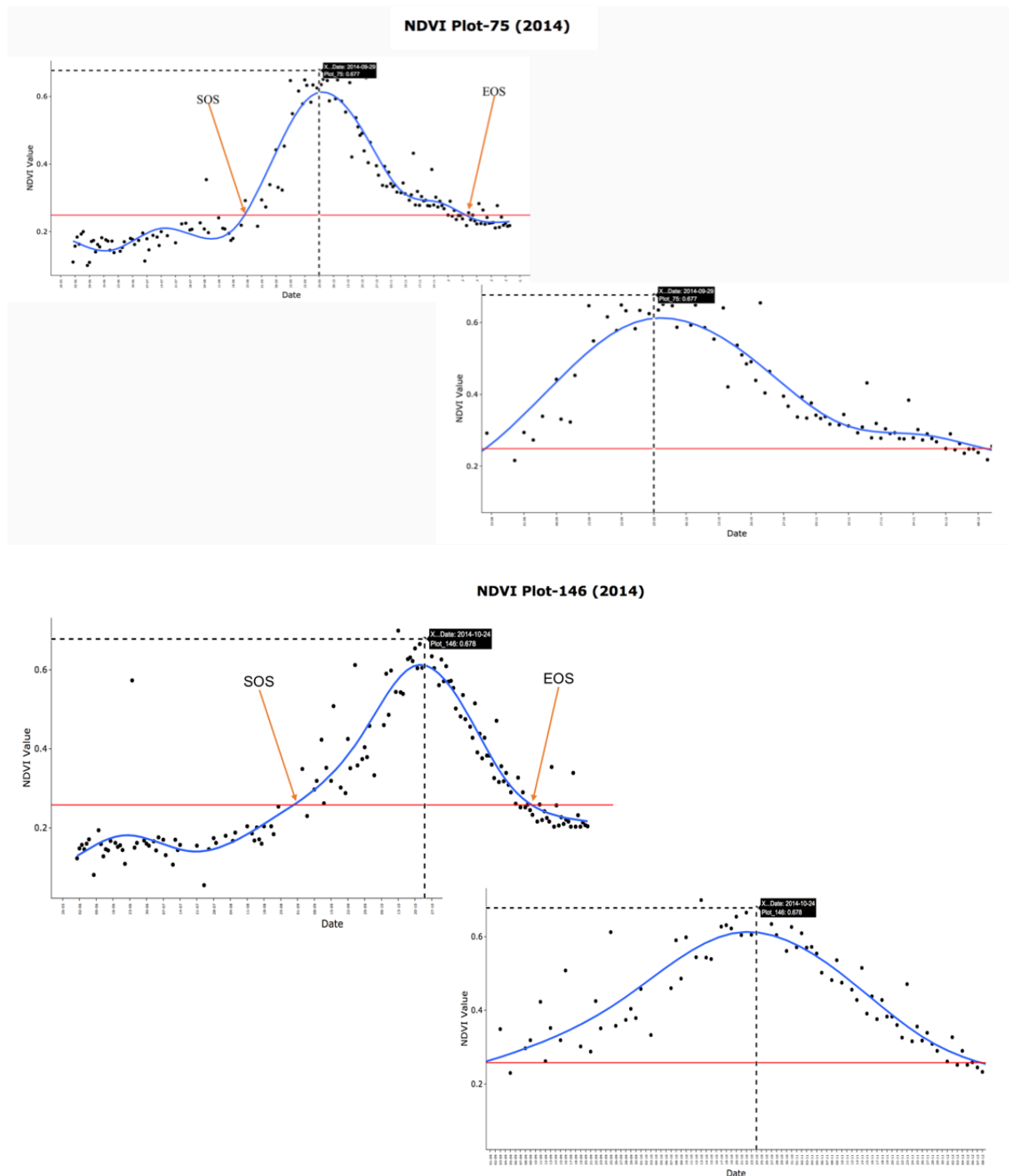


Figure 5.20: Growth season of a single plot.

Figure 5.20, shows the growth period of Plot-75 and Plot-146 from SOS (Start Of Season) indicating growth start to EOS (End Of Season) indicating the senescence stage. The dates for season start and end can be obtained from the graph. For Plot-75 the season starts from 24/08/2014 and ends at 09/12/2014 with NDVI peak of 0.677 on 29/09/2014. Plot-146 season starts from 30/08/2014 and ends at 07/12/2014 with NDVI peak of 0.678 on 24/10/2014.

## 5.6 Soil Moisture and NDVI:

The relationship between soil moisture and NDVI is calculated based on Lag, the underlying reason been the current state of NDVI can be affected by the state of soil moisture few periods back. It is important to understand the lag relationship, to identify the actual phase of soil moisture that has affected the growth of the crop. The change of NDVI is plotted against 6 Lag levels of soil moisture starting from 0 days up to 60 days with an increase of 10 days per lag so Lag0-0 days, Lag1-10 days, lag2-20 days, Lag3- 30 days, Lag4- 40 days, Lag-50 days, Lag6- 60 days.

Also CCF (Cross-Correlation Function) is calculated between the soil moisture and NDVI to identify if whether soil moisture “causes” changes in NDVI. CCF is defined as the set of sample correlations between  $x_{t+h}$  and  $y_t$  for  $h = 0, \pm 1, \pm 2, \pm 3$ , and so on where:  $t$ -time and  $h$ -deviation from time. A negative value for  $h$  is a correlation between the  $x$ -variable at a time before  $t$  and the  $y$ -variable at time  $t$ . The CCF value would give the correlation between  $x_{t-h/t+h}$  and  $y_t$  as:

- When one or more  $x_{t+h}$ , with  $h$  as negative are predictors of  $y_t$ , it is said that  $x$  leads  $y$ .
- When one or more  $x_{t+h}$ , with  $h$  as positive are predictors of  $y_t$ , it is said that  $x$  lags  $y$ .

The CCF is calculated based on Kendall Method, since it a non-parametric correlation coefficient that evaluates the degree of similarity between two sets non-related monotonic variables . the values can range between +1 and -1.

The soil moisture is recorded by SMAP in  $m^3$  (i.e., how much of a cubic metre is water out of the entire cubic metre of soil sample) For all the graphs showing the relationship between soil moisture lag and NDVI the Y-axis represents both soil moisture in  $m^3$  and NDVI values.

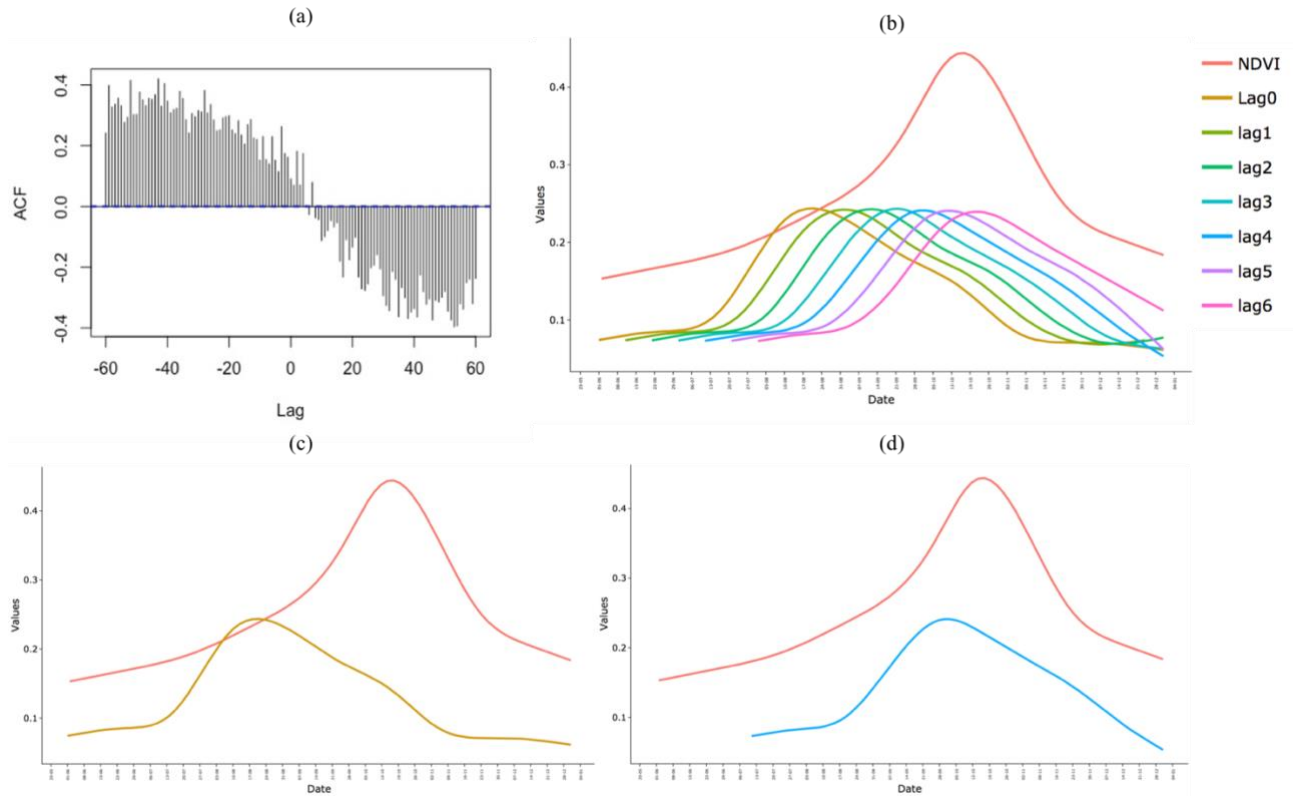


Figure 5.21: Soil Moisture Lags compared to NDVI value of Plot-75 for 2015. (a) CCF Graph, (b) Lag 0-Lag 6 compared to NDVI, (c) Lag 0 in relation to NDVI, (d) Lag 4 in relation to NDVI.

Lag0	Lag1	Lag2	Lag3	Lag4	Lag5	Lag6
0.05452639	0.166704	0.2670183	0.3199787	0.412927	0.3778785	0.2185119

Table 3: Kendall Correlation Values Plot-75 (2015)

Figure 5.21, shows the relationship between soil moisture and NDVI for Plot-75 in the year 2015. Figure 5.21(a) represents the CCF, it can be seen that x has positive correlation to t-h and can be concluded that soil moisture leads NDVI. Figure 5.21(b) shows all the lags compared to NDVI. Figure 5.21(c) shows lag0 which represents the soil moisture value on the same days as the NDVI and it can be seen from the Table 3 that Lag0 and NDVI has the least correlation and the effect of soil moisture is maximum with Lag4.

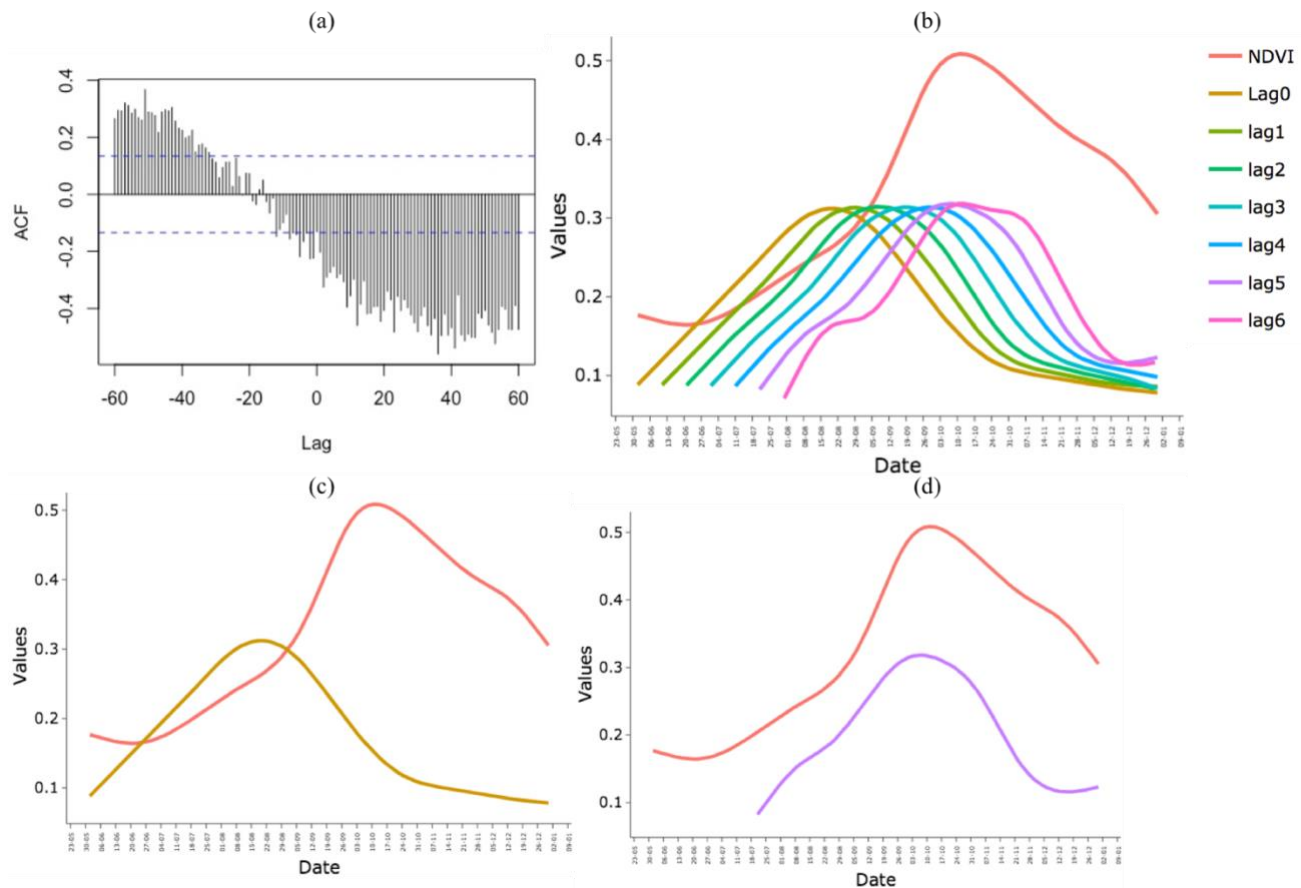


Figure 5.22: Soil Moisture Lags compared to NDVI value of Plot-75 for 2016. (a) CCF Graph, (b) Lag 0-Lag 6 compared to NDVI, (c) Lag 0 in relation to NDVI, (d) Lag 5 in relation to NDVI.

