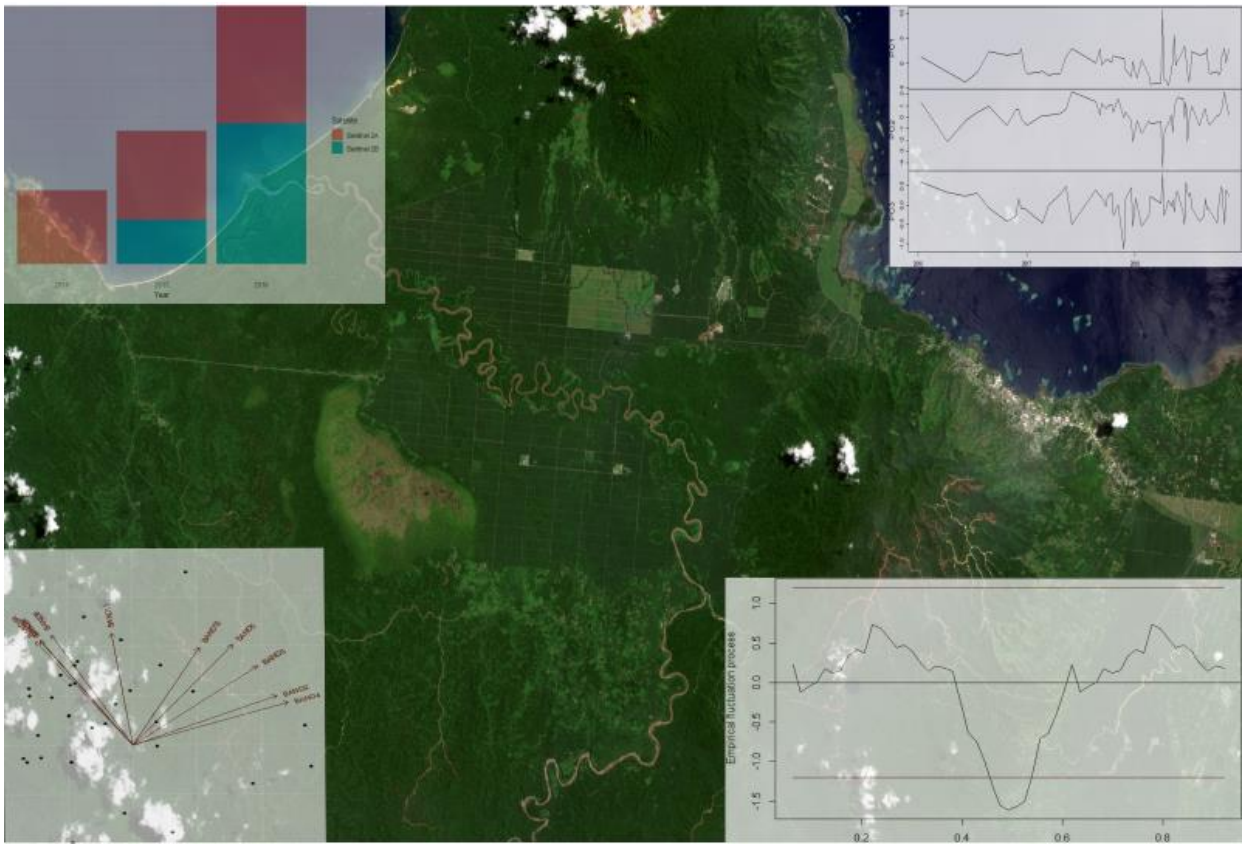
Deforestation detection using Sentinel-2 imagery.

Study case: West New Britain, Papua New Guinea.



Student details:

Name: Joel Koupermann
Student number: 5830605
Program: Master Earth Surface and Water, Utrecht University
E-mail: j.i.koupermannkuperman@students.uu.nl

Research details:

1st supervisor: Dr. Derek Karssenberg
2nd supervisor: Dr. Meng Lu
Credits: 37.5

ABSTRACT

Deforestation arises deep concerns about several issues related to the planet's future like biodiversity loss, erosion increase, land degradation, carbon emissions, temperature fluctuations, and changes in hydrological cycle, global warming, climate change, social impacts, among others. Remote sensing methods and techniques play a key rol on monitoring activities and assessing decision makers. Several methods have been developed over the last decades, principally based on vegetation indices analysis, which only uses a portion of the spectrum, while other methods less used have tried to gain information by dimension reduction techniques. The access to new Satellite Constellations like Sentinel-2, with a higher spatial resolution and richer spectral information than other Satellite programs as Landsat brings important opportunities to enhance the quality and accuracy to identify forest disturbances around the world. This study evaluates the performance of Sentinel-2 imagery detecting structural changes in forests based on a Empirical Fluctuation Process (e.g. OLS-MOSUM) over Normalized Difference Vegetation Index (NDVI) and Principal Component Analysis, to indicate which of both data sources is more feasibly. On the other hand, the present research evaluates if the method can perform with short time-series datasets in cloudy areas.

The results of the study depicts higher figure of merit's accuracy, user's accuracy, producer's accuracy, and overall accuracy using NDVI time series; however, though marginally lower accuracy was obtained using the PCA analysis, consider the PCA scores have 20 meter spatial resolution while the NDVI time series has 10 meters spatial resolution, the use of PCA still shows its potential.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Forests play a key role in life on Earth. The ecosystem goods and services that are produced by forests are fundamental for human well-being, ecological processes and climate regulation (Constanza et al., 1997; Krieger, 2001; Nasi et al., 2002; Chapin III, 2013). Forests can be classified by tree cover density in four big classes based on global climate domains as follows: Tropical (23.5°N – 23.5°S), Subtropical (25°N – 40°N, 25°S – 40°S), Temperate (~40°N-54°N, 40°S-54°S), and Boreal (50°N – 55°N to 65°N – 70°N) (Pan et al., 2013; Hansen et al., 2013). Tropical forests (Figure 1) gather the most bio-diverse ecosystems in the planet (Pimm & Raven, 2000; Bradshaw et al., 2009; Laurance et al., 2012), and the broader area of high-biomass (Pan et al., 2013). Hansen et al. (2013) found the largest forest loss in the Tropical domain with an estimated loss of 2100 km^2/year for the period 2000 to 2012 using Landsat data (30 m resolution). These forest loss findings were distributed among South America, Africa and Eurasia, as shown in Figure 2.

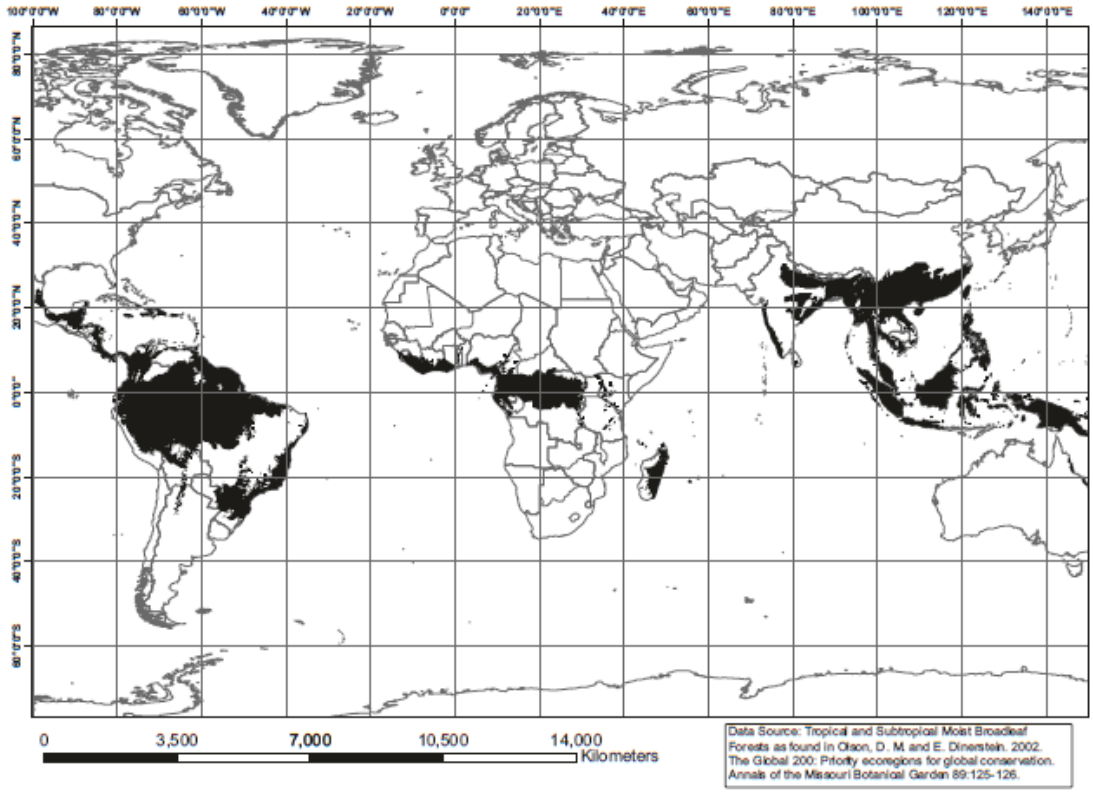


Figure 1: Tropical rainforests of the world (Credit: James M. Eaton) (McFarland, 2018)

Forest loss due to deforestation is mainly driven by economical activities such as expansion of commercial agriculture (e.g. palm oil, soybeans), logging (e.g. timber extraction), infrastructure (e.g. roads opening), and mining (e.g. gold, copper, iron ore) (Geist & Lambin, 2001; McFarland, 2018).

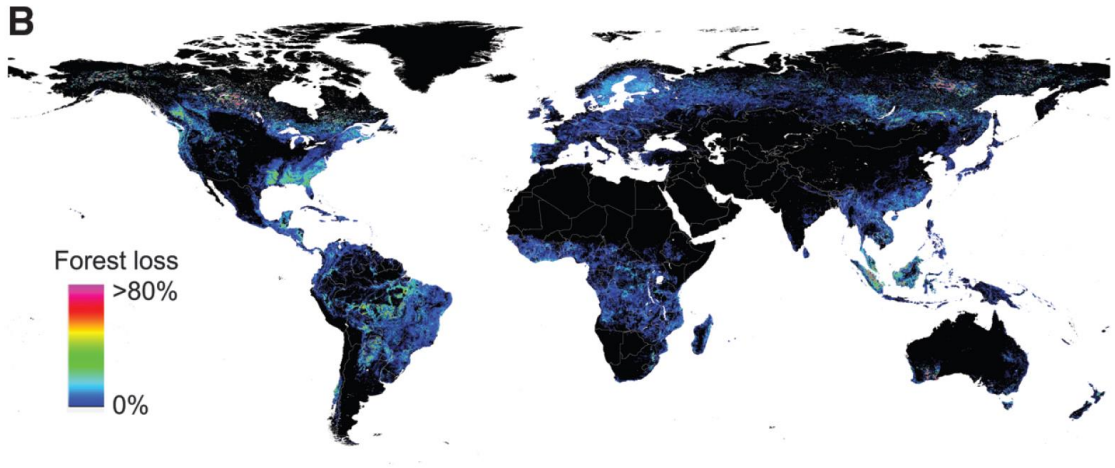


Figure 2: Global Forest Loss for the period 2000-2012. (Hansen et al., 2013)

For the present research it is important to define deforestation. Angelsen (1995) stated the following: "There is no clear definition of "deforestation," neither are there reliable estimates of its extent nor its primary causes, and - partly as a reflection of these - there is no consensus on the underlying causes" (p.1713). This statement can be proved when considering Lund (1999) research, where he surveyed 39 different definitions in which not all shared the same guidelines and parameters. Lund (1999) gives as example the three following definitions:

"**As a land cover change:** to clear an area of forests or trees, usually for commercial use of the lumber or agricultural use of the land (ITP Nelson, 1998).

As a land use change: the change of land use from forest to other land use (FAO, 1990).

As a land cover and land use change: the removal of a forest stand where the land is put to a non-forest use (Helms, 1998)" (p.128).

Later, the Food and Agriculture Organization of the United Nations (FAO) defined deforestation as: "The conversion of forest to another land use or the long-term reduction of the tree canopy cover below the minimum 10 percent threshold" (Schoene et al., 2007. p.8). Here it is important to emphasize the explanatory note of the same document that states the following: "The term specifically excludes areas where the trees have been removed as a result of harvesting or logging, and where the forest is expected to regenerate naturally or with the aid of silvicultural measures. Unless logging is followed by the clearing of the remaining logged-over forest for the introduction of alternative land uses, or the maintenance of the clearings through continued disturbance, forests commonly regenerate, although often to a different, secondary condition" (Schoene et al., 2007. p.8).

Deforestation arises deep concerns about several issues related to the planet's future (Verbesselt et al., 2012): biodiversity loss, erosion increase, land degradation, carbon emissions, temperature fluctuations, changes in hydrological cycle, global warming,

climate change, social impacts, among others (Buchanan et al., 2008; Bryan et al., 2010; Schultz et al., 2016b; Vergopolan & Fisher, 2016; Boyle, 2017; Benavent et al., 2018).

In the last years, open access to medium and high-resolution satellite data (e.g Landsat and Sentinel-2) has become an important contribution to assess deforestation detection and monitoring (Hamunyela et al., 2016; Lu et al., 2017). In addition, remote sensing products have improved considerably regarding spatial, temporal and spectral resolution. Spatio-temporal statistical modeling methods have been playing a fundamental role in natural resources management, ecosystem conservation research, and policy construction (Verbesselt et al., 2010; Lu & Hamunyela, 2016; Lu et al., 2016, 2017).

Remote sensing data have been strongly used in the last decades for deforestation monitoring (Coppin et al., 2002; Verbesselt et al., 2010b), however, the number of time series change detection methods developed is limited (Verbesselt et al., 2010b). Singh (1989) and Mas (1999) numbered a series of techniques and procedures for change detection (e.g. Univariate image differencing, Image regression, Principal Component Analysis, Vegetation index differencing, Post-classification comparison , among others). These reviewed techniques used very short time series datasets at that time, and the performance was not satisfactory for different reasons.

Furthermore new methods started to be developed, which had demonstrated to be robust for near real-time forest disturbance detection. The Breaks For Additive Season and Trend (BFAST) Monitor (Verbesselt et al., 2010) is a framework that detects changes from satellite image time series based on the Empirical Fluctuation Process (Kleiber et al., 2002; Zeileis, 2005, 2006; Zeileis & Hornik, 2007). Foregoing studies applying the BFAST monitor used as principal input indices two or three spectral bands (e.g. NDVI, NDMI, EVI), ignoring all other spectral information gathered by advanced earth observation sensors and accounting limitations because of seasonal signals, noise and other natural dynamics (Lu et al., 2017).

Lu et al. (2017) studied the BFAST monitor by dimension reduction using Principal Component Analysis (PCA) which is capable of creating a new index with reduced seasonality noise by using the complete spectral and temporal information of the satellite sensors. Thus, results have shown higher accuracy on one hand and reduced temporal detection delay of structural changes in forests on the other hand.

Moreover, several authors have stated the importance and the need for evaluating forest change detection based on Empirical Fluctuation Process with increased spatial and spectral resolution (Verbesselt et al., 2010a; Lu & Hamunyela, 2016; Hamunyela et al., 2016; Lu et al., 2016; Lu et al., 2017).

The present research will assess Forest Change Detection based on empirical Fluctuation Process using Sentinel-2 images over Normalized Difference Vegetation Index (NDVI) and dimension reduction by Principal Components Analysis (PCA) and evaluate the performance of the method using higher spatial resolution and richer spectral information imagery than previous studies. Also it will be evaluated if the

method can achieve reasonable accuracy using a short time-series dataset with high-cloud presence condition.

2. THEORETICAL BACKGROUND

2.1.Normalized Difference Vegetation Index

The Normalized Difference Vegetation Index (NDVI) is a well-known, widely studied and used vegetation index (Gandhi et al., 2015; Mandanici & Bitelli, 2016). NDVI is the resulting ratio of the difference among NIR (near-infrared band) and RED (red band) and posterior sum of the same two bands (Rouse Jr et al., 1974; Tucker 1979; Pettorelli et al., 2005; Yengoh et al., 2015; Lu et al., 2017):

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Equation 1

The values of NDVI are between -1 and 1, where higher values indicate healthy vegetation or a higher photosynthetic activity and low values unhealthy vegetation or the lack of photosynthetic activity (Pettorelli et al., 2005; Pettorelli et al., 2011, Yengoh et al., 2015). Therefore, higher disturbance in vegetation means lower productivity, hence lower NDVI value. Figure 3 depicts the hysteresis curve explained by Kinzig et al (2006), where resilient systems are able to reach the original state until certain disturbance threshold. If disturbance reaches critical points, restoration can achieve irreversible state of degradation (Kinzig et al., 2006; , Yengoh et al., 2015).

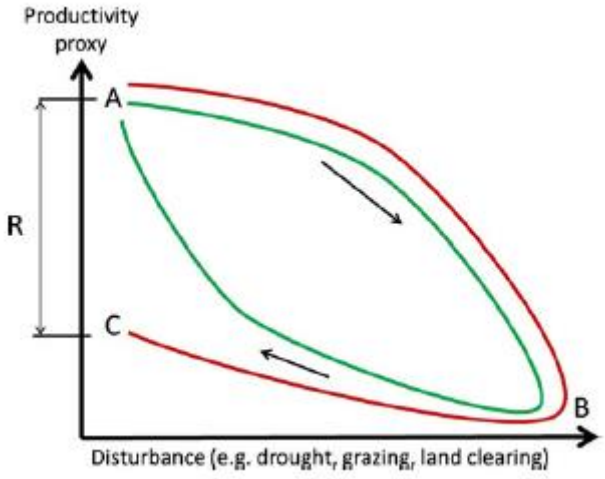


Figure 3: Stress cycle of vegetation related to productivity (principle of hysteresis). Point A, no stress, higher productivity. Point B, high stress, low productivity. Green curve shows a fully resilient system, capable of reaching its original state. A less-resilient system only recovers at a certain lower state (Point C – red curve). Resilience of a system, R, is related to the distance between A and C; the lower the value, the higher the resilience (taken from Yengoh et al, 2015)

NDVI considers the spectral properties of these bands and the reaction they have with vegetation. Green vegetation reflects more NIR and less visible light, while less green vegetation reflects more from the visible portion of the spectrum and less NIR (Yengoh et al., 2015). Therefore, higher index values represent the presence of vigorous vegetation and low values can represent degraded vegetation, bare soil and other

elements as snow and clouds, whereas water has negative values in most of the cases (e.g. suspended vegetation) (Yengoh et al., 2015), which can be observed in the reflectance curves on the RED and NIR wavelengths of the spectrum depicted in Figure 4. NDVI does not explain physical quantification of vegetation directly, though the correlation of physical properties (e.g. biomass, health condition, fractional vegetation cover, leaf area index), therefore are valuable measurements to study vegetation dynamics (Carlson & Ripley, 1997).

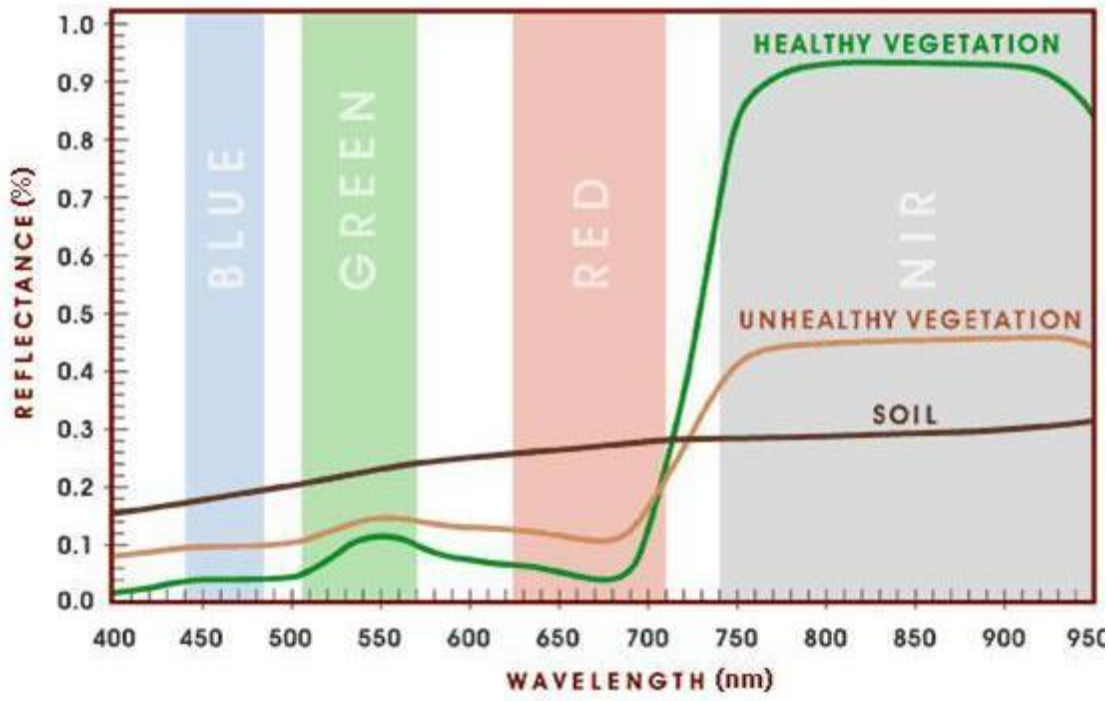


Figure 4: Spectral curve for healthy vegetation, unhealthy vegetation and soil (taken from www.gisresources.com)

2.2.Principal Component Analysis

Principal Component Analysis (PCA) is a multivariate dimension reduction data analysis technique used in a diverse group of scientific areas (Wold et al., 1987; Abdi & Williams, 2010; Jolliffe, 2011), remote sensing among them (Kwarteng & Chavez, 1989; Loughlin, 1991; Lu et al., 2017).

The aim of PCA is to extract the most important information from a group of observations depicting several variables, where abundant information results into a set of new orthogonal variables uncorrelated between them called Principal Components (PCs) (Kwarteng & Chavez, 1989; Abdi & Williams, 2010; Jolliffe, 2011). Thus, the largest amount of the total variance will be gathered in the first component, and it will be decreasingly distributed over the following components. The sum of the variance of all the components will be equal to the total variance of the original source (Kwarteng & Chanvez, 1989; Loughlin, 1991; Schowengerdt, 2006; Abdi & Williams, 2010; Van der Meer & De Jong, 2011; Jolliffe, 2011). The PCs are defined as the linear combinations of the original variables, the PC loadings or rotation are the coefficient of the linear combination and the PC scores the result of the linear combination among the

mean of the PCs and the PC loadings (Kwarteng & Chavez, 1989; Loughlin, 1991; Abdi & Williams, 2010).

2.3.Empirical Fluctuation Process

Empirical Fluctuation Process (EFP) is a method of structural change detection on linear regression models (Zeileis, 2005). This method has been used in several scientific disciplines, from which econometrics and statistics had paid much attention (Chu et al., 2005; Zeileis, 2005). In the last decade EFP has gained importance on environmental studies which focus their attention on forest disturbances detection (Verbesselt et al., 2010; Lu & Hamunyela, 2016; Lu et al., 2016; 2017). Fluctuation tests based on Ordinary Least Squares (OLS) have been found to be valuable as explorative tool and it is determined without a distinct pattern of deviation from stability parameter. It is preferred by their users because the feasibility at the time of computing and interpreting the results. The most common methods of EFP are OLS-based by Cumulative Sum (CUSUM) and Moving Sum (MOSUM) (Zeileis, 2005).

Former Structural Change Tests and the BFAST Monitor studies have proved that the OLS-MOSUM could detect structural change more accurately compared to OLS-CUSUM (Chu et al., 2005; Zeileis, 2005; Lu & Hamunyela, 2016; Lu et al., 2016; 2017). Foregoing, OLS-MOSUM test is more sensitive to parameters changes when the magnitude of residuals is larger, as residuals move in a window over the times series without considering a cumulative sum of previous observations (Chu et al., 2005; Lu et al., 2017).

OLS-MOSUM process is defined in Zeileis et al (2001) by

$$M_n^0(t|h) = \frac{1}{\hat{\sigma}\sqrt{n}} \left(\sum_{i=\lfloor N_n t \rfloor + 1}^{\lfloor N_n t \rfloor + \lfloor nh \rfloor} \hat{u}_i \right) \qquad (0 \le t \le 1-h)$$

Equation 2

$$= W_n^0 \left(\frac{\lfloor N_n t \rfloor + \lfloor nh \rfloor}{n} \right) - W_n^0 \left(\frac{\lfloor N_n t \rfloor}{n} \right),$$

Equation 3

where *h* is the data window's bandwidth (default value *h*=0.15), and N_n = (n- [nh])/(1 – h).

A structural change is identified and detected when the observations are deviated from the hypothesis of a Brownian Bridge – which is used in OLS-MOSUM processes, compared with Brownian motion, which is used in Recursive MOSUM- processes (Zeileis et al., 2001; Lu et al., 2017).

The null hypothesis of the structural change test is false (change detected) when the critical values of the OLS-MOSUM test intersect one of the two-sided boundary-crossing probabilities of the Brownian Bridge's hypothesis. The resulting calculation is

compared to the significance level (alpha), 0.05 is commonly used (Chu et al., 1995; Mangiafico, 2016).

2.4.Research Questions

RQ: Are Sentinel-2 satellite time-series capable of detecting deforestation by Empirical Fluctuation Process?

SQ1: To what extent does Structural Change Detection on Sentinel-2 imagery based on NDVI and PCA differ?

SQ2 : How does short amount of time-series datasets affect Structural Change Detection based on Empirical Fluctuation Process?

SQ3: It is possible to detect deforestation in areas with high amount of clouds?

3. METHODOLOGY

3.1. Study area

West New Britain province in Papua New Guinea was selected as study area. Historically this territory has been seriously affected by deforestation due to timber extraction and large scale commercial agriculture - principally coconut and palm oil plantations - (Buchanan et al., 2008).

The study site is located at the North of West New Britain (centered at: 5° 32' 37" S, 149° 59' 42" E, 625 km^2) (Figure 5). The Area of Interest (AOI) was delimited considering the study area and the ocean shore boundary in order to have only inland pixels during the analysis (Figure 6). The AOI has an area of 1.000 km^2.

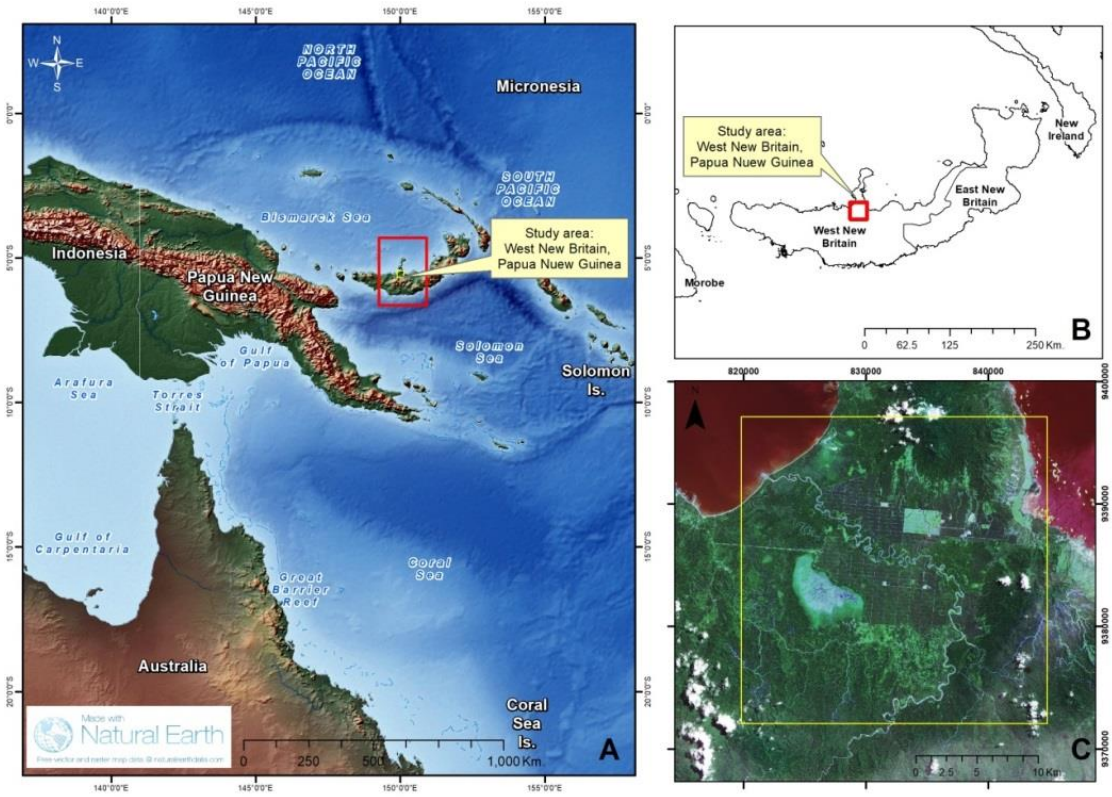


Figure 5: Map of the study area. (A) Location of study site in reference to Papua New Guinea country and its surroundings; (B) location of study site in reference to West New Britain province; (C) Study site (yellow square) using Sentinel-2 (MSI) image bands.