Lag0	Lag1	Lag2	Lag3	Lag4	Lag5	Lag6
0.006915708	0.01948238	0.1453802	0.1915342	0.3019469	0.3497944	0.3207433

Table 4: Kendall Correlation Values Plot-75 (2016)

Figure 5.22(a) shows the CCF between soil moisture and NDVI for the year 2016, but unlike 2015 the CCF value is negative until  $h=-20$  and then starts showing a positive correlation, in this case also soil moisture lead NDVI but the relationship is not very strong. Also the difference in correlation between lag 3 and lag 4 (seen in Table 4) is higher than the changes observed in 2015. This can be attributed to the reason that the initial few months of the growing season received more rainfall and had higher soil moisture.



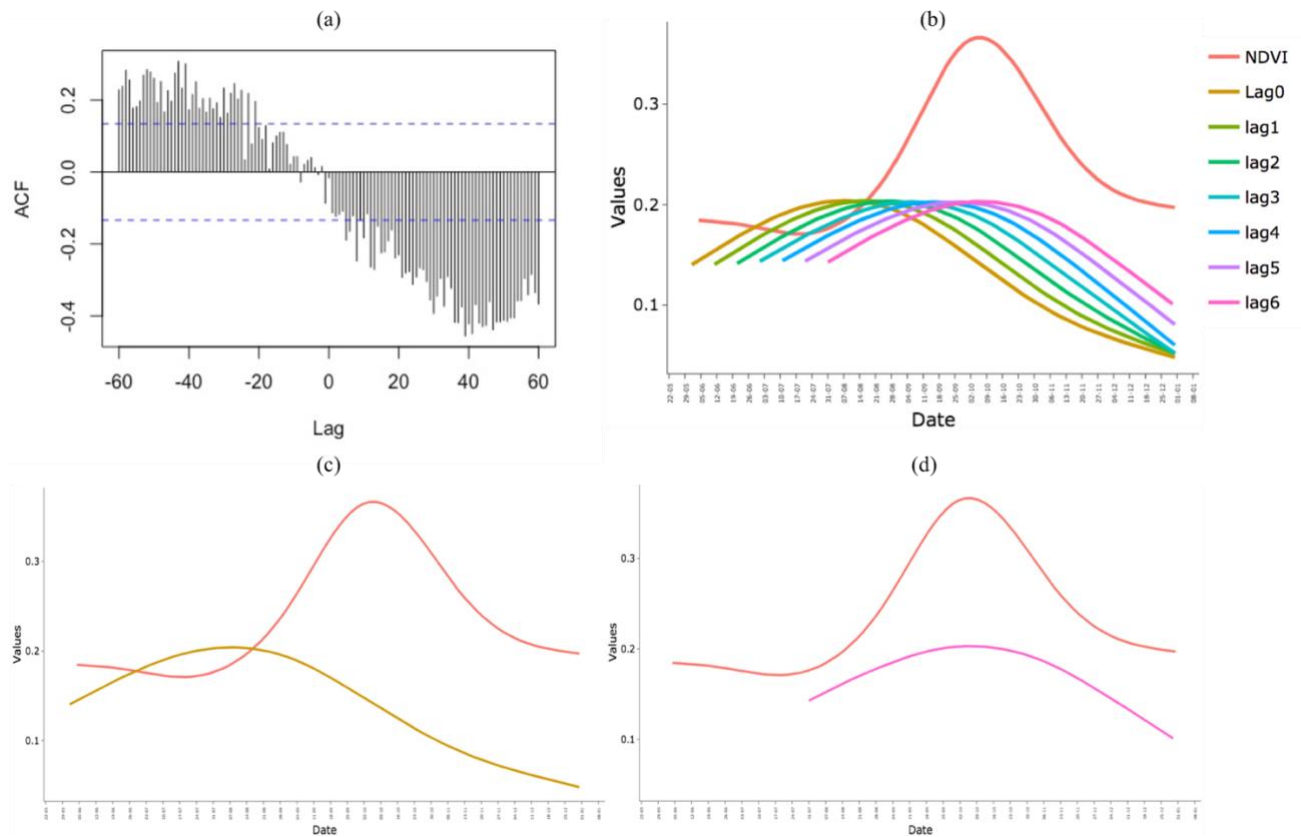


Figure 5.23: Soil Moisture Lags compared to NDVI value of Plot-75 for 2017. (a) CCF Graph, (b) Lag 0-Lag 6 compared to NDVI, (c) Lag 0 in relation to NDVI, (d) Lag 6 in relation to NDVI.

Lag0	Lag1	Lag2	Lag3	Lag4	Lag5	Lag6
-0.02019238	0.05214019	0.1109853	0.1841812	0.1998171	0.2205363	0.2305681

Table 5: Kendall Correlation Values Plot-75 (2017)

Figure 5.23(b), shows a lower level of soil moisture through all the lags as 2017 was a drought year Table 5, shows that the maximum correlation between NDVI and soil moisture is in Lag6 (60 days) with negative relation to Lag0. It is possible to have a negative relation in a drought year since the soil moisture in the current time is lower than the initial soil moisture. Based on the table and Figure 5.23(a), it can be concluded that soil moisture lead NDVI but only in the initial phase and doesn't have a strong relation throughout.

2015 and 2017 were drought years, but the relationship between the lag days and NDVI didn't get much affected for all the years (2015, 2016, 2017,) and a consistent lead of soil moisture to NDVI remained between Lag4-lag 6. It can be concluded that soil moisture is not much affected by rainfall as concluded by (Ahmed, Else et al. 2017) , and the NDVI is affected

by soil moisture up to a lag period of 60 days (2 months). Further plot based relationship between soil moisture and NDVI are in Annexure 10, Annexure 11 Annexure 12.

Following the Lag relationship between NDVI and soil moisture, it is important to identify the phenology phase at which the soil moisture has the maximum impact on NDVI. As seen from the above figures the soil moisture effects the overall NDVI at Lag4 - Lag6 with maximum correlation within this range but this necessarily doesn't have the same relation for the phenology phases.

Figure 5.24, shows the relationship between NDVI and soil moisture at time (t) without any Lag. The figure shows behaviour of 3 different plots Plot-75, Plot-146 and Plot-186 for the years 2015, 2016 and 2017. It can be seen that the crop growth starts after the soil moisture reaches it maximum for all the 3 plots in all the years with a distinct increase in NDVI values immediately after the soil moisture reaches maximum. Also drop in soil moisture affects the NDVI peak of the crop and the NDVI starts decreasing mostly after the soil moisture level falls below  $120 \text{ m}^3 - 150 \text{ m}^3$ . For the year 2017 the overall soil moisture levels were low, thus a gradual decline in the soil moisture along with existing climatic conditions restricted the maximum NDVI of the crops. It can be concluded that the maximum effect of the soil moisture is in the growth stage of the crop when it initiates the main development of the crop, and plays an important role in the crop health.

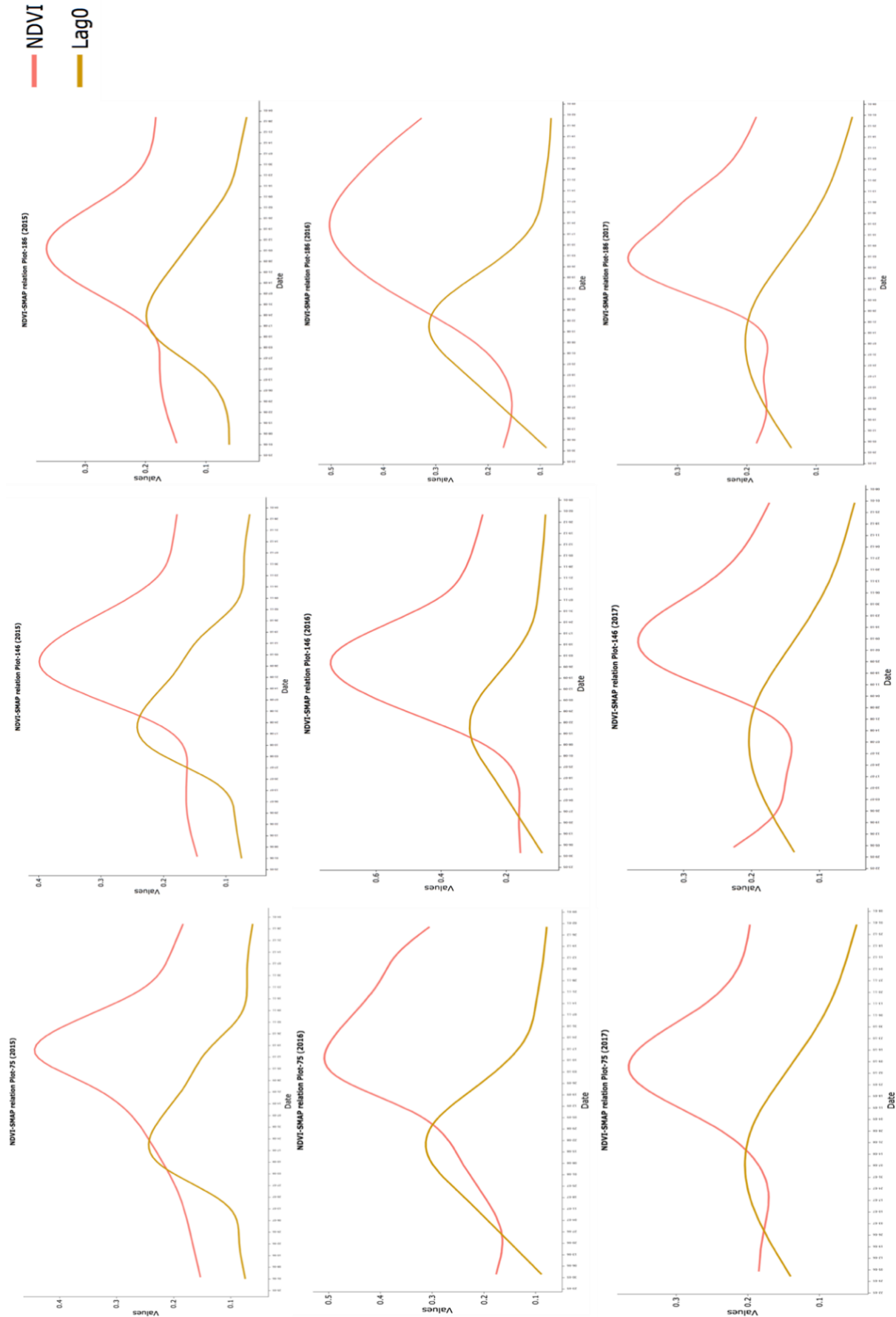


Figure 5.24: Effect of soil moisture at time ( $t$ ) on crop growth.

## 5.7 Tools for processing:

Purpose	Tool	Note
Data Acquisition:		
Data Format Conversion	Python Code	Source: Dr. Nikolas Prechtel (TUD)
Resampling	Python Code	Source: Dr. Nikolas Prechtel (TUD)
Clipping	Python Code	Source: Dr. Nikolas Prechtel (TUD)
Reprojection	QGIS	Batch processing tool
Digitisation	QGIS	Manually done.
Data Post-Processing:		
Cloud Masking	Python Code	Source: Dr. Nikolas Prechtel (TUD)
Accuracy Check	QGIS	Between actual image and masked image
Main Workflow:		
NDVI Calculation	Python Code	Source: Dr. Nikolas Prechtel (TUD)
Bayesian Data Smoothing	QGIS	Orfeo plugin in QGIS for image processing
Fusion	QGIS	Conditional merging using raster calculator
Accuracy Check	R	Code for R Squared r
Statistical Analysis	R	Code for CCF, Lag
Interactive NDVI visualization	R	Discussed in Chapter.5

## 6. Conceptual Tool:

This chapter will outline the conceptual design and workflow of the visualisation tool. It consists of a pipeline workflow along with a mock interface. As discussed in (section 2.6) the existing tools for visualisations are limited to a single sensor NDVI calculation and visualisation. The focus of this tool will be to create an uniform and continuous pipeline for calculating and viewing NDVI time series data, fuse sensors to increase coverage and NDVI generation. Furthermore, it is supposed to statistically establish and visualise relationship between NDVI and soil moisture (additional parameters can be incorporated later).

The interface can be created by combining QGIS plugin, R and Python as these were the tools used or developed for processing the data in the thesis. The advantage of combining these are, QGIS already has certain pre-defined tools and plugins for processing the data like the ORFEO Toolbox which was used for smoothening of the data by BDF. Similarly R has a dynamic interface of incorporating statistical calculation and analysis faster and convenient, and Python as the binder language for both the interfaces. The visualisation and processing tools are split into 3 tabs, which are shown as mock visualisations in Figure 6.1, Figure 6.3, and Figure 6.5 respectively with their workflow explained in the flowcharts.

### 6.1 Tab-1: Masking

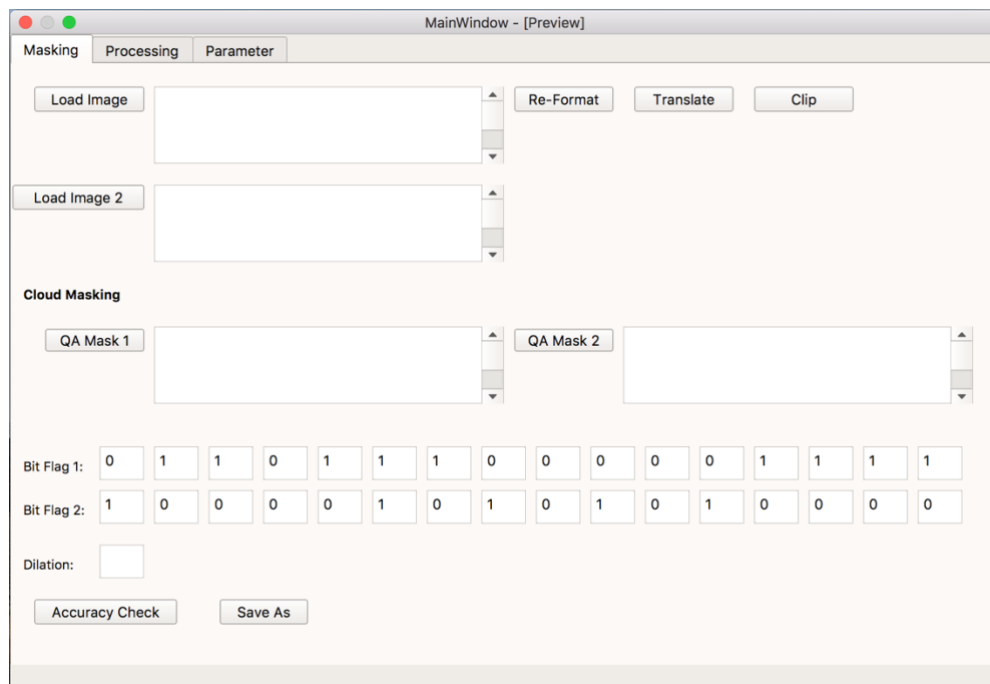


Figure 6.1: Mock Visualisation of Making tool (Tab-1).

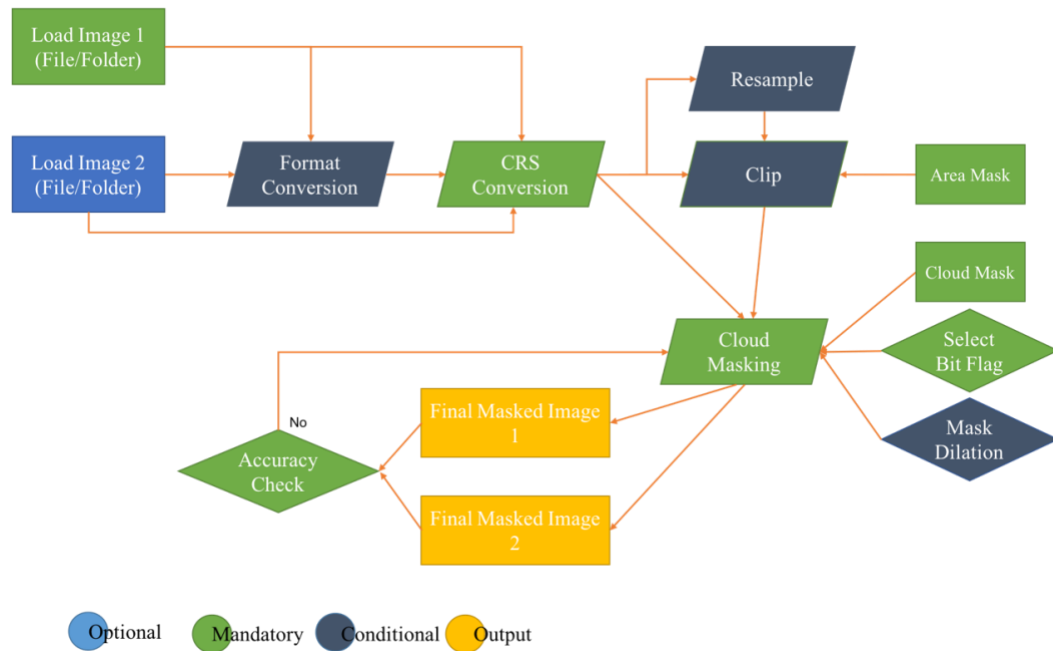


Figure 6.2: Workflow of Tab-1.

The boxes are coloured differently based on their priority. The ‘Optional’ are the inputs and processes that are voluntarily chosen by the user. The entire workflow can go ahead without these inputs as well. The ‘Mandatory’ are the inputs and processes that are required for a successful completion of the workflow. The ‘Conditional’, are the values and processes that are required to be provided depending on the type of input being used for the workflow.

As observed in the flowchart in Figure 6.2, the user can upload input images either as folder or a single file (image). The second image is optional in case the user wants to calculate NDVI of a single image set. Format Conversion is ‘Conditional’, if the images are in HDF then the conversion to geotiff is required. As the CRS of different images are different e.g. it is sinusoidal in MODIS and EASE in PROBA- V causing compatibility issues between images. CRS Conversion is ‘Mandatory’. Resampling is conditional as it is needed to be used if multiple images are being used for harmonizing the resolution. Clip is ‘Conditional’ depending if the user needs to study a particular region or not but if clip is being used then an Area Mask is ‘Mandatory’.

Cloud Masking is a ‘Mandatory’ part for the Image Processing whether single or multiple images are being used to get more accurate and complete values. The user has to upload the Quality layer as a mask layer –‘Mandatory’. The user has to select the bit flags to be used for masking data which is ‘Mandatory’ if the user has only one set of image then Bit Flag2 and QA flag2 are deactivated, but the dilation depends on the type of cloud mask, in certain cases

the mask are accurate unlike in others thus it's 'Conditional' depending on the type of cloud mask being used. The final output will be consisting of the cloud masked images as shown in Figure 4.11, the accuracy check is 'Mandatory' before obtaining the final images.

## 6.2 Tab-2: Processing

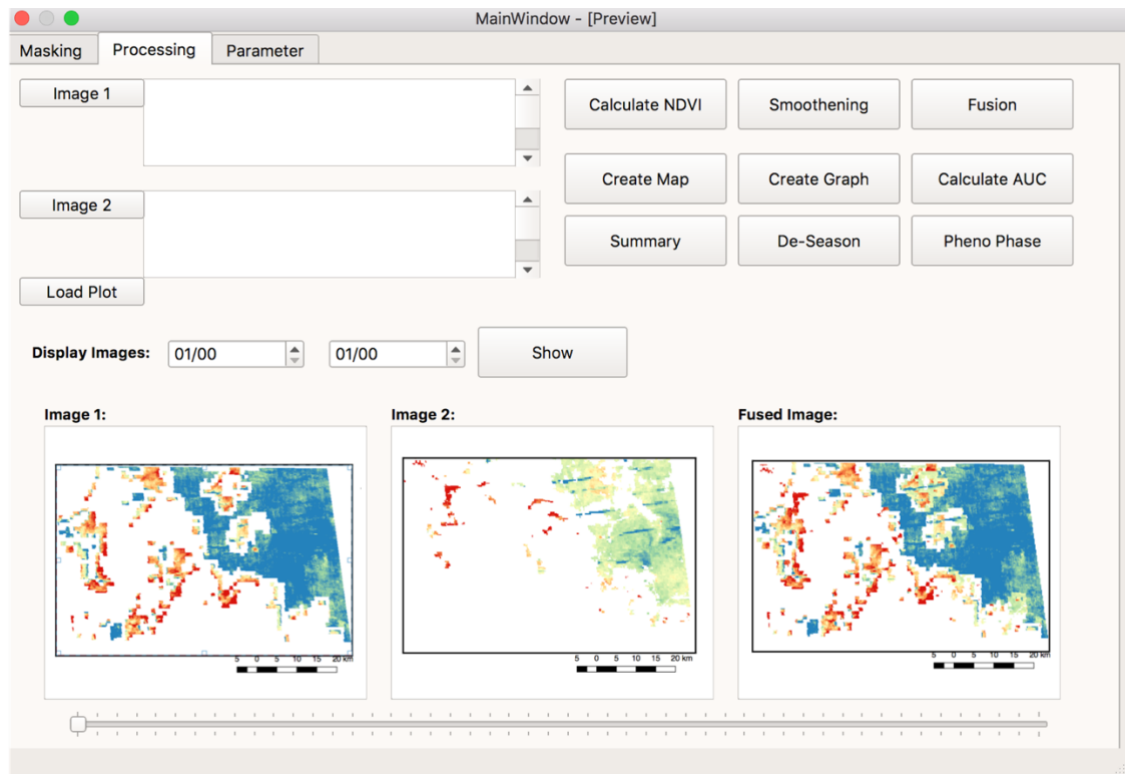


Figure 6.3: Mock Visualisation of Processing tool (Tab-2).

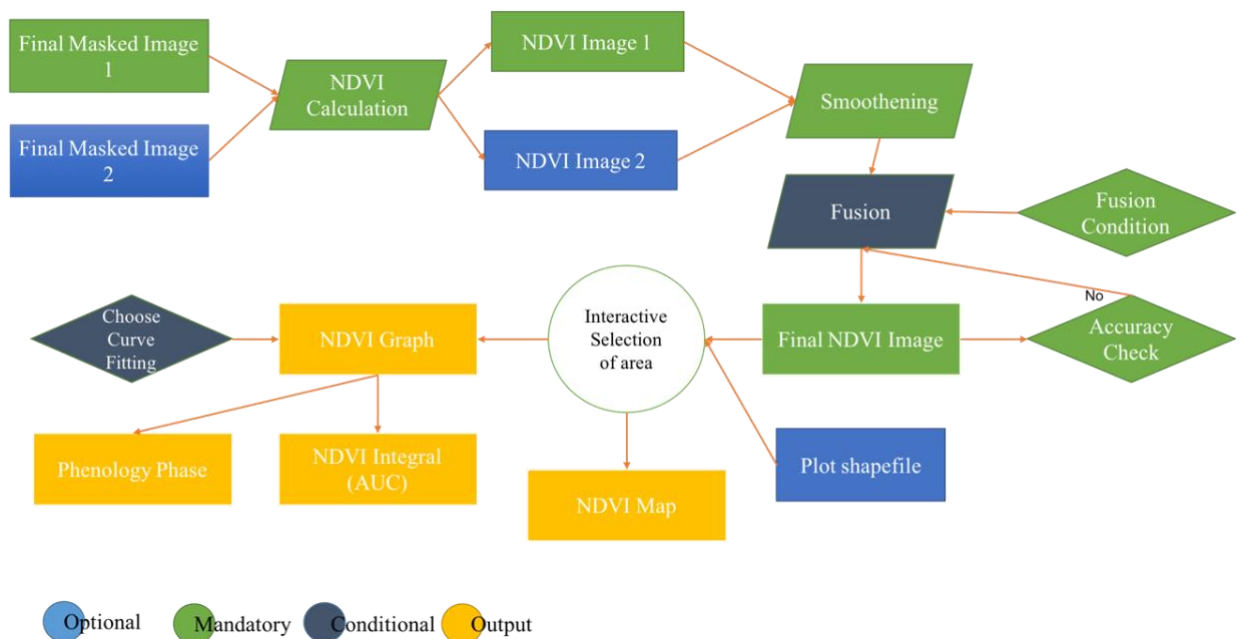


Figure 6.4: Workflow of Tab-2.