Figure 6: Sentinel-2A True Color Composite Image showing the Area of Interest, West New Britain, Papua New Guinea

3.2. Deforestation detection process

The deforestation detection process based on EFP in this study was constituted by five main stages. Firstly, the validation dataset creation from where the per-pixel analysis was implemented and later the validation was executed. Secondly the pre-processing of all Sentinel-2 imagery was carried out in order to prepare the data for the implementation of the method (e.g. cloud masking, cropping, resampling). Thirdly, data structuralizing and implementation of the method was done using ArcGIS desktop and R software (e.g. NDVI calculation, PCA calculation, EFP and Structural Change Detection test). Fourthly, the validation and accuracy assessment was performed, and finally, the results analysis and interpretation was carried out. The latter process is depicted in the method flow diagram in Figure 7.

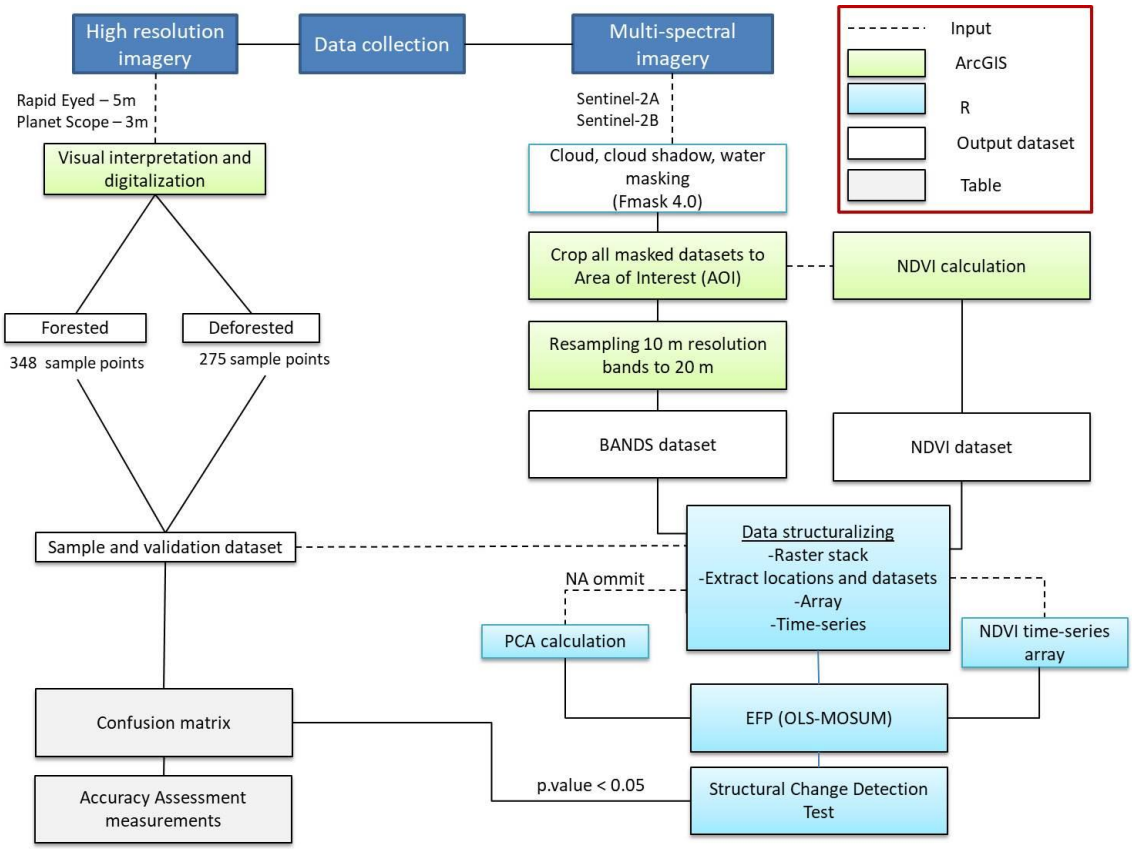


Figure 7: Deforestation Detection method flow diagram

3.3. Satellite Imagery

In our study, data from four satellites were used: Planet Scope, Rapid Eye, Sentinel-2A and Sentinel-2B.

For creating the sample and validation dataset, high resolution imagery was used from Planet Scope and Rapid Eye satellites. These satellites have a spatial resolution of 3m and 5m respectively.

For the deforestation detection analysis, Sentinel-2A and Sentinel 2-B were used. Ten of the thirteen spectral bands from Sentinel-2 datasets were used for PCA during this study. The discarded bands were the three 60m spatial resolution bands that correspond to coastal aerosols, water vapor and SWIR-Cirrus (See Table 1).

3.3.1. Sentinel – 2 Constellation

The Sentinel-2 mission of the European Space Agency (ESA) has been developed in the context of the European Union Copernicus programme (ESA, 2015; Wang et al., 2016; Wang & Atkinson, 2018). It contains two twin satellites (i.e. Sentinel-2 A and Sentinel-2 B), which are in the same orbit staged at 180° to each other, and the complete constellation has a temporal resolution of 5 days near the Equator. Sentinel-2 has an orbital swath width of 290 km and a multi-spectral imaging (MSI) sensor that covers 13

spectral bands in the visible (VIS), near infrared (NIR) and short wave infrared (SWIR) wavelengths. It has four bands at 10 meters (Figure 8), six bands at 20 meters (Figure 9) and three bands at 60 meters spatial resolution (Table 1) (Drusch et al., 2012; ESA, 2015; Wang et al., 2016; Wang & Atkinson, 2018). The Sentinel-2 mission has a global coverage and its principal objectives are to give continuance to SPOT and LANDSAT imagery time series and to contribute with observational data for land-cover, land-change detection and geophysical variables (Malenovský et al., 2012; Drusch et al., 2012; Sibanda et al., 2015; ESA, 2015) (See Figure 35A in the appendix, page 60).

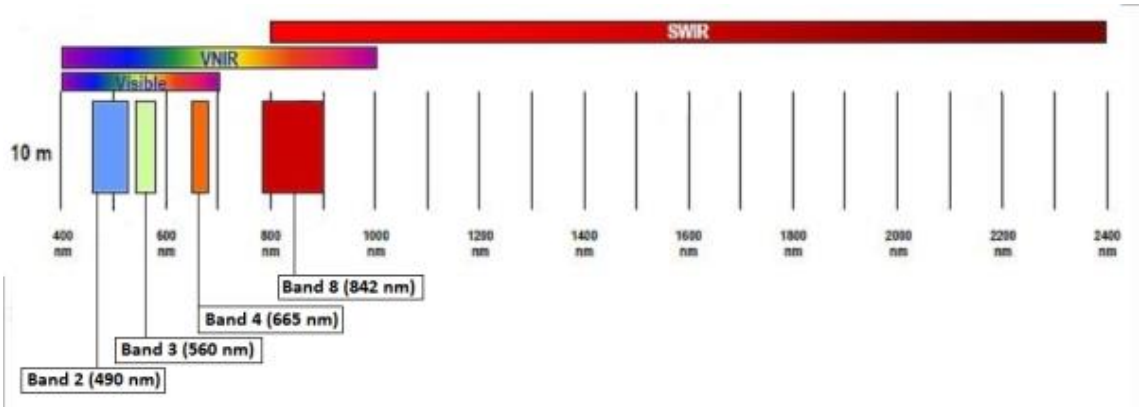


Figure 8: Sentinel-2 10 meters spatial resolution bands, and their location on the spectrum: Blue (Band 2), Green (Band 3), Red (Band 4) and NIR (Band 8) (taken from ESA, 2015; Mushoni et al., 2018)

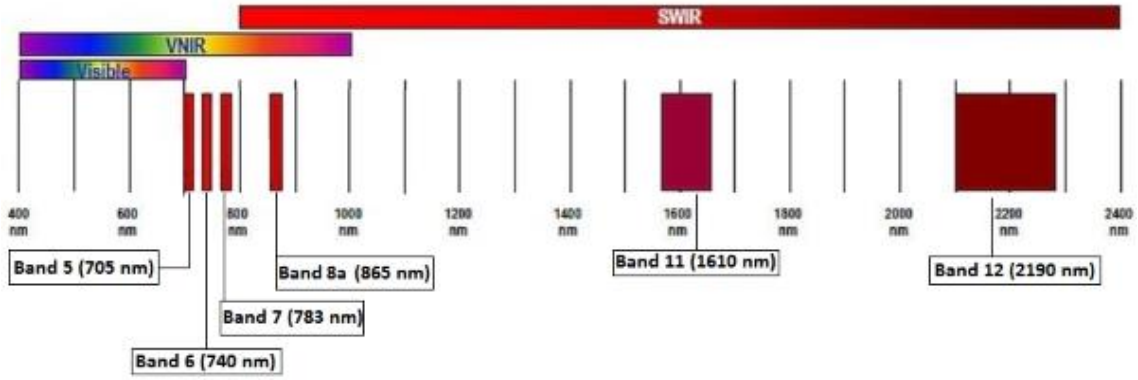


Figure 9: Sentinel-2 20 meters spatial resolution bands, and their location on the spectrum: Red Edge (Band 5), Red Edge (Band 6), Red Edge (Band 7), Narrow NIR (Band 8a), SWIR (Band 11) and SWIR (Band 12) (taken from ESA, 2015; Mushoni et al., 2018)

Band number	Band	Central wavelength (nm)	Bandwidth (nm)	Spatial resolution (m)
1	Coastal aerosol	443	20	60
2	Blue	490	65	10
3	Green	560	35	10
4	Red	665	30	10
5	Red Edge	705	15	20
6	Red Edge	740	15	20
7	Red Edge	783	20	20
8	NIR	842	115	10
8[a]	Narrow NIR	865	20	20
9	Water Vapour	945	20	60
10	SWIR - Cirrus	1375	30	60
11	SWIR	1610	90	20
12	SWIR	2190	180	20

Table 1: Sentinel-2 spectral and spatial resolution specifications (ESA, 2015)

For the present research Sentinel-2 mission Level-1C products were used. Level-1C products are ortho-images geo-referenced in UTM/WGS84 projection, containing Top of Atmosphere (TOA) reflectances, and radiometric and geometric corrections (ESA, 2015). Each Level-1C dataset constitutes 100 km^2 tiles (granules) which contain all available spectral bands. Each dataset is codified by the format Naming Convention for Sentinel-2 Level-1C products as shown in Figure 10 (ESA, 2015).

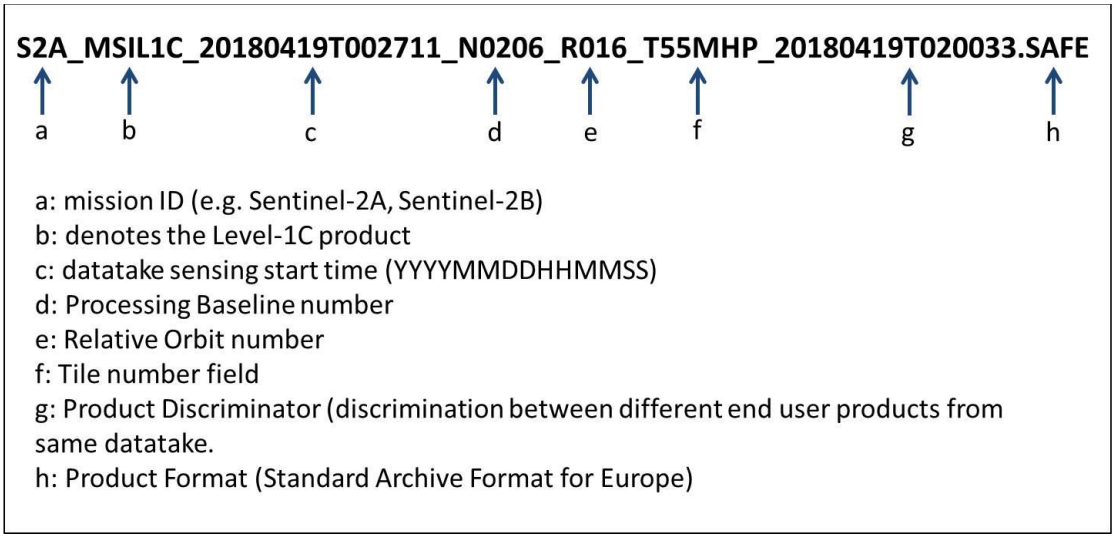


Figure 10: Format Naming Convention for Sentinel2 Level-1C products (ESA, 2015)

The Tile T55MHP covers the whole study area as shown in Figure 11. This facilitated the acquisition of all available 15% or higher cloud-free imagery datasets for the AOI, which are summarized in Figure 12.

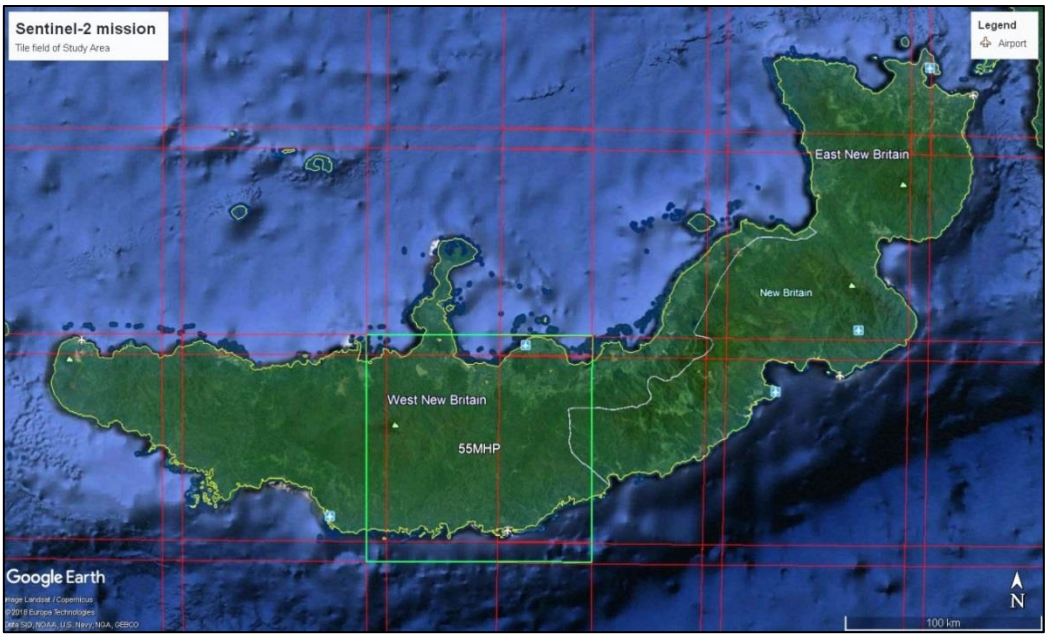


Figure 11: Tilling grid for the study area in reference with West New Britain Island (Google Earth, 2019; ESA, Sentinel Online)

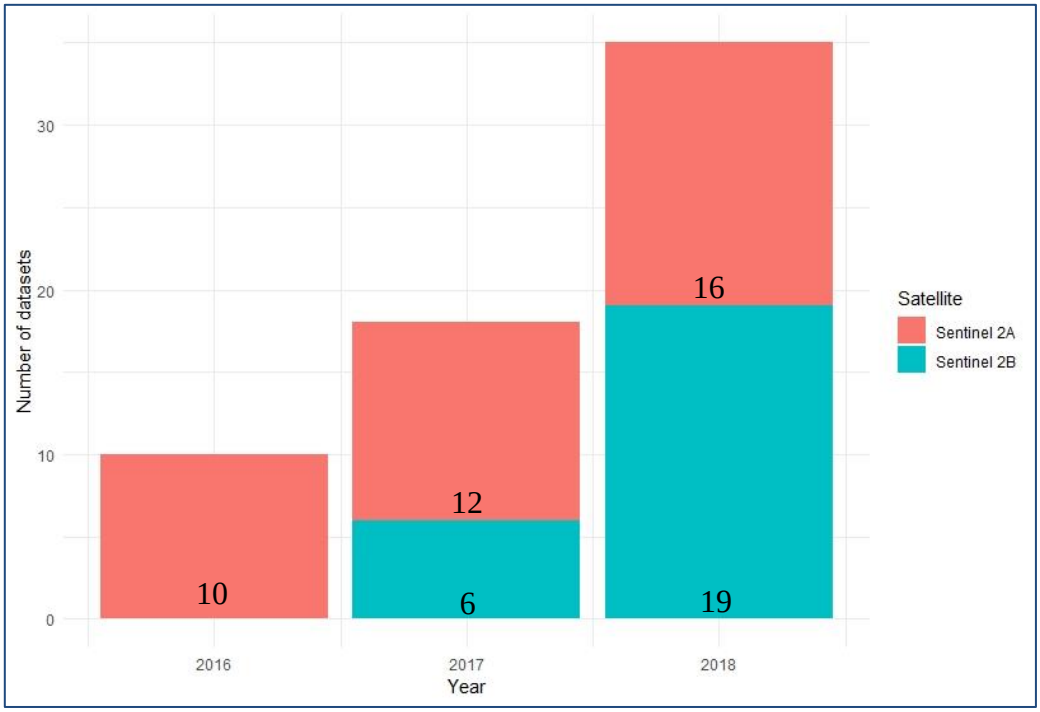


Figure 12: Number of Sentinel-2 time-series datasets per satellite.

Considering that the study area is near the Equator, a high cloud presence during the year means a limitation for remote sensing analysis. This is related to Research Question 3. A total of 63 datasets were found appropriate to be incorporated into the database to run the analysis, acquired from Sentinel Open Access Hub

(https://scihub.copernicus.eu/dhus/#/home). For the year 2016, 10 datasets had the required amount of cloud-free percentage, resulting in the 59% of all available datasets for this year. For year 2017, 18 datasets which represent the 60% of available imagery was used, and for year 2018, 35 datasets (56%) fulfilled the requisites for the analysis.

3.4. Validation dataset: RapidEye and PlanetScope constellations

For creating the validation dataset, high resolution satellite imagery was used. To access these datasets, a student account was created on the Planet's Education and Research Program (Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. (https://api.planet.com). This program gives researchers access to satellite imagery time-series of RapidEye and PlanetScope constellations datasets.

RapidEye Satellite Constellation is constituted by five satellites with a Multispectral push broom sensor with revisit time of 5.5 days at nadir at an altitude of 630 km in the Sun-Synhronous Orbit. Each satellite measures less than a cubic meter and weights 150 kg. The product used was the 5-band multispectral RapidEye Analytic Ortho Tile. Each tile scene has a size of approximately 25 km by 25 km. These datasets are orthorectified and geometrically and radiometrically corrected. The positional accuracy is less than 10m RMSE and the coordinate system is WGS-84. It has a spatial resolution of 5m, and spectral bands are shown in Table 2 (Planet Labs Inc., 2018).

PlanetScope Satellite Constellation is constituted of close to 120 satellites. Each satellite is a CubeSat 3U with a daily revisit time at an altitude of 475 km. The sensor used in these satellites is a four-bands frame Imager with a split-frame NIR filter. The product used in this case is also the PlanetScope Analytic Ortho Tile. Each tile scene has a size of approximately 25 km by 8 km. These datasets are orthorectified and geometrically corrected. The positional accuracy is also less than 10 m RMSE and the coordinate system is WGS-84. It has a spatial resolution of 3.125 m, and spectral bands are shown in Table 2. Both satellite products are delivered to users in GeoTIFF format and a comparison of the spatial resolution between them is showed in Figure 13 (Planet Labs Inc., 2018).

.Spectral Bands	PlanetScope	RapidEye
Blue	455 - 515 nm	440 - 510 nm
Green	500 - 590 nm	520 - 590 nm
Red	590 - 670 nm	630 - 685 nm
Red Edge	x	690 - 730 nm
NIR	780 - 860 nm	760 - 850 nm

Table 2: Spectral resolution and bandwidth of PlanetScope and RapidEye satellites

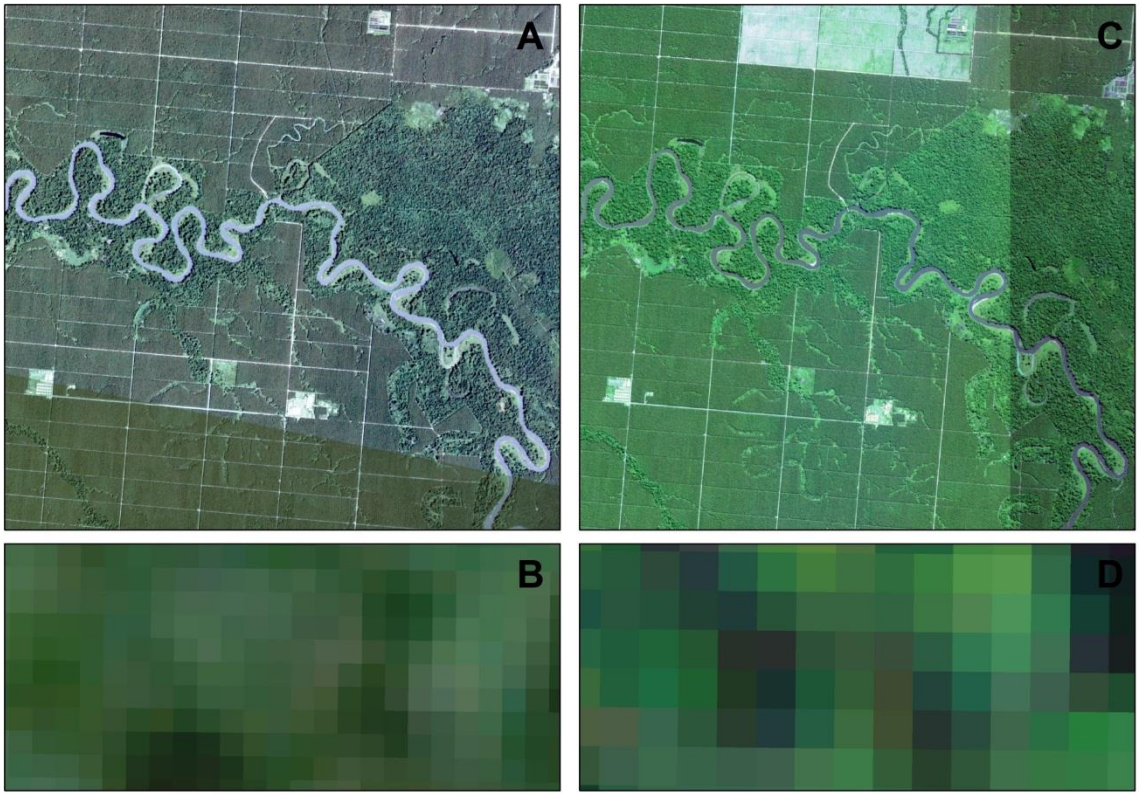


Figure 13: Spatial resolution comparison of PlanetScope and RapidEye. A) PlanetScope image, 3m resolution, sensed on January 4th, 2018, scaled to 1:50.000 B) Image A zoom to scale 1:500, pixel size 3m x 3m C) Rapid Eye image, 5m resolution, sensed on May 21, 20

RapidEye and PlanetScope images were used to create the validation dataset by visual interpretation of the time series available. Here it was necessary to download all cloud-free images from year 2016 to 2018. In total 93 tiles were used to complete the most of the AOI in different years as shown in Table 3.

Satellite	Resolution	Sense date	Tiles
PlanetScope	3m	10/18/2016	5
PlanetScope	3m	10/19/2016	3
PlanetScope	3m	1/18/2017	3
PlanetScope	3m	2/17/2017	5
RapidEye	5m	5/5/2017	6
PlanetScope	3m	7/15/2017	10
PlanetScope	3m	9/29/2017	8
PlanetScope	3m	10/3/2017	4
PlanetScope	3m	1/4/2018	11
RapidEye	5m	5/21/2018	6
PlanetScope	3m	8/26/2018	9
PlanetScope	3m	10/1/2018	4
PlanetScope	3m	10/2/2018	4
PlanetScope	3m	10/12/2018	3
PlanetScope	3m	10/30/2018	10
PlanetScope	3m	11/4/2018	2

Table 3: Cloud-free available high resolution imagery datasets

3.5. Creating the validation dataset

For the per-pixel analysis it was necessary to create a validation dataset in which further change detection process and results validation were performed take action. For this stage, RapidEye (5m resolution) and PlanetScope (3m resolution) imagery was used complemented in some cases with Sentinel-2 satellite imagery (10m resolution). The visual interpretation of high resolution imagery time series was executed with ArcGIS Desktop 10.5. The time period used here matched with the time series available in Sentinel-2 times series acquire for the change detection process. Therefore, as shown in Table 2, 16 high resolution times series were used in the period October 2016 to November 2018. Here is important to mention that Planet's program used for this purpose does not have complete time series data sets, thus all available cloud-free imagery was downloaded.

For validation dataset, 623 points were digitized and distributed among the whole AOI (Figure 14). The points were classified as "Forest" and "Deforestation". Deforestation class included a date attribute, which was identified by the visual interpretation of time series (Figure 15). In some cases, Sentinel-2 imagery was used to improve the deforestation detection date, as it has a more complete time series collection. For the "Forest" class, 348 points where identified which did not changed during the studied period, and for "Deforestation" class 275 points were created. Later, coordinates were assigned to each point.

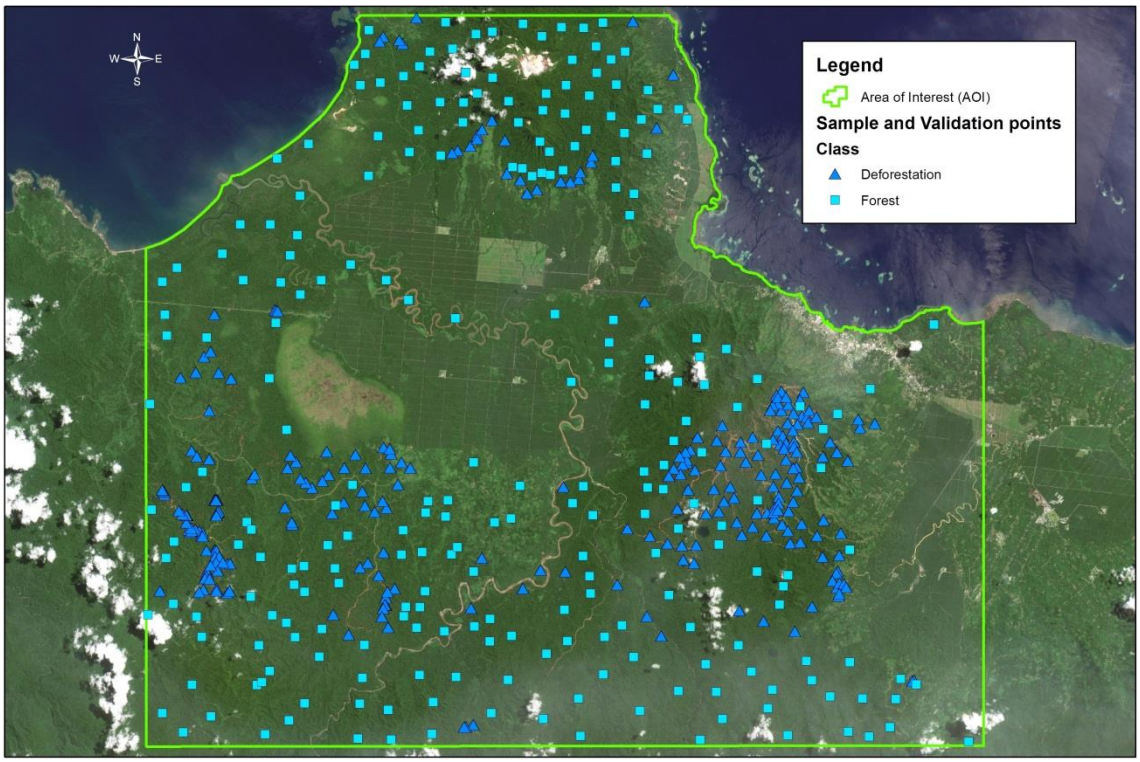


Figure 14: Sample and Validation points distributed over the Area of Interest. The "Deforestation" class is symbolized with a blue triangle and the "Forest" class by a light blue square.