Tab-2 (see: Figure 6.3) is the main processing toolset for producing the final maps and calculating the NDVI for each plot. As seen in the workflow (Figure 6.4) the user can load the masked images either as folder or single image. The second image is optional in case the user wants to calculate NDVI of a single image set. The NDVI calculation from the images is done based on the band positioning which can be provided as input to the NDVI calculator. Smoothing of the image is 'Mandatory' to reduce if any noise remains. The Fusion is 'Conditional' as this step is only required if two images are being used in the processing and in order to have a successful fusion the fusion condition is 'Mandatory'. An accuracy check should be done to check the final fused NDVI image or single NDVI image.

The final NDVI image generated is in an interactive window, where the user can select a plot (refer Figure 5.12) or choose a certain region (refer Figure 5.13) to generate the NDVI value of that region along with an NDVI map. The graph generated is interactive and provides information about the phenological phases and the NDVI integral value.

### 6.3 Tab-3: Parameter

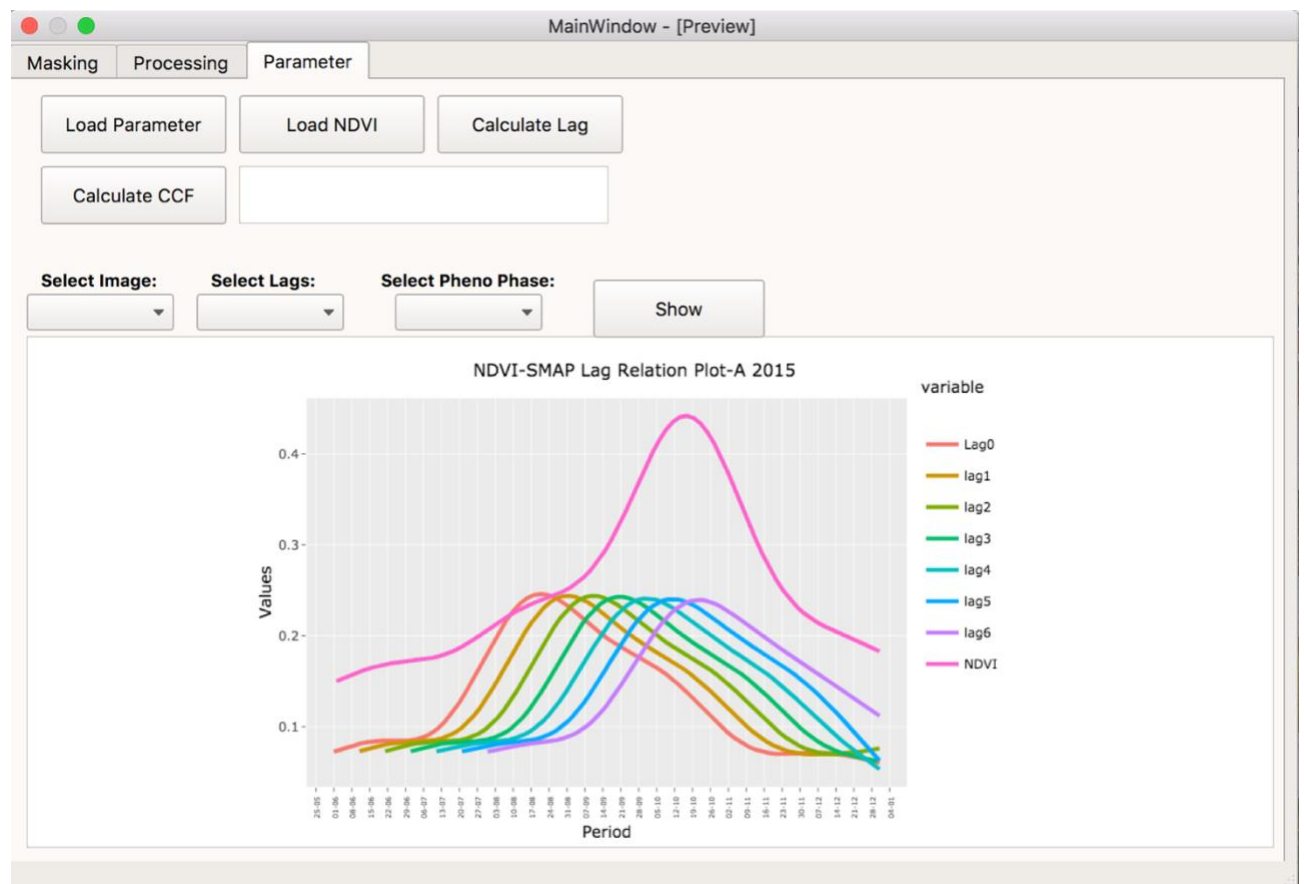


Figure 6.5: Mock Visualisation of Parameter tool (Tab-2)

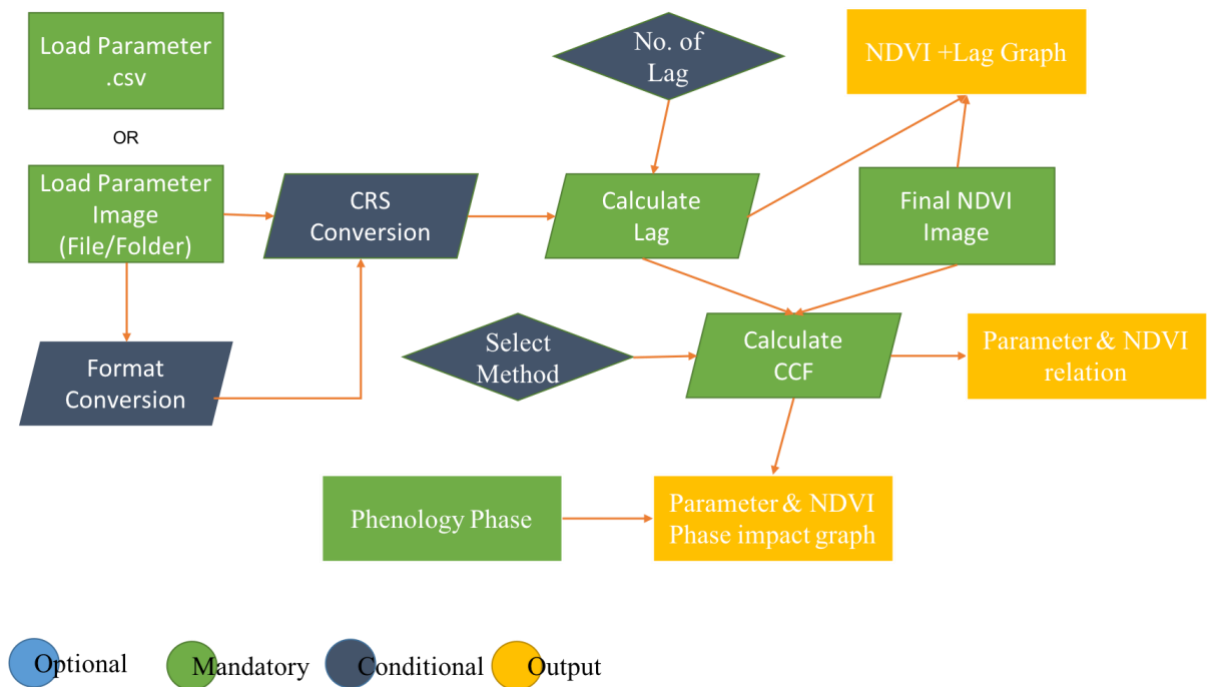


Figure 6.6: Workflow of Tab-3

Tab-3 (see: Figure 6.5) is created for using it with soil moisture. This tool can be developed more flexibly so other parameters like surface temperature or surface evaporation can be incorporated. The workflow requires the parameter data as a raster or .csv which is 'Mandatory'. The CRS conversion and format conversion are 'Conditional' as they are only required to be used if the dataset is a raster. The lag calculation is 'Mandatory' for establishing relation between NDVI and parameter, and the user has to provide the number of time lag in days, or it will consider the default. Combining the lag and NDVI data the relation graph can be created as output (refer Figure 5.21). The CCF ( Cross Correlation/Covariance coefficient) is 'Mandatory' as it provides the relationship between the variables used, additionally the user can select the method of calculation like Kendall, Spearman's or Pearson's. Combining the data will provide different outputs, which can be chosen from the R interface by the user.

## 7. Conclusion

This thesis calculates the NDVI values for the study area by fusing MODIS and PROBA-V data and establishes a relationship between soil moisture and NDVI using SMAP and the fused data. For data preparation it comprises of cloud masking, resampling, reprojection, smoothening and fusion. The data results and analysis comprises of mean monthly NDVI values of the study area, anomaly maps to identify the changes in NDVI values, creation of LOESS curve fitted NDVI Graphs for plots or regions with NDVI integral values, and calculation of lag based CCF based on Kendall method to establish a relationship between soil moisture and NDVI.

This research was guided by 5 main objectives, which are concluded below:

- Harmonisation and fusion of NDVI from MODIS and PROBA-V.

The first objective of this research was to harmonise and fuse MODIS and PROBA-V, which was successfully done as discussed in section 4.3. The main aim behind harmonising these two datasets were to increase and expand the image data availability for the study area.

As both the sensors have daily coverage the obstructions due to clouds were present in both the images. At the same time fusing of two sensors increased the spatial coverage for individual plots since the scan time for the sensors were different and consequently, the associated cloud covers too, which allowed an observation of different parts of the study area at different times on the same day. As a result an overall increase of 20-40% plot coverage was observed.

Thus it can be concluded that the use of two sensors increased the spatial coverage of the study area, as due to the fusion plot-based pixel information improved. But it had limited impact on the temporal coverage, on an overall availability PROBA-V had more data available than MODIS combined, days with heavy cloud covers were masked out for both MODIS and PROBA-V and information was missing from both the sensors,

- To produce a plot-based time series of vegetation development.

A plot based time series of vegetation development is created and discussed in 5.4. It gives information about a particular plot, and can be compared with other plots. Also the plot based analysis gives information about the NDVI integral for an individual plot or a region aiding the user to compare the performance across different years. The plot based analysis also

suggests the crop type by recognising the time for onset of growth , the senescence date and the date on which the NDVI peak is reached.

- Extrapolate vegetation signal as a function of soil moisture at a specific development stage.

A lag- based relationship is established between the soil moisture and NDVI data. It is observed that the correlation between the soil moisture and NDVI had the strongest positive relationship between lag 4-lag 6 depending on the year and the crop type. Plots which were growing sesame crop had a relation stronger in Lag 4, whereas plots growing Sorghum had the relation stronger during lag 5-lag 6. Sesame requires higher soil moisture when it starts to grow, until it reaches the impetus, whereas sorghum in comparison requires longer period. The NDVI can be extrapolated based on the Lag phase of the soil moisture, but the soil moisture is also affected by surface temperature, surface evaporation etc, and thus incorporating these parameters in future researches would increase the precision of the relationship.

- Analyse to what degree existing soil moisture recordings influence the vegetation signal at various crop development stages.

The main effect of soil moisture on NDVI is observed at the onset of growth. It was clearly observed that as Lagged time had a higher correlation with the overall NDVI, the current time of soil moisture majorly impacted the crop growth. As shown in Figure 5.24, most of the plots had the onset of growth immediately after the soil moisture reached it's peak. It was also observed that the amount of soil moisture affected the overall growth and health of the plant. In plots which had lower soil moisture during the onset of growth the NDVI values were lower compared to other plots with higher soil moisture within the same year.

- Discuss visualisation techniques synoptically displaying the source of information and the uncertainty involved.

In chapter 6, a conceptual model for the visualisation and processing of the images had been outlined. It aims to provide a plot-based time series NDVI visuals along with an anomaly map showing the changes and pipeline for processing and preparing dataset.

This thesis can be further adapted and extended at various levels to extract and predict more information on crop monitoring. Currently the thesis covers the calculation of an overall NDVI integral based on each plot, but this can be improvised to derive the NDVI integral at various phenology stages which can be further considered for crop yield prediction. Also the NDVI integral can be further used in relation to the soil moisture at various phenological stages

to establish a more consolidated relationship between soil moisture and crop growth aiding the identification of phases that are mostly affected. As, soil moisture is also affected by other factors like vegetation water content, surface temperature and evapotranspiration, incorporating these as additional parameters in the Soil Moisture-NDVI relationship will increase the certainty of the values. Considering relationship between other climatic parameters like temperature, rainfall with NDVI will contribute to a fine-grained estimation of productivity. An implementation of such an integrated tool will definitely help to monitor and predict the crop yield in a more dynamic and accurate manner.