After the completion of the sample and validation points in ArcGIS, the database was converted to *. CVS format and uploaded to R for further analysis.

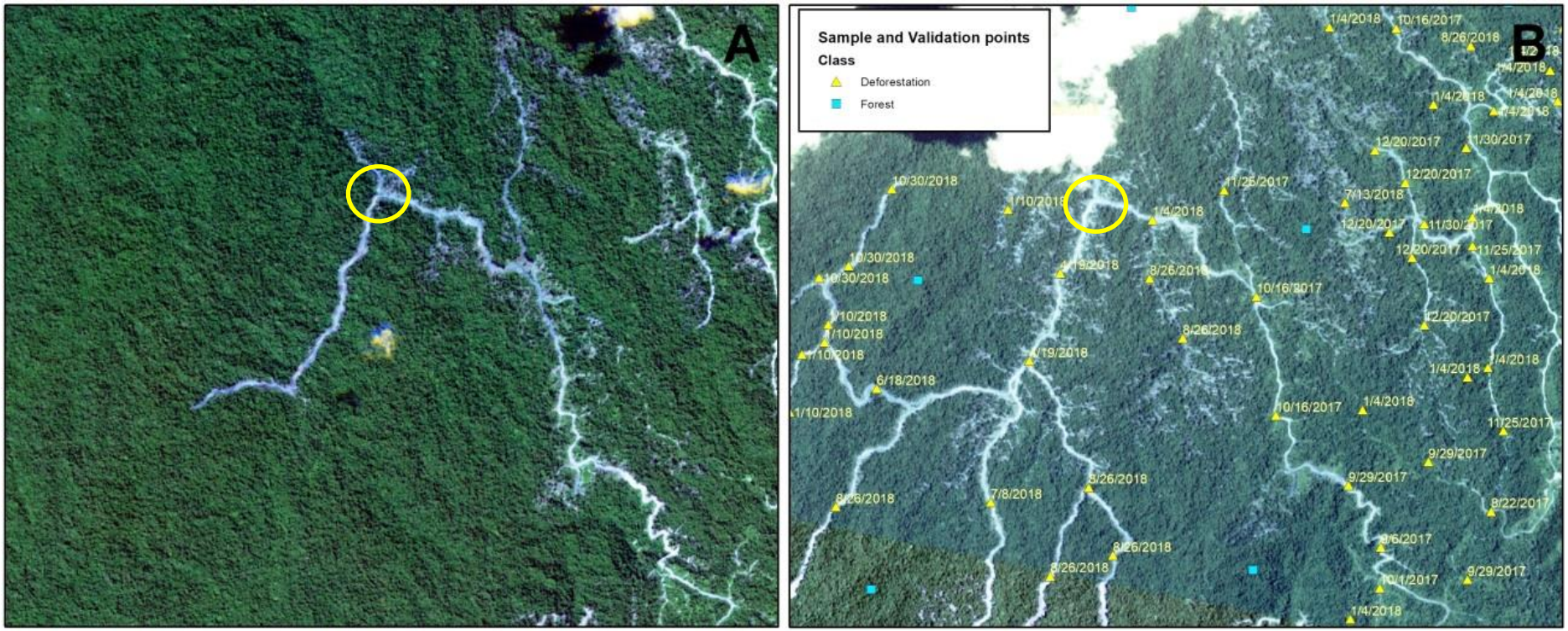


Figure 15: Example of "Deforestation" points with respective date when the change was detected. Both images correspond to PlanetScope satellite. Panel A shows a clean image from May 21st, 2018, Panel B shows an image dated from November 30th, 2018. In the latter, evolution of deforestation can be observed with the corresponding points showing the date when forest disturbances were detected by visualization procedures. Yellow circle as reference guide of the same location on image A and B.

3.6. Sentinel datasets pre-processing for Principal Component Analysis

3.6.1. Cloud, cloud shadow, water masking

In order to run the change detection process with Principal Component Analysis (PCA), first all Sentinel-2 time series datasets needed to be pre-processed. For all remote sensing analysis and procedures related to earth surface is fundamental to consider removing clouds, cloud shadows, snow and water bodies. These elements can substantially affect results in analysis like atmospheric correction, land cover classification, calculation of indexes or change detection processes (Zhu et al., 2015; Frantz et al., 2018). For this purpose Function of mask (Fmask) 4.0 tool with GUI was used over all Sentinel-2 datasets. The first version of this algorithm (Fmask 1.6) was created for Landsat images (Zhu & Woodcock, 2012), later version 3.3 of Fmask was developed to improve results for Landsat 4-7, to handle Landsat 8 datasets which came with a new cirrus band beneficial for cloud detection and finally, incorporate parameters to make Fmask work with Sentinel-2 images, which did not have a thermal band used for Landsat datasets, therefore, the Sentinel-2 cirrus band was used for this purpose and all steps where the Landsat thermal band worked were removed (Zhu et al., 2015). Fmask 4.0 follows the improvements achieved on the previous version and adds new characteristics as integrating auxiliary data (e.g. Global Surface Water Occurrence (GSWO), Digital Elevation Model (DEM)), new cloud probabilities, and spectral-contextual features that principally helps identifying better snow and ice from clouds (Zhu et al., 2018).

The most important parameter Fmask 4.0 uses is the "Cloud Probability Threshold" and its default value on the Fmask 4.0 tool with GUI is 20% for Sentinel-2 datasets. After a first tryout with default values, a test with different thresholds values was done using three Sentinel-2 datasets with varying amount of clouds in order to find the threshold that better fitted to the Study Area. The test is showed in Figure 16 and a 40% Cloud Probability Threshold was decided.

Fmask 4.0 tool with GUI automatically identifies all the data from the original Sentinel-2 Level-1C product, and safes the results on a new folder. The resulting values of the masking output are showed in Table 4.

Value	Class
0	clear land pixel
1	clear water pixel
2	cloud shadow
3	snow
4	cloud
255	no observation

Table 4: Resulting classification values of Fmask 4.0 output

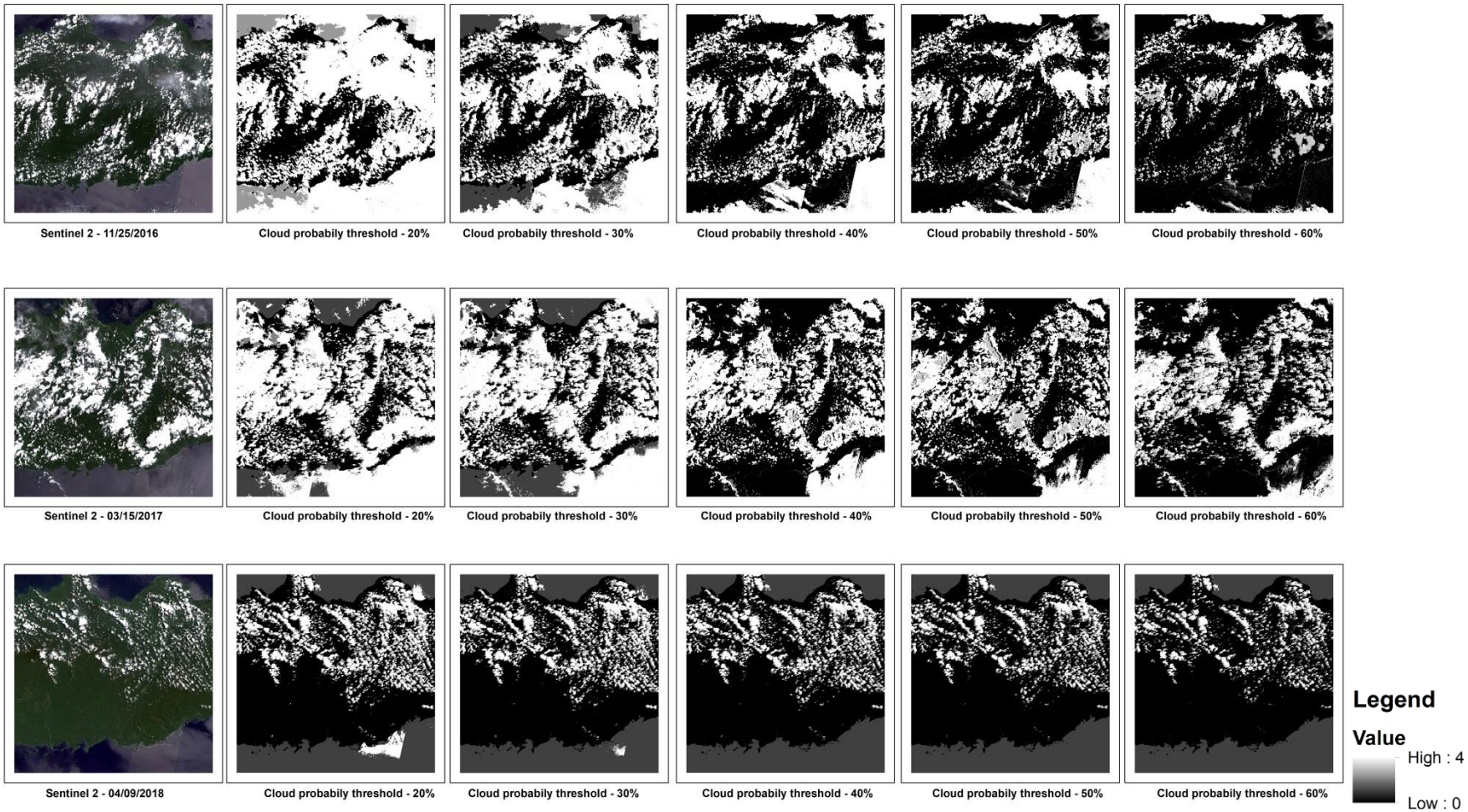


Figure 16: Fmask 4.0 - Cloud Probability Threshold test of three different images from the study area. Low values (black) corresponds to clear land pixels and high values (white) depicts cloud pixels

The resulting outputs corresponded to the entire granule of each image. Therefore all the outputs were cropped to the AOI using ArcGIS Desktop 10.5 and a calculation of the real percentages of clear land pixels inside the AOI was done as showed in Figure 17. Afterwards, the mask was applied to all datasets using the Conditional tool from Spatial Analyst Extension in ArcGIS Desktop 10.5, thus a new raster with only clear land pixels were created for every dataset.

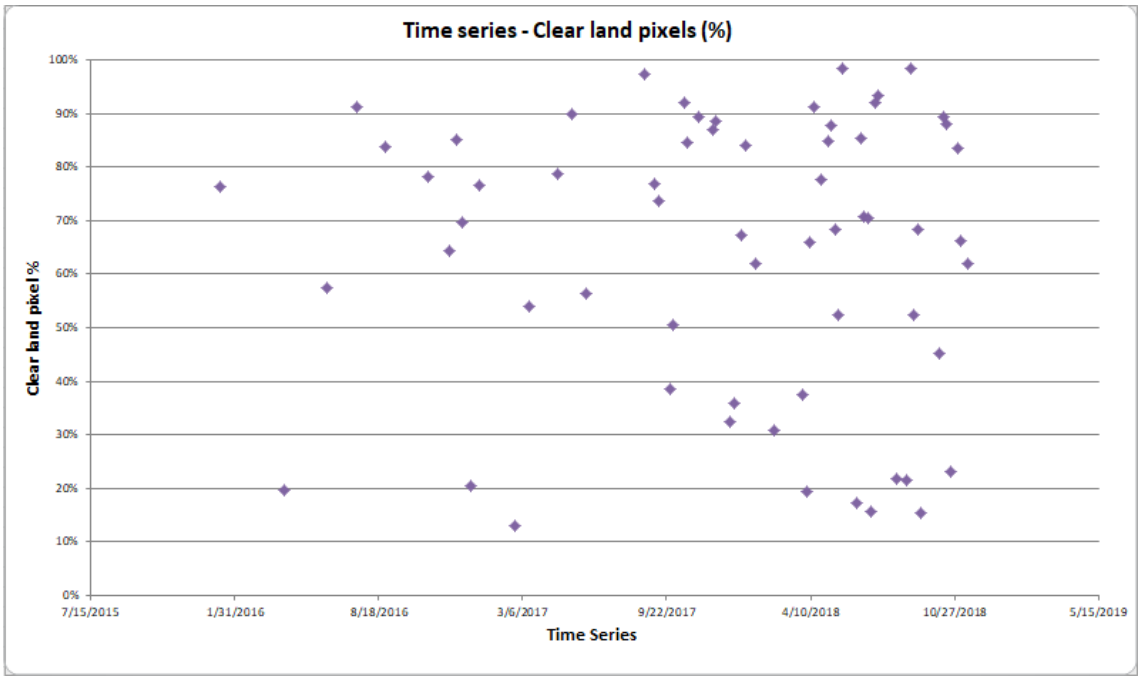


Figure 17: Clear land pixels percentage plot of all Sentinel-2 time-series datasets for the Area of Interest

Considering the low amount of valid datasets for the study area, every dataset with minimum clear land pixels is useful. However, images with less than 10% of clear land pixels where removed because the amount of information was too low, and in percentages between 10% and 30% there were areas that could have valid information.

3.6.2. Spectral bands resampling

Finally, in order to homolog the spatial resolution of all Sentinel-2 time-series datasets resampling the 10 meter resolution bands to 20 meters was necessary. This process was done with ArcGIS Desktop 10.5, using the Resample tool from Raster Processing toolbox. The resampling technique used was Bilinear Interpolation, which calculates a weighted average based on the four nearest input cell centers (ESRI,2019). An example is shown in Figure 18. This resampling method was executed to band 2 (blue), band 3 (green), band 4 (red) and band 8 (NIR) for all time-series (252 images in total).

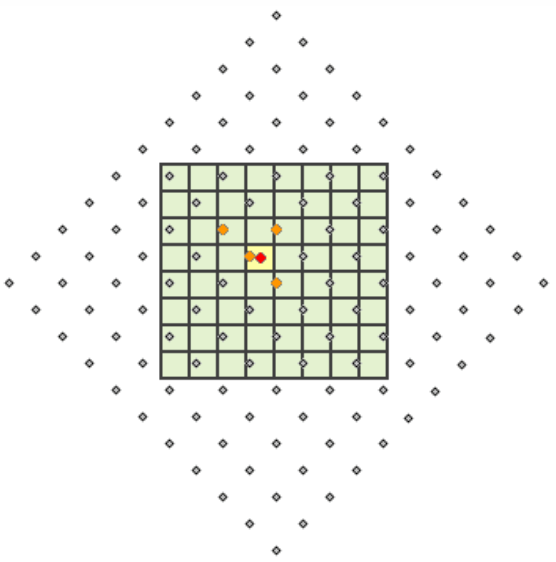


Figure 18: Example of Bilinear Interpolation Resampling method. The processing cell center is showed as a red point, cell centers are in gray, output cells are shaded in green, cell being processed shaded in yellow (ESRI, 2019)

With this last process, the Sentinel-2 dataset for running the Principal Component Analysis is finished.

3.7. Normalized Difference Vegetation Index calculation

In order to answer Research Question 1, Normalized Difference Vegetation Index (NDVI) was calculated for all time-series. For this purpose 10 m resolution bands 8 (NIR – 842nm) and band 4 (Red – 665nm) time series were used in the calculation done with ArcGIS Desktop 10.5 using Equation 1 as follows:

NDVI = (float(band8) – float(band4)) / (float(band8) + float(band4))

All NDVI outputs were masked applying the cloud, cloud shadow and water output done previously with Fmask 4.0.

3.8. Validation and accuracy assessment

Validation and accuracy assessment are the final results which determine how feasible the method is using Sentinel-2 time-series. To achieve this, a confusion matrix (error matrix) was elaborated, which is one of the most common methods to measure accuracy in remote sensing processes (Comber et al., 2012) . The confusion matrix is a cross tabulation (2x2) that compares the resulting values of certain process - which in this case will be the structural change detection on forest- with validation data from the validation dataset explained in Section 3.5 (Story & Congalton, 1986; Lewis & Brown, 2001; Carfagna & Gallego, 2005 Visa et al., 2011; Comber et al., 2012). The confusion matrix was elaborated for all overall Empirical Fluctuation Process results on NDVI and PCA datasets arrays for our study area.

Confusion matrix results where after placed in the Figure of Merit accuracy assessment method suggested by Pontius et al (2008) and taken from Lu et al (2017). The

confusion matrix results were classified as True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). TP refers to an agreement between forest disturbance detected by EFP and validation data. TN refers to unchanged forest during the whole time series that was not detected as change in EFP and that was classified as forest in the validation data. FP refers to a mismatch due to a forest disturbance detected by EFP when validation data pixel had a forest value. FN refers to a mismatch when EFP process did not detect any disturbance, while validation data indicated deforestation.

Once Figure of Merit is complete, accuracy assessment was measured using the suggested method by latter authors mentioned in this section. Here, four types of accuracy were calculated as shown in Table 5.

Accuracy measure	Formulae
Figure of merit	TP/ (TP + FN + FP)
User's accuracy	TP/ (TP+FP)
Producer's accuracy	TP /(TP+FN)
Overall accuracy	((TP+TN)/(TP+FN+FP+TN)

Table 5: Accuracy measurements applied for overall results of Empirical Fluctuation Process.

Pontius et al (2018) defines the latter accuracy measures as follows:

Figure of Merit is a statistical measurement that represents the ratio of intersection between the observed and predicted change, to the union of the observed and predicted change. User's accuracy depicts the proportion of positive predictions of the model as change, given that the model predicts a change, while Producer's accuracy represents the proportion of positive predictions as change, given that the validation dataset indicates a change.

Overall accuracy will permit to compare the results from present research with other similar studies (Lu et al., 2017).

4. IMPLEMENTATION

4.1.Data structuralizing

Once Sentinel-2 time-series spectral bands pre-processing and NDVI calculations were complete, they were migrated to R in order to accomplish the steps for the change detection analysis. At this stage, the Sentinel-2 spectral bands dataset and the NDVI dataset where processed in different R projects, here, we will explain the common steps done for both datasets.

4.1.1. Raster stack, sample dataset location extraction and array

Firstly, both datasets (Sentinel-2 spectral bands and NDVI calculations) where stacked (stack function), which results in a collection of RasterLayer objects with the same spatial extent and resolution. This step is needed to group the whole dataset, see Figures 19 and 20. Parallel to this, the sample database is uploaded to R, and by the use of the "extract" function from raster package in R, the locations (coordinates) of the sample dataset are related to the correspondent pixel of each image and to the specific reflectance value in the spectral bands and the NDVI pixel value on the NDVI dataset. This results in a two-dimensional matrix for each dataset. In the spectral bands the matrix dimension is [1:623, 1:630], which corresponds to the 623 sample points by 63 time-steps of 10 spectral bands each. The NDVI matrix dimension is [1:623, 1:63], corresponding to the 623 sample points and 63 time-steps.

At this point, all the data from the spectral bands and NDVI calculations are stored in matrixes. These two-dimensional matrixes are now converted to three-dimensional arrays. The array in R is a way to storage a vector with several attributes. For the spectral bands the array has the following dimensions: [1:623, 1:10, 1:63], meaning sample points (locations), spectral bands and time-steps respectively. For the NDVI dataset the array dimensions are [1:623, 1:1, 1:63], depicting sample points (locations), NDVI values and time-steps respectively. An example of the first 25 locations of time-step 1 of the spectral bands array can be seen in Figure 21.

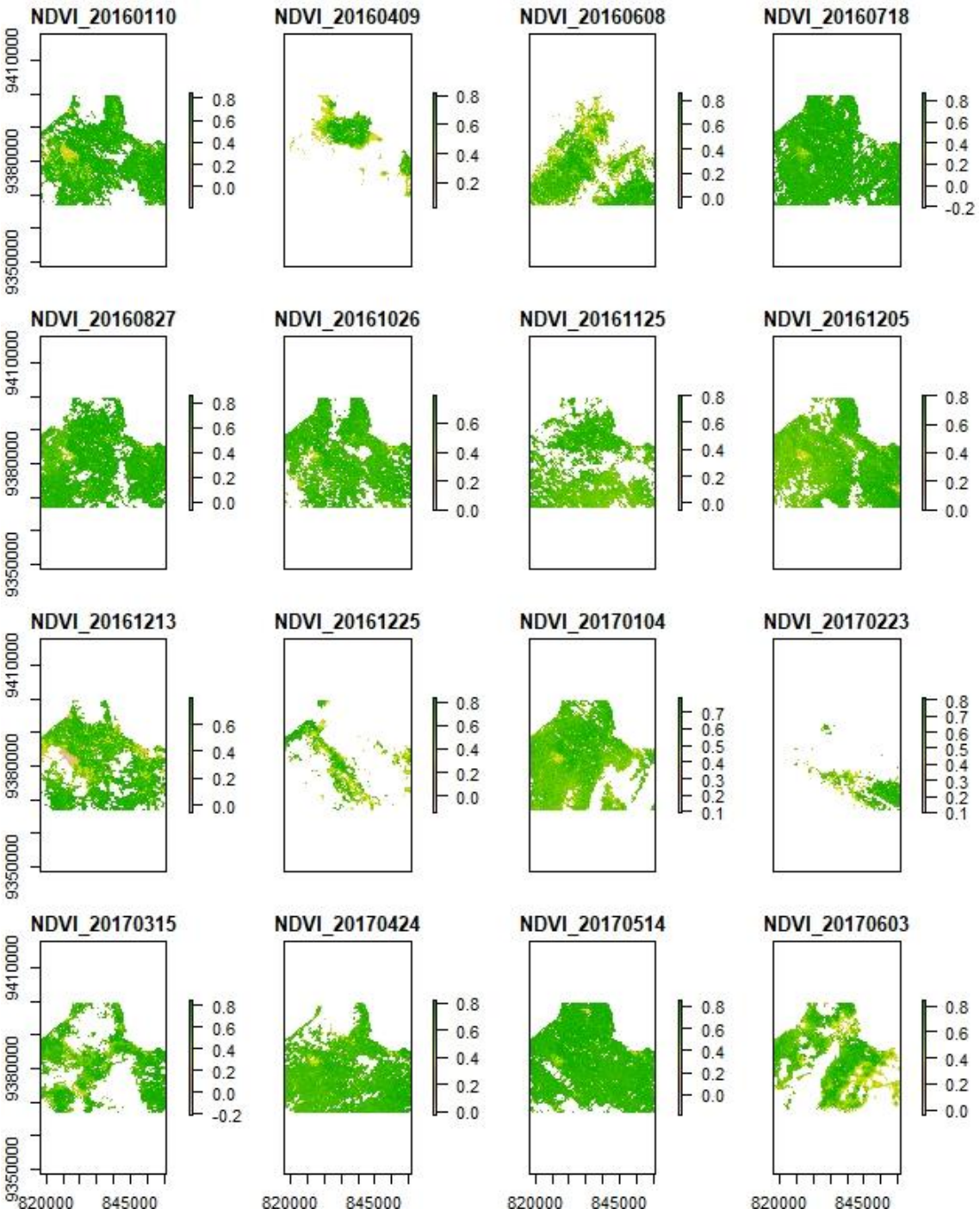


Figure 19: Masked NDVI time-series period 01/10/2016 to 06/03/2017. X and Y axis depicts UTM coordinates.

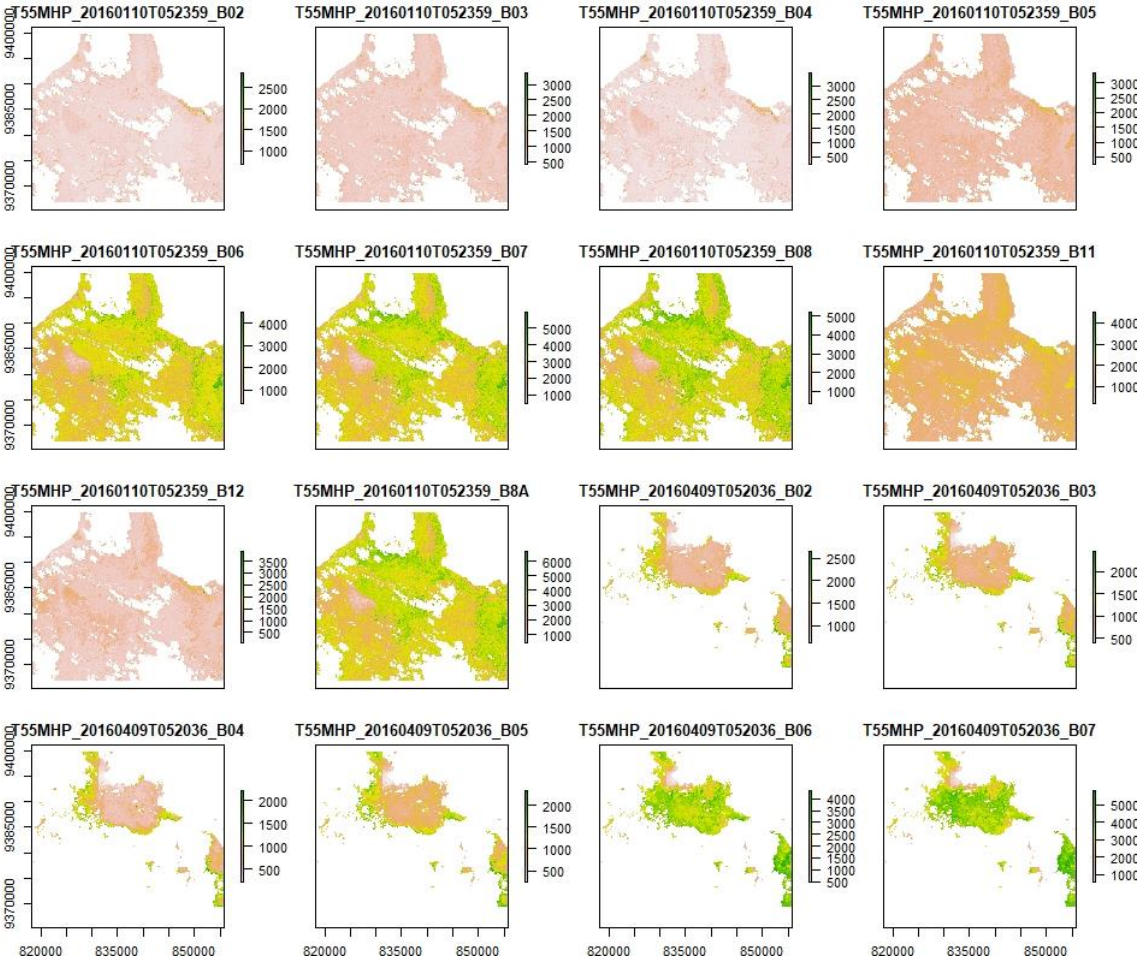


Figure 20: Masked spectral time-series period 01/10/2016 to 04/09/2016. X and Y axis depicts UTM coordinates.

```
, , t 1

     BAND2 BAND3 BAND4 BAND5 BAND6 BAND7 BAND8 BAND11 BAND12 BAND8A
1      779   628   381   645  1918  2435  2001   1148    376   2868
2      792   705   383   718  2251  2826  2860   1333    429   3248
3       NA    NA    NA    NA    NA    NA    NA     NA     NA     NA
4      858   876   468   992  2987  3756  3308   1802    643   4158
5      797   722   367   861  2391  3103  2639   1461    507   3385
6      805   747   395   962  2586  3366  2705   1782    631   3806
7      772   659   342   749  2387  3192  2427   1558    516   3581
8      762   625   320   604  1917  2571  2256   1311    428   3002
9      788   693   362   793  2608  3426  2761   1589    555   3757
10     796   707   344   660  2145  3071  3040   1285    426   3336
11     787   696   337   672  2490  3113  2916   1518    509   3844
12     751   574   303   404  1041  1384  1333    480    169   1469
13      NA    NA    NA    NA    NA    NA    NA     NA     NA     NA
14     778   638   347   558  1631  2136  1980    982    339   2444
15     734   558   290   461  1251  1724  1381    714    230   1779
16     800   711   366   679  2173  2783  2769   1401    506   3193
17     770   668   340   637  1925  2522  2406   1113    370   2740
18     780   665   340   633  2166  3199  2712   1240    413   3646
19    1039  1027   851  1210  2771  3441  3089   1831   1019   3847
20     826   766   381   817  2825  3661  3302   1794    659   4186
21     822   722   384   729  2170  2831  2720   1277    414   3334
22     823   739   448   854  2212  2825  2524   1376    680   3309
23     820   717   396   733  2191  2991  2527   1395    468   3543
24    1043   808   436   774  2238  2906  2741   1345    490   3279
25     817   757   384   781  2507  3200  3123   1357    454   3778
```

Figure 21: Example spectral bands array. Time-step 1, first 25 locations for the 10 spectral bands used on the research

4.2. Selection of training sample points

For analysis, 20 training sample points from the complete 623 sample and validation dataset were selected (deforestation = 10, forest = 10). It is important to mention that each point (training sample) represents one pixel. These samples were analyzed in detail following the general method. In Table 6, the training samples can be observed with their correspondent percentage of valid and omitted pixels due to the cloud masking process.

Sample	Valid time-steps	%	Omitted time-steps	Class
1	46	73%	17	Deforestation
33	37	59%	26	Deforestation
54	41	65%	22	Deforestation
101	41	65%	22	Deforestation
110	44	70%	19	Deforestation
119	46	73%	17	Deforestation
135	34	54%	29	Deforestation
177	41	65%	22	Deforestation
196	35	56%	28	Deforestation
239	38	60%	25	Deforestation
400	32	51%	31	Forest
410	23	37%	40	Forest
420	44	70%	19	Forest
430	45	71%	18	Forest
440	31	49%	32	Forest
450	37	59%	26	Forest
580	41	65%	22	Forest
590	37	59%	26	Forest
600	42	67%	21	Forest
620	35	56%	28	Forest

Table 6: Chosen samples (locations) showing the valid and omitted time-steps, and the correspondent class for each sample.

4.3. NDVI time-series

Time Series were created for the NDVI dataset. This was accomplished with "Zoo" package in R. In Figure 22 are shown the first half of the training samples time series where gaps due to the absence of time-series imagery can be clearly observed. These results are the inputs to run the EFP for NDVI.

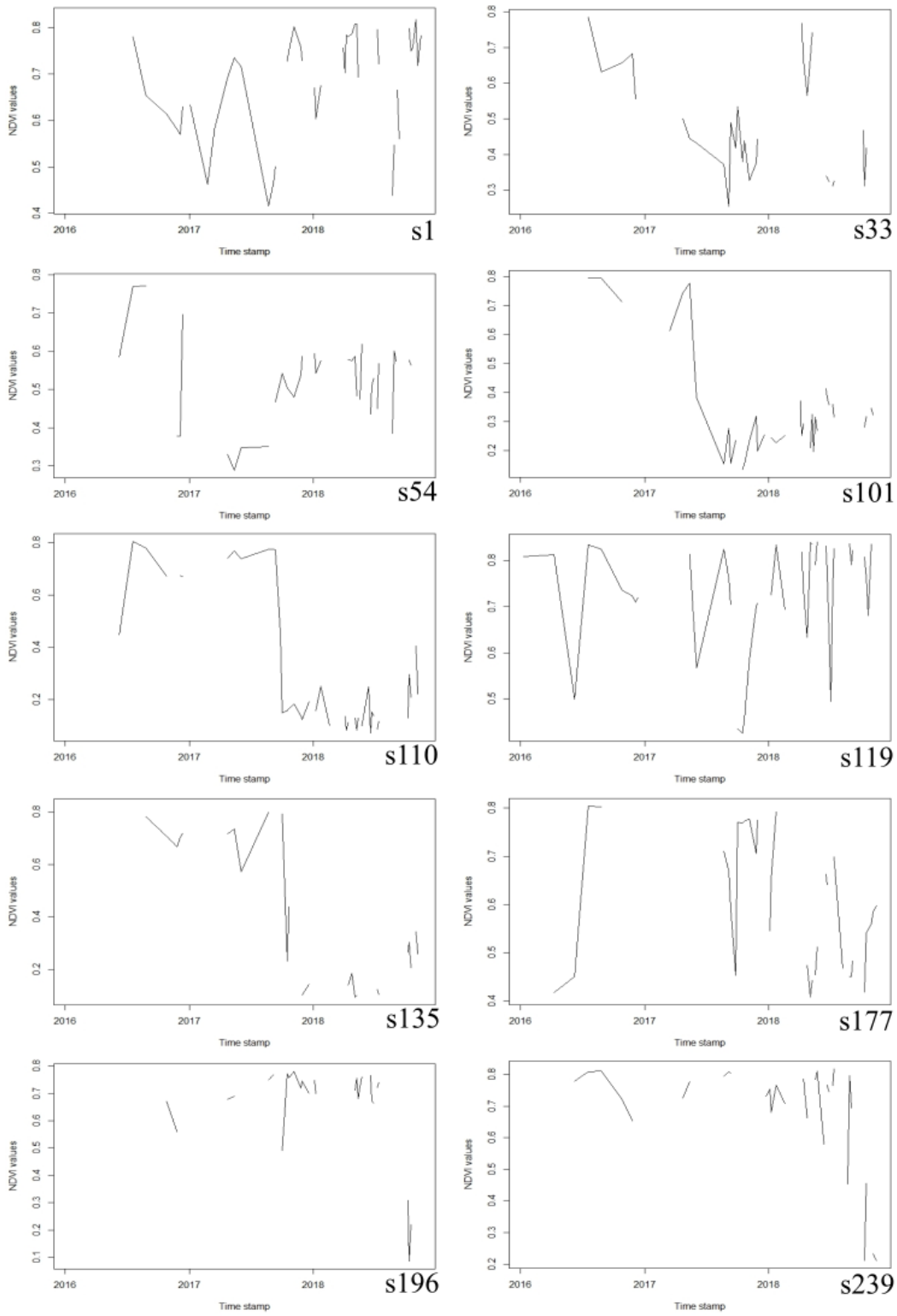


Figure 22: Example of NDVI time-series plots for different sample points (locations). "s" refers to sample and the consequently number

Table 7 depicts the dates when deforestation was observed in the validation dataset for each of the sample points in Figure 22.

Location	Time of real change
1	9/29/2017
33	4/24/2017
54	1/4/2017
101	8/22/2017
110	9/29/2017
119	8/22/2017
135	10/16/2017
177	8/26/2018
196	8/26/2018
239	1/10/2018

Table 7: Dates of observed (validation dataset) change for all samples depicted in Figure 22

4.4. Principal component Analysis

All the rows with missing values due to cloud masking are removed for computing the PCA. The 2D input matrix has reflectance values for each band as columns (variables) and the validation points as rows (observations) Once the array had only valid values, PCA was calculated for the Sentinel-2 spectral bands array (all bands and time-series) using the "prcomp" function in R.

The number of Principal Components (PC's) is equal to the number of variables used, which in this case are 10 spectral bands from Sentinel-2 satellite. Here we used several tools (data and plots) to determine which principal components should be used.

Is important to mention that on one hand, Principal Components are orthogonal to each other, thus there is no correlation between them (Kwarteng & Chavez, 1989) (see figures on Appendix on pages 61 to 64), and on the other hand, is possible to analyze the correlation between the variables (e.g. spectral bands) and their relation to the Principal Components, meaning which variables and observations (e.g. spectral data in a specific pixel along time-series) are better explained in certain PC.

To accomplish this, first PC loadings of all training datasets were observed in order to have a first income about their structure and the distribution of positive and negative values along the 10 spectral bands. Secondly the same was done with PC scores, and the correlation between bands in relation with the first two PC's was observed. Thirdly, variances and cumulative variances proportion were analyzed to decide which PC's to use for Empirical Fluctuation Process. Lastly, time-series were created for the chosen PC's.

The latter explained is represented in the next pages.

PC loadings or rotation values, are coefficients of the linear combination of the variables, and show the variance's magnitude between the variables. Here the variables of each training sample are all the available time-series spectral information for that specific pixel. The PC loadings values per band (training sample 1) are shown in Table 8 and Figure 23 and 24. More examples of PC loadings graphs can be found in the Appendix on pages 65 and 66.

	PC1	PC2	PC3	PC4	PC5
BAND2	0.26438	0.28391	-0.51494	0.04707	-0.21879
BAND3	0.38029	0.18614	-0.31926	-0.39470	0.01819
BAND4	0.27125	0.41861	-0.13381	0.19386	-0.29342
BAND5	0.42040	0.17659	-0.00474	0.22348	0.61294
BAND6	0.32019	-0.37256	-0.06905	0.20529	0.41234
BAND7	0.28704	-0.41050	-0.03948	0.22970	-0.08313
BAND8	0.33001	-0.31705	0.08445	-0.74219	-0.06532
BAND8A	0.28553	-0.40476	-0.02839	0.32866	-0.52418
BAND11	0.29291	0.15547	0.59143	0.02046	-0.17961
BAND12	0.27207	0.28723	0.50093	0.01672	-0.04225
	PC6	**PC7**	**PC8**	**PC9**	**PC10**
BAND2	0.53859	-0.25487	0.07152	-0.09726	-0.40575
BAND3	-0.08311	-0.10985	-0.03404	0.46914	0.56785
BAND4	-0.33166	0.46022	-0.11731	-0.48882	0.18734
BAND5	-0.35341	0.01094	0.30166	0.14660	-0.35904
BAND6	0.22773	-0.24416	-0.29768	-0.46657	0.35208
BAND7	0.34533	0.67085	0.03030	0.34251	-0.01838
BAND8	-0.12275	0.12175	0.01324	-0.29853	-0.33037
BAND8A	-0.42604	-0.40358	0.00723	0.13954	-0.08454
BAND11	0.29509	-0.13537	0.56930	-0.11743	0.25351
BAND12	0.11550	-0.07332	-0.68928	0.22939	-0.20987

Table 8: PC loadings (rotation) for training sample 1

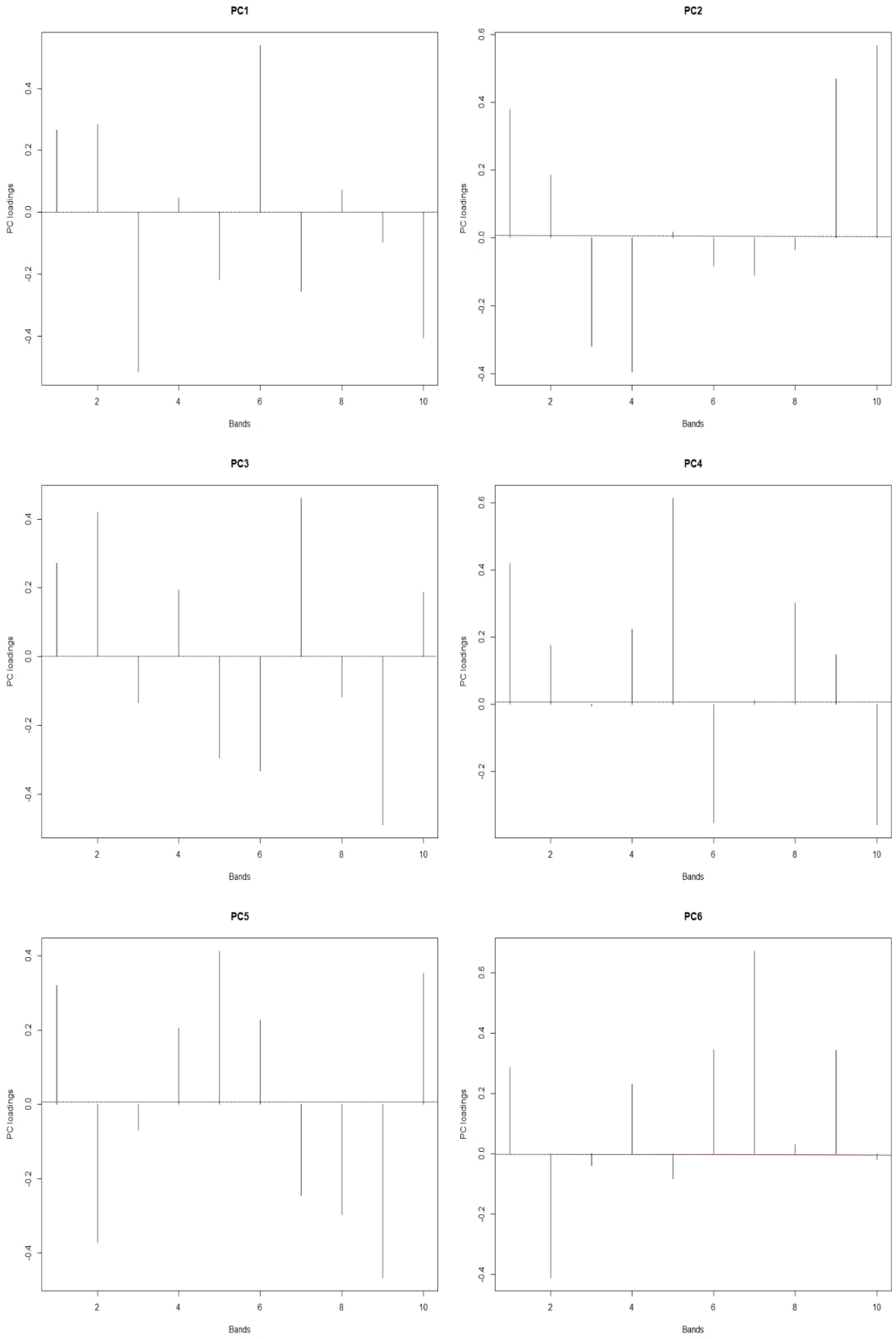


Figure 23: PC loadings (rotation) of all bands for PC's 1-6 of training sample 1. The sequence of the bands (1-10) corresponds to the following: Bands 1-3 to visible; Bands 4-6 to Red Edge; Bands 7 and 10 to NIR; and Bands 8 and 9 to SWIR.

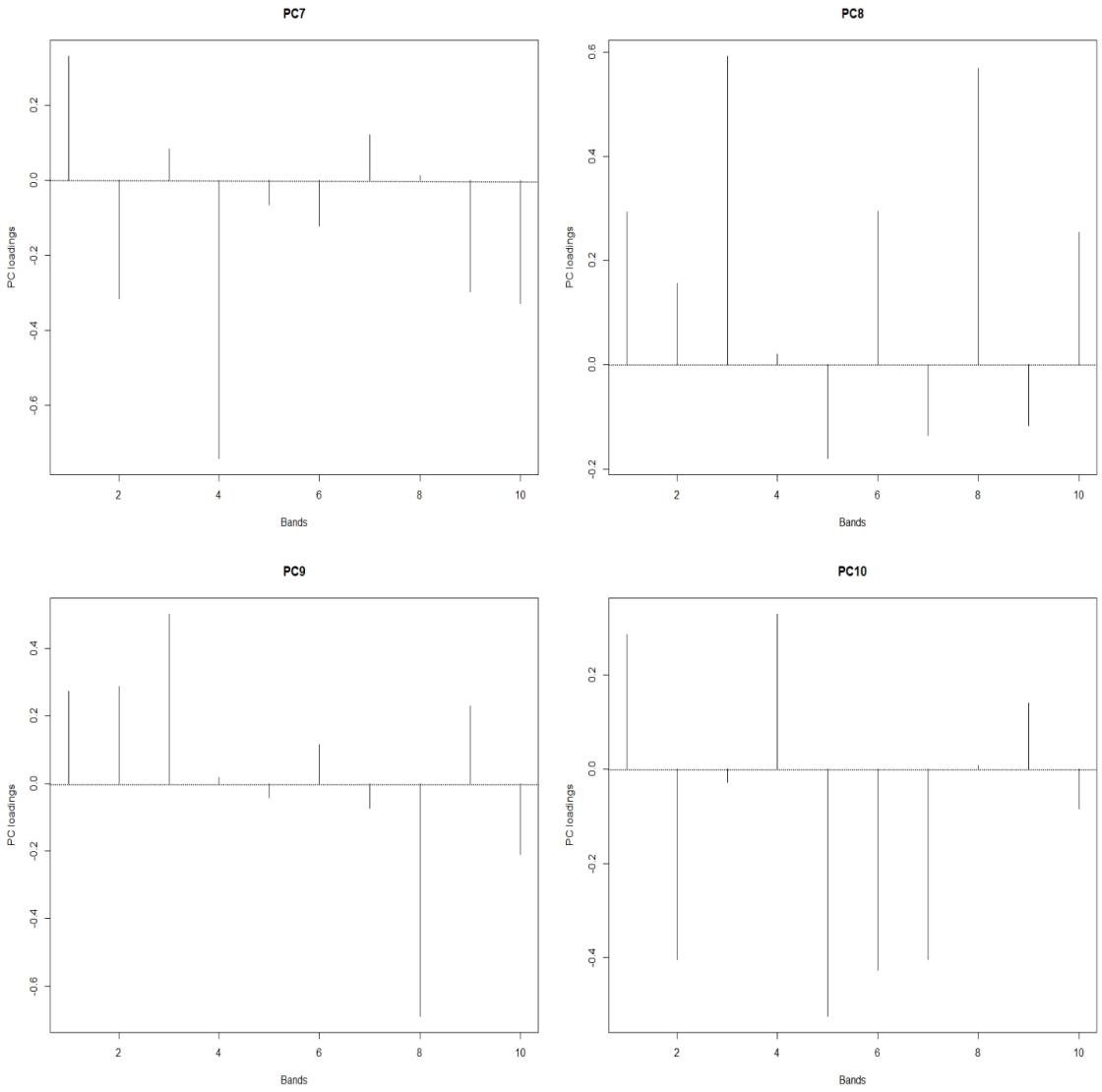


Figure 24: PC loadings (rotation) of all bands for PC's 7-8 of training sample 1. The sequence of the bands (1-10) corresponds to the following: Bands 1-3 to visible; Bands 4-6 to Red Edge; Bands 7 and 10 to NIR; and Bands 8 and 9 to SWIR.

PC scores values are the linear combination of standardized variables weighted by PC loadings. Therefore, data was centered (Table 9) and standardized before applying PCA. PC scores are useful to analyze which Principal Components represent the higher variability of the data by dimension reduction, and be able to define which PC are the most feasible to work with as depicted in Table 10.

BAND2	BAND3	BAND4	BAND5	BAND6
930.2286	839.8571	597.4857	851.6571	2283.2000
BAND7	**BAND8**	**BAND11**	**BAND12**	**BAND8A**
3042.0000	2637.3714	1408.5714	577.3143	3397.6000

Table 9: Center values for training sample 196

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
t 1	-1.41692260	0.56726812	-0.08933349	0.197045112	-0.04452689	0.1938802561	0.068690589	0.091865316	0.078786725	-2.356184e-02
t 4	-1.44238015	0.98779504	0.17471882	-0.078748317	-0.03683009	-0.2499158388	0.026204240	-0.028884920	0.082189653	1.006135e-01
t 6	1.40916186	1.47155537	-0.26552255	-0.057680716	0.08657742	0.0605464523	0.052834746	0.232209600	0.027120867	-3.790303e-04
t 7	1.22558168	-0.34244720	-0.97473463	-0.019938366	0.05024363	-0.0045213268	-0.121274074	0.121790605	0.018761689	-3.028944e-02
t 12	6.35303658	1.52615952	-0.97192954	-0.013872014	-0.06668724	0.0005583335	0.207973225	-0.076464654	-0.058774379	1.014081e-01
t 14	0.29412357	0.71980032	0.03786842	0.286725170	0.15029094	0.2191992320	0.165708140	-0.084431174	-0.021957547	2.020932e-02
t 15	-1.45129415	-0.10692239	-0.07185174	0.026560745	0.04679610	-0.0734810165	0.007399925	-0.065115583	-0.080130189	-4.562545e-02
t 17	-1.76776423	0.30477329	0.21417903	0.007830033	-0.16017138	-0.2538591082	0.080361527	-0.171115644	0.012306157	4.344949e-02
t 18	-1.13952961	0.93311968	0.44254276	-0.434726118	-0.17930553	-0.1462324019	0.086699094	-0.046608569	-0.052599929	-6.041367e-02
t 21	-2.76506491	-4.16774394	-1.58120996	0.213256543	-0.05827774	-0.1381022576	0.295045088	-0.108368582	0.046464921	-3.621012e-02
t 22	-0.26067264	1.82286564	0.51398664	0.113849822	0.14341311	0.2285018348	0.046317109	-0.056112843	0.048632367	-8.810428e-03
t 23	0.49380016	2.26467678	0.50046032	-0.047894759	0.07723958	0.0208754908	-0.006133774	-0.032906832	0.192061798	1.114912e-02
t 24	-0.18249989	2.10377292	0.31174173	-0.151123839	0.05577412	0.1995797333	-0.136364638	0.030435593	0.006653092	-7.757057e-03
t 25	-0.45107852	0.79383782	0.06911623	-0.159063047	0.02686143	-0.0292205360	-0.069516600	0.002580386	-0.044051716	-5.210262e-02
t 26	-0.44129066	1.13249477	0.23563685	-0.027029810	-0.05526529	0.3017832280	-0.002289833	-0.177822047	-0.039065728	-6.862709e-03
t 27	1.45781498	2.49692194	-0.54624931	-0.015159307	-0.19988925	0.0577672632	-0.284032271	0.162621077	-0.035087147	2.178134e-02
t 29	-0.05600982	1.32206379	0.41320323	0.210884535	0.40371006	-0.0059824440	0.048705027	-0.071480376	-0.117648137	3.995455e-02
t 30	-0.35481766	0.44427921	-0.17201338	0.193050087	0.19225056	0.2270032651	0.127332757	0.050923270	0.004941772	-1.750234e-05
t 36	-1.09292863	0.85485012	0.47652456	-0.146251559	0.14792266	-0.3066061289	0.154643490	-0.018242774	-0.047999538	4.980623e-02
t 39	-0.52060033	1.10130291	0.34804419	0.070433427	0.16225580	-0.0885046628	0.193293112	0.211733141	-0.135154367	-4.876088e-02
t 40	-1.77629428	0.55676619	0.10726909	-0.227538561	-0.26541083	0.0151079718	0.164659895	0.144130587	0.011411530	-2.557429e-02

Table 10: PC scores of training sample 196 for the first 40 time series. The lacking time-series values are due because they have been omitted as NA values.

Figure 25 and 26 show the correlation between spectral bands and their values in respect to the Principal Components. Figure 25 depicts the plot for sample 196, where the first two Principal Components (PC1 and PC2) explains the 95.7% of the variance. Also, on one hand shows a strong correlation between band 6 and band 8, and band 8a and band 7. The correlation is high between all those bands. Though these bands have higher values on PC2, also shows positive values on PC1. The other correlated group of bands is bands 2, 3, 5 and 12, even though the stronger correlation is observed in bands 2 and 3. This group of bands shows their higher values on PC1 and negative values in PC2. It can be seen that band 11 and band 4 has a low correlation with other bands.

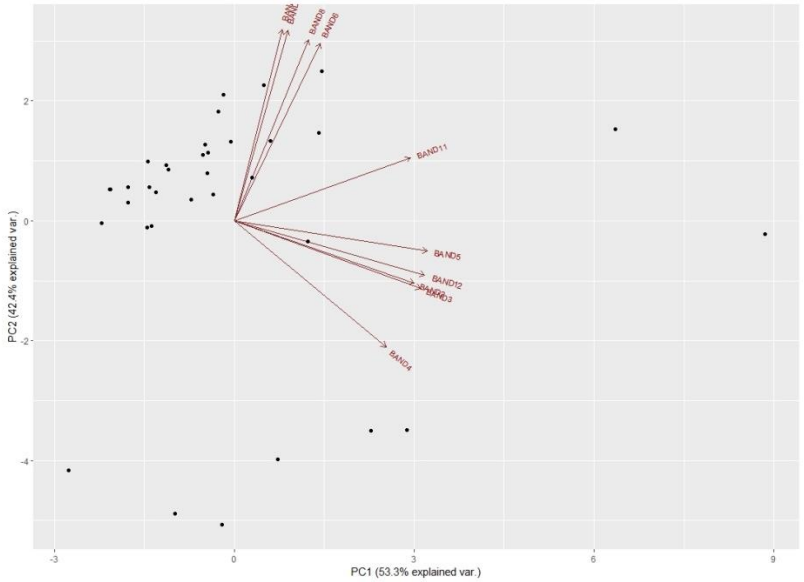


Figure 25: Bi-plot of PC scores of sample 196

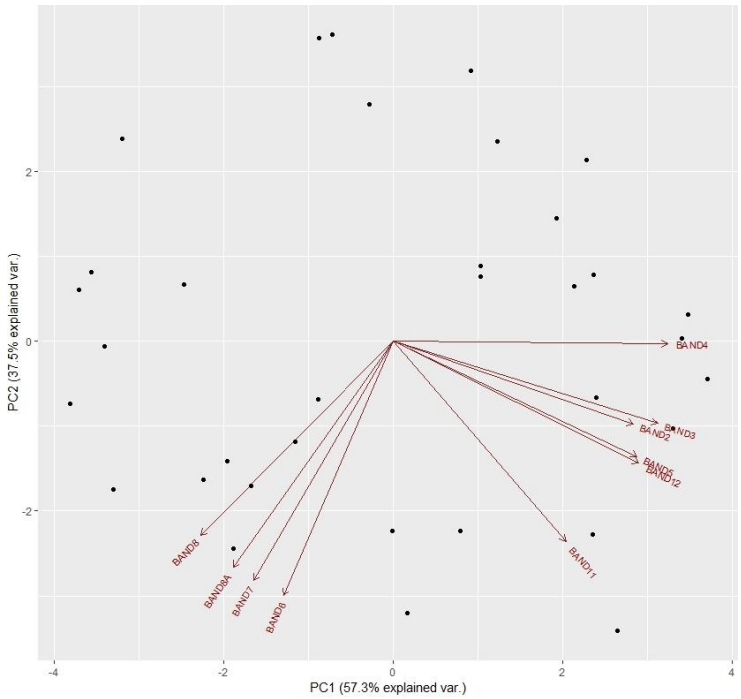


Figure 26: Bi-plot of PC scores of sample 135

Figure 26 shows a different case. Here, sample 135 shows correlation between bands, though, is interesting to notice that bands 6, 7, 8 and 8a have negative values in both PC's. These could tell that for those bands, that pixel is better explained in other PC's rather than PC1 or PC2.

After observing the bi-plots for all the samples, which only explain the relation of values towards the first two PC's, shoulder-plots can graphically explain the cumulative variance proportion of PC's for each sample. Figure 27 shows examples for 4 training samples. In the shoulder-plots can be observed that the first three PC's (PC1, PC2 and PC3) are the ones with largest variance, though, PC4 shows a boundary condition from where most of PC's start having negative values and a sort of equilibrium is founded. The PC's that explains the most of the data for each sample are PC1, PC2, PC3 and PC4, which are the PC's that will be used for the Empirical Fluctuation Process. As depict on Table 11, the first four PC's of all training samples explains almost 99 percent of the data.