## References

- Ackerman, S. A., R. E. Holz, R. Frey, E. W. Eloranta, B. C. Maddux and M. McGill (2008). "Cloud Detection with MODIS. Part II: Validation." Journal of Atmospheric and Oceanic Technology **25**(7): 1073-1086.
- Adam, E., O. Mutanga and D. Rugege (2010). "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review." Wetlands Ecology and Management **18**(3): 281-296.
- Ahmed, M., B. Else, L. Eklundh, J. Ardo and J. Seaquist (2017). "Dynamic response of NDVI to soil moisture variations during different hydrological regimes in the Sahel region." International Journal of Remote Sensing **38**(19): 5408-5429.
- Biro, K., B. Pradhan, M. Buchroithner and F. Makeschin (2013). "Land Use/Land Cover Change Analysis and Its Impact on Soil Properties in the Northern Part of Gadarif Region, Sudan." Land Degradation & Development **24**(1): 90-102.
- Biro, K., B. Pradhan, H. Sulieman and M. Buchroithner (2012). "Exploitation of TerraSAR-X Data for Land use/Land Cover Analysis Using Object-Oriented Classification Approach in the African Sahel Area, Sudan." Journal of the Indian Society of Remote Sensing **41**(3): 539-553.
- Chen, J., P. Jönsson, M. Tamura, Z. Gu, B. Matsushita and L. Eklundh (2004). "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter." Remote Sensing of Environment **91**(3-4): 332-344.
- Chen, T., R. A. M. de Jeu, Y. Y. Liu, G. R. van der Werf and A. J. Dolman (2014). "Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia." Remote Sensing of Environment **140**: 330-338.
- Commission), J. R. C. E. (2015). "Comparison of PROBA-V, METOP and eMODIS NDVI anomalies over the Horn of Africa.pdf." EU Publications.
- Das, N. N. D., R.S. (2015). "Level 3 Active/Passive Soil Moisture
- Product Specification Document."
- Doraiswamy, P. (2004). "Crop condition and yield simulations using Landsat and MODIS." Remote Sensing of Environment **92**(4): 548-559.
- Doraiswamy, P. C. and T. Hodges (1979). Crop phenology literature review for corn, soybean, wheat, barley, sorghum, rice, cotton, and sunflower. United States, Lockheed Electronics Co.; Houston, TX, United States.



- Duveiller, G., R. Lopez-Lozano and A. Cescatti (2015). "Exploiting the multi-angularity of the MODIS temporal signal to identify spatially homogeneous vegetation cover: A demonstration for agricultural monitoring applications." Remote Sensing of Environment **166**: 61-77.
- Eilers, P. H. (2003). "A perfect smoother." Anal Chem **75**(14): 3631-3636.
- Entekabi, D. D., N. et.al (2014). "SMAP\_Handbook."
- Fasbender, D., J. Radoux and P. Bogaert (2008). "Bayesian Data Fusion for Adaptable Image Pansharpening." IEEE Transactions on Geoscience and Remote Sensing **46**(6): 1847-1857.
- Fensholt, R., I. Sandholt and M. S. Rasmussen (2004). "Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements." Remote Sensing of Environment **91**(3-4): 490-507.
- Hermance, J. F. and H. M. Sulieman (2018). "Adequacy of the daily TMPA 3B42 high-resolution satellite precipitation product for monitoring hydrometeorological hazards in the Southeast Sahel of Africa." International Journal of Remote Sensing **39**(8): 2579-2596.
- Inglada, J., O. Hagolle and D. Dedieu (2011). "LOW AND HIGH SPATIAL RESOLUTION TIME SERIES FUSION FOR IMPROVED LAND COVER MAP PRODUCTION.pdf." 6th International Workshop on the Analysis of Multi-temporal Remote Sensing Images (Multi-Temp): 77-80.
- Kang, S. (2003). "A regional phenology model for detecting onset of greenness in temperate mixed forests, Korea: an application of MODIS leaf area index." Remote Sensing of Environment **86**(2): 232-242.
- Manakos, I., I. Gitas, M. Schardt, K. C., E. Dragozi and H. Gallaun (2012). "Evaluating the Potential of the PROBA-V Sensor in Estimating Forest Cover Change Over a Range of European Biogeographical Regions: The FM@PROBA-V Project." Advances in Geosciences.
- Matsushita, B., W. Yang, J. Chen, Y. Onda and G. Qiu (2007). "Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest." Sensors (Basel, Switzerland) **7**(11): 2636-2651.
- Meroni, M., D. Fasbender, R. Balaghi, M. Dali, M. Haffani, I. Haythem, J. Hooker, M. Lahlou, R. Lopez-Lozano, H. Mahyou, M. Ben Moussa, N. Sghaier, T. Wafa and O. Leo (2016). "Evaluating NDVI Data Continuity Between SPOT-VEGETATION and

- PROBA-V Missions for Operational Yield Forecasting in North African Countries." Ieee Transactions on Geoscience and Remote Sensing **54**(2): 795-804.
- R. Reichle, G. D. L., R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu. (2012). "SMAP Level 4 Surface and Root Zone Soil Moisture (L4\_SM) Data Product Algorithm Theoretical Basis Document."
  - Raghavendra, B. R. and M. A. M. Aslam (2017). "Sensitivity of vegetation indices of MODIS data for the monitoring of rice crops in Raichur district, Karnataka, India." Egyptian Journal of Remote Sensing and Space Sciences **20**(2): 187-195.
  - Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka and H. Ohno (2005). "A crop phenology detection method using time-series MODIS data." Remote Sensing of Environment **96**(3-4): 366-374.
  - selwar, A. y. (2014). Technical Efficiency and Productivity Measurement Using Data Envelopment Analysis Approach in Sudanese Agricultural Sector. Masters, Sudan University of Science and Technology.
  - Simpson, J. J., T. McIntire, Z. Jin and J. R. Stitt (2000). "Improved Cloud Top Height Retrieval under Arbitrary Viewing and Illumination Conditions Using AVHRR Data." Remote Sensing of Environment **72**(1): 95-110.
  - Strabala, K. "MODIS Cloud Mask User's Guide."
  - T. J. Farrar, S. E. N., \* and A. R. Lare\* (1994). "The Influence of Soil Type on the Relationships between NDVI, Rainfall, and Soil Moisture in Semiarid Botswana. II. NDVI Response to Soil Moisture." Remote Sensing Environment **50**: 121-133.
  - Toan, T. L., F. Ribbes, W. Li-Fang, N. Floury, D. Kung-Hau, K. Jin Au, M. Fujita and T. Kurosu (1997). "Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results." IEEE Transactions on Geoscience and Remote Sensing **35**(1): 41-56.
  - Vermote, E., Wolfe, R. (2015). MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V006 U. E. R. O. a. S. E. C. NASA EOSDIS Land Processes DAAC. Sioux Falls, South Dakota.
  - Vermote, E., Wolfe, R. (2015). MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V006 U. E. R. O. a. S. E. C. NASA EOSDIS Land Processes DAAC. Sioux Falls, South Dakota.
  - Wang, X., H. Xie, H. Guan and X. Zhou (2007). "Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions." Journal of Hydrology **340**(1): 12-24.

- Wardlow, B., S. Egbert and J. Kastens (2007). "Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains." Remote Sensing of Environment **108**(3): 290-310.
- Wolters, E. D., W. Iordache, M.D. Swinnen. E. (2018). "Proba-V user manual."
- Zhang, D. and G. Zhou (2016). "Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review." Sensors (Basel) **16**(8).
- Zhao, H., Z. Yang, L. Di and Z. Pei (2012). Evaluation of Temporal Resolution Effect in Remote Sensing Based Crop Phenology Detection Studies. Computer and Computing Technologies in Agriculture V, Berlin, Heidelberg, Springer Berlin Heidelberg.

## Table

**Table 6: PROBA-V Cloud Mask (State Map)**

Bit Number	Parameter	Bit Combination	Interpretation
0-2	Cloud/Ice or Snow/Shadow Flag	000	Clear
		001	Shadow
		010	Undefined
		011	Cloud
		100	Ice
3	Land/Sea	0	Sea
		1	Land
4	Radiometric Quality SWIR Flag	0	Bad
		1	Good
5	Radiometric Quality NIR Flag	0	Bad
		1	Good
6	Radiometric Quality RED Flag	0	Bad
		1	Good
7	Radiometric Quality BLUE Flag	0	Bad
		1	Good
8	SWIR Coverage	0	No
		1	Yes
9	NIR Coverage	0	No
		1	Yes
10	RED Coverage	0	No
		1	Yes
11	BLUE Coverage	0	No
		1	Yes

**Table 7: MODIS (Aqua/Terra) Spectral Bands.**

<b>Band</b>	<b>Bandwidth*</b>	<b>Primary Use</b>
1	620 - 670	Land/Cloud/Aerosols Boundaries
2	841 - 876	
3	459 - 479	Land/ Cloud/ Aerosols Properties
4	545 - 565	
5	1230 - 1250	
6	1628 - 1652	
7	2105 - 2155	
8	405 - 420	Ocean Colour/ Phytoplankton/ Biogeochemistry
9	438 – 448	
10	483 – 493	
11	526 – 536	
12	546 – 556	
13	662 – 672	
14	673 – 683	
15	743 - 753	
16	862 - 877	
17	890 - 920	Atmospheric Water Vapour

18	931 - 941	
19	915 - 965	
20	3.660 - 3.840	Surface/ Cloud Temperature
21	3.929 - 3.989	
22	3.929 - 3.989	
23	4.020 - 4.080	
24	4.433 - 4.498	Atmospheric Temperature
25	4.482 - 4.549	
26	1.360 - 1.390	Cirrus Clouds/ Water Vapour
27	6.535 - 6.895	
28	7.175 - 7.475	
29	8.400 - 8.700	Cloud Properties
30	9.580 - 9.880	Ozone
31	10.780 - 11.280	Surface/ Cloud Temperature
32	11.770 - 12.270	
33	13.185 - 13.485	Cloud Top Altitude
34	13.485 - 13.785	
35	13.785 - 14.085	
36	14.085 - 14.385	



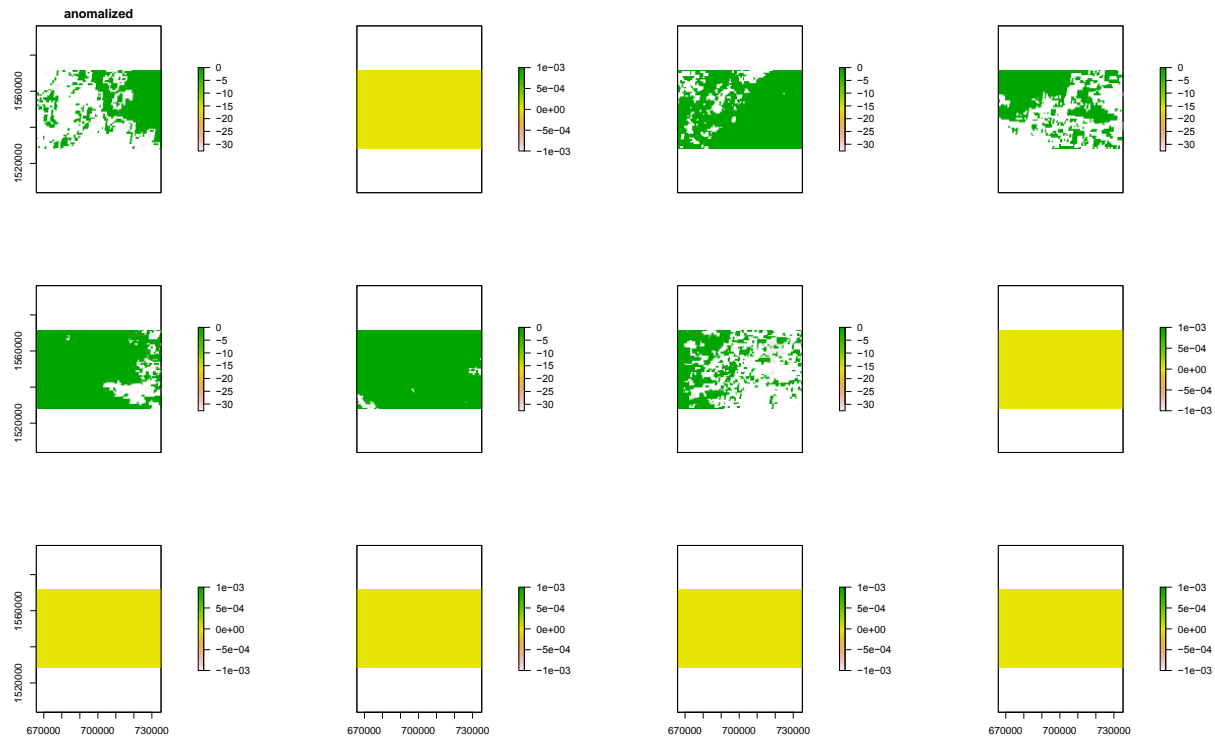
\*Bands 1 to 19 are in nm; Bands 20 to 36 are in  $\mu\text{m}$