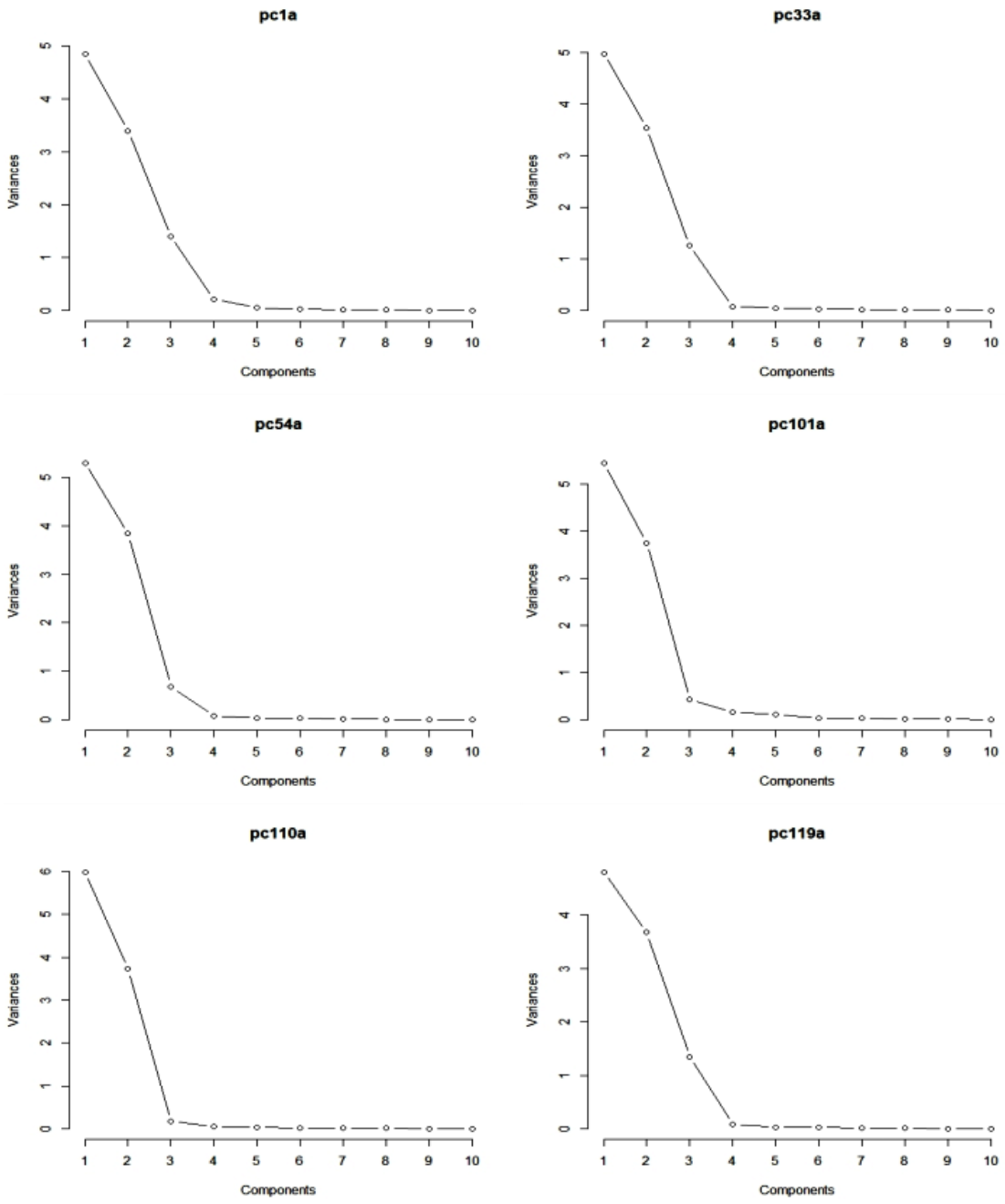


Figure 27: Shoulder plots of a portion of training samples, showing the variance explained from each Principal Component

Sample	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
1	0.484	0.823	0.964	0.985	0.991	0.994	0.996	0.998	0.999	1.000
33	0.497	0.852	0.978	0.986	0.990	0.994	0.996	0.998	0.999	1.000
54	0.529	0.912	0.981	0.989	0.993	0.995	0.997	0.998	0.999	1.000
101	0.544	0.919	0.962	0.978	0.989	0.993	0.996	0.998	0.999	1.000
110	0.598	0.970	0.988	0.992	0.995	0.997	0.998	0.999	0.999	1.000
119	0.480	0.847	0.981	0.989	0.993	0.995	0.997	0.999	0.999	1.000
135	0.573	0.947	0.977	0.985	0.993	0.995	0.997	0.998	0.999	1.000
177	0.689	0.878	0.980	0.992	0.994	0.996	0.997	0.998	0.999	1.000
196	0.533	0.957	0.982	0.988	0.992	0.995	0.997	0.999	0.999	1.000
239	0.559	0.906	0.963	0.978	0.986	0.991	0.995	0.997	0.999	1.000
400	0.831	0.955	0.973	0.986	0.992	0.995	0.998	0.999	0.999	1.000
410	0.898	0.985	0.990	0.995	0.996	0.998	0.999	0.999	0.999	1.000
420	0.873	0.964	0.982	0.988	0.992	0.995	0.997	0.999	0.999	1.000
430	0.827	0.976	0.989	0.994	0.995	0.997	0.998	0.999	0.999	1.000
440	0.714	0.958	0.985	0.991	0.995	0.997	0.998	0.999	0.999	1.000
450	0.831	0.974	0.991	0.995	0.997	0.998	0.999	0.999	0.999	1.000
580	0.681	0.937	0.984	0.991	0.995	0.997	0.998	0.999	0.999	1.000
590	0.628	0.931	0.964	0.984	0.992	0.995	0.997	0.999	0.999	1.000
600	0.811	0.965	0.982	0.989	0.993	0.996	0.997	0.998	0.999	1.000
620	0.706	0.948	0.973	0.983	0.990	0.996	0.998	0.999	0.999	1.000

Table 11: Cumulative variance proportion of all PC's for every sample. Marked red depicts the four first PC's which represent the most of the data for all samples

Finally, PC scores time-series were created for top four Principal Components. The PC scores time-series have been created using "zoo" package in R. The PC scores time-series are the input for applying the Empirical Fluctuation Process, and thereby, the Change Detection Test, aim of this research. Figure 28 shows as example the PC scores time-series for training sample 1.

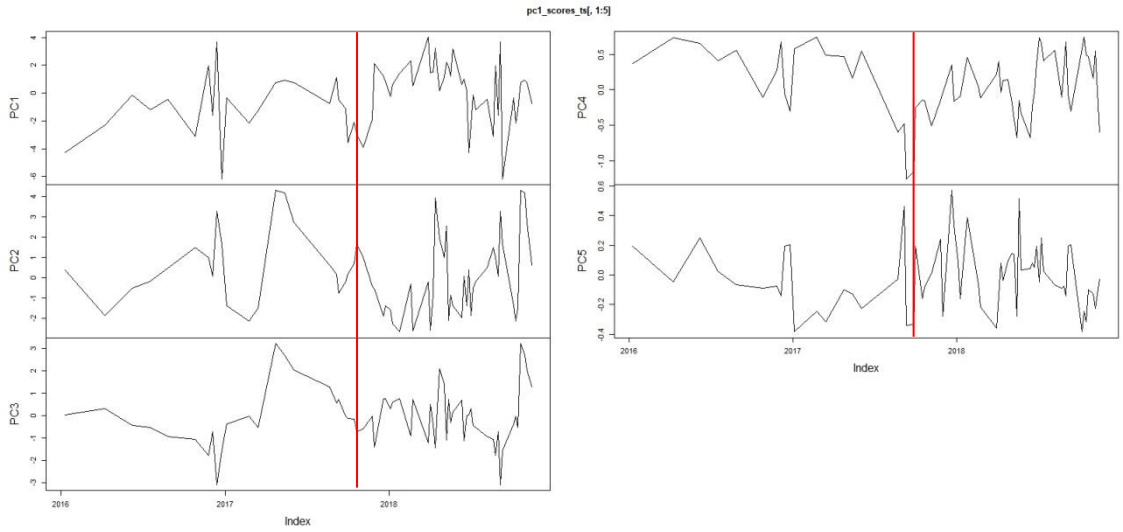


Figure 28: PC scores time-series for training sample 1 (five first PC's). The red line depicts the time change observed in the validation dataset.

5. RESULTS

This chapter first shows the results on change detection for the Normalized Difference Vegetation Index (NDVI) on the 20 training samples (section 4.1), followed by the results on change detection for the Principal Component Analysis (PCA) of the 20 training samples (section 4.2), succeeded by the overall results of the complete NDVI and PCA datasets (Section 4.3), and finally validation and accuracy assessment on the overall results will be provided (section 4.4).

After computing the OLS-MOSUM (Ordinary Least Squares – Moving Sum of residuals) test on each of the chosen samples, a Structural Change Test (SCT) is ran by the use of function "sctest" from the strucchange package in R (Zeileis et al., 2001). The OLS-MOSUM test plots show a change is detected when the plotted line intersects the margin boundaries lines in both extremes (superior and inferior) of the plot.

5.1. Change detection over NDVI time-series samples

For NDVI training sample pixels, in 60% of the known deforestation points were detected. As observed in Table 12, six training samples have p-values of less than 0.05. The MOSUM test plots considered the time-series as an index, and represents when a change has been detected. Some samples have EFP values that approximates considerably to the boundary line (e.g. s33), which could indicate a structural change that was not detected, and that will be evaluated after the validation analysis. On the other hand, for training samples known as forest (forest have not change during the time-series) no changes have been detected. Figure 29 and Figure 30 depicts the MOSUM-OLS test plots for all training samples.

Sample	MO	p-value	Sample	MO	p-value
1	0.85762	0.3115	400	0.98227	0.2115
33	1.1757	0.06783	410	0.47089	0.622
54	1.3041	**0.02182**	420	1.1419	0.08772
101	1.6677	**0.01**	430	0.9086	0.2706
110	1.4464	**0.01**	440	0.80854	0.3509
119	1.2741	**0.02831**	450	0.92545	0.2571
135	1.0616	0.1478	580	1.1218	0.09958
177	1.0849	0.1291	590	0.91641	0.2643
196	1.714	**0.01**	600	0.80865	0.3508
239	1.601	**0.01**	620	0.60781	0.5121

Table 12: MOSUM and p-values computed for NDVI dataset samples. Numbers in bold depict values below alpha, therefore, indicates a structural change have been detected.

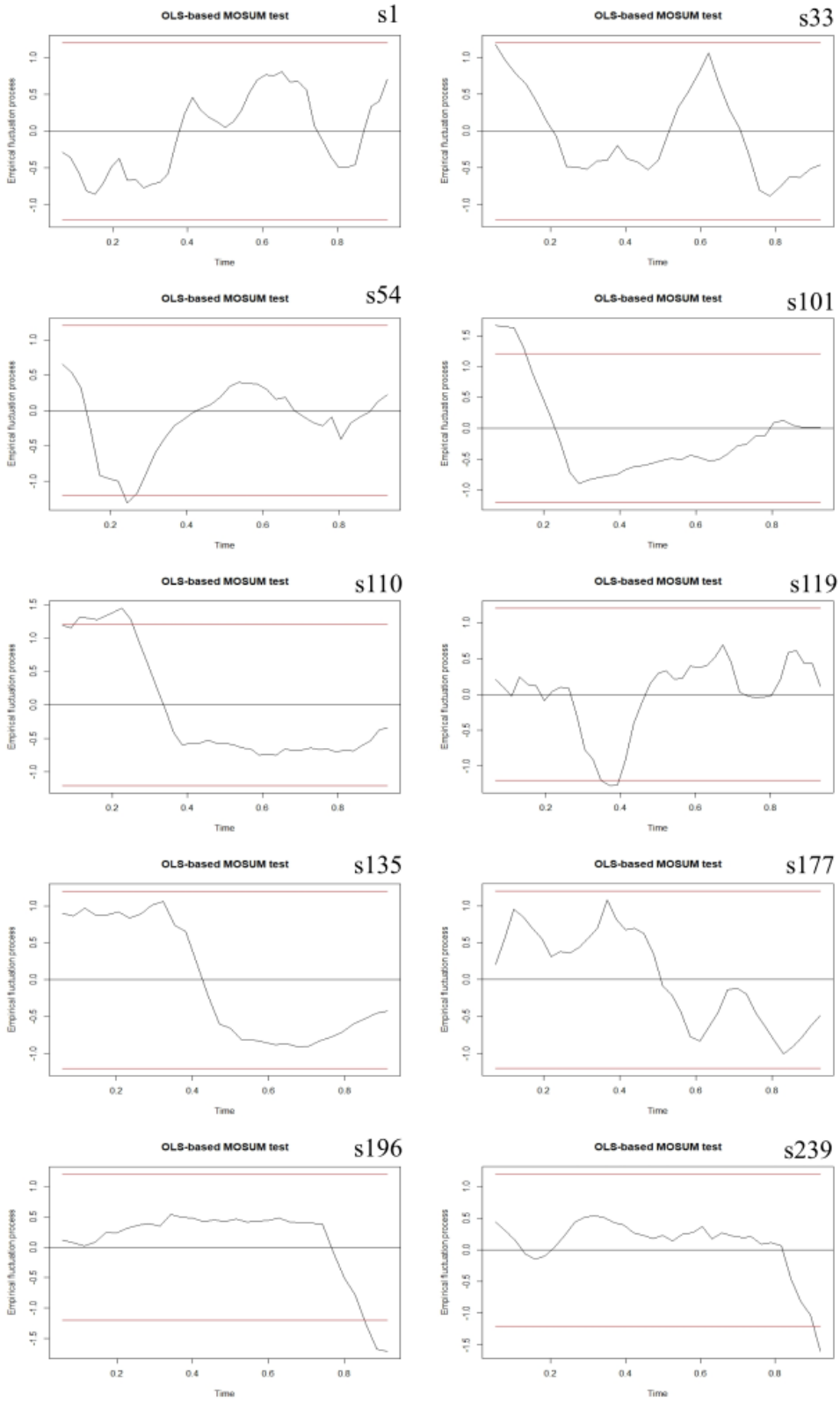


Figure 29: OLS-MOSUM test plot for first half of NDVI dataset samples

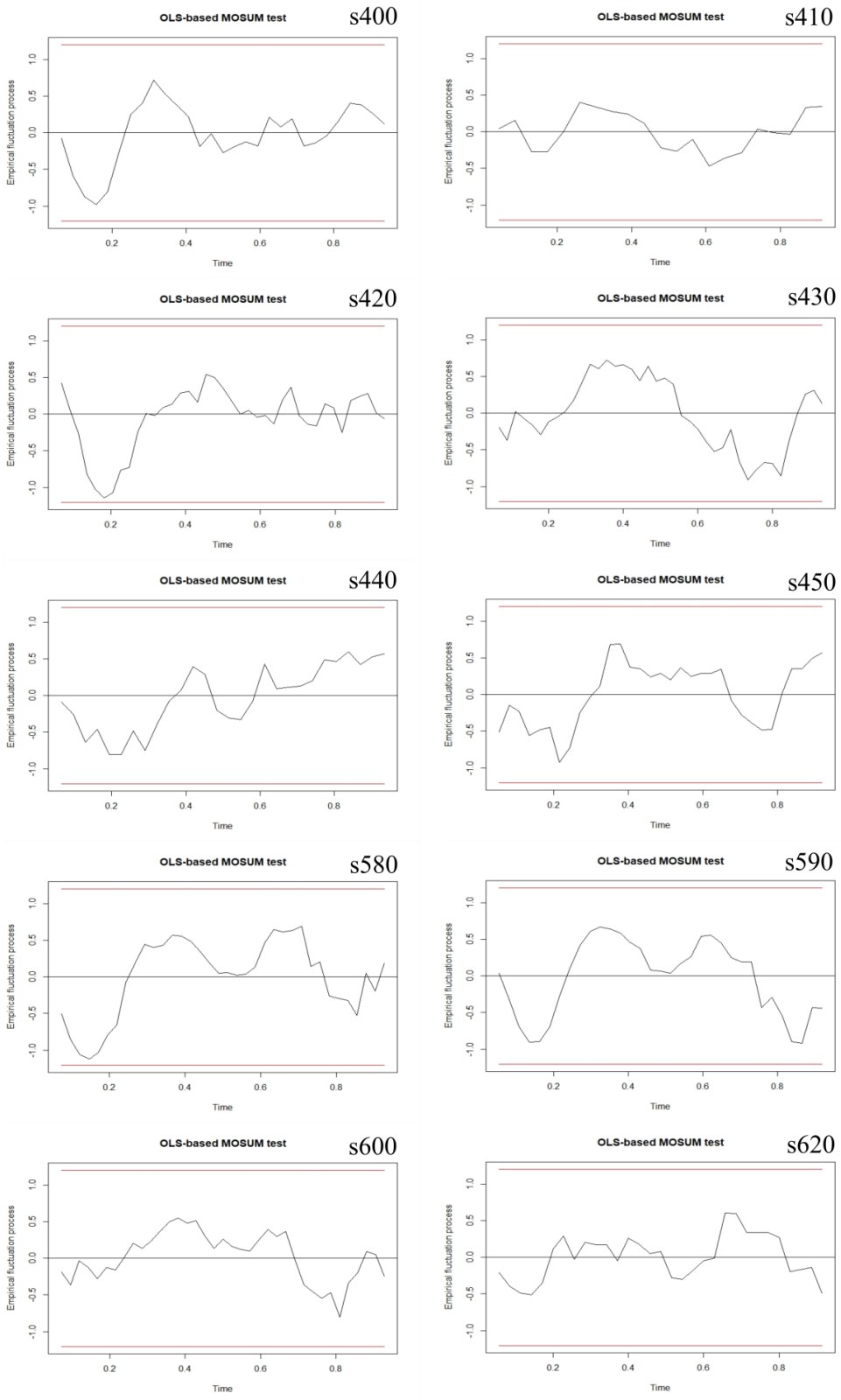


Figure 30: OLS-MOSUM test plot for second half of NDVI dataset samples

5.2. Change detection based on EFP over PCA time-series

The change detection method PCA based has several findings among each PC (Table 13). For training samples, PC1 have detected more structural changes. 50% of the known deforested pixels have resulting p-values less than 0.05. PC2 and PC4 detected four changes each on deforested known pixel, and PC2 detected one structural change on forest known pixel, while PC4 detected two changes. Few matches in structural change were found. In training sample 101 PC1 and PC4 detected a structural change, en training sample 110 PC1 and PC2, in training sample 135 PC1 and PC2, and finally training sample 177 in PC1, PC2 and PC3.

Also, PC2 and PC4 detected structural change on forest known pixels (e.g. training samples 430, 440 and 580)

Samples	PC 1	PC 2	PC 3	PC 4
1	0.142	0.08273	1.0144	**0.01**
33	0.5436	**0.03221**	0.08613	0.0799
54	0.08248	0.1664	0.3245	0.2544
101	**0.0208**	0.1558	0.4219	**0.01**
110	**0.01**	**0.04307**	0.5078	0.3575
119	0.2325	0.3779	0.1258	**0.01**
135	**0.02929**	0.303	0.3566	**0.01**
177	**0.4923**	**0.02976**	**0.04815**	0.1119
196	0.01008	**0.01**	0.43	0.5649
239	**0.01**	0.2535	0.3805	0.2396
400	0.5618	0.428	0.305	0.5323
410	0.4766	0.4302	0.3156	0.1951
420	0.4	0.08389	0.4782	0.64
430	0.2855	**0.01**	0.4519	0.1375
440	0.4945	0.3225	0.396	**0.04156**
450	0.4234	0.3395	0.5484	0.367
580	0.4769	0.2012	0.3694	**0.0203**
590	0.4684	0.2601	0.2646	0.2002
600	0.3435	0.4198	0.548	0.5539
620	0.6201	0.5341	0.2189	0.5772

Table 13: P-values computed for the four first PC's in all samples. Numbers in bold depict values below alpha, therefore, indicates a structural change have been detected.

Figures 31 and 32 shows the OLS-MOSUSM test plots for all training samples computed for PC1. The plots for the resting three chosen PC's can be found in the Appendix on pages 67 to 72.

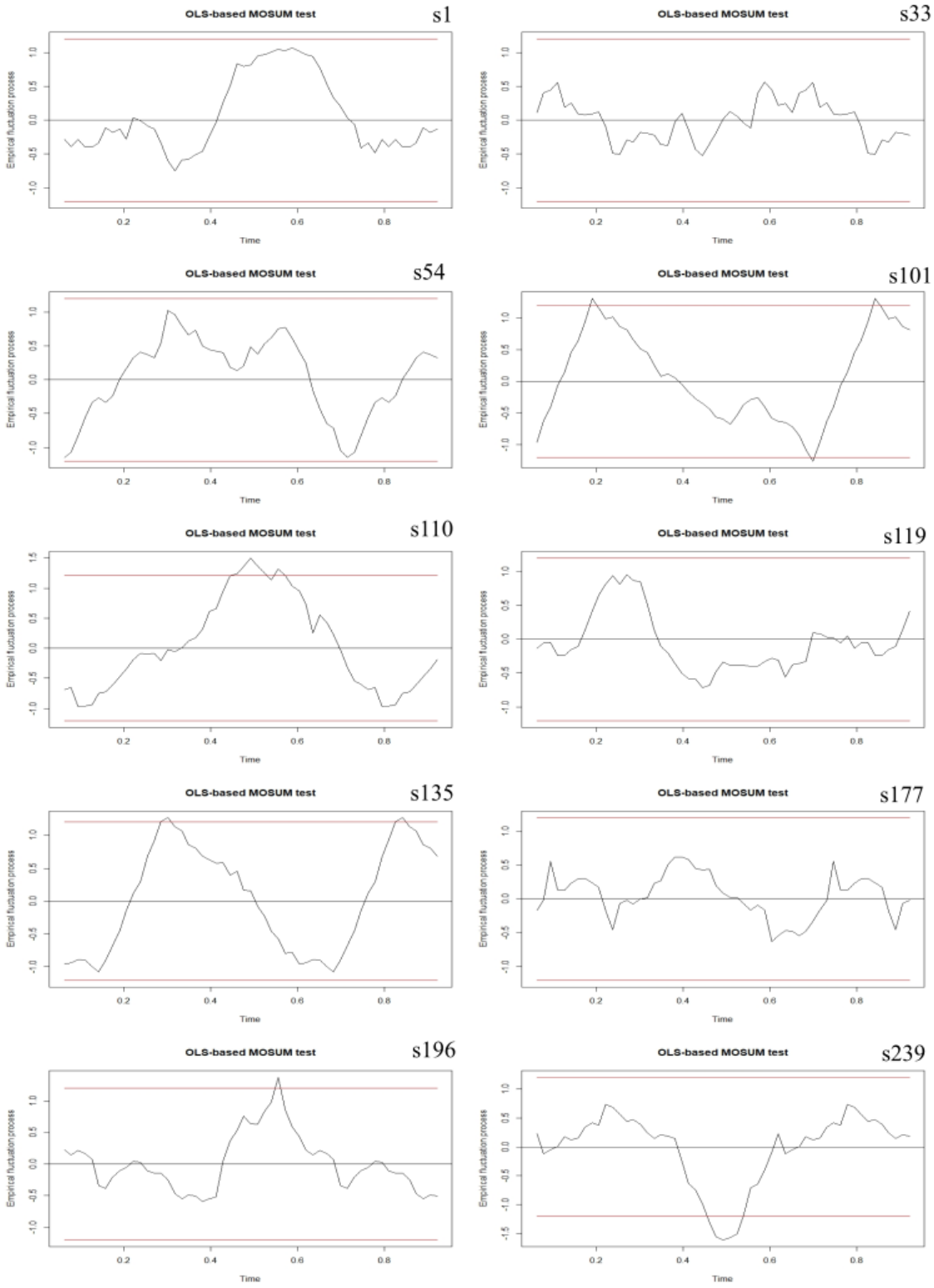


Figure 31: OLS-MOSUM test plot for known deforested training samples computed with PC1 loadings

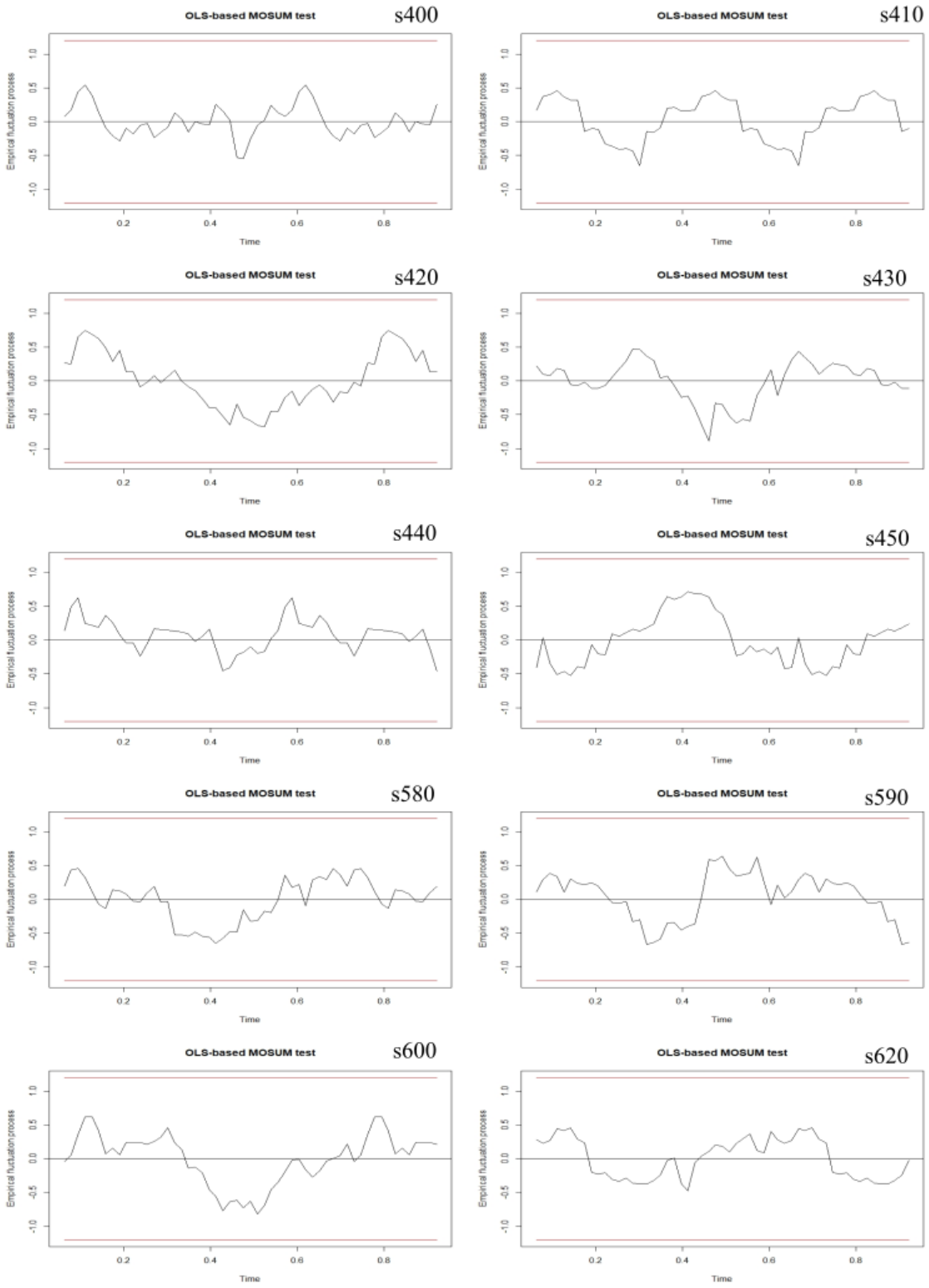


Figure 32: OLS-MOSUM test plots for known forest training samples computed with PC1 loadings

5.3. General results for complete NDVI and PCA datasets

In this section, general results for NDVI dataset and PCA dataset are explained. Here, Empirical Fluctuation Test and Structural Change Detection Test were applied to each dataset arrays. The dimensions of NDVI array are 1:623, 1:63 (two-dimensional), which represents the sample/validation points and time-series respectively. Dimensions for the PCA array are 1:623, 1:10, 1:63 (three dimensional), which represents sample/validation points, spectral bands and time-series respectively.