**Table 8: MODIS Quality assessment parameters for the QA SDS Layer.**

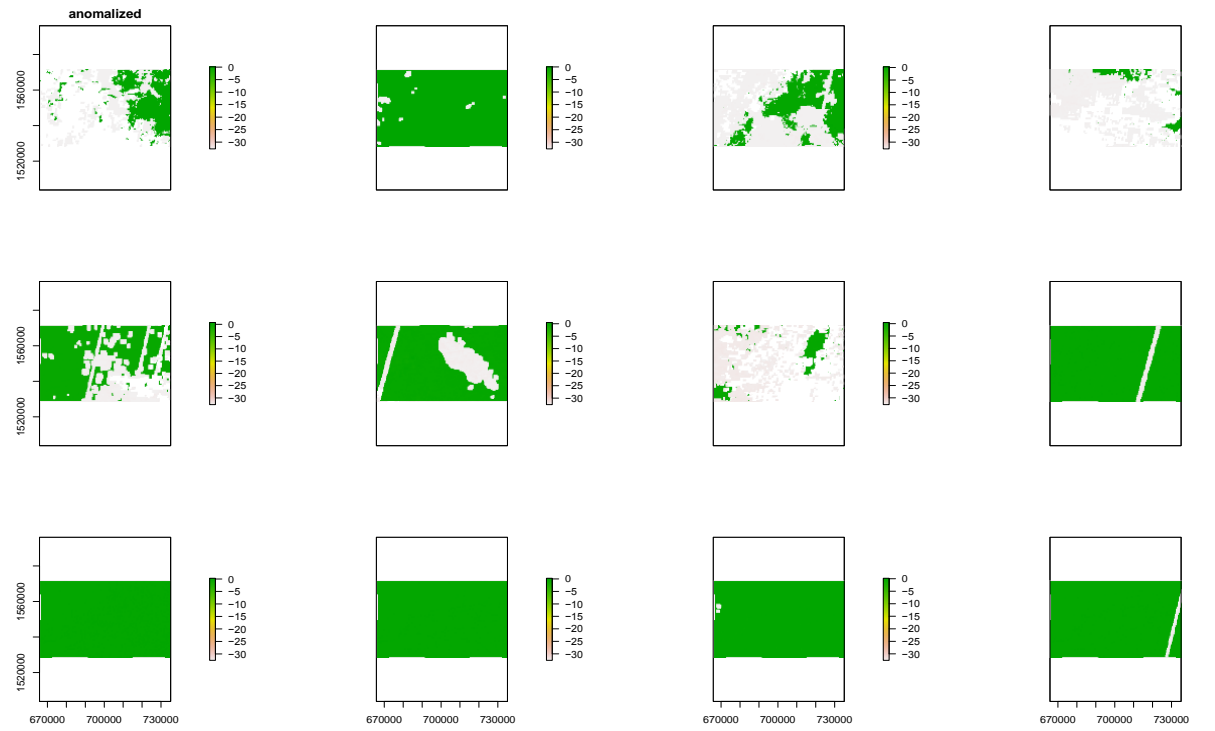
Bit No.	Parameter Name	Bit Combination	Interpretation
0-1	Cloud State	00	Clear
		01	Cloud
		10	Mixed
		11	Not set, assumed clear
2	Cloud Shadow	1	Yes
		0	No
3-5	Land/Water Flag	000	Shallow Ocean
		001	Land
		010	Ocean coastlines and lake shorelines
		011	Shallow inland water
		100	Ephemeral water
		101	Deep inland water
		110	Continental/moderate ocean
		111	Deep ocean
6-7	Aerosol Quantity	00	Climatology
		01	Low
		10	Average
		11	High
8-9	Cirrus Detected	00	None
		01	Small
		10	Average
		11	High
10	Internal cloud algorithm flag	1	Cloud
		0	No cloud
11		1	Fire

	Internal fire algorithm flag	0	No fire
12	MOD35 Snow/Ice Flag	1	Yes
		0	No
13	Pixel is adjacent to cloud	1	Yes
		0	No
14	Salt pan	1	Yes
		0	No
15	Internal Snow Mask	1	Snow
		0	No snow

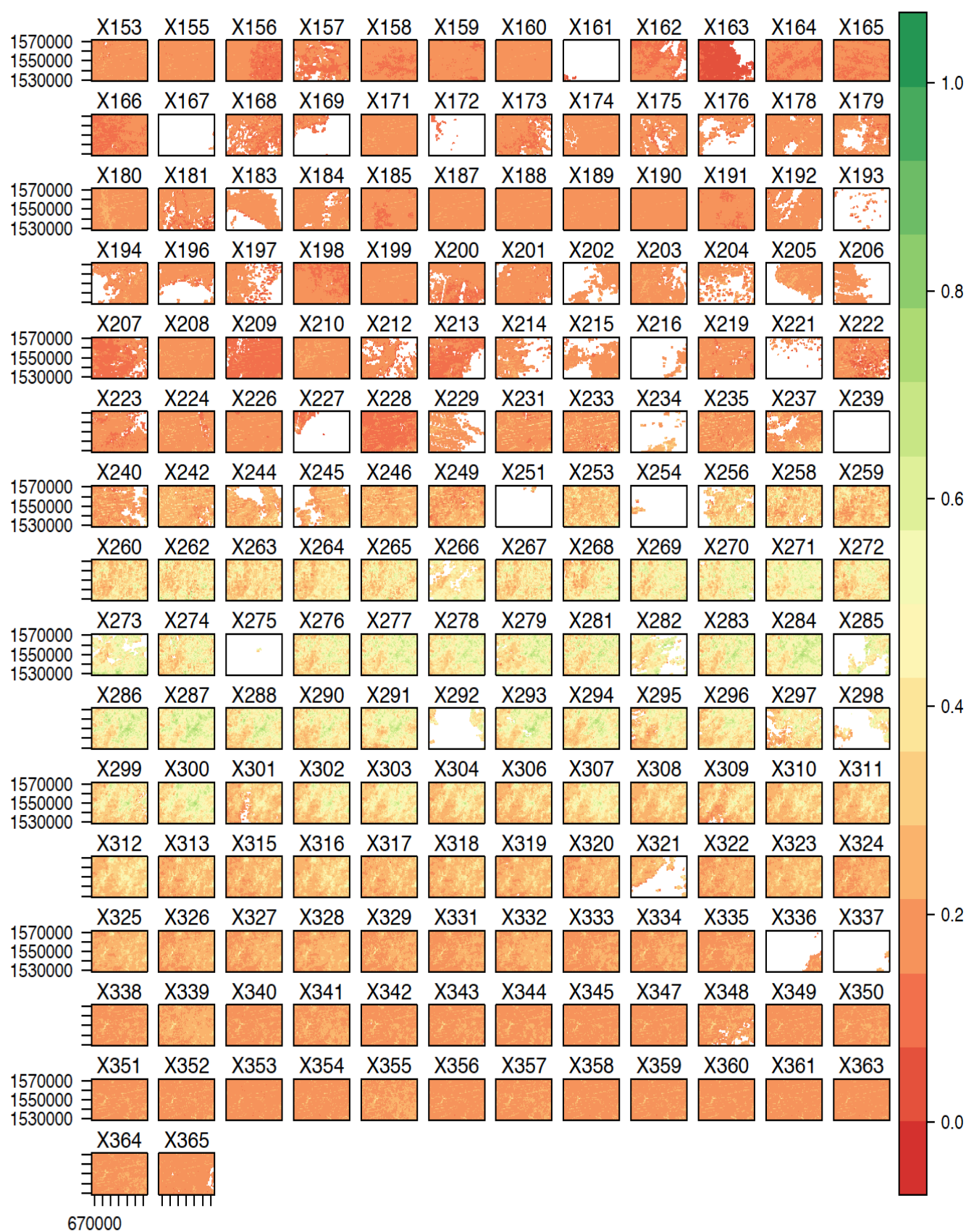
Annexure



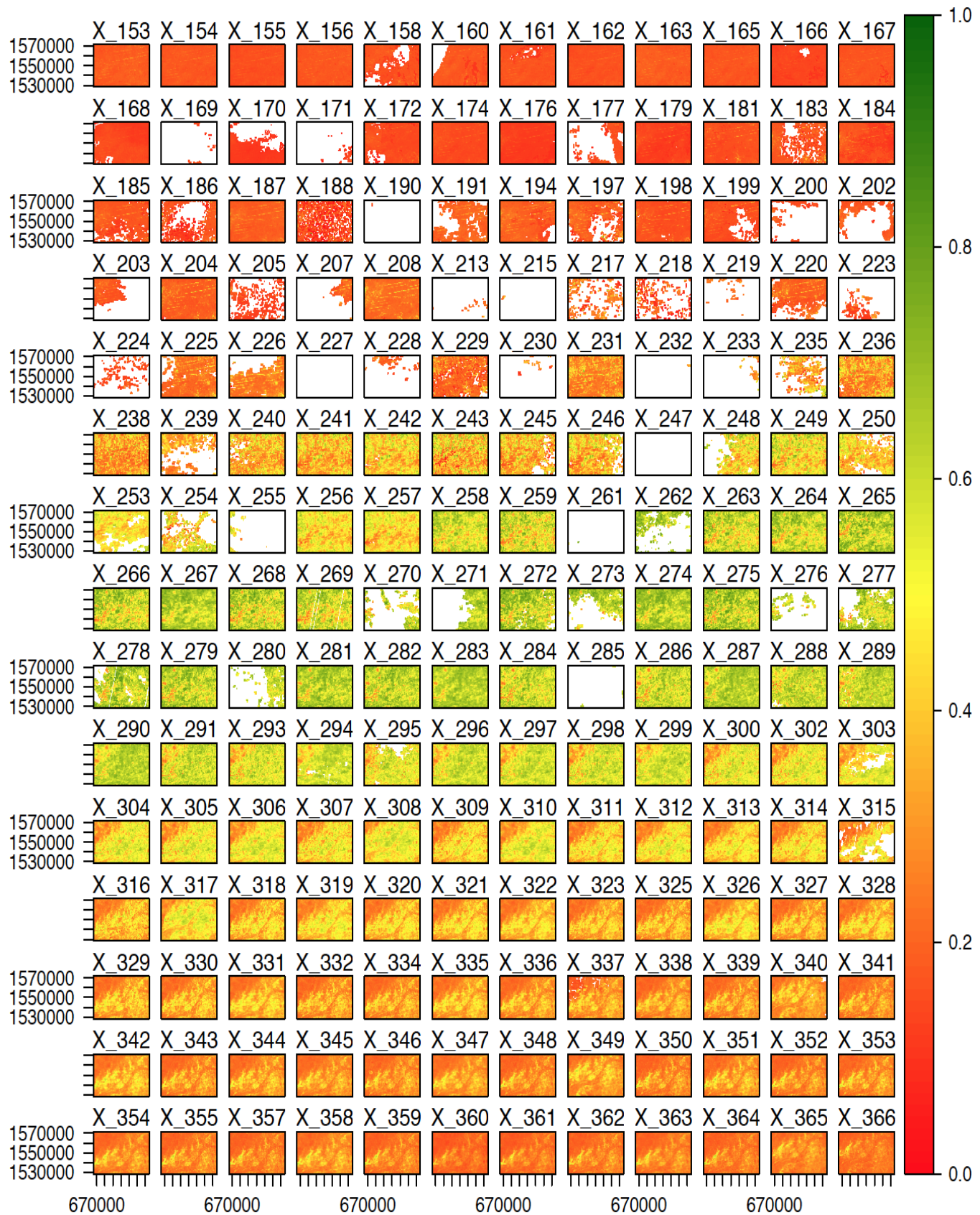
Annexure 1: Anomaly between Final Fused Image and MODIS