Figure 33 shows the findings of structural change detection over all datasets. In latter Figure, is clearly observable that have decided to use the first four PC's was good decisions due the majority positive structural change detection are founded there. Here is important to mention that there were 275 sample/validation points that corresponded to deforestation and 348 to forest. When a structural change detection test results True (p-value < 0.05), the supposition is that deforestation happened, though this is only confirmed after the validation process (section 4.4). A higher amount of findings happened on the NDVI dataset, 240 true values, while PC2 is the second higher in amount of findings with 188. PC4 and PC1 have 112 and 111 respectively.

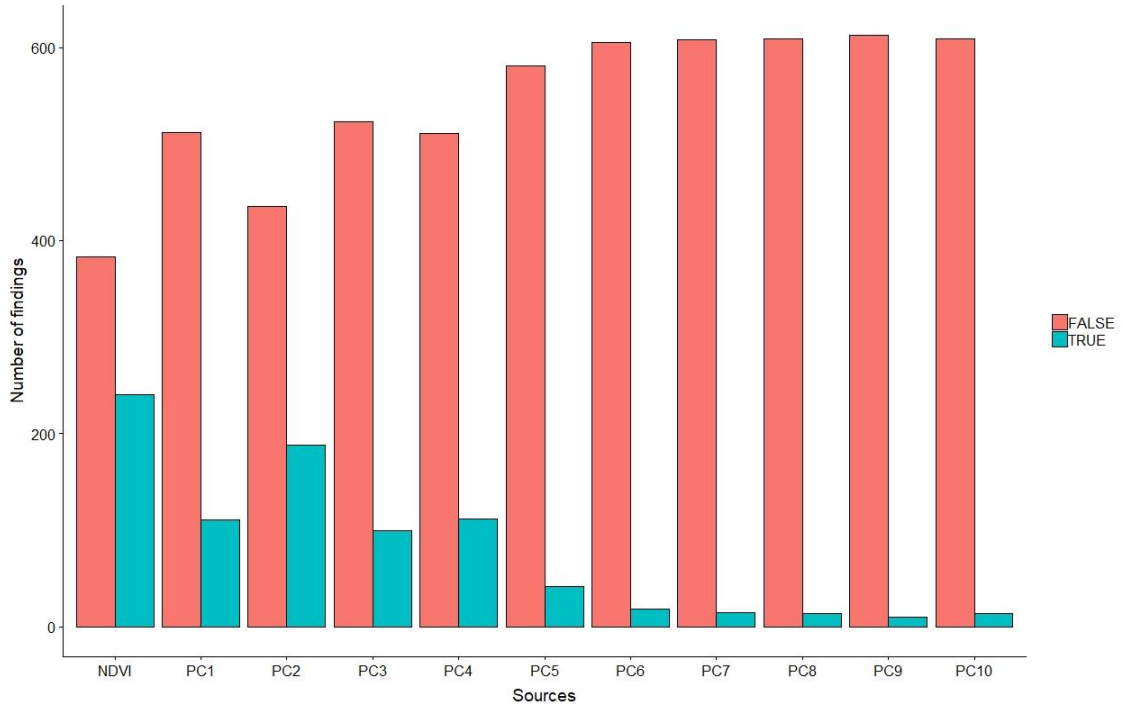


Figure 33: Overall Structural Change Detection findings by EFP in NDVI and PCA datasets. True depicts a structural changed was detected, and False depicts no structural change was detected

Considering only PC's findings, in Figure 34 can be observed that PC2 is more sensitive to forest disturbances in Sentinel-2 dataset. PC1, PC3 and PC4 have almost the same amount of true values, while PC2 is closer to NDVI findings.

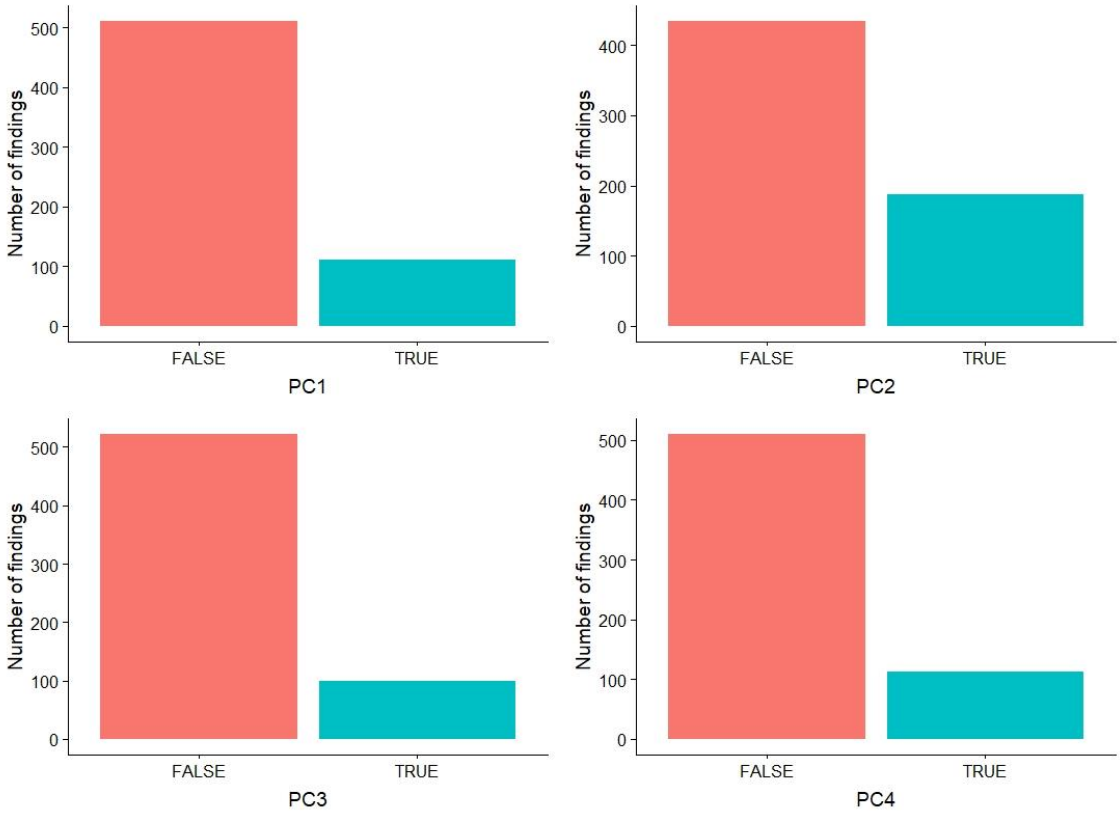


Figure 34: Structural Change Detection findings by EFP on the first four PC's. True depicts a structural changed was detected, and False depicts no structural change was detected

5.4. Validation and accuracy assessment

5.4.1. Validation

The validation and accuracy assessment method used here is explained in Section 3.9. The figure of merit was elaborated using the findings of Section 4.3 and the validation dataset (section 3.3). Table 12 shows the abbreviation and description for each class.

CLASS	ABBR	DESCRIPTION
TRUE POSITIVE	TP	Deforestation matches between EFP and validation dataset
TRUE NEGATIVE	TN	Forest unchanged matches between EFP and validation dataset
FALSE NEGATIVE	FN	Forest unchanged in EFP results are identified as Deforestation in validation dataset
FALSE POSITIVE	FP	Deforestation in EFP results are identified as forest unchanged in validation dataset

Table 14: Figure of merit's classes and description for validation and accuracy assessment process

Here is important to observe TP values in order to see which source data was able to find a larger amount of forest disturbances (structural change), even though considering the TN values is also important because the actual unchanged forest condition was not confused with other state. The results of the figure of merit showing deforestation and unchanged forest are reported on Table 15. NDVI was able to detect more TP than all

the PC's with 188 matches over the validation dataset. PC2 is the PC to have more positive matches of the four PC chosen, with 145 TP matches over the validation dataset. PC1 and PC4 have relatively the same amount of TP's matches with 78 and 77 respectively. PC1 was able identify a larger amount of True Negative matches with 315, and PC2 and PC4 had 306 and 313 respectively. NDVI's TN matches are below the before mentioned PC's with 296 matches.

	NDVI	PC1	PC2	PC3	PC4
TP	188	78	146	47	77
TN	296	315	306	295	313
FN	87	197	129	228	198
FP	52	33	42	53	35

Table 15: Figure of merit matrix results for NDVI and PC's. TP(True Positive), TN (True Negative), FN (False Negative), FP (False Positive)

5.4.2. Accuracy assessment

To assess accuracy, method explained in Section 3.9 was used. Here Figure of Merit (FOM), User's Accuracy (UA), Producer's accuracy (PA) and Overall Accuracy (OA) was measured. Table 16 shows the results of the latter mentioned measurements for all NDVI and PC's outputs. Following with the results in previous section, NDVI achieved higher accuracy in general; having a 57% in FOM, while PC2 with the second higher FOM accuracy achieved a 46%. PC1 and PC4 achieved 25% each while PC3 has the lower FOM with 14%. UA achieved by NDVI and PC2 resulted the same with 78%, while PC1 and PC2 achieved close percentages with 70% and 69% respectively. PC3 also achieved the lower value here with 47%. Producer accuracy's higher values were achieved once again by NDVI and PC2 with 68% and 53% respectively while PC1 and PC4 achieved significantly lower PA with 28%. PC3 achieved the lowest with 17%. Respect with OA, NDVI achieved the higher percentage with 78% followed by PC2 with 73%. PC1 and PC4 achieved the same OA with 63%, while PC3 achieved the lower OA with 55%.

	NDVI	PC1	PC2	PC3	PC4
Figure of merit	57%	25%	46%	14%	25%
User's accuracy	78%	70%	78%	47%	69%
Producer's accuracy	68%	28%	53%	17%	28%
Overall accuracy	78%	63%	73%	55%	63%

Table 16: Accuracy assessment on structural change detection results for NDVI and PC. Figure of Merit (FOM), User's Accuracy (UA), Producer's Accuracy (PA) and Overall Accuracy (OA)

6. DISCUSSION

This chapter will discuss in a first instance the importance of sample and validation dataset for the method used in the present research (section 5.1). Subsequently, we will discuss the findings obtained for both datasets (NDVI and PCA) on deforestation detection through Empirical Fluctuation Process (section 5.2).

6.1. Importance of sample and validation dataset

Sample and validation dataset plays a fundamental role in the method used in the present research. The method here proposed is a pixel-based (Desclée et al, 2006; Aguirre-Gutierrez et al., 2012; Chen et al, 2012; Hussain et al, 2013) change detection method, thus, the data gathered from satellite datasets is based on the accuracy of the sample and validation points.

Validation data is often a limitation for land-cover and change detection studies due to the lack of it or because the good quality of the data, the costs that represents to elaborate new datasets (Cohen et al., 2010; Lu et al, 2015). Previous studies in the same topic generated the validation dataset based on Landsat images, complementing them with images from Google Earth or Bing Maps (Hamunyela et al., 2016; Lu et al., 2017). Obviously this concerns limitations due to spatial resolution of Landsat which can result in misclassified pixels. The use of web engines as Google Earth and Bing Maps do help, though a posterior quality revision must be done in order to match the points generated there with the Landsat image pixels.

In the present research we approach this issue by getting access to high resolution imagery from PlanetScope (3m resolution) and RapidEye (5m resolution) satellites as explained in Section 3.3. This free-resource available for researches means a great opportunity to improve sample and validation datasets. Even though, a quality assessment needs to be done as well. In the present research, spatial displacements of 15 to 30 meters where found when observing the validation dataset made with high resolution imagery over the Sentinel-2 (10 m and 20 m resolution) images. Consequently, manual alignment of the points was done in order to make them match with the pixels identified as reference pixels.

This findings bring up the that sample and validation datasets are still a critic step in change detection process, and which needs to be consciously developed and managed in the oncoming researches, thus the quality of results depends on the quality of sample and validation data.

6.2. Seasonality

West New Britain province in Papua New Guinea, where the study area is located, has a Humid Tropical Climate. Mean annual rainfall in the study area is between 4000 and 6000 mm and there is not a clear seasonality over the year (McAlpine, 1983). Food and Agriculture Organization (FAO) establishes for the study area vegetation as Lowlands rainforest. This explains the above where evergreen rainforest is not conditioned by seasonality. Therefore, having an overall accuracy of 78% on fitting the mean model to NDVI time-series tells how well can the method perform in areas where seasonality does not need be addressed directly.

6.3. Deforestation detection results on NDVI and PCA data

In section 4, the findings for forest change detection based on Empirical Fluctuation Process are explained. Considering the training sample points (40 samples) that were selected to analyze in detail the method, we can observe that a small portion of the data can confuse the overall performance of the method. This condition does not apply to NDVI dataset, as NDVI results showed from the beginning to have a better performance in what refers to change detections. As seen in Table 12, EFP-based change detection on NDVI achieved more findings on the "Deforestation" known samples (first ten) than the four first PC's. In the training samples, the PC that achieved a better performance was PC1 with 50% of matches, while PC2 and PC4 achieved 40% matches. PC3 always showed to have a lower performance than the others. This is important to discuss, because on the overall results with complete datasets, PC2 achieved the best performance for all PC's with an overall accuracy of 73%, just 5 percent below NDVI overall accuracy, and 10% higher than PC1 and PC4. Also, PC2 achieved the same user's accuracy than NDVI with 78% which is 8% higher than the achieved by PC1 (Table 16).

Here is important to mention that NDVI dataset was calculated from Sentinel-2 (10 m spatial resolution), which means the spectral information of bands 8 and 4 did not change from its original, while PCA calculations were effectuated over upscale 20 meters resolution images (see Section 3.4.2), which can change spectral information, or discard valuable spectral information from pixels (Marceau & Hay, 1999; Du et al., 2016). Therefore, the results obtained by PC2 shows a great performance of the method by using PCA, and would be important for future studies to downscale the six 20 meters spatial resolution to 10 meters. Several fusion methods have been developed in the last decade to downscale remote sensing products (Li, 2000; Ranchin & Wald, 2000; Svab & Ostir, 2006; Thomas et al., 2008; Choi & Kim, 2010; Metwalli et al., 2010; Wang et al., 2015; Wang et al., 2016a), and specific fusion methods has been developed in recent years for Sentinel-2 images with very encouraging results in keeping spectral information (Wang et al., 2016b; Wang & Atkinson, 2018). Thus,

research on forests change detection based on EFP over Sentinel-2 imagery is promissory and brings opportunity to improve accuracy.

Considering previous researches that can be somehow comparable to the present research, it is worth to highlight the study made by Lu et al (2017), which used "Breaks For Additive Season and Trend" (BFAST) Monitor framework. The BFAST method is also based on Empirical Fluctuation Process, and in the cited study a comparison between the performances of the method was done over several vegetation indices (NDVI between them) and PCA. The BFAST method conducted by Lu et al (2017) considers other parameters that were not used for the present research as the creation of a Seasonality Reduced Index (SRI). The latter study was based on Landsat imagery of 30 m spatial resolution and six spectral bands were used for PCA, and a 30 years' time-series dataset was used. The study evaluated two study sites: a dry tropical forest in Bolivia, and a moist tropical forest in Brazil. Because the atmospheric, weather and vegetation conditions of Brazilian study site are most alike to the study area of the present research (West New Britain, Papua New Guinea), those findings will be compared will the results here explained.

Lu et al (2017) accuracy assessment's findings using Landsat imagery was done comparing time, and some differences can be found with the present research. Lu et al (2017) achieved an overall accuracy of 63.2% and 47.4% for NDVI and PCA respectively, compared with a 78% and 73% for NDVI and PCA respectively in the present research. In Lu et al (2017), user's accuracy achieved 23.9% for NDVI and 19.6% for PCA, while the present research achieved 78% for NDVI and PCA. On one hand producer's accuracy in Lu et al (2017) for NDVI was of 44.3% and for PCA 64.6, and on the other hand the present research achieved 68% and 53% for NDVI and PCA respectively. Lastly, in what refers to Figure of Merit accuracy, Lu et al (2017) achieved 18.4% and 17.7% for NDVI and PCA respectively, while in the present research 57% was achieved for NDVI and 46% for PCA. Considering the cited research used time to measure accuracy, also other causes can be discussed; a larger spectral resolution of Sentinel-2 and a higher spatial resolution can improve the performance of EFP, which would need to be studied in order to understand the behavior of both satellites respecting the method. The validation dataset on the present research was done with high-resolution imagery, which lowers the risk of having False Negatives and False Positives in the final results. Here, is necessary to mention once again that sample and validation dataset is critical; as a good quality validation dataset gives you a certain staring point, as you know that in that specific pixel deforestation happened, and the method only tries to prove what you already know if you have a good quality sample and validation dataset. Another reason for the difference in accuracy results can be the geographical location, considering that the present research study site is closer to the equator and Atmospherical dynamics can differ.

Sentinel-2 constellation has been gathering information since 2015 with one satellite and both satellites working at the same time since 2017. This makes that Sentinel-2 time-series are not extensive as other satellites used for this purpose (e.g. Landat satellites). Even though, this study shows that inside the time-frame where the study

was developed, was possible to detect structural changes in the forest of the study area. One of the challenges here, and considering the high cloud presence in the area, is to perform the best cloud-masking possible, in order to take advantage of every available pixel that can match with the analysis. Considering that in the present study only 63 time-series where used, and from those the 51% had above 70% of clear land pixels, the results for the time-period used in the study are satisfactory. Comparing time-series availability in other similar studies, Lu et al (2017) used Landsat imagery nad the study was developed in two study sites, where in one of the – Bolivian site – 444 images where available, and on the second – Brazilian site – 225 images were used. This shows the considerably difference in time-series availability between satellites, though Sentinel-2 have the advantage of having a 5 days temporal resolution with the two satellites (e.g. Sentinel-2A and Sentinel-2B) and further in time Sentinel-2 time series will be more abundant, and therefore better contribute to forest dynamics studies. The 5 days temporal resolution of Sentinel-2 constellation also contributes to have a higher probability of free-cloud images.

7. CONCLUSION

The present research had two main objectives. The first objective was to assess the method (Forest Change Detection based on Empirical Fluctuation Process) and evaluate the performance of it over NDVI and PCA data using Sentinel-2 imagery. The first objective was characterized on the first research sub-question: 1) To what extent Structural Change Detection on Sentinel-2 imagery based on NDVI and PCA differs? The second objective was to evaluate if the method can achieve significant accuracy results by working on short time-series with high amount of cloud presence. The second objective was described by the second and third research sub-questions: 2) How does short amount of time-series datasets affect Structural Change Detection based on Empirical Fluctuation Process? 3) How does high amount of cloud-pixels in the time-series dataset affect the results of Structural Change Detection based on Empirical Fluctuation Process?

Concerning the first objective, was observed that NDVI results achieved higher accuracy levels in all the realized accuracy measurements, having detected more true deforestation events in the Area of Interest than the PCA results. Here is important to consider that NDVI dataset had a higher resolution than the PCA dataset, and could affect the overall results. In despite of the latter explained, PCA results, specifically PC2 results, got significantly high accuracy levels than other PC's and compared to NDVI. The PCA assessment on the method showed the importance of studying all PC's before making a decision on which one should be used to detect forest disturbances. At the beginning of the study was hypothesized that PCA was going to perform better, though overall results showed NDVI performed better in this case by a marginal difference and with better spatial resolution than PCA, as was previously mentioned.

For the second objective regarding to time-series and cloud presence on the study area, the hypothesis done at the beginning of the research established that the short amount of time-series and high amount of cloud-pixels was not going to affect the findings for the time period over study. The final results shows a level of accuracy of 78% on NDVI and 73% on PCA, which can be considered as a feasible accuracy level, considering that the study area is highly clouded over the year and that the time period under study was of 3 years. Therefore, the method of Forest Change Detection based on Empirical Fluctuation Process has a good performance with short time-series in cloudy areas.

The results of this study add knowledge on forest change detection methods based on Empirical Fluctuation Process (EFP) over Normalized Difference Vegetation Index and dimension reduction using Principal Component Analysis.

8. APPENDIX

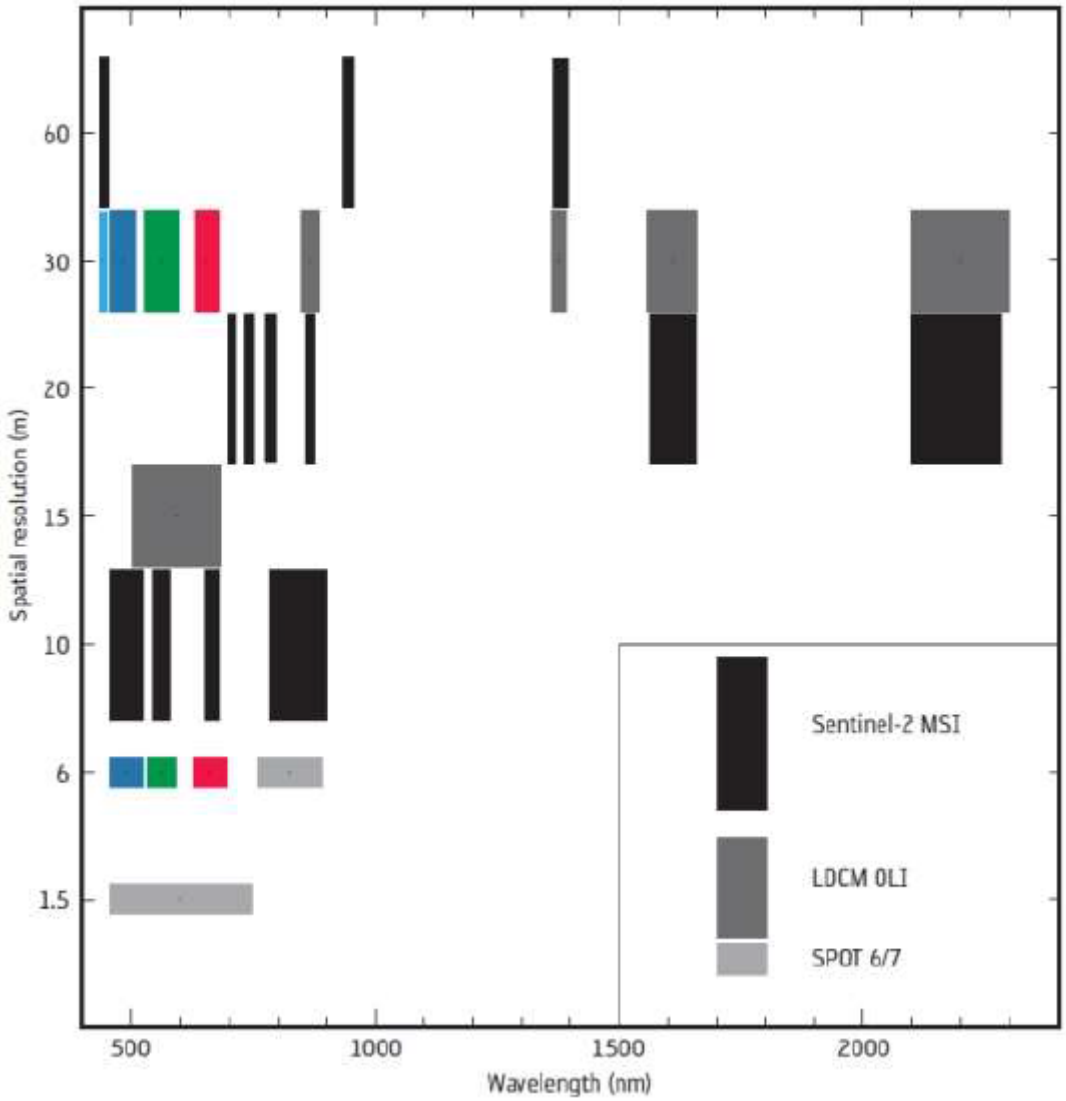


Figure 35A: Comparison of Spatial Resolution and Wavelength Characteristics of SENTINEL-2 Multispectral Instrument (MSI), the Operational Land Imager (OLI) On-Board LANDSAT-8, and SPOT 6/7 Instruments (taken from ESA, 2015)

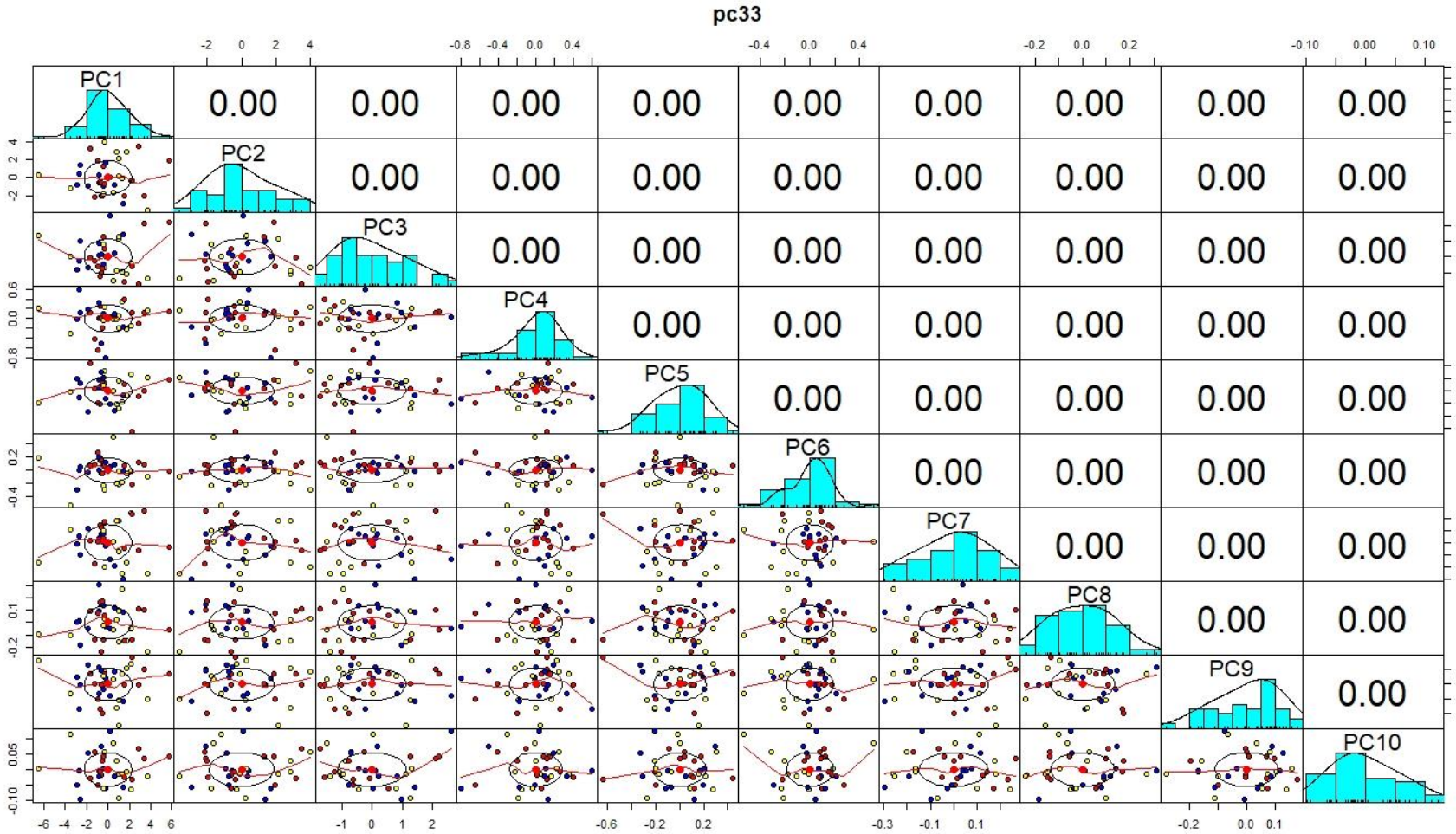


Figure 36A: All possible scatter plots between the Principal Components for training sample 33