Annexure 2: Anomaly between Final Fused Image and PROBA-V

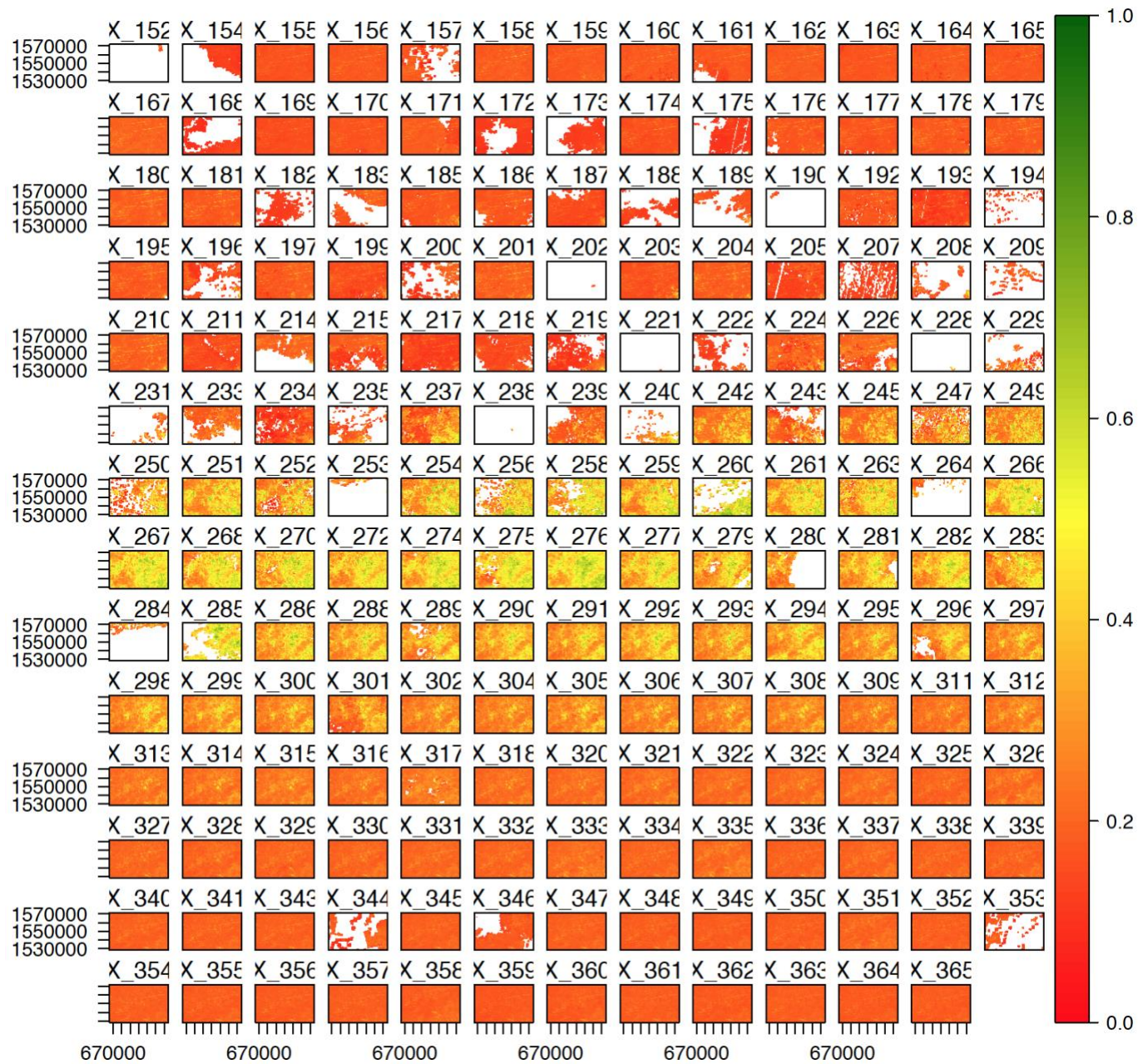


**Annexure 3 : Daily NDVI Value- 2015 from June - December**



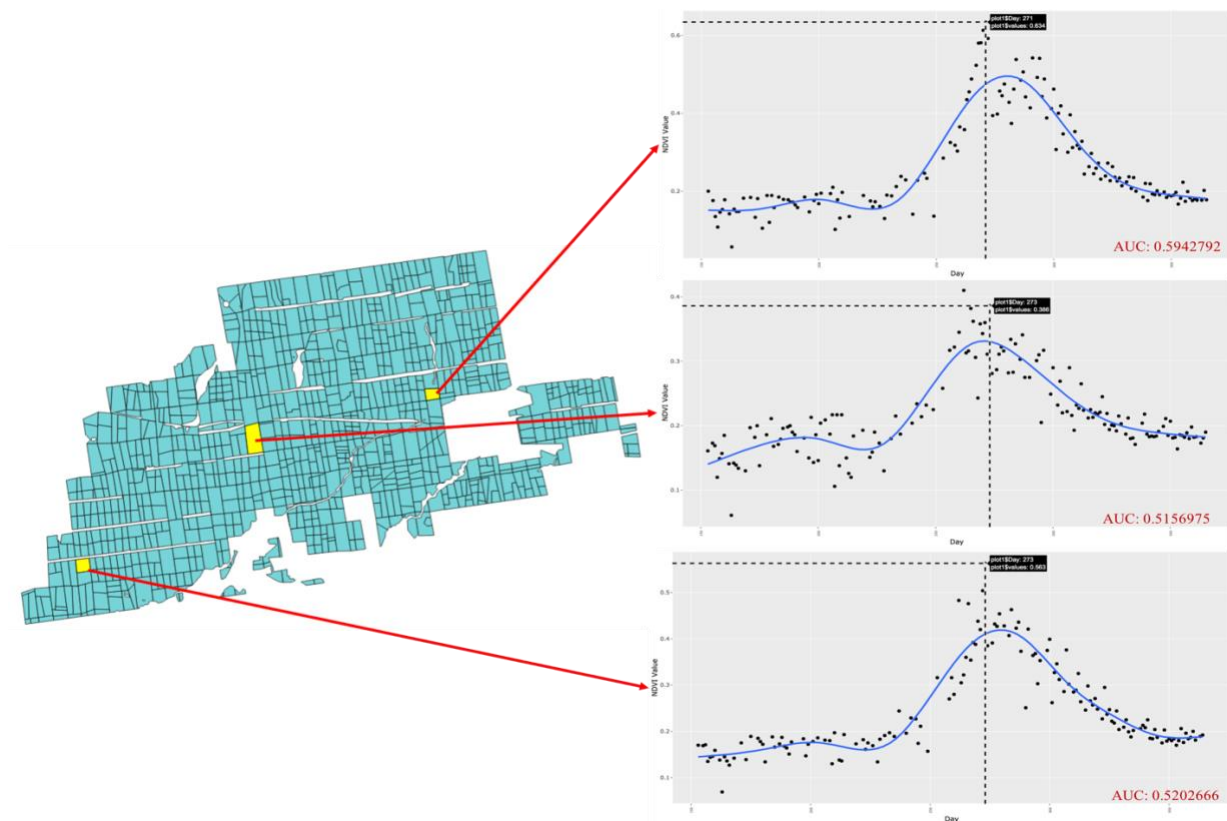
**Annexure 4: Daily NDVI Value- 2016 from June - December**