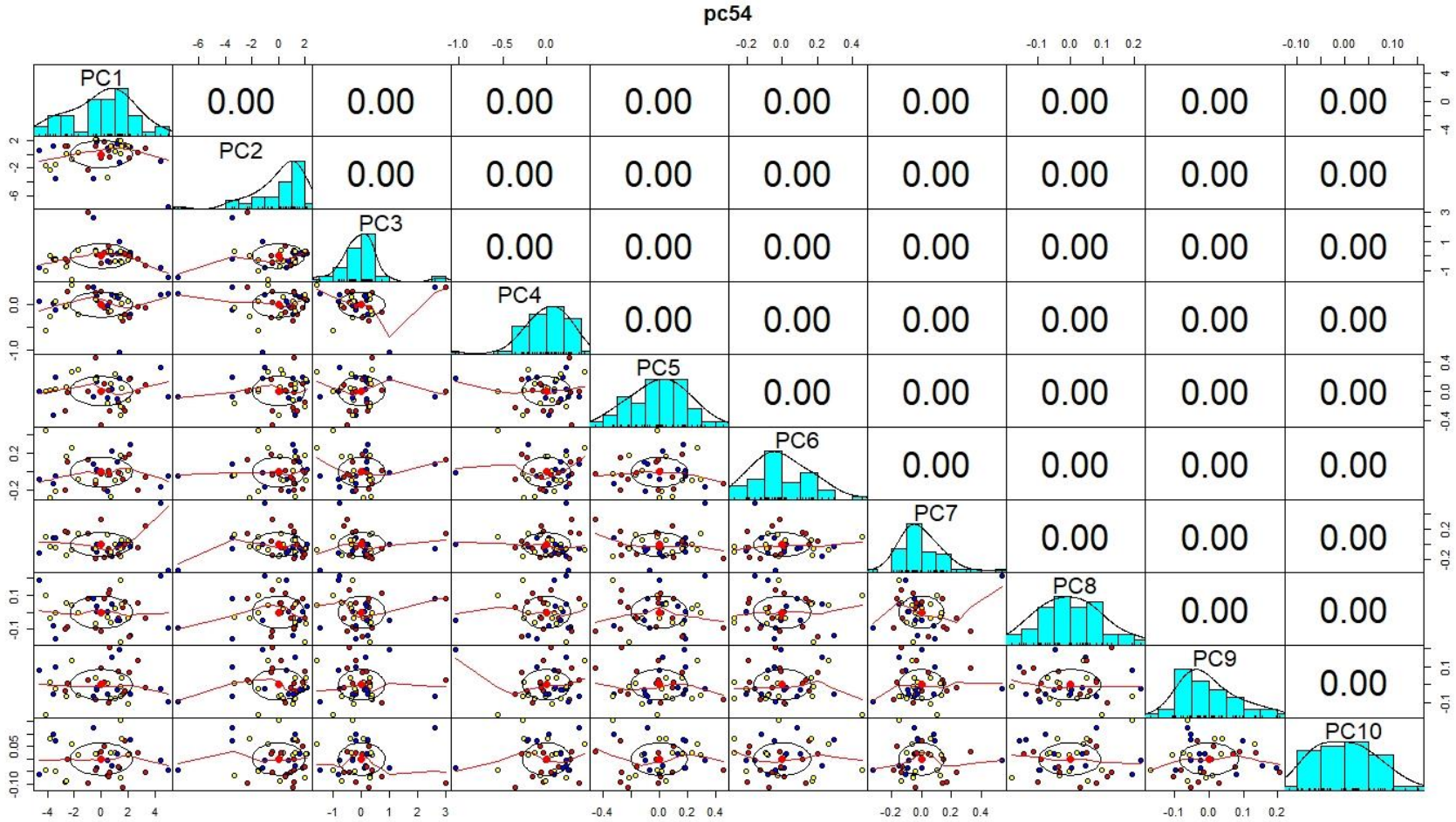


Figure 37A: All possible scatter plots between the Principal Components for training sample 54

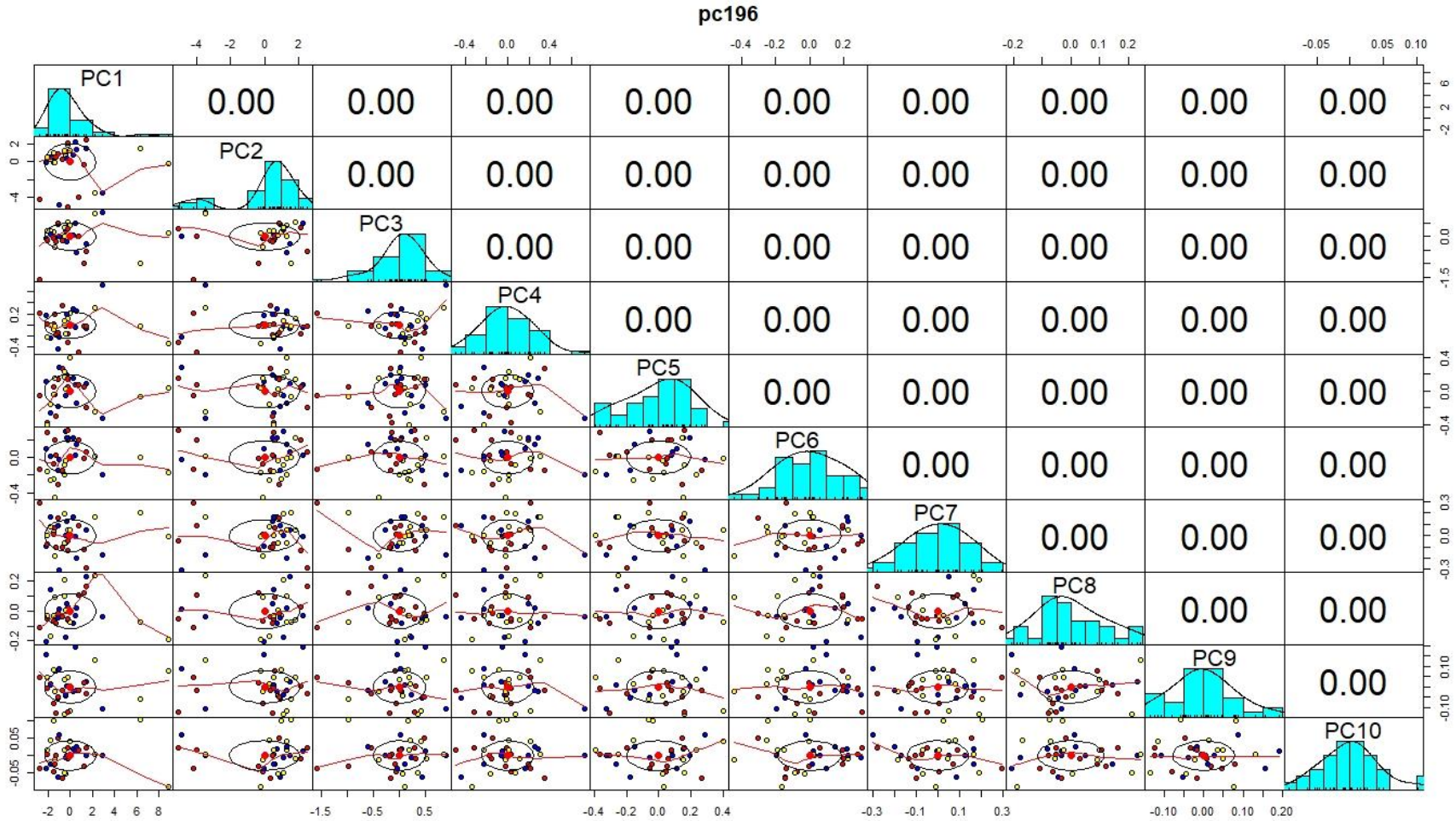


Figure 38A: All possible scatter plots between the Principal Components for training sample 196

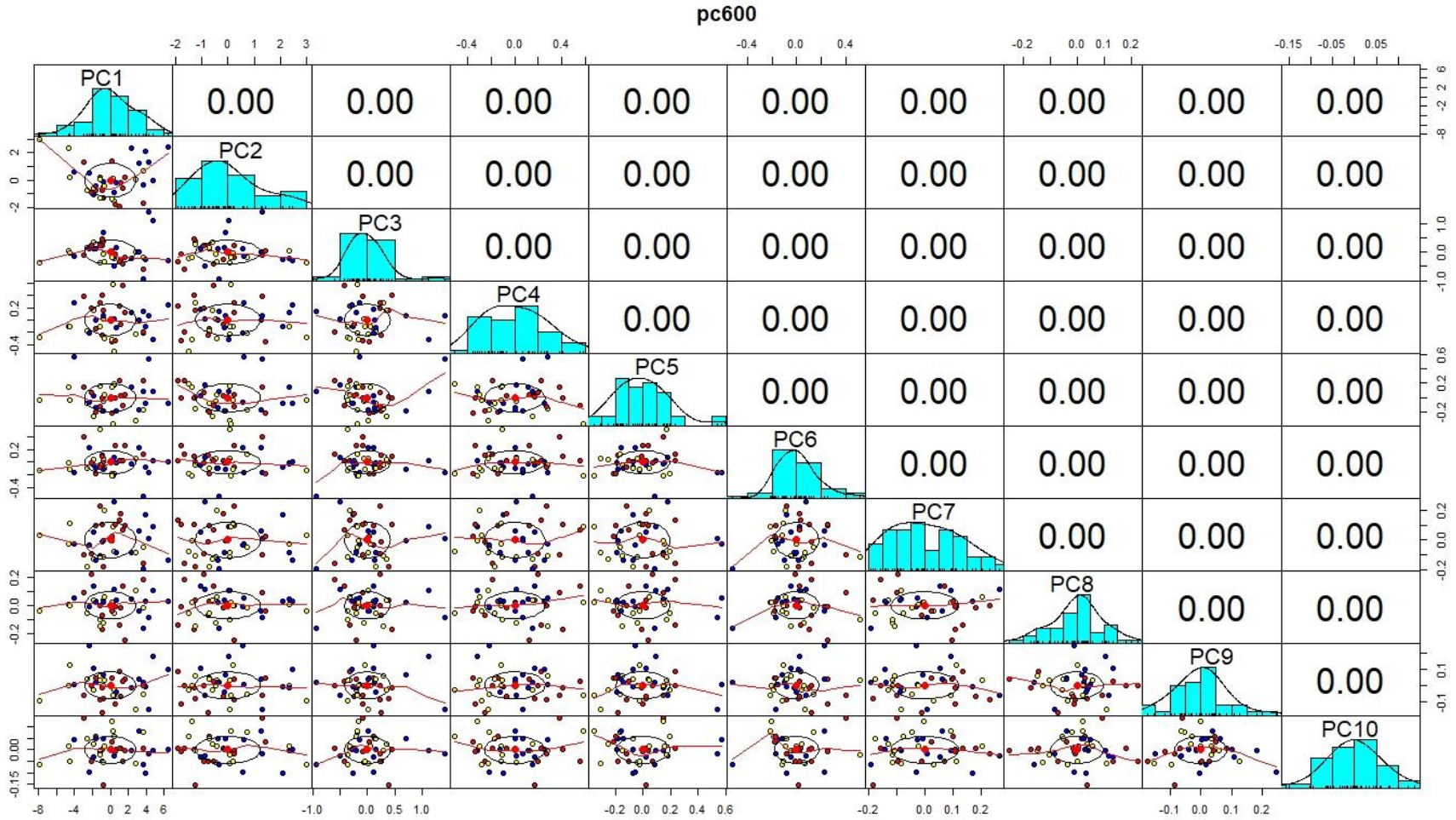


Figure 39A: All possible scatter plots between the Principal Components for training sample 600

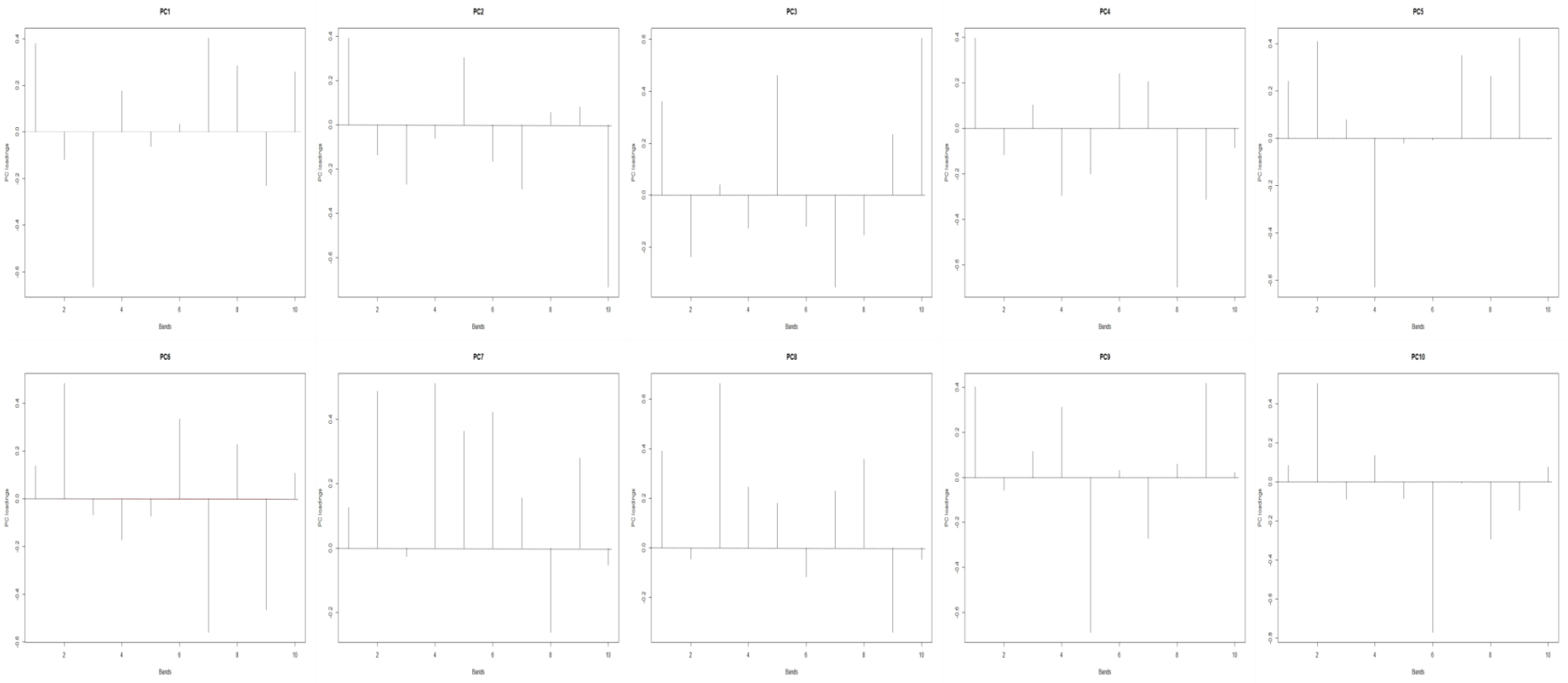


Figure 40A: PC loadings (rotation) of all bands of training sample 110. The sequence of the bands (1-10) corresponds to the following: Bands 1-3 to visible; Bands 4-6 to Red Edge; Bands 7 and 10 to NIR; and Bands 8 and 9 to SWIR.

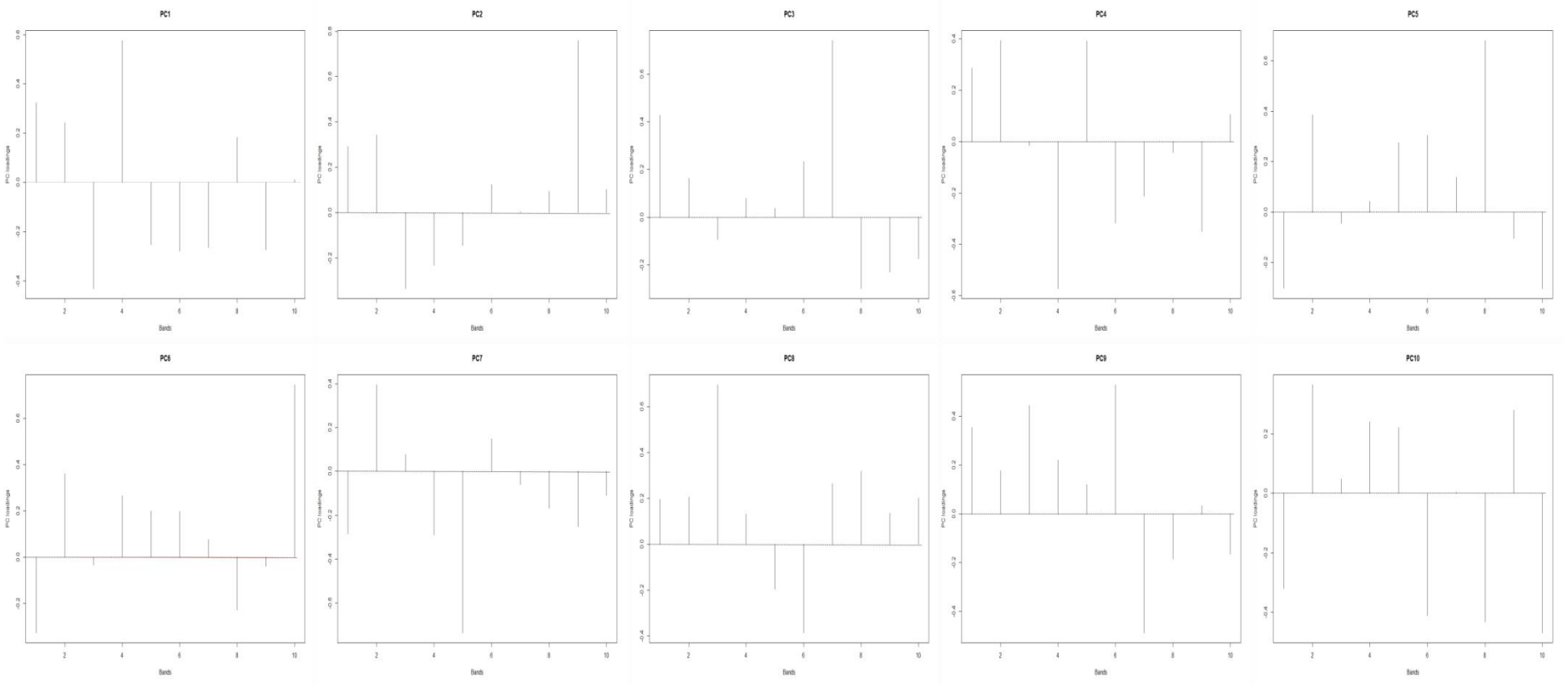


Figure 41A: PC loadings (rotation) of all bands of training sample 119. The sequence of the bands (1-10) corresponds to the following: Bands 1-3 to visible; Bands 4-6 to Red Edge; Bands 7 and 10 to NIR; and Bands 8 and 9 to SWIR.

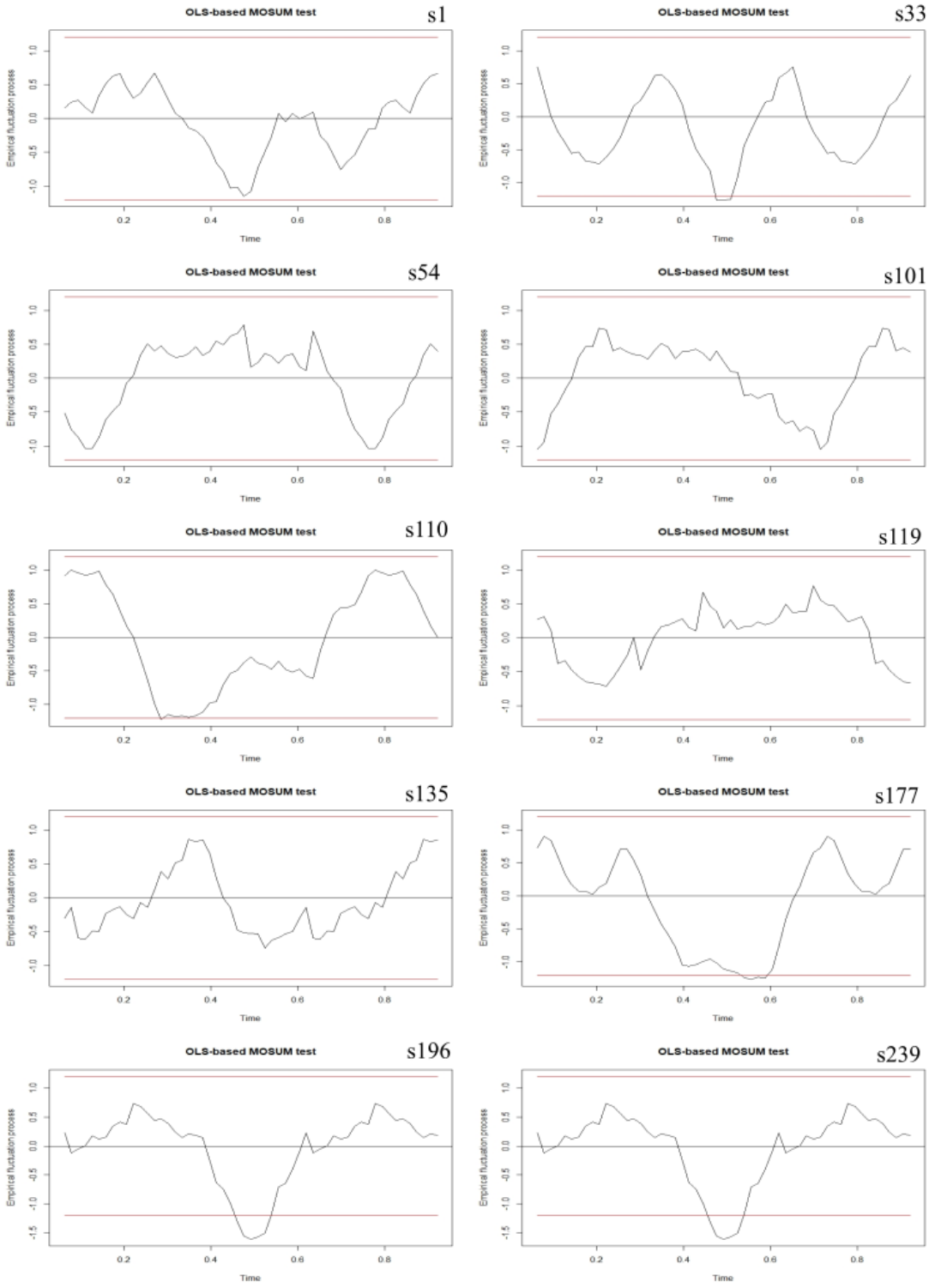


Figure 42A: OLS-MOSUM test plot for known deforested training samples computed with PC2 loadings

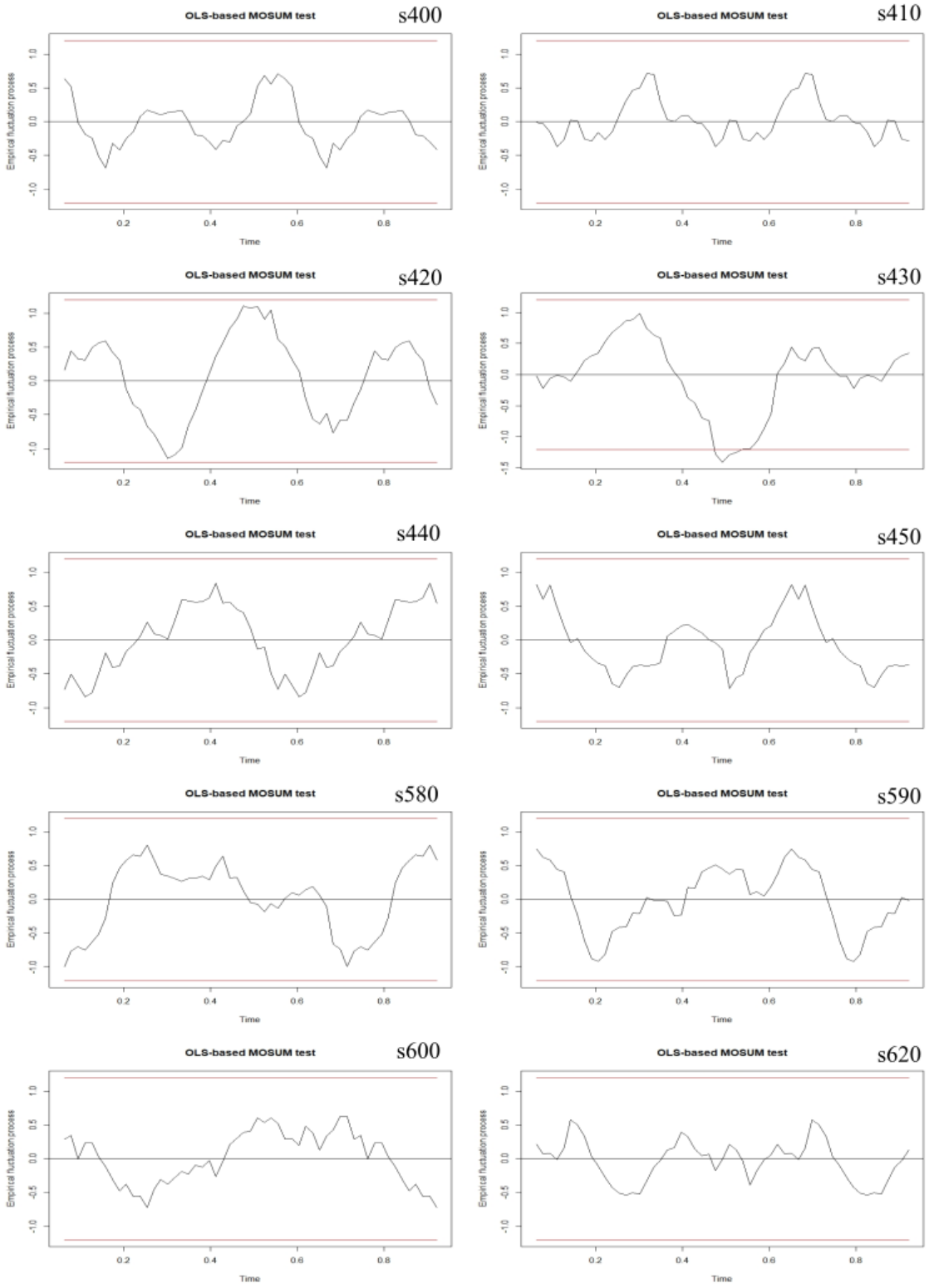


Figure 43A: OLS-MOSUM test plot for known forest training samples computed with PC2 loadings