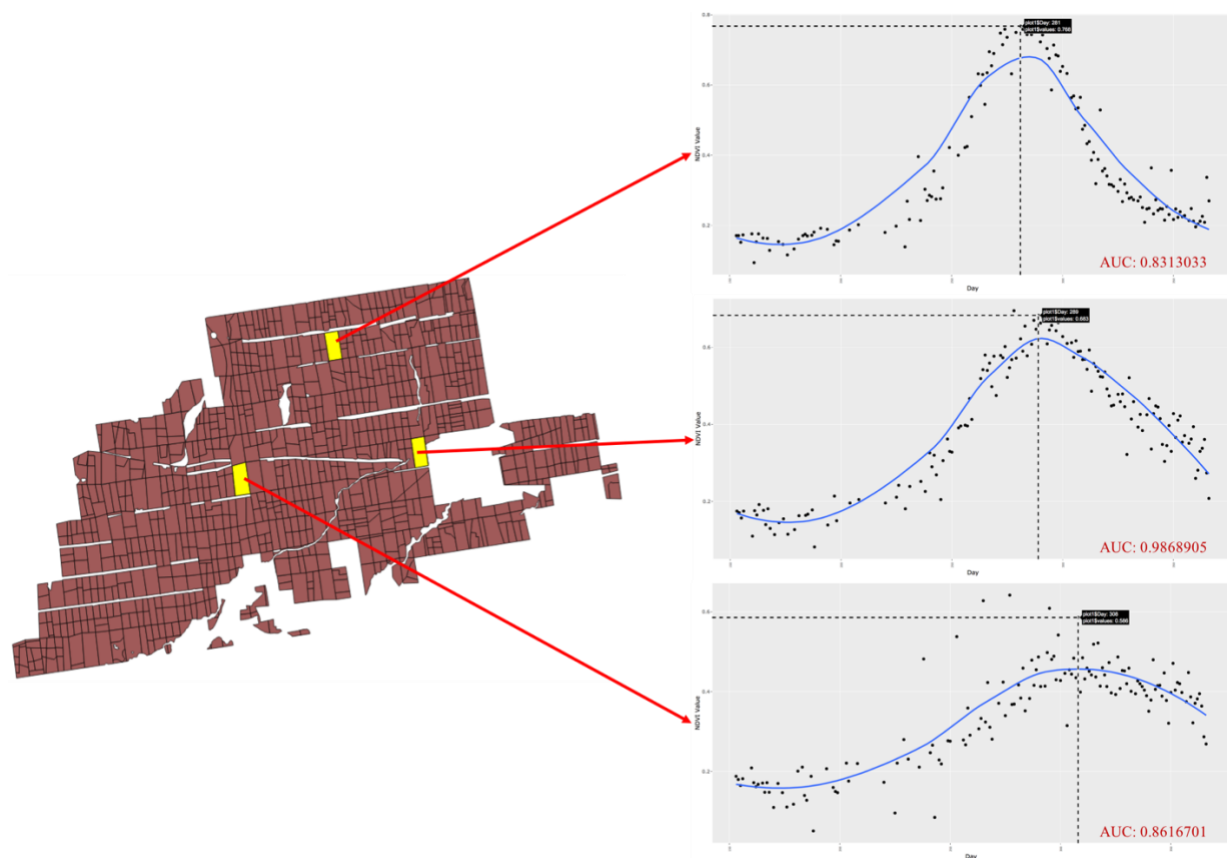


**Annexure 5: Daily NDVI Value- 2017 from June-December**

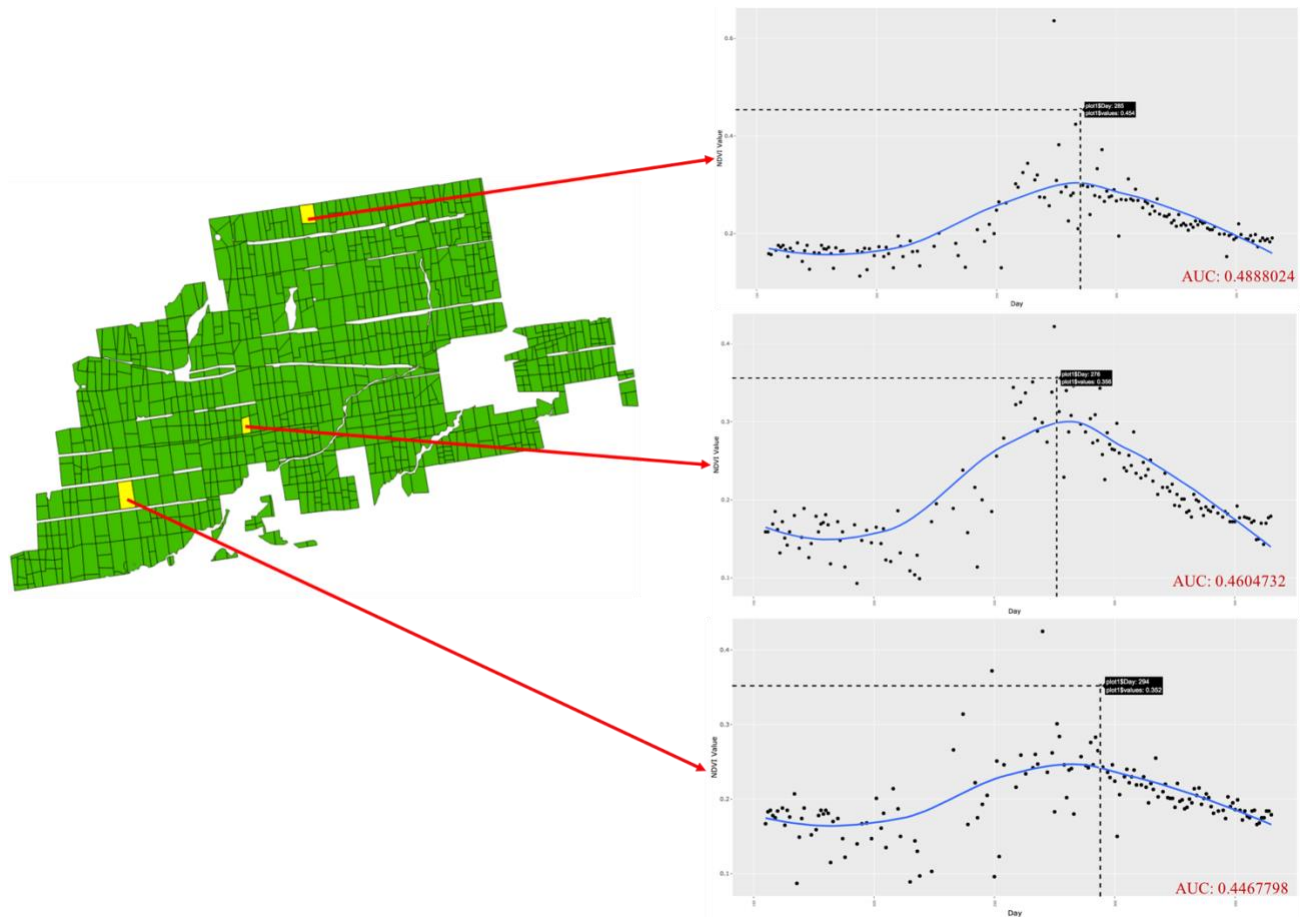




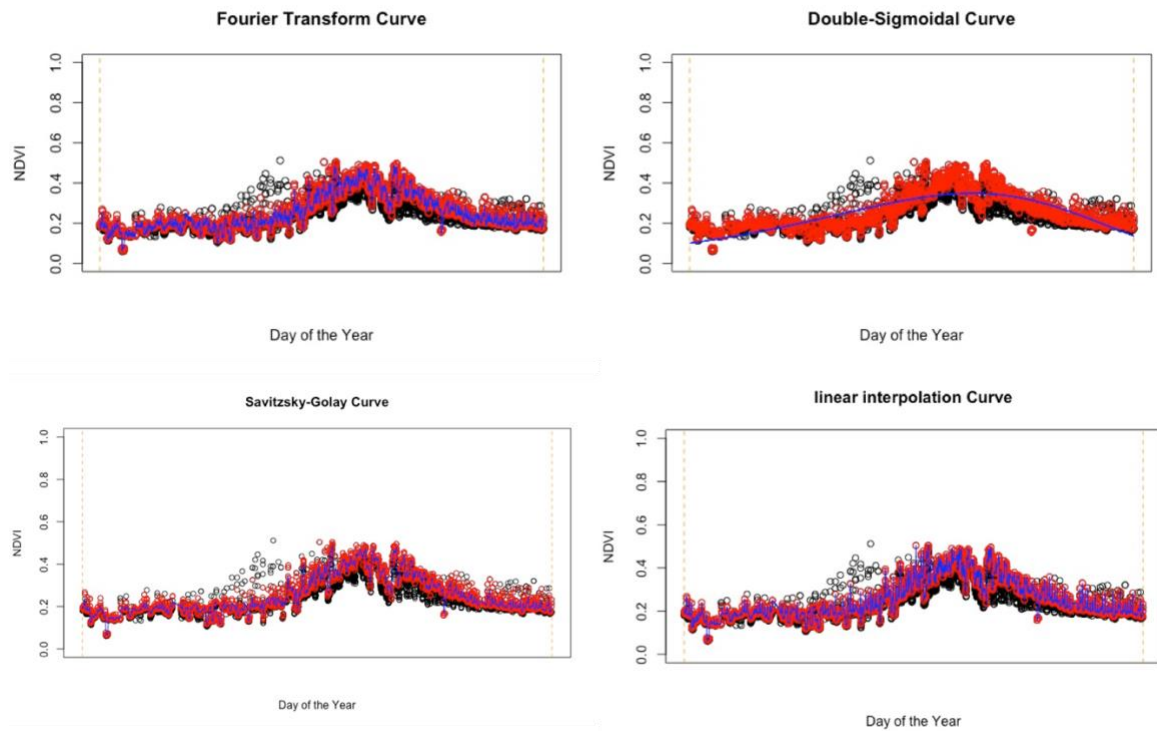
**Annexure 6: NDVI values plot basis for the year 2015**



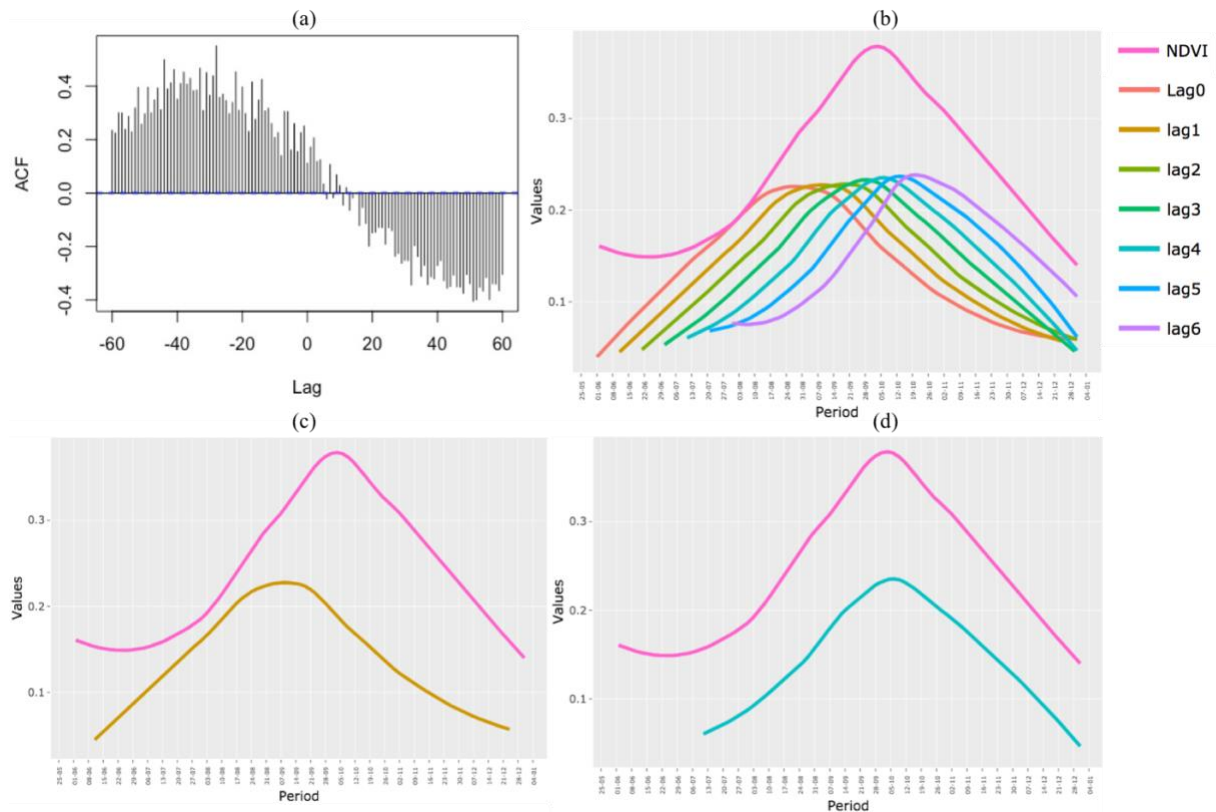
**Annexure 7: NDVI values plot basis for the year 2016.**



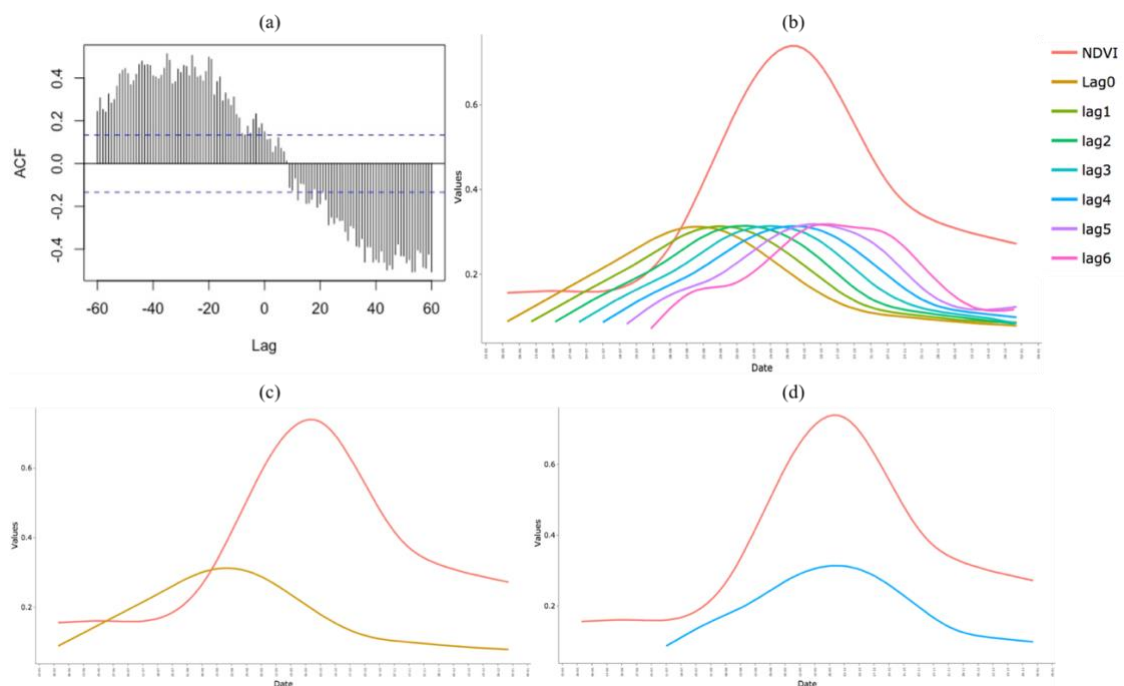
**Annexure 8: NDVI values plot basis for the year 2017**



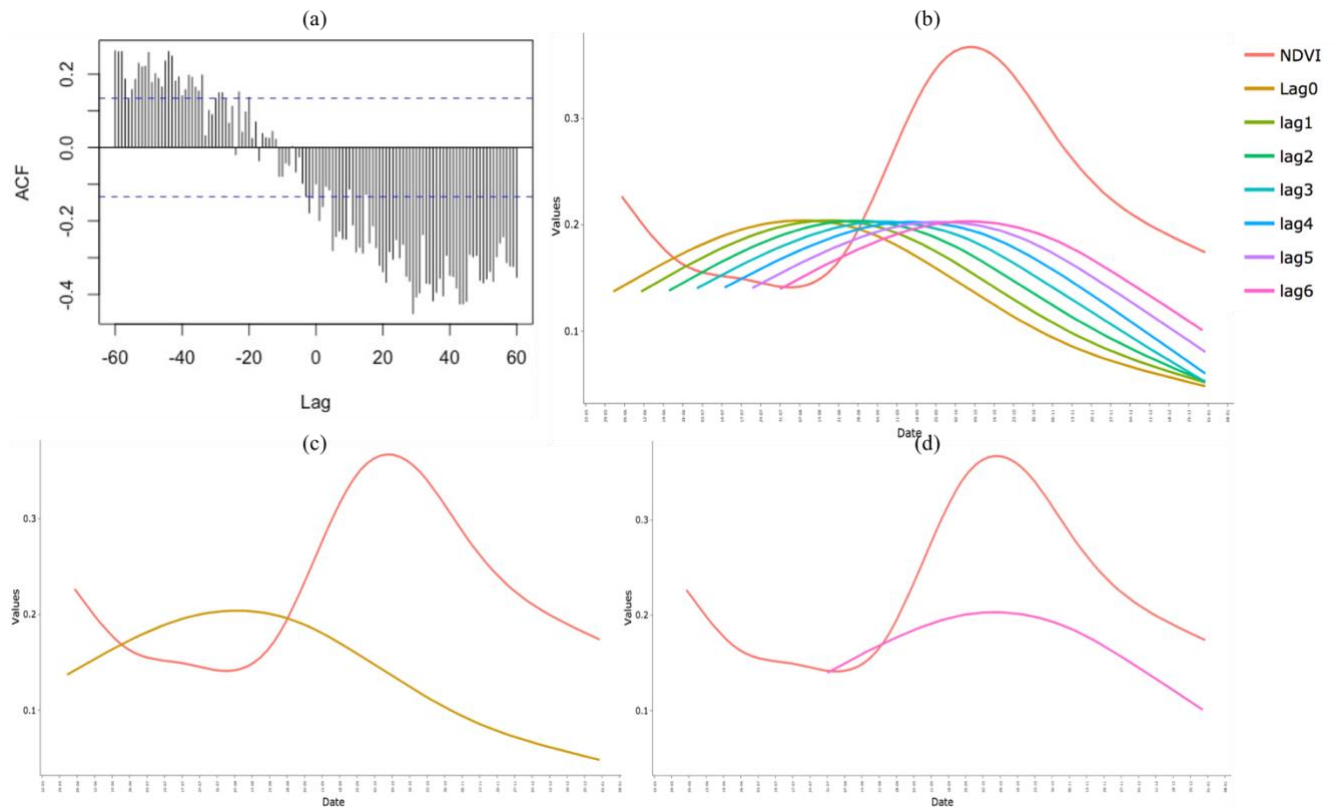
**Annexure 9: Different Curve Fitting Methods for NDVI**



**Annexure 10: Soil Moisture Lags compared to NDVI value of Plot-146 for 2015. (a) CCF Graph, (b) Lag0-Lag6 compared to NDVI, (c) Lag0 in relation to NDVI, (d) Lag4 in relation to NDVI**



**Annexure 11: Soil Moisture Lags compared to NDVI value of Plot-146 for 2016. (a) CCF Graph, (b) Lag0-Lag6 compared to NDVI, (c) Lag0 in relation to NDVI, (d) Lag4 in relation to NDVI**



**Annexure 12: Soil Moisture Lags compared to NDVI value of Plot-146 for 2017. (a) CCF Graph, (b) Lag0-Lag6 compared to NDVI, (c) Lag0 in relation to NDVI, (d) Lag6 in relation to NDVI**