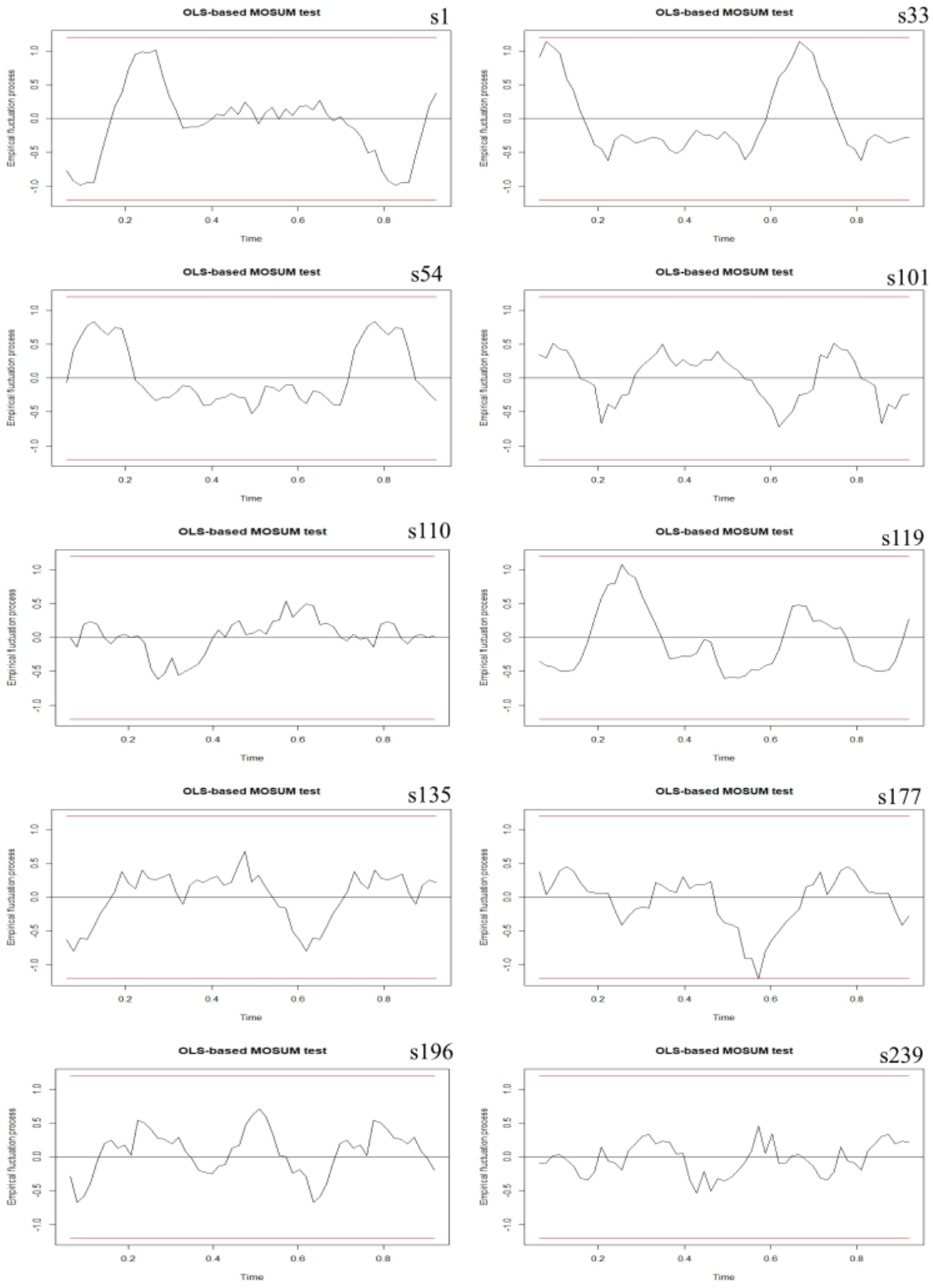


Figure 44A: OLS-MOSUM test plot for known deforested training samples computed with PC3 loadings

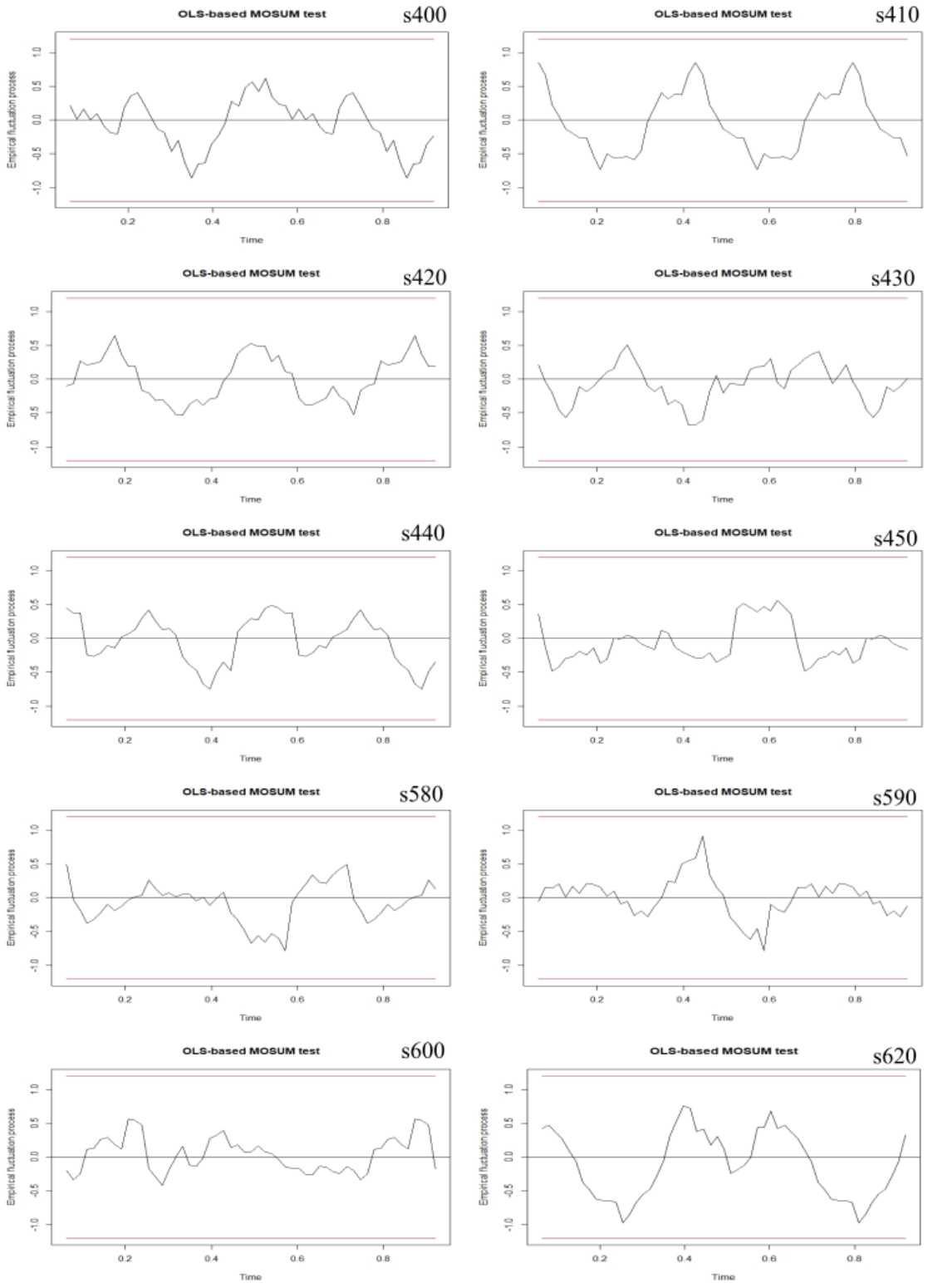


Figure 45A: OLS-MOSUM test plot for known forest training samples computed with PC3 loadings

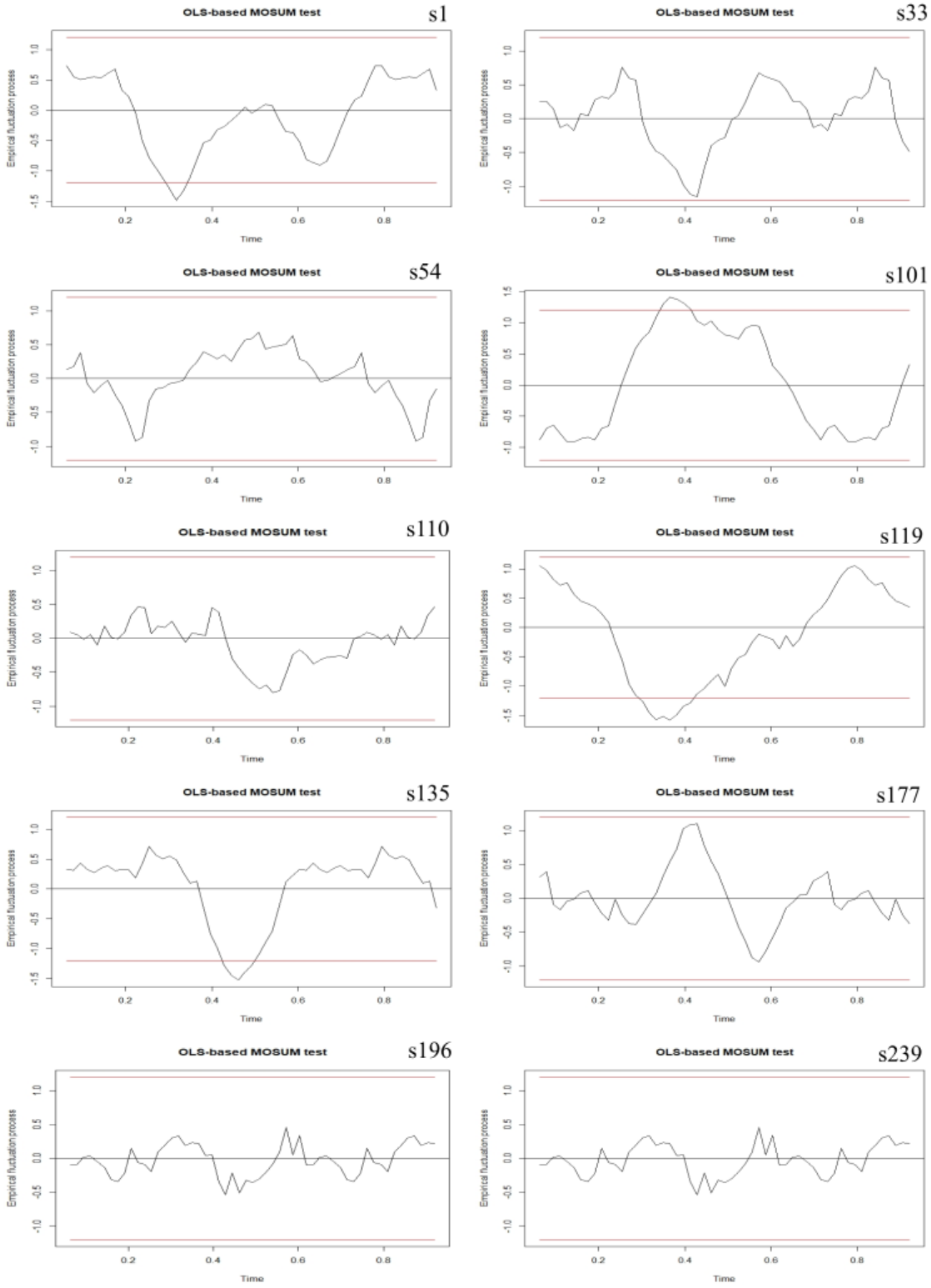


Figure 46A: OLS-MOSUM test plot for known deforested training samples computed with PC4 loadings

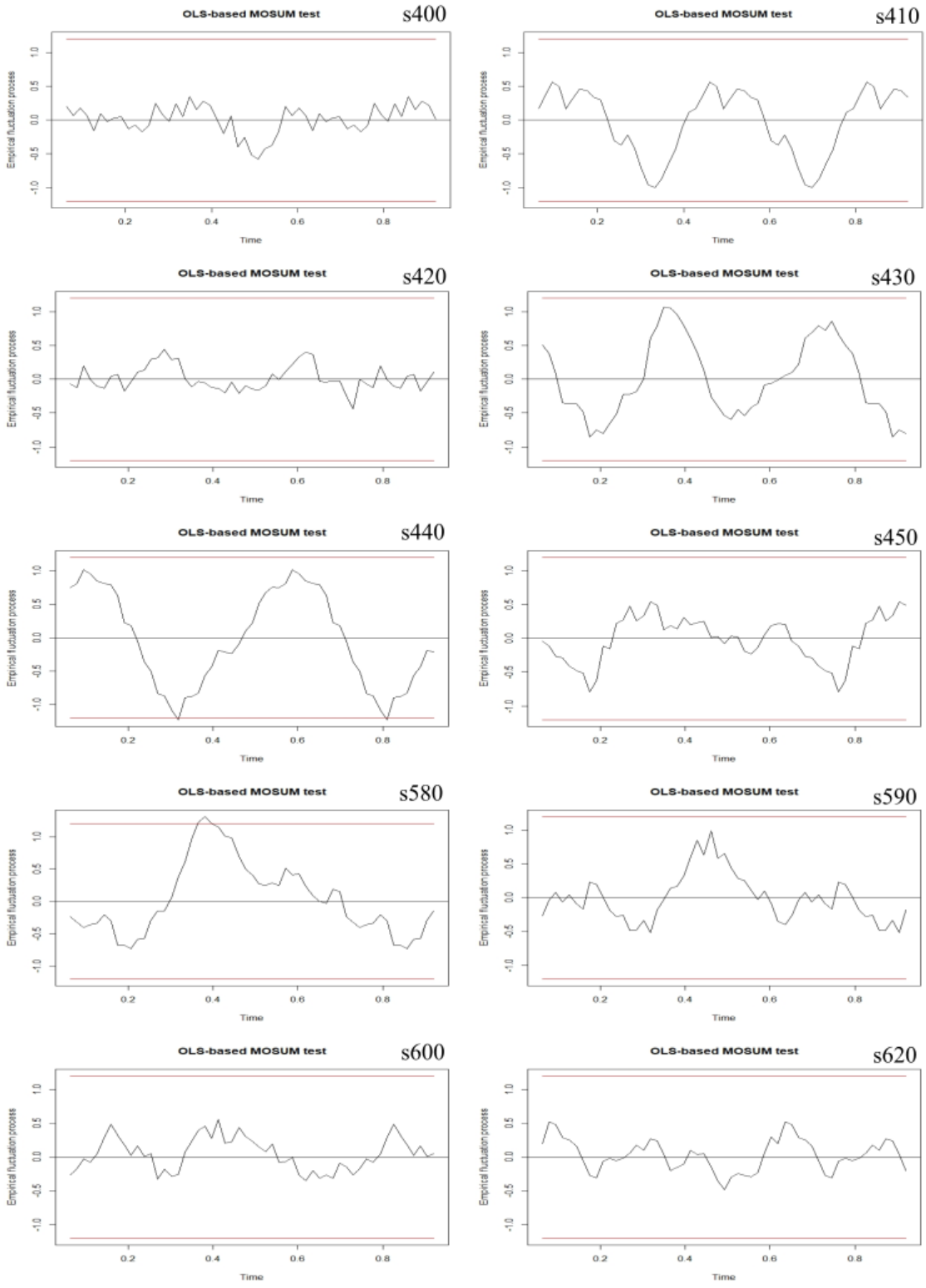


Figure 47A: OLS-MOSUM test plot for known forest training samples computed with PC4 loadings

9. REFERENCES

Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.

Aguirre-Gutiérrez, J., Seijmonsbergen, A. C., & Duivenvoorden, J. F. (2012). Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. Applied Geography, 34, 29-37.

Aleixandre-Benavent, R., Aleixandre-Tudó, J. L., Castelló-Cogollos, L., & Aleixandre, J. L. (2018). Trends in global research in deforestation. A bibliometric analysis. *Land Use Policy*, *72*, 293-302.

Angelsen, A. (1995). Shifting cultivation and "deforestation": a study from Indonesia. World Development, 23(10), 1713-1729.

Bradshaw, C. J., Sodhi, N. S., & Brook, B. W. (2009). Tropical turmoil: a biodiversity tragedy in progress. Frontiers in Ecology and the Environment, 7(2), 79-87.

Bryan, J., Shearman, P., Ash, J., & Kirkpatrick, J. B. (2010). Estimating rainforest biomass stocks and carbon loss from deforestation and degradation in Papua New Guinea 1972–2002: best estimates, uncertainties and research needs. *Journal of Environmental Management*, *91*(4), 995-1001.

Boyle, S. A. (2017). Overview of the World's Forests.

Buchanan, G. M., Butchart, S. H., Dutson, G., Pilgrim, J. D., Steininger, M. K., Bishop, K. D., & Mayaux, P. (2008). Using remote sensing to inform conservation status assessment: estimates of recent deforestation rates on New Britain and the impacts upon endemic birds. *Biological Conservation*, *141*(1), 56-66.

Carfagna, E., & Gallego, F. J. (2005). Using remote sensing for agricultural statistics. International statistical review, 73(3), 389-404.

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote sensing of Environment, 62(3), 241-252.

Chapin III, F. S. (2013). Overview of Ecosystem Functions and Services: Their Importance and Vulnerability.

Chen, G., Hay, G. J., Carvalho, L. M., & Wulder, M. A. (2012). Object-based change detection. International Journal of Remote Sensing, 33(14), 4434-4457.

Choi, J., Yu, K., & Kim, Y. (2010). A new adaptive component-substitution-based satellite image fusion by using partial replacement. IEEE transactions on geoscience and remote sensing, 49(1), 295-309.

Chu, C. S. J., Hornik, K., & Kaun, C. M. (1995). MOSUM tests for parameter constancy. *Biometrika*, *82*(3), 603-617.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync—Tools for calibration and validation. Remote Sensing of Environment, 114(12), 2911-2924.

Comber, A., Fisher, P., Brunsdon, C., & Khmag, A. (2012). Spatial analysis of remote sensing image classification accuracy. Remote Sensing of Environment, 127, 237-246.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., ... & Raskin, R. G. (1997). The value of the world's ecosystem services and natural capital. nature, 387(6630), 253.

Coppin, P., Lambin, E., Jonckheere, I., & Muys, B. (2002). Digital change detection methods in natural ecosystem monitoring: A review. In *Analysis of multi-temporal remote sensing images* (pp. 3-36).

Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. Remote Sensing of Environment, 102(1-2), 1-11.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Meygret, A. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote sensing of Environment, 120, 25-36.

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sensing, 8(4), 354.

Frantz, D., Haß, E., Uhl, A., Stoffels, J., & Hill, J. (2018). Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*.

Gandhi, G. M., Parthiban, S., Thummalu, N., & Christy, A. (2015). NDVI: vegetation change detection using remote sensing and GIS–a case study of Vellore District. Procedia Computer Science, 57, 1199-1210.

Geist, H. J., & Lambin, E. F. (2001). What drives tropical deforestation. LUCC Report series, 4, 116.

Hamunyela, E., Verbesselt, J., & Herold, M. (2016). Using spatial context to improve early detection of deforestation from Landsat time series. *Remote Sensing of Environment*, *172*, 126-138.

Hansen, B. E. (1997). Approximate asymptotic p values for structuras-change tests. Journal of Business & Economic Statistics, 15(1), 60-67.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A. A., Tyukavina, A. & Kommareddy, A. (2013). High-resolution global maps of 21st-century forest cover change. S*cience*, *342*(6160), 850-853.

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS Journal of photogrammetry and remote sensing, 80, 91-106.

Jolliffe I. (2011) Principal Component Analysis. In: Lovric M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg

Kinzig, A. P., Ryan, P. A., Etienne, M., Allison, H. E., Elmqvist, T., & Walker, B. H. (2006). Resilience and regime shifts: assessing cascading effects. Ecology and society, 11(1).

Kleiber, C., Hornik, K., Leisch, F., & Zeileis, A. (2002). strucchange: An R package for testing for structural change in linear regression models. Journal of statistical software, 7(2), 1-38.

Krieger, D. J. (2001). Economic value of forest ecosystem services: a review.

Kwarteng, P., & Chavez, A. (1989). Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens*, *55*, 339-348.

Lambert, J., Drenou, C., Denux, J. P., Balent, G., & Cheret, V. (2013). Monitoring forest decline through remote sensing time series analysis. *GIScience & remote sensing*, *50*(4), 437-457.

Laurance, W. F., Useche, D. C., Rendeiro, J., Kalka, M., Bradshaw, C. J., Sloan, S. P., ... & Arroyo-Rodriguez, V. (2012). Averting biodiversity collapse in tropical forest protected areas. Nature, 489(7415), 290.

Lewis, H. G., & Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. International Journal of Remote Sensing, 22(16), 3223-3235.

Li, J. (2000). Spatial quality evaluation of fusion of different resolution images. International Archives of Photogrammetry and Remote Sensing, 33(B2; PART 2), 339-346.

Loughlin, W. P. (1991). Principal component analysis for alteration mapping. Photogrammetric Engineering and Remote Sensing, 57(9), 1163-1169.

Lu, M., & Hamunyela, E. (2016). On-Line Change Monitoring with Transformed Multi-Spectral Time Series, a Study Case in Tropical Forest. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *41*, 987.

Lu, M., Pebesma, E., Sanchez, A., & Verbesselt, J. (2016). Spatio-temporal change detection from multidimensional arrays: Detecting deforestation from MODIS time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, *117*, 227-236.

Lu, M., Hamunyela, E., Verbesselt, J., & Pebesma, E. (2017). Dimension reduction of multi-spectral satellite image time series to improve deforestation monitoring. *Remote Sensing*, *9*(10), 1025.

Lund, H. G. (1999). A ‘forest' by any other name…. Environmental Science & Policy, 2(2), 125-133.

McAlpine, J. R. (1983). Climate of Papua New Guinea. Canberra, ACT: Commonwealth Scientific and Industrial Research Organization in association with Australian National University Press.

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R., & Berger, M. (2012). Sentinels for science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land. Remote Sensing of Environment, 120, 91-101.

Mandanici, E., & Bitelli, G. (2016). Preliminary comparison of sentinel-2 and landsat 8 imagery for a combined use. Remote Sensing, 8(12), 1014.

Mangiafico, S. S. (2016). Summary and Analysis of Extension Program Evaluation in R, version 1.10. 0.

Marceau, D. J., & Hay, G. J. (1999). Remote sensing contributions to the scale issue. Canadian journal of remote sensing, 25(4), 357-366.

Mas, J. F. (1999). Monitoring land-cover changes: a comparison of change detection techniques. International journal of remote sensing, 20(1), 139-152.

Metwalli, M. R., Nasr, A. H., Allah, O. S. F., El-Rabaie, S., & El-Samie, F. E. A. (2010). Satellite image fusion based on principal component analysis and high-pass filtering. JOSA A, 27(6), 1385-1394.

McFarland, B. J. (2018). The Context of Tropical Rainforest Deforestation and Degradation. In Conservation of Tropical Rainforests (pp. 7-58). Palgrave Macmillan, Cham.

Muhsoni, F. F., Sambah34, A., Mahmudi, M., & Wiadnya, D. (2018). Comparison of different vegetation indices for assessing mangrove density using sentinel-2 imagery. Int. J. Geomate, 14, 42-51.

Nasi, R., Wunder, S., & Campos, J. J. (2002). Forest ecosystem services: can they pay our way out of deforestation?.

Pan, Y., Birdsey, R. A., Phillips, O. L., & Jackson, R. B. (2013). The structure, distribution, and biomass of the world's forests. Annual Review of Ecology, Evolution, and Systematics, 44, 593-622.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, *20*(9), 503-510.

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jędrzejewska, B., Lima, M., & Kausrud, K. (2011). The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. Climate Research, 46(1), 15-27.

Pimm, S. L., & Raven, P. (2000). Biodiversity: extinction by numbers. Nature, 403(6772), 843.

Planet Labs Inc. (2018), PLANET IMAGERY PRODUCT SPECIFICATIONS

Pontius, R. G., Boersma, W., Castella, J. C., Clarke, K., de Nijs, T., Dietzel, C., ... & Koomen, E. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science*, *42*(1), 11-37.

Qiu S., Zhu Z., He B. May 22, 2018. Fmask 4.0 Handbook

Ranchin, T., & Wald, L. (2000). Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. Photogrammetric engineering and remote sensing, 66(1), 49-61.

Rouse Jr, J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS.

Schoene, D., Killmann, W., von Lüpke, H., & Wilkie, M. L. (2007). Definitional issues related to reducing emissions from deforesation in developing countries (Vol. 5). Rome: Food and Agriculture Organization of the United Nations.

Schowengerdt, R. A. (2006). Remote sensing: models and methods for image processing. Elsevier.

Schultz, M., Clevers, J. G., Carter, S., Verbesselt, J., Avitabile, V., Quang, H. V., & Herold, M. (2016a). Performance of vegetation indices from Landsat time series in deforestation monitoring. International journal of applied earth observation and geoinformation, 52, 318-327.

Schultz, M., Verbesselt, J., Avitabile, V., Souza, C., & Herold, M. (2016b). Error sources in deforestation detection using BFAST monitor on landsat time series across three tropical sites. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(8), 3667-3679.

Sentinel-2 User Handbook, ESA, July 24, 2015. Accessed on: Oct. 1, 2018. [Online]. Available: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook.

Sibanda, M., Mutanga, O., & Rouget, M. (2015). Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. ISPRS Journal of Photogrammetry and Remote Sensing, 110, 55-65.

Singh, A. (1989). Review article digital change detection techniques using remotely-sensed data. International journal of remote sensing, 10(6), 989-1003.

Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. Photogrammetric Engineering and remote sensing, 52(3), 397-399.

Švab, A., & Oštir, K. (2006). High-resolution image fusion: Methods to preserve spectral and spatial resolution. Photogrammetric Engineering & Remote Sensing, 72(5), 565-572.

Thomas, C., Ranchin, T., Wald, L., & Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. IEEE Transactions on Geoscience and Remote Sensing, 46(5), 1301-1312.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, *8*(2), 127-150.

Van der Meer, F. D., & De Jong, S. M. (Eds.). (2011). *Imaging spectrometry: basic principles and prospective applications* (Vol. 4). Springer Science & Business Media.

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010a). Detecting trend and seasonal changes in satellite image time series. *Remote sensing of Environment*, *114*(1), 106-115.

Verbesselt, J., Hyndman, R., Zeileis, A., & Culvenor, D. (2010b). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. Remote Sensing of Environment, 114(12), 2970-2980.

Verbesselt, J., Zeileis, A., & Herold, M. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, *123*, 98-108.

Vergopolan, N., & Fisher, J. B. (2016). The impact of deforestation on the hydrological cycle in Amazonia as observed from remote sensing. *International journal of remote sensing*, *37*(22), 5412-5430.

Visa, S., Ramsay, B., Ralescu, A. L., & Van Der Knaap, E. (2011). Confusion Matrix-based Feature Selection. MAICS, 710, 120-127.

Wang, Q., Shi, W., Atkinson, P. M., & Zhao, Y. (2015). Downscaling MODIS images with area-to-point regression kriging. *Remote sensing of environment*, *166*, 191-204.

Wang, Q., Shi, W., Atkinson, P. M., & Pardo-Igúzquiza, E. (2016a). A new geostatistical solution to remote sensing image downscaling. *IEEE transactions on geoscience and remote sensing*, *54*(1), 386-396.

Wang, Q., Shi, W., Li, Z., & Atkinson, P. M. (2016b). Fusion of Sentinel-2 images. *Remote sensing of environment*, *187*, 241-252.

Wang, Q., & Atkinson, P. M. (2018). Spatio-temporal fusion for daily sentinel-2 images. *Remote Sensing of Environment*, *204*, 31-42.

Watson, R. T., Noble, I. R., Bolin, B., Ravindranath, N. H., Verardo, D. J., & Dokken, D. J. (2000). Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1-3), 37-52.

Yengoh, G. T., Dent, D., Olsson, L., Tengberg, A. E., & Tucker III, C. J. (2015). Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales: Current Status, Future Trends, and Practical Considerations. Springer.

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2001). strucchange. An R package for testing for structural change in linear r

Zeileis, A. (2005). A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals. *Econometric Reviews*, *24*(4), 445-466.

Zeileis, A. (2006). Implementing a class of structural change tests: An econometric computing approach. *Computational Statistics & Data Analysis*, *50*(11), 2987-3008.

Zeileis, A., & Hornik, K. (2007). Generalized M-fluctuation tests for parameter instability. *Statistica Neerlandica*, *61*(4), 488-508.

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, *118*, 83-94.

Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269-277.

Zhu, Z., Qiu, S., He, B., Shang, R., & Lin, Y. (2018, December). Fmask 4.0: Cloud and Cloud Shadow Detection in Landsats 4-8 and Sentinel-2. In AGU Fall Meeting Abstracts. (Accepted)

Statement of originality of the MSc thesis

I declare that:

1. this is an original report, which is entirely my own work,

2. where I have made use of the ideas of other writers, I have acknowledged the source in all instances,

3. where I have used any diagram or visuals I have acknowledged the source in all instances,

4. this report has not and will not be submitted elsewhere for academic assessment in any other academic course.

Student data:

Name: Joel Isaac Koupermann Kuperman

Registration number: 5830605

Date: July 9, 2019

Signature: