

Estimating vegetation phenology at 30m resolution with multi- temporal optical imagery for a rangeland site in Kenya

STELLA GACHOKI MUTHONI

February 2018

SUPERVISORS:

Dr. A. Vrieling

Dr. M. Marshall



Estimating vegetation phenology at 30m resolution with multi- temporal optical imagery for a rangeland site in Kenya

STELLA GACHOKI MUTHONI

Enschede, The Netherlands, February 2018

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

Specialization: Natural resources

SUPERVISORS:

Dr. A. Vrieling

Dr. M. Marshall

THESIS ASSESSMENT BOARD:

Dr. C.A.J.M. de Bie (Chair)

Dr. F. Fava (External Examiner, International Livestock Research Institute)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author and do not necessarily represent those of the Faculty.

ABSTRACT

Vegetation phenology is the study of ‘tracking’ life-cycle events of plants such as the onset, offset and length of a growing season in relation to climatic variables. Tracking vegetation growth in savannas is important to have timely knowledge on the availability of forage for cattle and wild animals that depend on this for their food intake. Satellite-based vegetation phenology studies for the savannas exist at spatial resolutions of 250m and coarser. Recently vegetation phenology studies are performed using optical satellite data at spatial resolutions of 10-30m. However, such studies have not been performed yet for savanna ecosystems.

The goal of this study was to evaluate if vegetation phenology can be effectively retrieved from 30m resolution data for a savanna site in Kenya that has two vegetation growth periods per year (long and short rains). The first method combined NDVI observations from three sensors (Landsat 7, Landsat 8 and Sentinel-2A) to model NDVI seasonal dynamics for each season and year separately (2015-2017). The second method combined NDVI observations from multiple years of Landsat data (1999-2017) into a single ‘synthetic year’ to model the multi-annual average vegetation dynamics for each season. In both cases, a double hyperbolic tangent model was used to model the NDVI seasonal dynamics. A threshold approach was then applied to extract start- and end-of-season (SOS/EOS), length of the growing season (LGS), cumulative NDVI (cumNDVI), maximum NDVI (maxNDVI), growth phase (AMP1) and decay phase (AMP2) for each season.

Despite the various challenges in the region such as persistent cloud cover, inter-annual variation in vegetation greenness and two short (~3months) seasons in a single year, vegetation phenology retrievals at 30m resolution were possible. The single-season retrievals for the long rains of 2017 had the highest percentage of pixels (~53%) having a successful retrieval while the long rains of 2016 had the lowest percentage (~15%). These differences may be explained partially by the maximum temporal gap between the observations where the long rains of 2017 had ~33days and the long rains of 2016 had ~74days. Comparison of the multi-annual average retrievals from aggregated Landsat and eMODIS showed an RMSD of ± 14 days (SOS), ± 12 days (EOS), and an MSD of ± 8 days for both EOS and SOS on average. Contrary to the expectations, these values were considered high for a region that has quick green-up and decay phase. This comparison was carried out to assess how multi-annual average phenology from a less frequent sensor (Landsat) compared to similar retrievals derived from a more frequent sensor (MODIS).

Although ground-based validations are required for this study, estimation of vegetation phenology at a finer scale in the savanna ecosystem is possible. For multi-annual average retrievals, future research should focus on better ways of combining the multiple-years into a ‘synthetic year’ that produces less scattered NDVI profiles in each season. For the single-season persistent cloud cover remains the main issue limiting the frequent in-season observation of the vegetation status and thus for estimating phenology. The recent availability of Sentinel-2B images increases the chances of getting more cloud-free observations and a possibility to estimate vegetation phenology at a 10m resolution scale which was not possible in this study.

KEYWORDS: Vegetation phenology, multi-temporal imagery, spatial resolution, model fitting, seasonality, savanna, rangeland

ACKNOWLEDGEMENTS

I would like to thank Netherlands Fellowship Programme (NFP) for giving me the financial opportunity to study my MSc at University of Twente (ITC). To the whole of ITC administration, thank you for nurturing me into realizing a profession.

To my supervisors, Dr. Anton Vrieling and Dr. Michael Marshall thank you for being patient with me through the understanding of this research and all the accomplishments. Special thanks to Dr. Anton Vrieling for the time he sacrificed to help with the pre-processing of a large number of satellite data used in this research. To my external examiner, Dr. Francesco Fava, thank you for the facilitation of my stay at ILRI and your assistance in organizing the fieldwork. Special thanks also to Dr. Thomas Groen for the statistical advice given throughout the research.

To my family thank you for being there for me. Julia Wambui (Mom), Edith Wanjiku (Sister), Caxton Moses (Brother) and Edyrian Mwangi (Nephew), you have all made priceless sacrifices in your own unique ways to see the success of my studies here at The Netherlands.

Above all, if it weren't for you God we could not have made it. Thank you, God, for good health, knowledge and the strength that You gave each one of us. The success of this Thesis is all for Your glory.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1. Background.....	1
1.2. Research objectives	3
1.3. Research questions	4
2. STUDY AREA AND DATA	5
2.1. Study area.....	5
2.2. Data.....	7
2.3. Software	10
3. METHODS.....	11
3.1. Phenology retrieval model	11
3.2. Phenology from combined Landsat and Sentinel-2 data.....	12
3.3. Phenology from combining multi-year Landsat acquisitions	13
3.4. Explaining spatial variability of phenology from differences in vegetation composition	14
4. RESULTS	15
4.1. Phenological parameters from fused Landsat and Sentinel-2 data.....	15
4.2. Phenological parameters from multi-year Landsat and eMODIS data.....	19
4.3. Retrieved patterns in relation to vegetation composition on ground	24
5. DISCUSSION.....	27
6. CONCLUSION.....	29
7. APPENDICES	35
7.1. Appendix 1	35
7.2. Appendix 2	36
7.3. Appendix 3	37
7.4. Appendix 4	38
7.5. Appendix 5	38

LIST OF FIGURES

Figure 1: Average monthly rainfall. The averages were derived from rainfall data of Kapiti Farm of 1999-2016. The error bars indicate the standard deviations between the years.	5
Figure 2: location of the study area. (a) – Grey faint boundaries represent county boundaries of Kenya, black boundaries represent counties where the Athi-Kapiti Plains are located, blue polygon represent the approximate Athi-Kapiti extent, the red polygon is Kapiti Farm. Similar colours explained for (a) applies to (b) and (c). (c) – the background is a Sentinel-2A RGB image acquired on 20th December 2016. The black triangle in (c) is the rainfall station inside Kapiti Farm	6
Figure 3: Total annual rainfall in millimetre of Kapiti Farm per season. Short rain season runs from September to February while long rains is from March to August	8
Figure 4: Spatial distribution of vegetation community within Kapiti Farm. The background is a Sentinel-2A RGB image acquired on 20th December 2016.	9
Figure 5: representative pixel for each vegetation structure average dekadal NDVI trajectory from multi-year eMODIS data (2001-2017). These locations were randomly selected from the six-sample location in each group. Black vertical lines represent the approximate transition dates. The red line is the temporal range for the long rains while the purple line is for the short rains	11
Figure 6: A summary of estimated vegetation cover per vegetation community using LandPKS protocol.	14
Figure 7 Number of observation with a valid NDVI value per season and the maximum temporal gap between these observations from fused Landsat 7- 8 and Sentinel-2 data between september 2015 - September 2017	15
Figure 8: Short rains of 2015/2016 and 2016/2017 from the fused Landsat 7, 8 and Sentinel-2 observation. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs and (D) Acacia shrubs. The black dots indicate SOS ₅₀ and EOS ₅₀ respectively	16
Figure 9: Long rains of 2016 and 2017 from the fused Landsat 7, 8 and Sentinel-2 observation. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs and (D) Acacia shrubs. The black dots indicate SOS ₅₀ and EOS ₅₀ respectively	16
Figure 10: detailed legend for different figures used in this study. a) colours for each month; the colours are divided into 10 days in a month, the first ring (1) represent day 1-10, (2) day 11-20 and (3) day 21-30, this legend is used for SOS ₅₀ and EOS ₅₀ . b) days of the LGS ₅₀ . c) the cumNDVI values. d) shows theAMP1 and AMP2 values. e) represents the maxNDVI values	17
Figure 11: Phenological parameters and cumNDVI retrieved from fused Landsat 7, 8 and Sentinel-2 within Kapiti Farm. They are from the four different seasons. Short rains 2015/2016, short rains 2016/2017, long rains 2016 and long rains 2017. Detailed legend is in Figure 10	18
Figure 12: Variograms of phenological and seasonal greenness parameters retrieved from fused Landsat 7, 8 and Sentinel-2 NDVI time series of Kapiti Farm	19
Figure 13: Combined years of Landsat archive NDVI time series for the representative vegetation community. The black dots represent SOS ₅₀ and EOS ₅₀ respectively. The locations are the same but different vegetation season. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs, (D) Acacia shrubs.	20
Figure 14: Phenological and seasonal greenness parameters retrieved from combined Landsat archive within Kapiti Farm. For the legend see Figure 10.	20
Figure 15: Variograms of phenological and seasonal biomass parameters retrieved from combined years of Landsat archive NDVI time series of Kapiti Farm	21
Figure 16: Aggregated combined years of Landsat(250m) and eMODIS v5 phenological parameters for the long and short rains season of Kapiti Farm. For detailed legend refer to Figure 10.....	22
Figure 17: Density scatter plots of aggregated phenological parameters of combined years of Landsat against eMODIS v5 phenological parameters. The X and Y axis for SOS ₅₀ and EOS ₅₀ are the DOY while for the cumNDVI, maxNDVI, AMP1 and AMP2 are the NDVI value. The dashed line in grey represent the slope while the red line represent the regression line	23
Figure 18: shows the combination of all 1999-2017 Landsat-derived NDVI values combined in a single synthetic year for two points. The colour of each observation relates to the amount of seasonal precipitation in each specific year.	24

LIST OF TABLES

Table 1: A summary of the remotely sensed data used.....	7
Table 2: Phenological parameter retrievals from fused Landsat 7, 8 and Sentinel-2 NDVI time series for the vegetation communities of Kapiti Farm. These are means of the six points collected for each group.	17
Table 3: Retrieved phenological and seasonal greenness parameters averaged from the 24 sample points from combined years of Landsat archive.....	21
Table 4: Average phenological parameters and seasonal greenness for Kapiti Farm from eMODIS v5 NDVI time series	22
Table 5: ANOVA analysis of the vegetation community within Kapiti Farm. Results shaded in grey are those that rejected null hypothesis	25
Table 6: Tukey HSD analysis carried for the parameters that rejected the ANOVA null hypothesis. A = grasslands, B = Denser tree cover, C = Diverse shrubs and D = Acacia shrubs. Results shaded in grey are those that rejected the Tukey HSD test	25

ABBREVIATIONS

NDVI – Normalized Difference Vegetation Index

MODIS - Moderate Resolution Imaging Spectroradiometer

eMODIS - EROS Moderate Resolution Imaging Spectroradiometer

AVHRR – Advanced very-high Resolution radiometer

SPOT - Satellite Pour l'Observation de la Terre

LPA – Landsat Phenology Algorithm

ENVI – Environment for visualising Images

IBLI – Index-Based Livestock insurance

NASA – National Aeronautics and Space Administration

IDL – Interactive Data Language

DOY – Day of the Year

RMSD – Root Mean Squared Deviation

MSD – Mean Signed Deviation

ANOVA – Analysis of Variance

HSD – Honest Significant Difference

EOS – End of Season

SOS – Start of Season

LGS – Length of Growing Season

GCC – Green Chromatic coordinate

EVI – Enhanced Vegetation Index

1. INTRODUCTION

1.1. Background

Savannas are a major ecosystem covering one-sixth of the global land surface with the largest area (~15.1million km²) situated in Africa (Grace et al., 2006; Sankaran et al., 2005). The savanna landscape has a unique heterogeneous vegetation distribution of patchy trees with understory shrubs as well as grasses (Walker et al., 2014; Whitecross et al., 2017). The co-existence of these vegetation types makes savannas rather unique compared to other ecosystems that have a single vegetation type (Baudena et al., 2015).

Savannas are known as habitats for a diversity of wildlife populations (Cardoso et al., 2002). Other than that, a large portion of the savanna is under the rangelands. These rangelands have grass-like plants or shrubs as the predominant vegetation cover and are mainly used for grazing or browsing by domesticated and wildlife animals (Werner, 2009). For the case of the African savanna, much of the land is for rearing livestock as well as hosting a wide variety of wild animals. Livestock rearing in Africa savannas is usually either privately or communally owned and is a key socio-economic activity for communities living in these landscapes (Sandhage-Hofmann et al., 2015).

Despite the importance of the savannas, the seasonal dynamics of its vegetation is still poorly understood (Sankaran et al., 2004). For instance, the practicality of how grass and trees co-exist competing for the same resources such as water, without one vegetation type outdoing the other remains unclear (Sankaran et al., 2004; Scholes et al., 1997; Jeltsch et al., 1996; House et al., 2003). It has been argued that the seasonal onset and offset of grasses and trees may differ (Higgins et al., 2011; Boke-Olén et al., 2016), and this temporal separation could explain the co-existence of the two vegetation types (Scholes et al., 1993). However, only a few studies (Higgins et al., 2011; Moore et al., 2017) have explored and discussed savanna vegetation dynamics in relation to this temporal separation.

Vegetation productivity and plant survival within savannas depend largely on how well plants can adapt to the variability of moisture availability (Ma et al., 2013; Moore et al., 2017; Cipriotti et al., 2008). The current changes and variations in climate are threatening the vegetation productivity of the savannas. For instance, Githeko et al. (2007) indicated that Africa is one of the most vulnerable continents to climate change and climate variability. These climate changes and variations are inevitably affecting moisture availability, vegetation productivity, forage quality, land use systems and rangeland-based livelihoods in the savannas (Hoffman et al., 2008). Githeko et al. (2007) estimated that by the year 2080 the arid and the semi-arid lands of Africa will have increased by 5-8% and indicated that a variety of other African ecosystems are experiencing changes at a faster rate than expected, e.g., the ecosystems located in the southern part of Africa.

Other than the climatic shift, human pressures such as land use change, human encroachment, wood harvesting, human bushfires, land fragmentation and overexploitation of vegetation (Macharia & Ekaya, 2005) have altered the overall vegetation composition in most of the African savanna. Based on a study done by Giglio et al. (2006), 75% of the total burned areas in Africa, mostly initiated by humans, were inside the savannas. These forms of land degradation threaten the wildlife and livelihoods of people depending on the savanna. To be able to maintain a sustainable structure and health of the rangelands, there is need to find better, up-to-date and robust methods for monitoring the savanna vegetation (IPCC, 2007; King, 2008; Hoffman et al., 2008).

Different methods exist to study the seasonal dynamics of vegetation, often referred to as *vegetation phenology*. By definition, *vegetation phenology* is the study of ‘tracking’ the life-cycle events of plants such as the onset, offset and length of a growing season in relation to climatic variables (Adole et al., 2016). The definitions of these life cycles may differ between studies. For instance, onset might be the moment when the increase in vegetation greenness reaches some threshold or the moment that flowering starts. Besides regular visual observation of plants in the field (Elmendorf et al., 2016), automated approaches exist that ‘track’ vegetation greenness over time. These include the use of digital repeat cameras placed in the field (Studer et al., 2007) or use of satellite remote sensing data (Vrieling et al., 2017a). Field observations and digital cameras are capable of characterizing individual plant species phenology but are limited in spatial coverage and for the case of field observations, they are time-consuming (Richardson et al., 2013). On the other hand, while satellite-based phenology has proven to be a powerful tool for studying the seasonal greenness at the landscape scale, it is sometimes limited by the presence of clouds at satellite overpass time and by the coarse spatial resolution (generally $\geq 250\text{m}$) of the satellite images. While finer resolution imagery is available, in most cases, such data have a low temporal frequency due to the trade-off between spatial and temporal resolution (Richardson et al., 2013).

Other than the mentioned issues affecting the use of satellite data to study vegetation phenology, use of this data requires a complex stepwise methodology which in most cases is location based (Reed et al., 2009). The first step involves calculation of vegetation indices such as the Normalized Vegetation Difference Index (NDVI) which is then followed by employing different phenology techniques such as threshold, curve-derived and model fitting (de Beurs et al., 2010). These phenological techniques can be used together at times, but they depend mostly on the nature of the satellite data and the purpose of the study. For instance, model fitting methods such as double logistic model tend to explain remotely sensed data better and to extract the phenological parameters normally a threshold approach is applied (Ming et al., 2011). Nonetheless, the model fitting technique requires estimation of multiple parameters and it is still unclear how estimation of these parameters is affected by the temporal resolution of remotely sensed data (Ahl et al., 2006).

A review done by Adole et al. (2016) showed that satellite-based phenology using threshold and or model fitting techniques is the most common way of studying vegetation phenology in Africa. Most of these satellite-based studies uses coarse ($\geq 250\text{m}$) satellite remote sensing data such as the Advanced Very High Radiometer (AVHRR: e.g. Zhang et al., 2014), Moderate-resolution Imaging Spectroradiometer (MODIS: e.g. Zhang et al., 2005) and Satellite Pour l’Observation de la Terre (SPOT vegetation: e.g. Meroni et al., 2013). However, only a few (O’Farrell et al., 2007; Yamagiwa et al., 2008) field-based studies that exist in Africa using either ground observations or digital cameras and currently they only exist in the southern part of Africa. Lack of financial support is the main reason prohibiting ground-based phenological studies in most African countries (Blom et al., 2015) as well as political instability and physical accessibility in these countries (Laurance et al., 2006).

Recently, vegetation phenology is possible to retrieve at finer spatial resolution using satellites such as Landsat (30m) and Sentinel-2 (10m). The revisit period of these two sensors (Sentinel-2 and Landsat) is not as frequent as those sensors with coarser resolution (e.g., AVHRR, MODIS). To counter for this trade-off, studies have tried fusing fine spatial resolution data with more frequent but coarse resolution data (Frantz et al., 2016; Gao et al., 2017; Walker et al., 2014) or combining years of Landsat archive data into a ‘synthetic year’ (Melaas et al., 2016; Melaas et al., 2013; Nijland et al., 2016). ‘Synthetic year’ basically means using all the available cloud-free Landsat images to generate seasonal curves for each pixel by assuming that these pixels are repeated annually within the long-term Landsat time-series (Nijland et al., 2016). Other than fusing sensors data or use of synthetic year from finer spatial resolution satellite data, recently a study that was done by, Vrieling et al. (2017b) showed that in some cases, data acquired from sensors like Sentinel-2 could

be used alone to estimate phenology for a single season. The single season here means that the satellite data used is from individual single years. Nevertheless, this depends on various factors such as the presence of clouds and whether the location is near the orbits that happen to have more images than the normal three images per month.

Studies that have been using finer spatial-resolution satellite imagery to estimate vegetation phenology had until present existed for the temperate environments. For example, multi-annual average vegetation phenology derived from a Landsat ‘synthetic year’ (Melaas et al., 2016, 2013; Nijland et al., 2016; Vrieling et al., 2017a) or single season vegetation phenology derived from single use of Sentinel-2 images (Vrieling et al., 2017b). It is still unknown if similar methods can work in a more complex ecosystem such as the Africa savanna. The heavy clouds contamination in the tropical climates of the Africa savanna is one of the major challenge affecting the quality of satellite-based images (Huete et al., 2002). A study by Hashim et al. (2014) showed that for the satellite images acquired over the equator in the tropics in a single year, 75% of the images had heavy cloud contamination. Other than that, a large part of the Africa savanna (e.g., East Africa savannas) has two relatively short vegetation seasons (~3 months each) within the year. These short seasons have high inter-annual variations in vegetation greenness between years, as opposed to the existing studies that have a single season which is relatively stable within single years. These issues coupled with frequent droughts (Maitima et al., 2003) makes it even harder to estimate its vegetation phenology.

This study attempts retrieval of vegetation phenology from multi-temporal high-resolution Landsat and Sentinel-2 imagery and to assess how these phenological metrics relate to different vegetation types for a savanna grassland ecosystem in the semi-arid rangelands of Kenya. A double hyperbolic tangent model was fitted on NDVI values derived from two sets of data separately to achieve the objectives of this study. The first dataset was fused Landsat 7, Landsat 8 and Sentinel-2 images from 2015-2017 at Landsat resolution and the second dataset was combined years (1999-2017) of Landsat archive (5,7&8) into a synthetic year. Assessment of how Landsat ‘synthetic year’ retrievals agreed with a commonly used sensor were performed using EROS Moderate Resolution Imaging Spectroradiometer (eMODIS v5). Using coarser resolution to assess the ‘validity’ of the retrievals was not the best way but as mentioned ground-based vegetation phenological studies are a challenge in Africa.

1.2. Research objectives

The main objective of this research is to evaluate if vegetation phenology is possible to retrieve from multi-temporal high-resolution Landsat and Sentinel-2 imagery for a savanna grassland ecosystem in the semi-arid rangelands of Kenya. To achieve the main objective, the research aims

- I. to evaluate if the combined use of Sentinel-2 and Landsat satellite imagery provides sufficient temporal detail for retrieving phenological parameters for a single season;
- II. to assess if fine-resolution spatial patterns of vegetation phenology can be extracted by combining years of Landsat data into a single synthetic year and assess how these spatial patterns compare to those obtained from multiple years of eMODIS data;
- III. to explain the retrieved spatial patterns of vegetation phenology based on differences in vegetation composition within the study area.

1.3. Research questions

Objective (I) – to evaluate if the combined use of Sentinel-2 and Landsat satellite imagery provides sufficient temporal detail for retrieving phenological parameters for a single season

- a) Do the combined use of Sentinel-2 and Landsat imagery provide sufficient observations to enable retrieval of phenological parameters within the temporal range of each season?

Hypothesis: Despite the heavy cloud contamination, there is a sufficient number of observations to allow fitting of a phenology retrieval model within the temporal range of each season.

- b) Do the fused Sentinel-2 and Landsat images allow for a spatially-consistent retrieval of vegetation phenology across the study site for each of the four growing seasons contained between September 2015 and September 2017?

Hypothesis: Derived vegetation phenological parameters shows a spatially-consistent pattern within the study area.

Objective (II) - to assess if fine-resolution spatial patterns of vegetation phenology can be extracted by combining years of Landsat data into a single ‘synthetic year’ and assess how these spatial patterns compare to those obtained from multiple years of eMODIS data

- a) Does the combination of multi-year Landsat imagery into a single synthetic year allow retrieval of spatially-consistent phenological parameters for the entire landscape?

Hypothesis: Despite the inter-annual variability in vegetation greenness, combining multiple years of Landsat into a ‘synthetic year’ allows estimation of spatially-consistent multi-annual average vegetation phenological parameters for the long and the short rains seasons.

- b) How do these phenological parameters compare to those estimated from a coarser spatial-resolution eMODIS but with finer temporal detail?

Hypothesis: Despite the differences in temporal and spatial resolution between the sensors, aggregated Landsat phenological parameters are almost similar to those estimated from combined years of eMODIS data.

- c) Would grouping multi-year Landsat NDVI data by use of seasonal precipitation first result in a ‘synthetic year’ that has less scattered NDVI values?

Hypothesis: Grouping years of Landsat NDVI based on the amount of seasonal precipitation experienced in that year, results in a ‘synthetic year’ that has less scattered NDVI temporal trajectories.

Objective (III) - to explain the retrieved spatial patterns of phenology based on differences in vegetation composition within the study area

- a) Do different vegetation types translate into significant differences in retrieved phenological parameters from combined years of Landsat?

Hypothesis: different vegetation types show significant differences in retrieved phenological parameters in long and short rains seasons

2. STUDY AREA AND DATA

2.1. Study area

Kenyan rangelands cover approximately 80% of Kenya's land surface and are home to millions of pastoralist and agro-pastoralist communities (Mathu, 2009). These rangelands are mainly used for livestock grazing and as habitats for wildlife (Mganga et al., 2011; Mathu, 2009). Statistics from the Government of Kenya (1993) showed that the arid and semi-arid lands accommodate 60% of Kenya total livestock population, 25% of human population and a large population of wild animals.

The Athi-Kapiti Plains are an example of semi-arid rangelands in Kenya which are within Machakos, Kajiado and Nairobi counties. They cover a land area of approximately ~2010 km² which is mainly a mixture of open grasslands (72%) and bush and woodlands (28%; Bekure et al., 1991). Grazing intensity, seasonal rainfall variation and human disturbances such as land conversions and fire are the main controls of the amount of vegetation productivity per moment in time in these rangelands (Bekure et al., 1991). The soil type found here is predominantly deep black vertisols and the area drains towards Athi River basin (Leeuw et al., 1991).

The Kamba community is the main occupant of the region and is predominantly agro-pastoralists. Other than livestock rearing, these plains host a diverse wildlife population. During the wet seasons, these plains fill in as transitory territories for wild animals from Amboseli national park. They are also the favored grounds for calving by wildebeest population relocating from southern plains of Tsavo and Chulu (Conservation of Kenya, 2010).

Kapiti Farm is within the Athi-Kapiti plains. The Farm is owned privately by the International Livestock Research Institute (ILRI) and covers ~130 km². The region has bi-modal precipitation seasons that approximately run from October to December (locally referred to as short rains) and March to May (long rains). *Figure 1* shows the average monthly rainfall as observed from the single rainfall station located in Kapiti Farm since 1940. The average seasonal precipitation (1999-2016) from the same rainfall station is 275mm for the 'short rains' and 270mm for the 'long rains.'

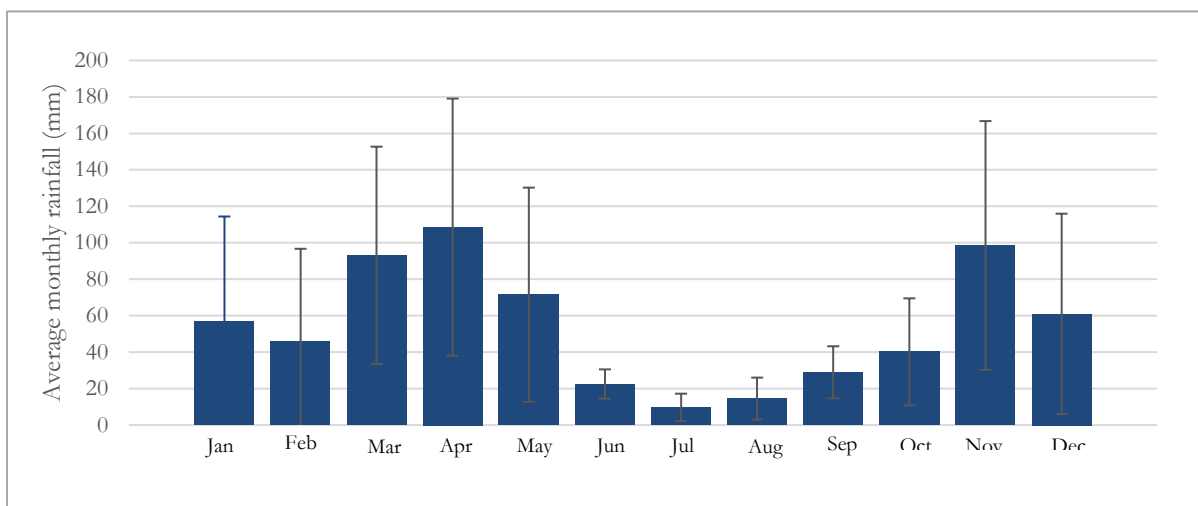


Figure 1: Average monthly rainfall. The averages were derived from rainfall data of Kapiti Farm of 1999-2016. The error bars indicate the standard deviations between the years.

Kapiti Farm is home to approximately 2500 cattle (Boran breed and a couple of Boran-Friesian crossbreds), 1200 sheep (Dorper, Kenyan Red Maasai and crossbreds), 250 Galla goats and wild animals such as lions, hyenas, jackals, zebras, impalas, giraffes, and ostriches among others. Breeding of disease-free cows and performance of research about the East coast fever and African animal trypanosomiasis (GEOGLAM RAPP, 2017) is also carried out on this Farm. *Figure 2* shows a map of the study area.

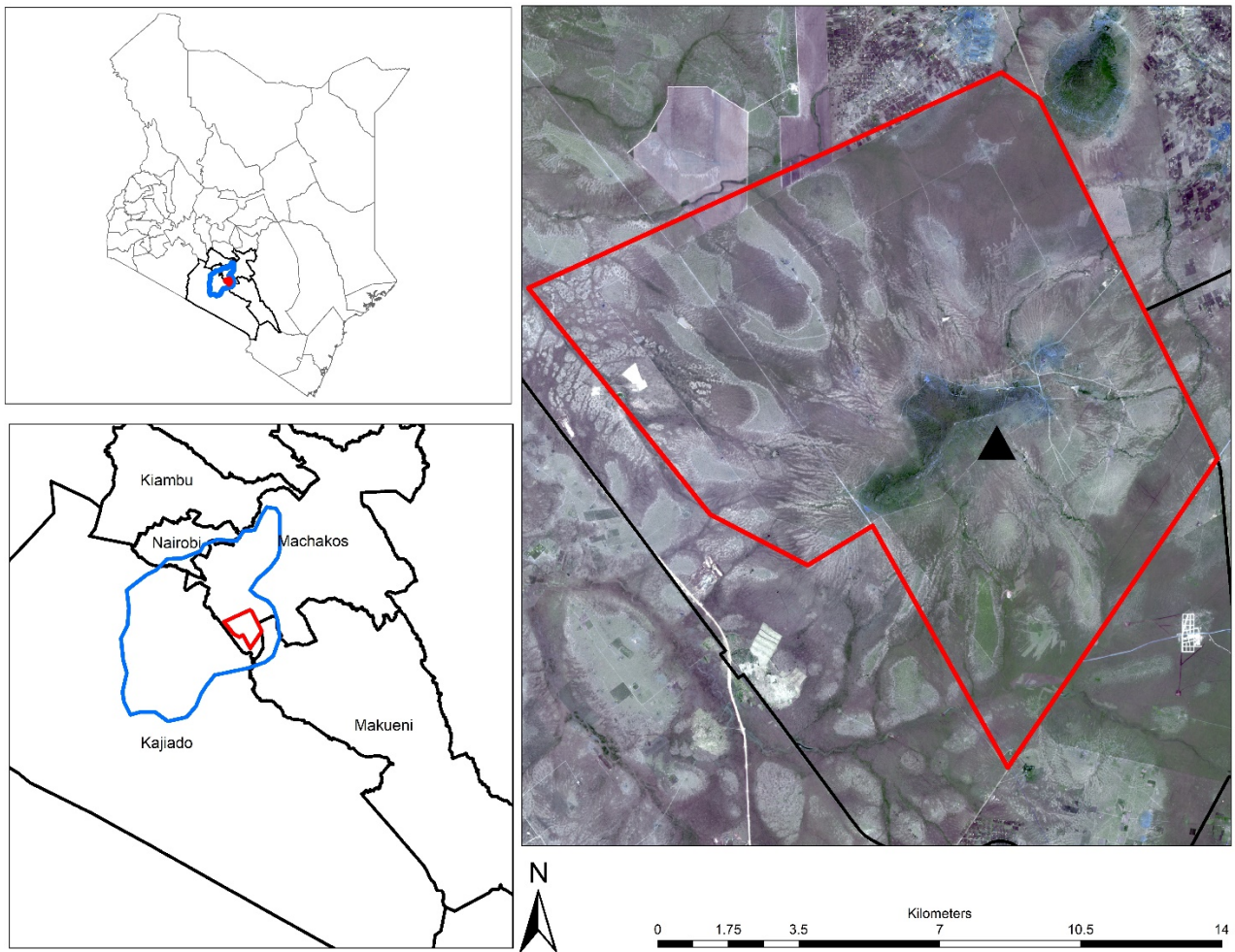


Figure 2: location of the study area. (a) – Grey faint boundaries represent county boundaries of Kenya, black boundaries represent counties where the Athi-Kapiti Plains are located, blue polygon represent the approximate Athi-Kapiti extent, the red polygon is Kapiti Farm. Similar colours explained for (a) applies to (b) and (c). (c) – the background is a Sentinel-2A RGB image acquired on 20th December 2016. The black triangle in (c) is the rainfall station inside Kapiti Farm

2.2. Data

Table 1 summarizes the remotely sensed data that were used to achieve the objectives of this study.

Table 1: A summary of the remotely sensed data used

Product	Sensor	Spatial resolution	Spectral resolution	Temporal resolution	Source	Years
Landsat 5	MSS	30m	NDVI, pixel_qa blue band	16 days	https://earthexplorer.usgs.gov/ https://espa.cr.usgs.gov/	1999 - 2017
Landsat 7	TM & ETM+	30m	NDVI, pixel_qa blue band	16 days	https://earthexplorer.usgs.gov/ https://espa.cr.usgs.gov/	1999 - 2017
Landsat 8	OLI	30m	NDVI, pixel_qa blue band	16 days	https://earthexplorer.usgs.gov/ https://espa.cr.usgs.gov/	1999 - 2017
Sentinel-2	MSI	10m	B8 (Red) B4 (NIR)	10 days	https://scihub.copernicus.eu/dhus/ https://earthexplorer.usgs.gov/	2015 - 2017
eMODIS	MODIS	250m	NDVI	10 days	USGS	2001 - 2017

Additional data used: Twenty-four sample locations of vegetation collected within Kapiti Farm and rainfall data (1940-2016) sheet from the single rain station in Kapiti Farm.

2.2.1. Landsat 1999-2017

The complete archive of Landsat Collection-1 imagery that based on visual observation of the quick looks contained at least partial cloud-free coverage over the Kapiti Farm was downloaded. Image searching was performed using USGS Earth Explorer (<https://earthexplorer.usgs.gov/>) whereas the ordering of surface reflectance and vegetation index products was done through the EROS Science Processing Architecture (ESPA) of USGS (<https://espa.cr.usgs.gov/>). Landsat 4-7 reflectance and vegetation index data are derived using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS algorithm; Masek et al., 2006) which was developed by NASA. This algorithm applies MODIS atmospheric correction routines to obtain Top of Atmosphere (TOA) reflectance, surface reflectance and masks for clouds, cloud shadows, land, adjacent clouds and water (Department of the Interior U.S. Geological Survey, 2017). For Landsat 8 a similar algorithm, Landsat Surface reflectance Code (LaSRC), is used to generate the same products as LEDAPS. From ESPA, for the whole Landsat archive the Normalized Difference Vegetation Index (NDVI), quality layers, and the blue band reflectance were obtained.

From the quality layer (pixel_qa), all pixels that contained clouds, cloud shadows, or snow (even though snow does not occur in the study area) were masked out from the NDVI images. In addition, pixels with a blue-band reflection of less than 0.01 were masked out, as it was observed for many locations cloud shadows were not effectively detected by the quality layer. A buffer of 60m was also created around the combined mask to account for edge-effects. This masking and buffering process inevitably resulted in a loss of some useful pixels, but the erroneous inclusion of the vegetation indices that do not relate to the real vegetation conditions was considered a more serious problem because it may have a negative impact on the effective retrieval of phenology.

2.2.2. Sentinel-2 2015-2017

Sentinel-2 images were downloaded from ESA's Copernicus Open Access Hub (<https://scihub.copernicus.eu/dhus/>) and some through the USGS Earth Explorer, given that a few good scenes were identified that seemingly were unavailable on the Copernicus hub. The images were atmospherically corrected using the Sen2Cor programme, which besides surface reflectance also provides a scene classification file containing cloud and cloud shadows. Based on the scene classification file, all pixels not classified as dark area pixels, vegetation, bare soils and water were masked out (<https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm>). In addition to this, all pixels with less than 0.01 reflectance in the blue band were masked out and a buffer of 50m was created around all masked pixels to cater for edge effects around cloud and cloud shadows.

NDVI was then calculated using this equation in IDL:

$$NDVI = \frac{Band\ 8 - Band\ 4}{Band\ 8 + Band\ 4} \quad \text{Equation 1}$$

where: *Band 8* = Near infra-red band (0.842 μm)

Band 4 = Red band (0.665 μm)

2.2.3. eMODIS 2001-2017

NDVI eMODIS v5 images from 2001 – February 2017 were used. The images are a product derived from measurements of the Moderate Resolution Imaging Spectroradiometer onboard the Terra satellite and comprises of 10-day maximum NDVI value composites at 250m resolution (Vrieling et al., 2016). Originally these composites are delivered every five days however just the 1-10, 10-20 and 21 to last day images were used. During NDVI pre-processing a piecewise linear regression Swets algorithm (Swets et al., 1999) is applied for each pixel time series for data smoothening while still maintaining the vegetation maxima. The three images per month were assumed to be from the following dates; 6th, 16th, 26th for every month other than February which was set at 25th since information about the exact date of acquisition per pixel are not provided with the data.

2.2.4. Monthly rainfall data 1940-2017

Monthly rainfall data for the Kapiti Farm was provided by ILRI from the year 1940-2017 from the single rainfall station in the area. This data was used to understand and relate the seasonal precipitation with multi-year annual deviations of Landsat. It was assumed that the single rainfall station (*Figure 2*) in Kapiti could represent the temporal variability of rainfall for the entire Kapiti Farm. *Figure 3* shows an overview of the annual seasonal rainfall over the years. Only data from 1999 -2016 are shown because multi-year Landsat data used was within that range.

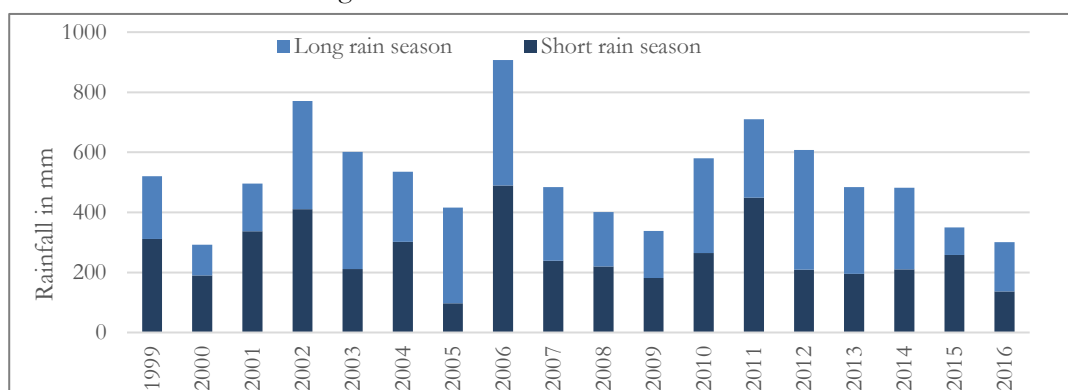


Figure 3: Total annual rainfall in millimetre of Kapiti Farm per season. Short rain season runs from September to February while long rains is from March to August

2.2.5. Vegetation sample points

Based on a reconnaissance field survey in Kapiti Farm four different vegetation composition were discerned:

- (A) = Grassland - purely grass
- (B) = Denser tree cover – tall trees (>2m) and understory shrubs and grass
- (C) = Diverse shrublands – more than one shrub type and understory grass
- (D) = Acacia shrublands – acacia shrubs and understory grass

A total of twenty-four sample points, six points for each group, was collected during field data collection using the global Land-Potential Knowledge System (LandPKS) protocol. LandPKS protocol is a cloud-based system used for monitoring vegetation in the rangelands (Corinna et al., 2011). This sampling method involves finding a central location inside a homogenous vegetation structure and then measuring 25 meters in the North, East, West and South direction. After every 5 meters, a one-meter bar (with five marks at 10cm, 30cm, 50cm, 70cm and 90cm) is steadily dropped down vertical and the vegetation under or over the marked position is noted down in the LandPKS form (Appendix 1). For this study, some parts in the original form were removed and other sections added, for instance, visual estimation and the photo sections. The visual estimation at each location was done to assess how well the LandPKS protocol agreed with a simple visual vegetation estimation while the photos were taken at an arm-angle in four directions (NE, SE, NW, SW).

The twenty-four sample points were equally distributed within the four vegetation structure strata leaving each group with six points and their spatial distribution can be seen in *Figure 4*. These data were used to relate the vegetation composition on the ground with the phenological parameters retrieved from multi-year Landsat data.

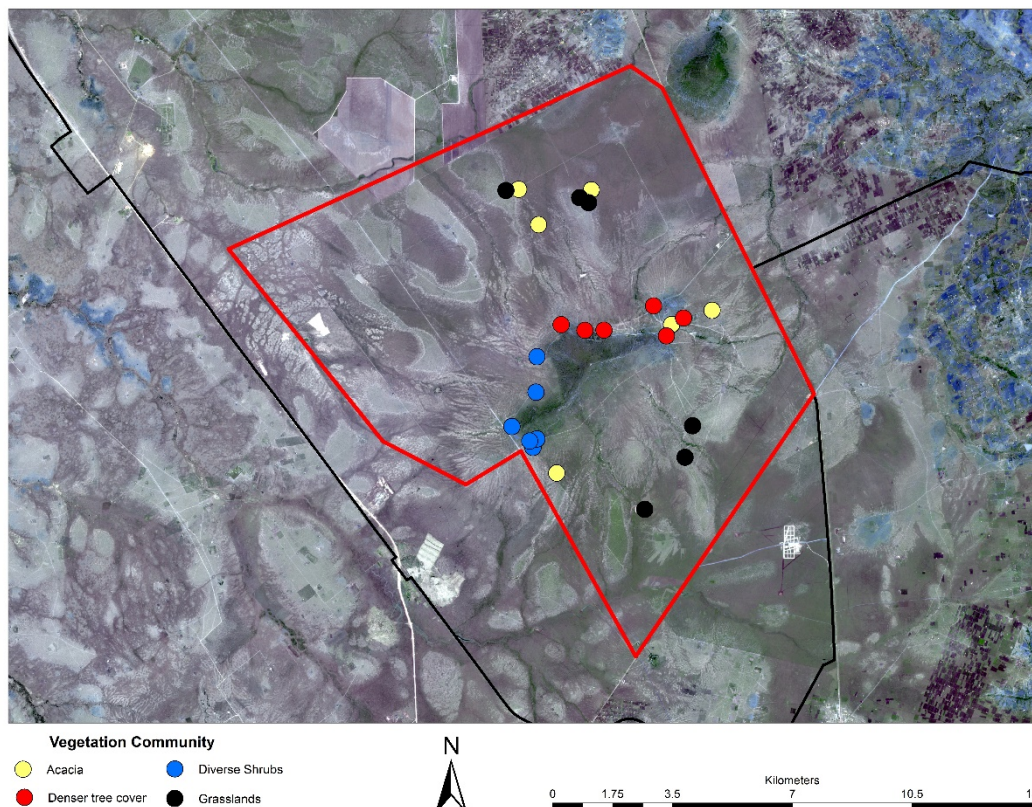


Figure 4: Spatial distribution of vegetation community within Kapiti Farm. The background is a Sentinel-2A RGB image acquired on 20th December 2016.

2.3. Software

The following software was used for the accomplishment of this research:

- ENVI + IDL – data processing and phenology parameters estimation from Landsat, Sentinel-2 and eMODIS data. This included an adaptation of existing IDL-code written by Michele Meroni (JRC) and Anton Vrieling (University of Twente).
- Excel 2016 – Statistical analysis
- ArcMap 10.5 – Raster polygon conversion and cartographic production
- R version 3.4.1 – spatial-autocorrelation analysis and density scatter plots

3. METHODS

3.1. Phenology retrieval model

For the NDVI trajectories derived from various data sources (Chapter 2), a model was fitted to the data. Here the double hyperbolic tangent model was selected (Meroni et al., 2014) for estimating vegetation phenology and other seasonal greenness parameters, which can be written as:

$$\text{NDVI}(t) = a_0 + a_1 \frac{\tanh[(t-a_2)*a_3]+1}{2} + a_4 \frac{\tanh[(t-a_5)*a_6]+1}{2} - a_4 \quad \text{Equation 2}$$

where t = time in days and a_0 to a_6 are the seven model parameters that need to be fitted. The function was fitted using MPFIT as implemented in IDL (Markwardt, 2009). The model parameters relate to features of the NDVI trajectory and are initiated as written in *italics* (Vrieling et al., 2017):

- a_0 : the minimum NDVI value. *The minimum NDVI in the first half of the time series;*
- a_1 (a_4): the NDVI amplitude of the green-up (senescence) phase. *The difference between the maximum NDVI and the minimum NDVI in the first (second) half of the time series;*
- a_2 (a_5): the inflection point (days) for the green-up (senescence phase). *The midpoint between the start (end) of the time series and the time of maximum NDVI;*
- a_3, a_6 : control the slope at the inflection point for both phases (day⁻¹). *Initialized at 0.02 [-0.02].*

The double hyperbolic tangent model, whose functions are similar to a double logistic model, was chosen for this study because of the ability to retrieve more accurate and successful fits in regions having multiple vegetation types (Vrieling et al., 2017a). The study area has two separate vegetation seasons within one year (short and long rains) and to fit the model for each season ‘breakpoints’ needed to be set. These breakpoints or rather transition dates that define the time-frame for each season were guided by the average NDVI trajectory as observed from multi-year eMODIS v5 data from 2001 to 2017 (Figure 5). Though the minimum NDVI value between the end of short rains and start of long rains was observed to be around 6th of March 15th of March was taken as the transition date based on user preference. Between the end of the long rains and the start of short rains, transition dates were set at 15th October and 15th September. The overlap between the two seasons was because of the more extended drier period observed (Figure 5).

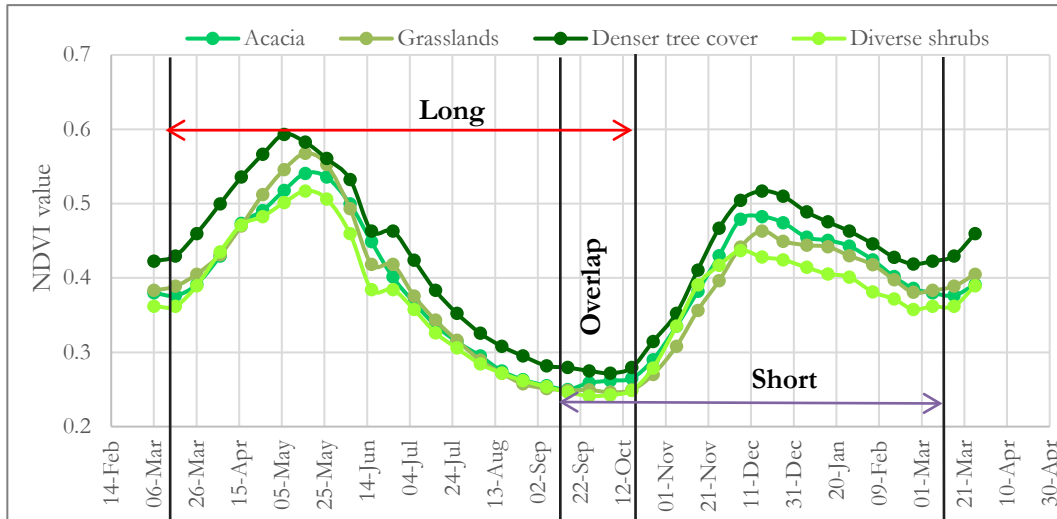


Figure 5: representative pixel for each vegetation structure average dekadal NDVI trajectory from multi-year eMODIS data (2001-2017). These locations were randomly selected from the six-sample location in each group. Black vertical lines represent the approximate transition dates. The red line is the temporal range for the long rains while the purple line is for the short rains

After fitting the double hyperbolic tangent model, a threshold approach (White et al., 1997), which uses annually redefined maximum and minimum NDVI values, was used to estimate the following parameters:

- SOS_{50} : start of season, defined as the first moment when the fitted function reaches 50% of the amplitude between maximum VI and the fitted minimum value (growth) for each season;
- EOS_{50} : end of season, defined as the first moment of the senescence phase when the fitted function reaches 50% of the amplitude between maximum VI and the fitted minimum value (senescence) for each season;
- LGS_{50} : length of the season, EOS_{50} minus SOS_{50} ;
- $maxNDVI$ – the maximum value of the model;
- $cumNDVI$ – Integral of the fitted model between SOS_{50} and EOS_{50} ;
- amplitude 1 & 2 –the difference between the maximum and the fitted NDVI value (growth and decay phase respectively).

3.2. Phenology from combined Landsat and Sentinel-2 data

To assess if phenological parameters can be effectively retrieved using finer spatial resolution data of individual-year observations, optical images acquired between 2015 to 2017 by three satellites, i.e., Landsat 7 & 8 and Sentinel-2, were fused together, whereby the Sentinel-2 data was resampled to match the 30m Landsat resolution. These sensors were combined because the number of cloud-free observations from each sensor alone could not allow retrieval of vegetation phenological parameters in a single season.

When examining the spatial overlay between imagery, it was found that for the study area Sentinel-2 has a 20m shift in the northward direction and a 5m shift in the westward direction in relation to Landsat. In this case, before aggregation Sentinel-2 was manually shifted by updating the header information describing the tie point for the upper-left pixel. While in the y-direction the 10m Sentinel-2 resolution cells precisely fitted within a Landsat pixel, in the x-direction this fit is not precise. To account for the 30m Landsat containing two full 10m resolution Sentinel-2 pixels and two pixels for 50%, a weighted average was implemented in IDL for the spatial aggregation. For the 10m resolution Sentinel-2A pixels that did not completely cover the output Landsat definitions, they received a 0.5 weight. To avoid incorporating pixels with a valid NDVI value, due to the clouds and cloud shadows, such pixels were not used to calculate the spatial average NDVI. If less than 70% of the 30x30m output pixel had a valid NDVI, a no data value was assigned to that pixel.

A layer stack of the two Landsat and 30m resampled Sentinel-2 datasets was made and used as input for the phenology retrievals with the double hyperbolic tangent model. To appreciate the number of valid NDVI data points for each season (2015-2016 short, 2016 long, 2016-2017 short, 2017 long), maps reflecting this number were generated. A maximum temporal gap map between the observations at each location was generated for the entire landscape for each season separately. The number of valid NDVI observation map and the maximum temporal gap map was used to assess for which individual season sufficient observations were available to attempt fitting of the double hyperbolic tangent model.

Model fitting was first tested for the pixels corresponding to the vegetation sample points and later applied over the whole region to generate spatial patterns of the phenological parameters. Variograms were generated using R packages (gstat, stats4: Gräler et al., 2016) to test for spatial consistency of the retrieved spatial patterns.

3.3. Phenology from combining multi-year Landsat acquisitions

Combining years of Landsat meant that for each observation only the day of the year (DOY, ranging from 1 to 366) was retained, and not the year (Melaas et al., 2013). In this way, a so-called ‘synthetic year’ was created that combines all NDVI observations for each pixel. This transformation leads to increased number of observations within the synthetic year, allowing to overcome issues of few cloud-free observations that any individual year may face. To understand the seasonal behavior and the year-to-year variability of NDVI for sampled vegetation locations, the NDVI (for all non-masked dates) was plotted against the day-of-year (DOY) while still allowing to identify which observation belongs to which year.

To avoid scattered Enhanced Vegetation Index (EVI) values observed between different years, Melaas et al. (2016) applied some smoothing on the data. For the long-term Landsat images, Melaas et al. (2016) calculated the 10th and 90th percentile average of all EVI values at each pixel contained within a moving window of three years successively before constructing the ‘synthetic year.’ This procedure was not carried out in this study, as it was observed that the high inter-annual variability combined with often missing observations particularly for higher NDVI values could have resulted to a not so well successful data smoothing. The double hyperbolic tangent model was fit through the ‘synthetic year’ for the sampled locations and then applied to all pixels within Kapiti Farm. To test for spatial consistency, variograms of the retrieved spatial patterns were generated using R packages (gstat, stats4; Pebesema et al., 2017).

To assess how finer spatial resolution phenological retrievals from Landsat agreed with better temporal resolution phenological retrievals, eMODIS data were introduced as a benchmark. Studies such as Vrieling et al. (2016) used averages of annual phenological parameters extracted from eMODIS time series but for this study, all years of eMODIS were used as a synthetic year. To assess to what extent using eMODIS as a ‘synthetic year’ deviated from annual averages of phenological parameters estimated from eMODIS time series the double hyperbolic tangent model was fit for every single year for the twenty-four sample points and the averages calculated. The model was then fit for the whole of Kapiti Farm using all years of eMODIS as a synthetic year. The phenological parameters retrieved from both methods were recorded for the twenty-four sample points together with their standard deviations.

To assess if retrievals from multi-year Landsat were almost similar to those retrieved from multi-year eMODIS, Landsat retrievals were aggregated to the eMODIS resolution in IDL. The aggregation process was carried out in IDL and to cater for the cells having a no data value; a condition was set that at least 50% of the coarser resolution pixel needed to have a valid data value. Density scatter plots were generated using R-Packages (ggplot2, hexbin) where, SOS, EOS, cumNDVI, AMP1 and AMP2 retrievals from Landsat were plotted against the same retrievals extracted from eMODIS. The coefficient of determination (R^2), Root Mean Squared Deviation (RMSD) and Mean Signed Deviation (MSD) were calculated for each phenological parameter. The R^2 value, in this case, explains how well the spatial variability of the mentioned retrievals from eMODIS is explained by the variability of the same retrievals from Landsat; the RMSD value shows how retrievals from eMODIS deviate from those extracted from Landsat while the MSD value summarises how well each retrieval from the two-dataset matches.

To help understand if grouping multi-year Landsat NDVI observations using seasonal precipitation would result to a ‘synthetic year’ having almost similar NDVI profiles, two sample locations from areas having grassland (A) and denser tree cover (B) were used for testing in each season separately. The seasonal precipitation was divided into five classes based on the minimum and maximum value seen for each season. These classes were as follows: 0 -100, 100-200, 200-300, 300-400 and 400-500.

3.4. Explaining spatial variability of phenology from differences in vegetation composition

Phenological parameters and greenness measures that were generated from multi-year Landsat were used to explain the spatial variability of vegetation composition in Kapiti Farm. *Figure 6* shows a summary of the average estimated vegetation cover using LandPKS protocol (Herrick et al., 2013) for each group.

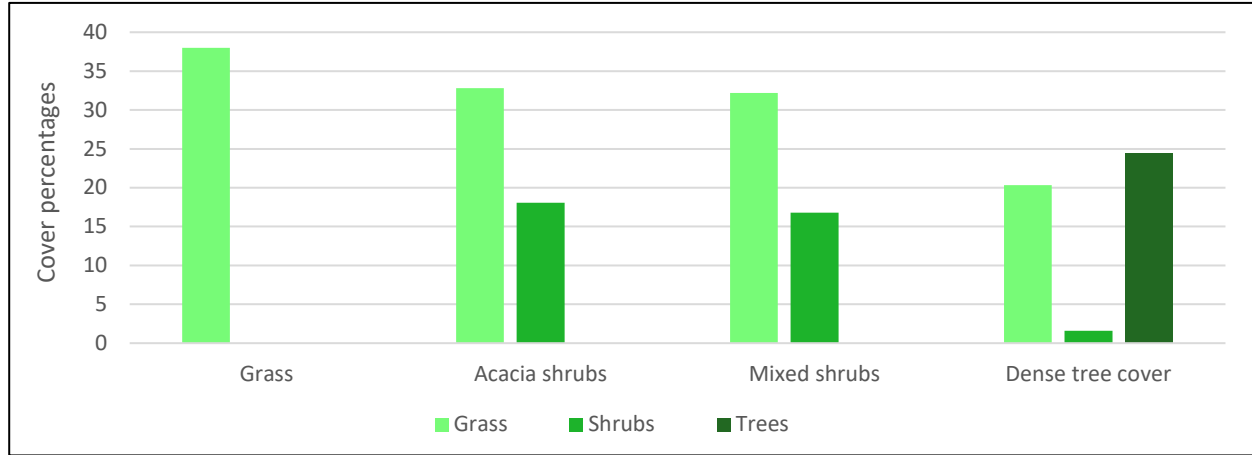


Figure 6: A summary of estimated vegetation cover per vegetation community using LandPKS protocol.

To understand if there were significant differences in vegetation phenological parameters and or seasonal greenness measure in these four groups, single factor analysis of variance (ANOVA) was carried out in Excel separately for each parameter. The null hypothesis tests if the means of the groups are the same and to reject the null hypothesis one needs to check whether the calculated F-value is greater than the critical F-value that is calculated by setting a certain level of significance (in this case 0.05). A rejected null hypothesis only indicates that at least one of the means is significantly different from the other but does not indicate which that is. To understand which means were significantly different the Tukey Honest Significant Difference (HSD) test was carried out. The HSD test involved making pairwise comparison among the means of different groups. *Equation 3* was used to calculate the HSD-statistic. The score was then checked from Tukey critical value table. If the HSD statistic was greater than the HSD critical value, it was concluded that the two means were significantly different (Glen, 2016).

$$HSD = \frac{M_i - M_j}{\sqrt{\frac{MSE}{n}}} \quad \text{Equation 3}$$

Where

$M_i - M_j$ – the difference between the means of the pairs in question

MSE – Mean Square error between the two pairs

n – number observations

4. RESULTS

4.1. Phenological parameters from fused Landsat and Sentinel-2 data

The first set of analyses examined the number of observations per pixel and the maximum temporal gap between the observations for each season tested between September 2015 and September 2017 (*Figure 7*). Maximum temporal gap calculations were done for locations having more than eight observations because the model requires at least eight observations to initiate the fitting. It was noted that locations with few number of observations showed a larger maximum temporal gap in most cases across the different seasons.

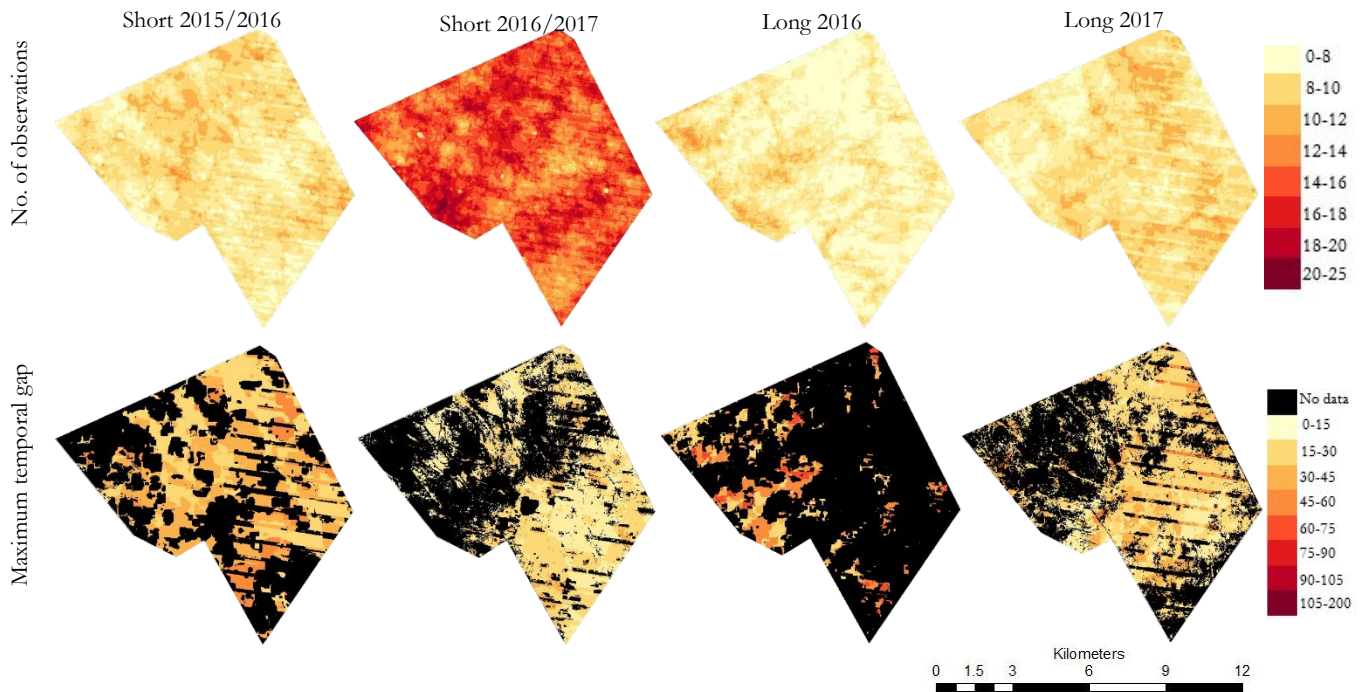


Figure 7 Number of observation with a valid NDVI value per season and the maximum temporal gap between these observations from fused Landsat 7- 8 and Sentinel-2 data between September 2015 - September 2017

Temporal profiles from fused Landsat 7- 8 and Sentinel-2 NDVI for the different vegetation communities locations are shown in *Figure 8*. These locations were picked purposively from the six sample points per each group, ensuring that the location had retrievals from each season. Nevertheless, Acacia shrublands (D) did not have any sample with retrievals in all seasons meaning the profiles are from different locations (XY). The short rains of 2015/2016 have a relatively higher maximum NDVI (~ 0.6) compared to the short rains of 2016/2017 (~ 0.45). This likely links to the smaller amount of precipitation during 2016/2017 short season (137mm) as compared to 2015/2016 (275mm). SOS and EOS for the short rains are approximately in November and January (of the following year) respectively. The LGS for the short rains of 2015 is 84 ± 11 days while the short rains of 2016 have 45 ± 7 days. For the long rains of 2016, there were no retrievals from the representative vegetation samples. As *Figure 9* shows, the long rains 2016 has observations gaps existing between April and July which could have caused the unsuccessful retrievals because one of the phenology model conditions is only to try fitting if there are at least four observations before and after the midpoint. The temporal profiles for the long rains of 2017 are shown in *Figure 9*. On average the start and end of the season for the long rains within Kapiti Farm are within April and July respectively. The LGS for the long rains of 2017 is 66 ± 23 days on average.

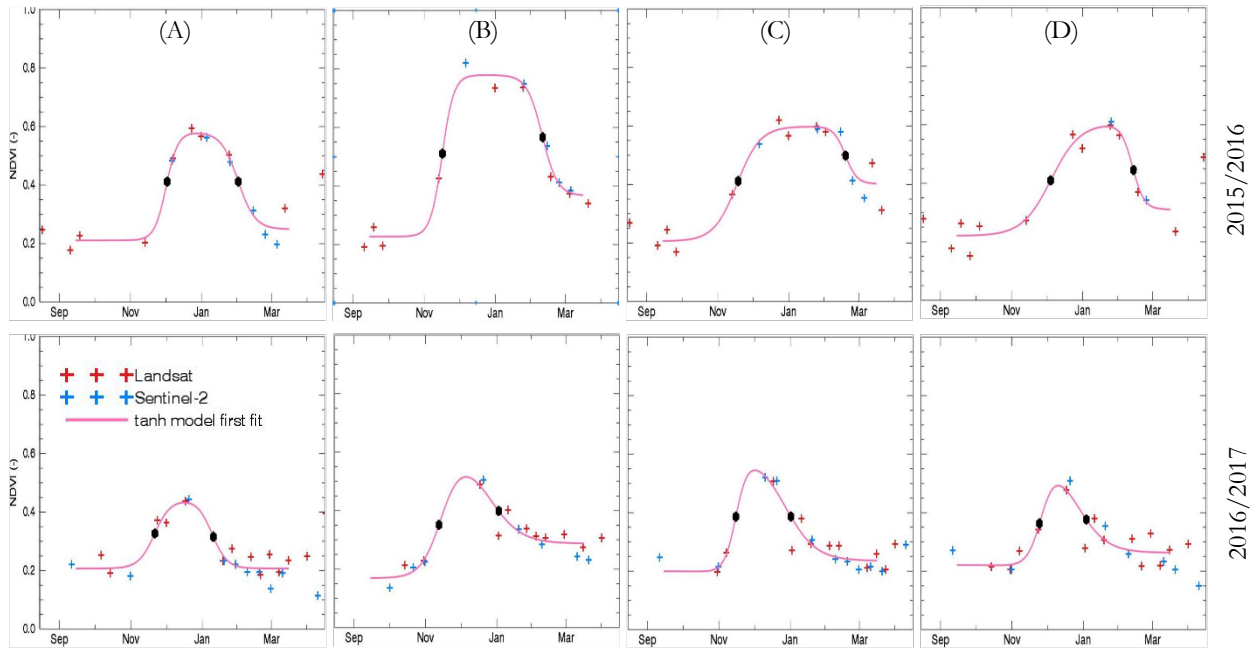


Figure 8: Short rains of 2015/2016 and 2016/2017 from the fused Landsat 7, 8 and Sentinel-2 observation. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs and (D) Acacia shrubs. The black dots indicate SOS_{50} and EOS_{50} respectively

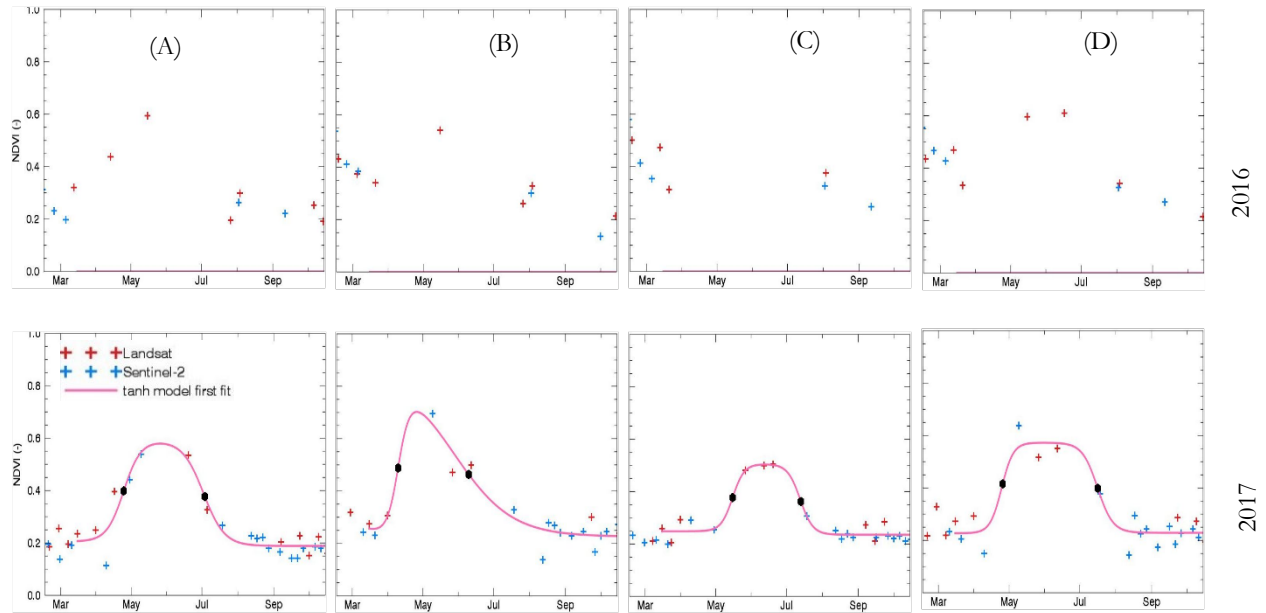


Figure 9: Long rains of 2016 and 2017 from the fused Landsat 7, 8 and Sentinel-2 observation. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs and (D) Acacia shrubs. The black dots indicate SOS_{50} and EOS_{50} respectively

To avoid repetition of a similar legend within all the derived phenological and seasonal greenness spatial patterns, a detailed legend was made. This Legend can be seen in Figure 10.

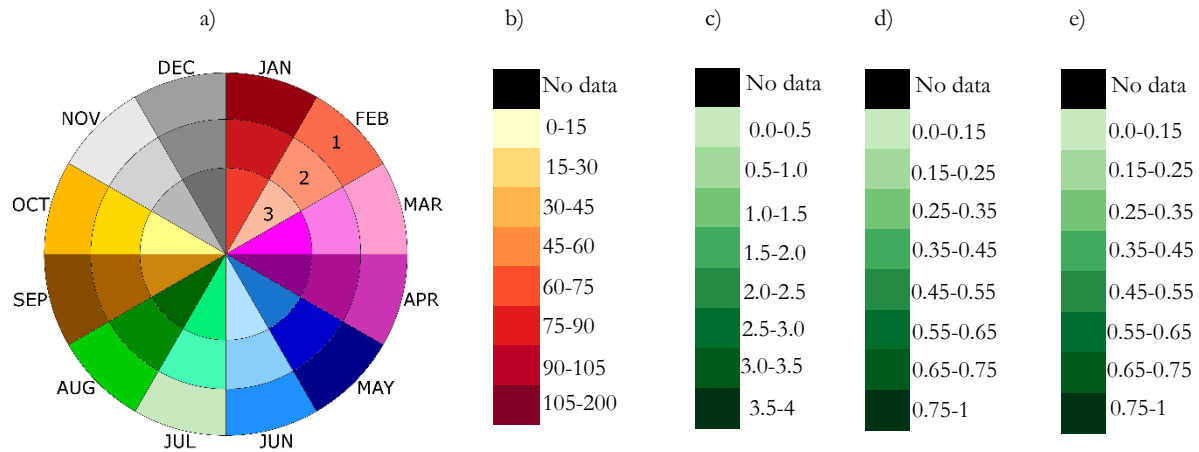


Figure 10: detailed legend for different figures used in this study. a) colours for each month; the colours are divided into 10 days in a month, the first ring (1) represent day 1-10, (2) day 11-20 and (3) day 21-30, this legend is used for SOS_{50} and EOS_{50} . b) days of the LGS_{50} . c) the $cumNDVI$ values. d) shows the $AMP1$ and $AMP2$ values. e) represents the $maxNDVI$ values

Figure 11 shows phenological parameters retrieved after applying the phenology retrieval model to all pixels of fused Landsat 7,8 and Sentinel-2 within Kapiti Farm. For the short rainy seasons, 49.10% (2015/2016) and 51.04% (2016/2017) of the total pixels had successful phenological retrievals while for the long rains of 2016 this was 15.96% and 53.57% for the long rains 2017. The average phenological parameters per vegetation community from the twenty-four sample locations are shown in Table 2. The short rains of 2015/2016 have an early SOS_{50} and a later EOS_{50} compared to those from 2016/2017. These differences could be because of the differences in nature of the model curve as seen in Figure 8 where the short rains 2015/2016 curves are more spread than those from 2016/2017.

Table 2: Phenological parameter retrievals from fused Landsat 7, 8 and Sentinel-2 NDVI time series for the vegetation communities of Kapiti Farm. These are means of the six points collected for each group.

	SOS_{50}	EOS_{50}	LGS_{50}	$cumNDVI$	$maxNDVI$	$AMP1$	$AMP2$	No.of samples
Vegetation type	2015/2016 short season							
(A)	11 Nov \pm 14	2 Feb \pm 5	82 \pm 19	1.47 \pm 0.31	0.57 \pm 0.02	0.37 \pm 0.02	0.33 \pm 0.04	3
(B)	16 Nov \pm 0	12 Feb \pm 3	88 \pm 3	1.98 \pm 0.19	0.70 \pm 0.09	0.46 \pm 0.08	0.30 \pm 0.09	4
(C)	19 Nov \pm 1	15 Feb \pm 2	88 \pm 5	1.62 \pm 0.12	0.58 \pm 0.01	0.38 \pm 0.01	0.26 \pm 0.07	2
(D)	4 Dec	13 Feb	71	1.296	0.59	0.37	0.28	1
	2016/2017 short season							
(A)	21 Nov \pm 2	6 Jan \pm 4	45 \pm 4	0.64 \pm 0.07	0.46 \pm 0.07	0.26 \pm 0.05	0.25 \pm 0.05	4
(B)	23 Nov \pm 6	10 Jan \pm 6	47 \pm 4	0.79 \pm 0.07	0.54 \pm 0.04	0.31 \pm 0.05	0.25 \pm 0.05	5
(C)	23 Nov \pm 8	29 Dec \pm 2	36 \pm 14	0.61 \pm 0.17	0.55 \pm 0.007	0.34 \pm 0.003	0.30 \pm 0.001	2
(D)	22 Nov \pm 5	8 Jan \pm 5	47 \pm 8	0.67 \pm 0.06	0.47 \pm 0.05	0.27 \pm 0.02	0.24 \pm 0.04	4
	2017 long season							
(A)	4 May \pm 18	11 Jul \pm 7	68 \pm 21	0.96 \pm 0.29	0.46 \pm 0.09	0.25 \pm 0.08	0.28 \pm 0.09	4
(B)	10 Apr \pm 2	2 Jul \pm 25	83 \pm 25	1.53 \pm 0.28	0.62 \pm 0.09	0.35 \pm 0.08	0.37 \pm 0.11	4
(C)	6 May \pm 16	12 Jul \pm 1	67 \pm 16	1.08 \pm 0.27	0.50 \pm 0.03	0.29 \pm 0.04	0.29 \pm 0.04	6
(D)	9 May \pm 20	9 Jul \pm 13	62 \pm 14	0.55 \pm 0.07	0.31 \pm 0.09	0.32 \pm 0.09	0.34 \pm 0.07	4

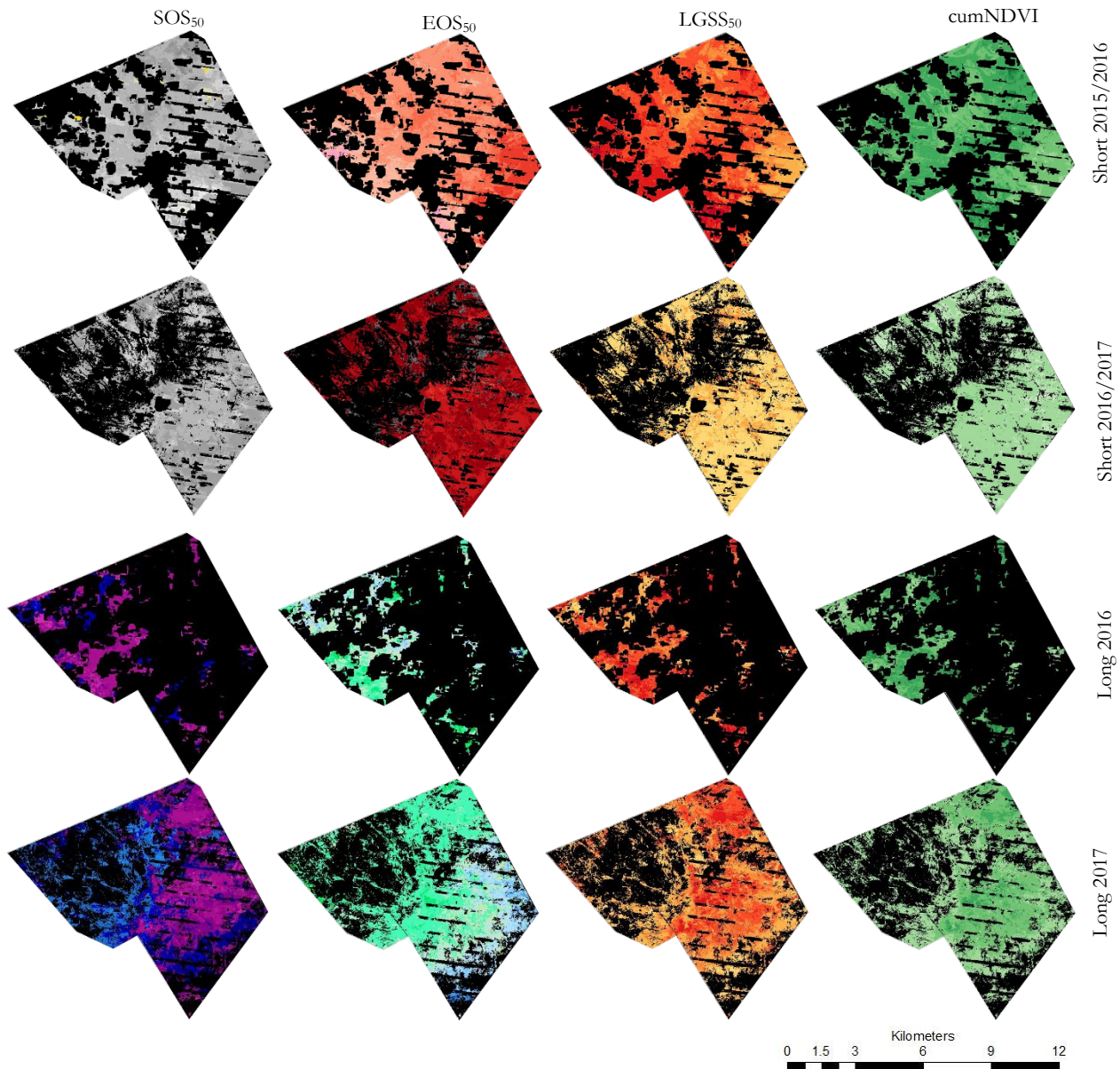


Figure 11: Phenological parameters and cumNDVI retrieved from fused Landsat 7, 8 and Sentinel-2 within Kapiti Farm. They are from the four different seasons. Short rains 2015/2016, short rains 2016/2017, long rains 2016 and long rains 2017. Detailed legend is in Figure 10

Further statistical analysis carried out to assess the spatial consistency of the retrieved phenological patterns are shown in Figure 12 in form of variograms. The differences in variograms forms could not be explained, but it can be concluded that at nearby locations the semi-variance is lower than locations further apart indicating that the further the pixels are from each other the less correlated they are.

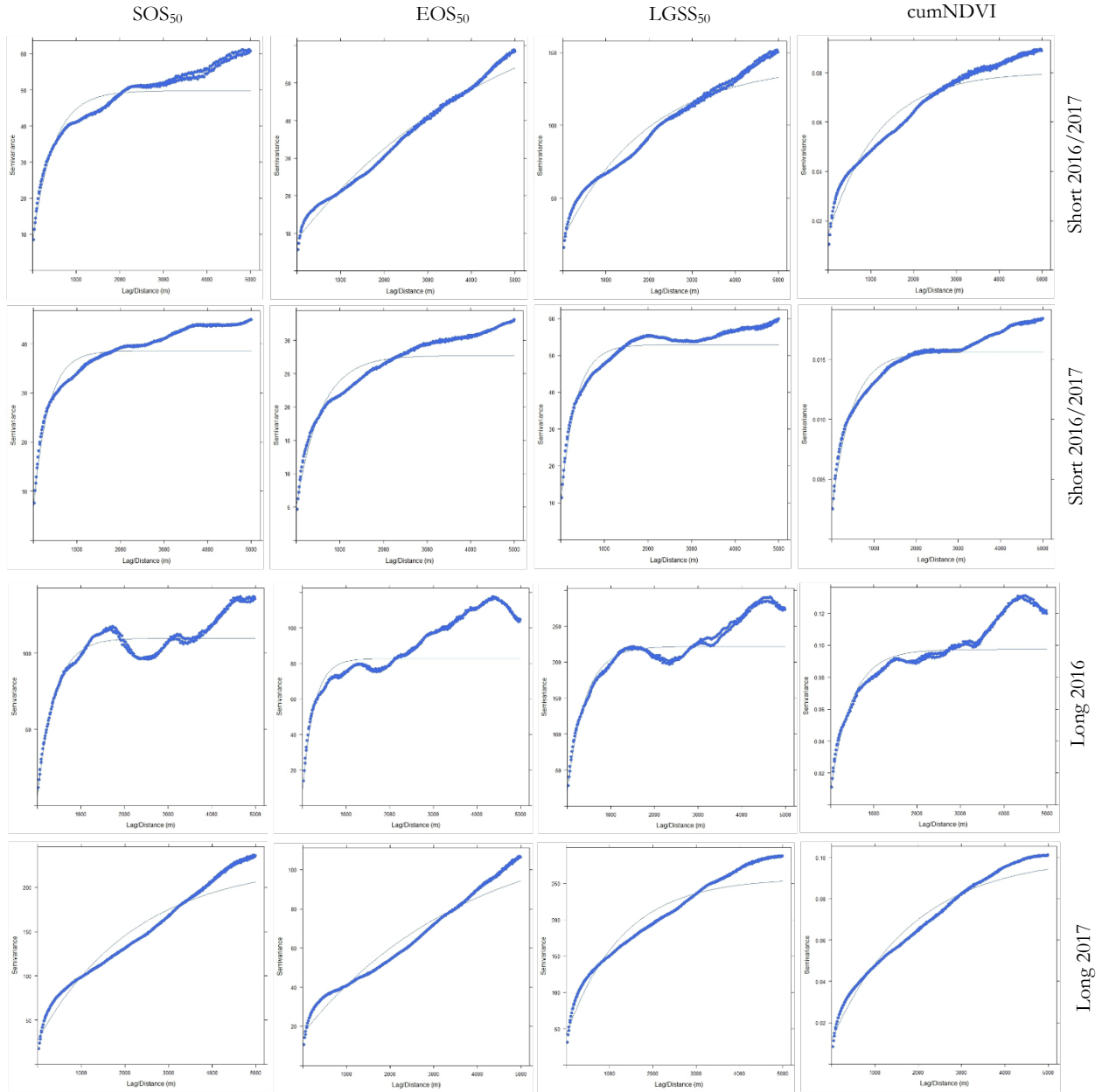


Figure 12: Variograms of phenological and seasonal greenness parameters retrieved from fused Landsat 7, 8 and Sentinel-2 NDVI time series of Kapiti Farm

4.2. Phenological parameters from multi-year Landsat and eMODIS data

Temporal profiles from multi-year Landsat NDVI time series as a ‘synthetic year,’ for the representative vegetation community are illustrated in Figure 13 for both seasons. A clearer temporal pattern emerges from the plots of the long rains while the short rains plots shows observations that appear to be somewhat noisy hence lacking clear spatial patterns.

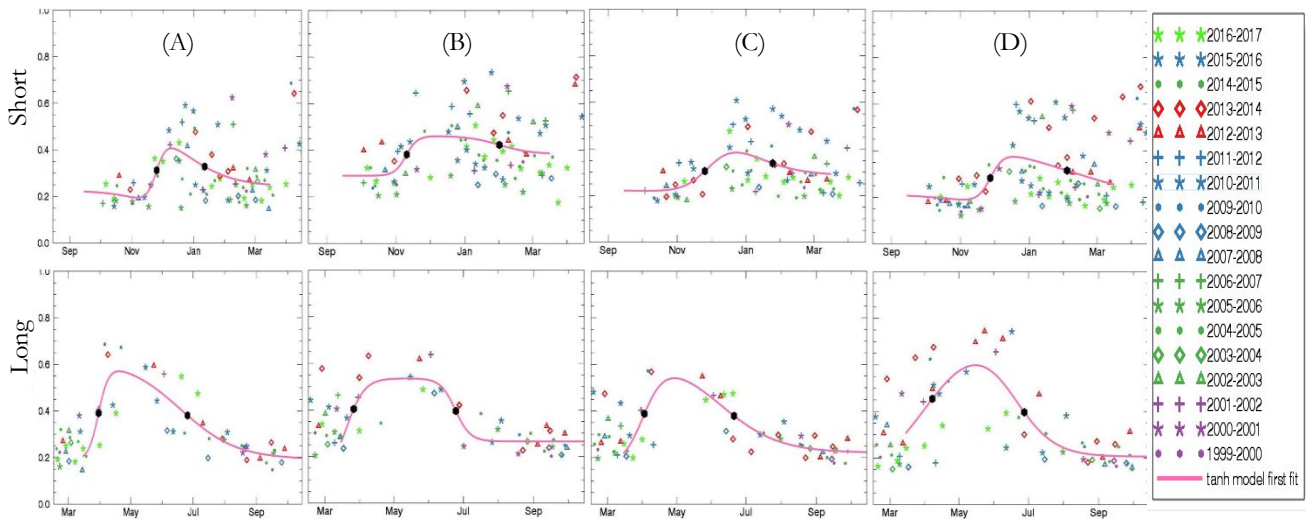


Figure 13: Combined years of Landsat archive NDVI time series for the representative vegetation community. The black dots represent SOS_{50} and EOS_{50} respectively. The locations are the same but different vegetation season. (A) Grasslands, (B) Denser tree cover, (C) Diverse shrubs, (D) Acacia shrubs.

Figure 14 shows the spatial patterns of selected phenological parameters after fitting the model through each pixel. For the long rains, parameters could successfully be retrieved for 99.1% of the pixels while for short rains season this was 64.22%. As seen from the profiles (Figure 13) the short rains observations were very noisy which could have translated to the low percentage of successful fits. On average the long rains experiences SOS_{50} on 4 (± 9 days) April, EOS_{50} on June 25 (± 9 days). The short rains have SOS_{50} on 21 (± 7 days) November, EOS_{50} on January 18 (± 12 days). The length of the season is 57 ± 14 (short) and 82 ± 15 (long) days on average. Figure 15 shows variograms generated to test the spatial consistency of the retrieved pattern. The nature of the forms from these variogram depicts that the semi-variance is lower for observations near each other than those far from each other. Table 3 provides the average summary of the phenological and seasonal greenness parameters as observed from the twenty-four sample locations.

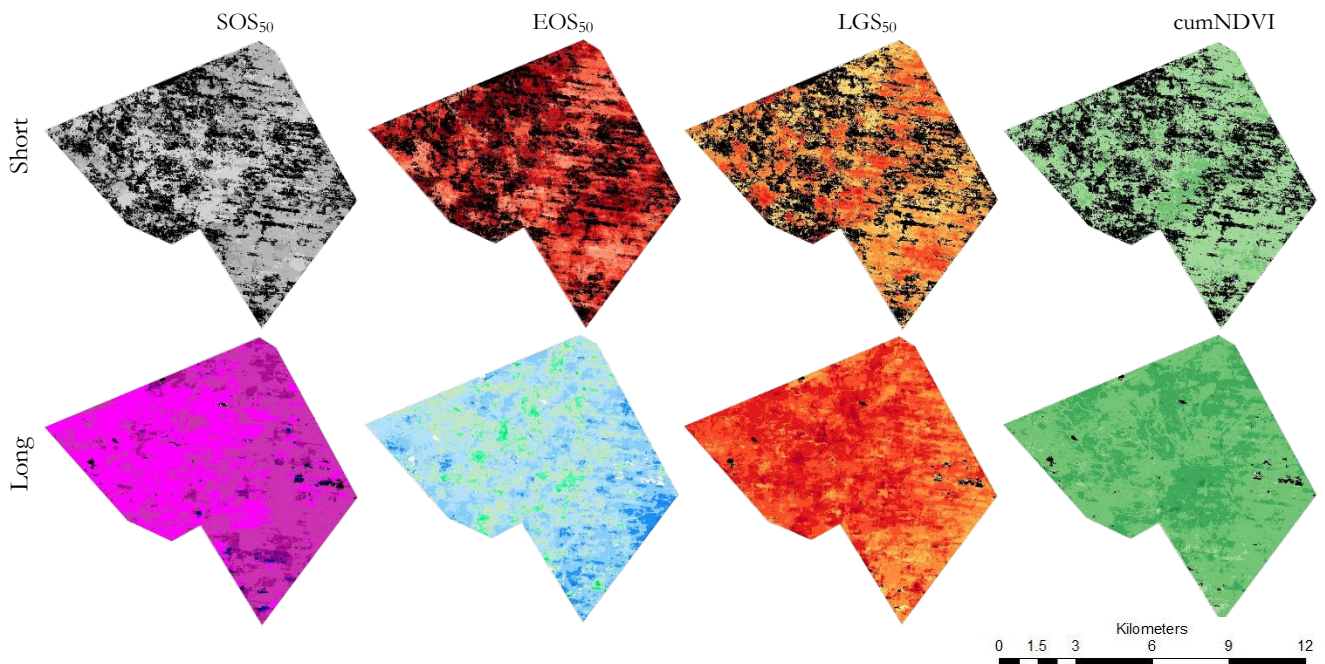


Figure 14: Phenological and seasonal greenness parameters retrieved from combined Landsat archive within Kapiti Farm. For the legend see Figure 10.

Table 3: Retrieved phenological and seasonal greenness parameters averaged from the 24 sample points from combined years of Landsat archive

	SOS ₅₀	EOS ₅₀	LGS ₅₀	cumNDVI	maxNDVI	AMP1	AMP2	No.of samples
Vegetation type	Combined Landsat archive – short rains							
(A)	27 Nov \pm 7	15 Jan \pm 10	49 \pm 12	0.63 \pm 0.17	0.41 \pm 0.02	0.20 \pm 0.03	0.14 \pm 0.03	6
(B)	10 Nov \pm 1	22 Jan \pm 12	73 \pm 12	1.12 \pm 0.15	0.47 \pm 0.02	0.21 \pm 0.06	0.09 \pm 0.02	2
(C)	18 Nov \pm 6	25 Jan \pm 16	68 \pm 13	0.83 \pm 0.16	0.39 \pm 0.02	0.17 \pm 0.03	0.10 \pm 0.01	4
(D)	23 Nov \pm 4	16 Jan \pm 12	53 \pm 10	0.73 \pm 0.08	0.44 \pm 0.05	0.34 \pm 0.05	0.14 \pm 0.02	4
	Combined Landsat archive – long rains							
(A)	6 Apr \pm 14	23 Jun \pm 6	78 \pm 18	1.34 \pm 0.24	0.57 \pm 0.08	0.35 \pm 0.05	0.36 \pm 0.08	6
(B)	30 Apr \pm 3	2 Jul \pm 7	94 \pm 8	1.72 \pm 0.15	0.59 \pm 0.07	0.32 \pm 0.06	0.33 \pm 0.07	6
(C)	3 Apr \pm 6	19 Jun \pm 4	77 \pm 8	1.41 \pm 0.16	0.61 \pm 0.06	0.39 \pm 0.06	0.36 \pm 0.05	6
(D)	8 Apr \pm 9	28 Jun \pm 12	81 \pm 19	1.46 \pm 0.22	0.61 \pm 0.06	0.38 \pm 0.06	0.39 \pm 0.05	6

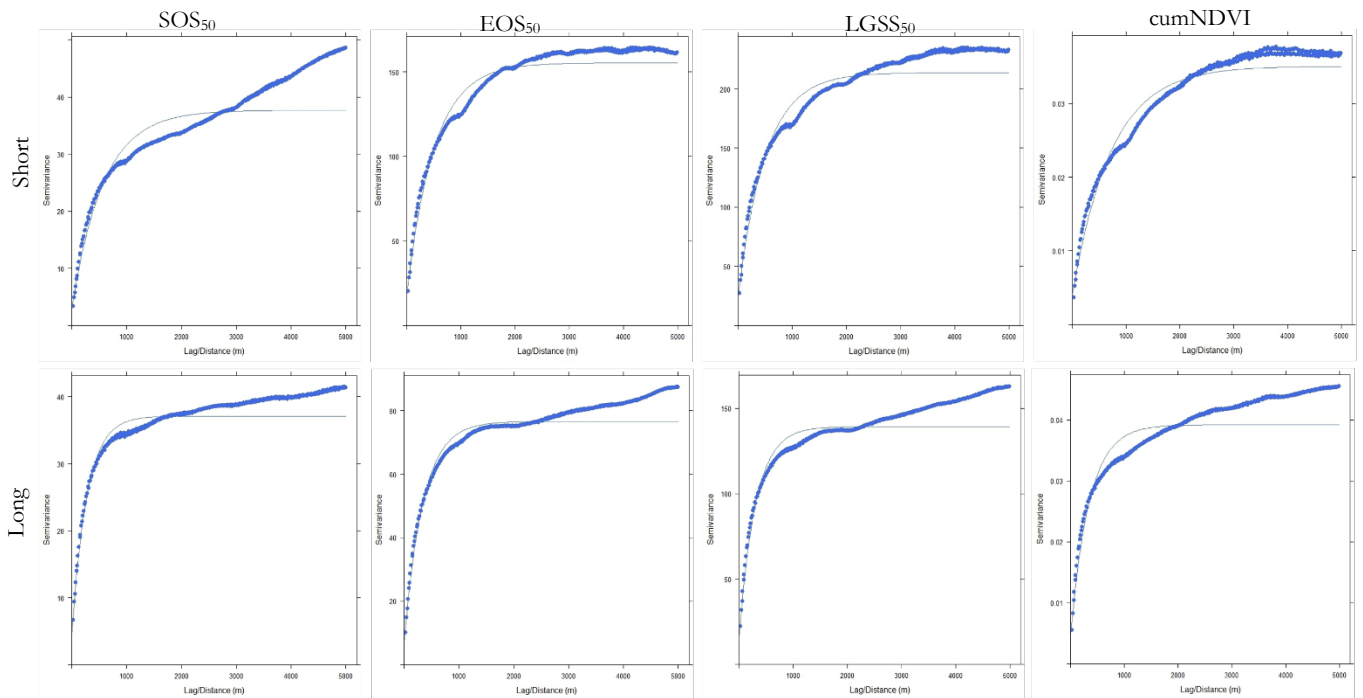


Figure 15: Variograms of phenological and seasonal biomass parameters retrieved from combined years of Landsat archive NDVI time series of Kapiti Farm

Phenological retrievals from combined years of Landsat were aggregated to 250m resolution for both seasons and compared to those retrieved from eMODIS. Figure 16 shows the spatial patterns of both aggregated Landsat and eMODIS. As earlier mentioned for eMODIS first the retrievals for the twenty-four sample points were estimated for each year (2001–2017) independently. For each retrieval, annual averages were calculated and are summarised in Table 4. This was to assess how combining years of eMODIS into a ‘synthetic year’ before fitting the model deviated from when the model is fitted in single years and later averages of the retrievals calculated. On average the SOS is within April and EOS early July for the long rains while for the short rains, SOS is within November and EOS in February. LGS is 53 \pm 7 (short) and 80 \pm 9 (long) days on average.

Table 4: Average phenological parameters and seasonal greenness for Kapiti Farm from eMODIS v5 NDVI time series

	SOS ₅₀	EOS ₅₀	LGS ₅₀	cumNDVI	maxNDVI	AMP1	AMP2	No. of samples
Combined eMODIS v5 Long Rains	13 Apr \pm 3	2 Jul \pm 3	81 \pm 8	1.38 \pm 0.11	0.56 \pm 0.02	0.39 \pm 0.05	0.35 \pm 0.02	24
Single years averages Long Rains	14 Apr \pm 3	25 Jun \pm 8	79 \pm 14	1.43 \pm 0.076	0.63 \pm 0.013	0.34 \pm 0.08	0.30 \pm 0.05	24
Combined eMODIS v5 Short Rains	16 Nov \pm 4	6 Feb \pm 5	54 \pm 7	1.36 \pm 0.15	0.52 \pm 0.034	0.27 \pm 0.03	0.12 \pm 0.06	19
Single years averages Short rains	15 Nov \pm 6	6 Feb \pm 7	52 \pm 11	1.26 \pm 0.08	0.45 \pm 0.022	0.24 \pm 0.03	0.10 \pm 0.05	19

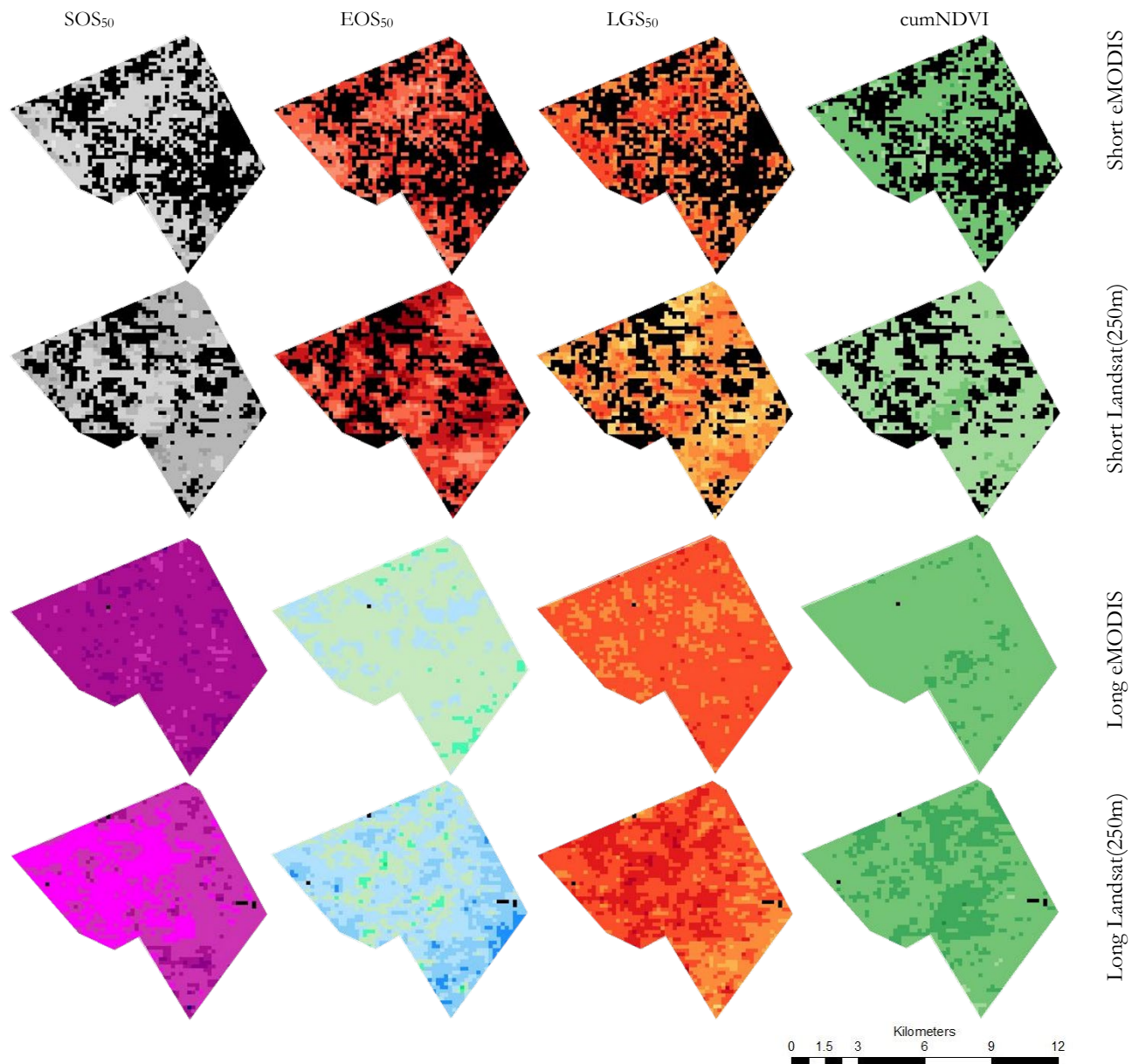


Figure 16: Aggregated combined years of Landsat(250m) and eMODIS v5 phenological parameters for the long and short rains season of Kapiti Farm. For detailed legend refer to Figure 10.

To assess how well aggregated Landsat agreed with eMODIS retrievals; density scatter plots were generated (Figure 17). The SOS_{50} and EOS_{50} of the short rains have an RMSD of ± 11 days while for the long rains has ± 14 days as well as an MSD value of ± 5 days (short rains) and ± 10 days for the long rains. The small R^2 exists within a small study area (130km^2) and therefore the small variability in eMODIS retrievals could not be accurately explained by the Landsat retrievals.

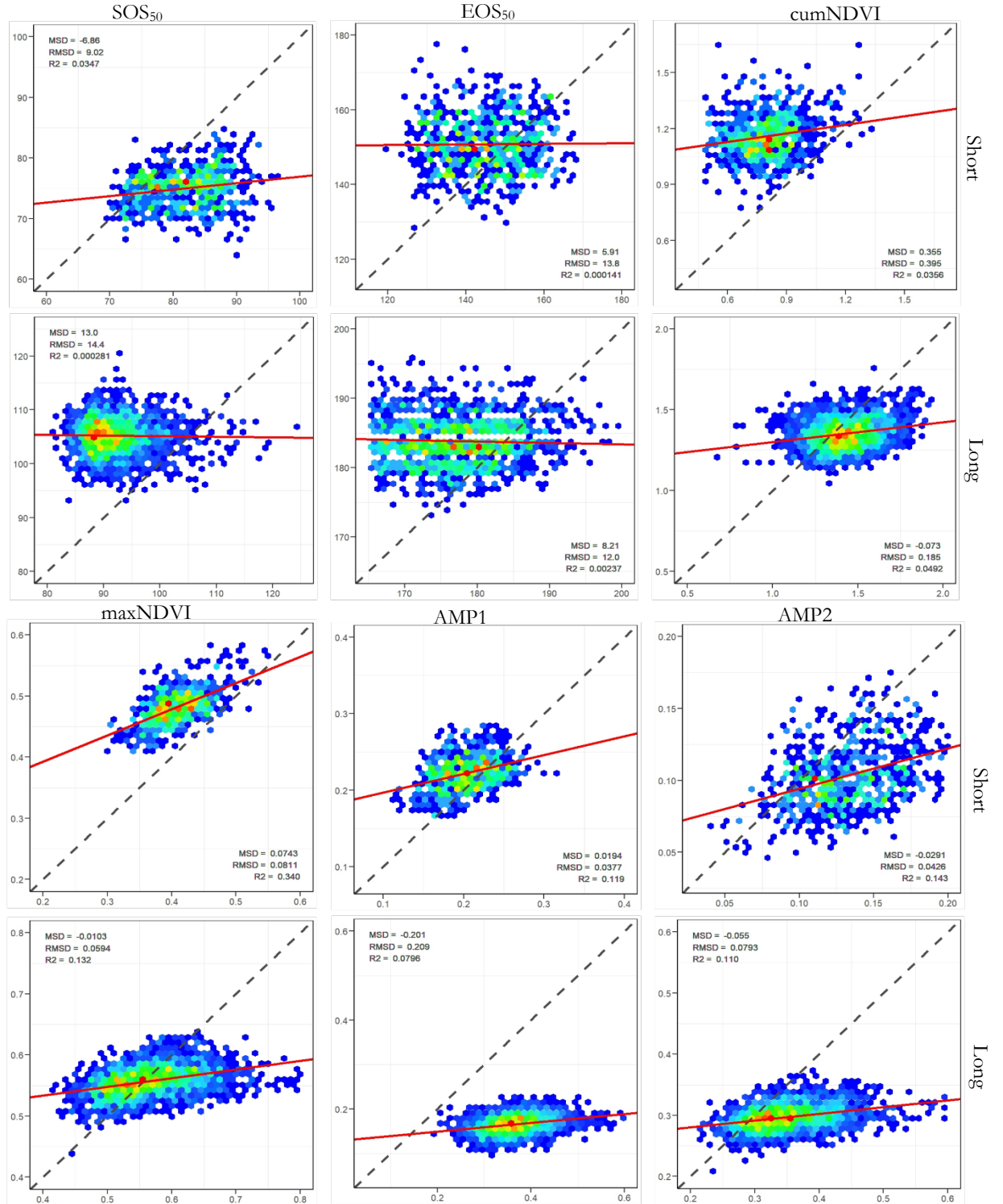


Figure 17: Density scatter plots of aggregated phenological parameters of combined years of Landsat (X-axis) against eMODIS v5 (Y-axis) phenological parameters. The X and Y axis for SOS_{50} and EOS_{50} are the DOY while for the cumNDVI, maxNDVI, AMP1 and AMP2 are the NDVI value. The dashed line in grey represent the slope while the red line represents the regression line

To assess if grouping years of Landsat with respect to the amount of precipitation in that year would result to a synthetic year having almost similar NDVI profiles two sample locations were used for testing (*Figure 18*). While the gradient of the NDVI profile is not clear, it was observed that for drier years (below 100mm) the NDVI values always resulted to very low values in all seasons while for the wetter years (above 100mm) no clear patterns in the NDVI values could be seen. In general, total seasonal precipitation alone does not seem to explain the amount of vegetation greenness in a given location.

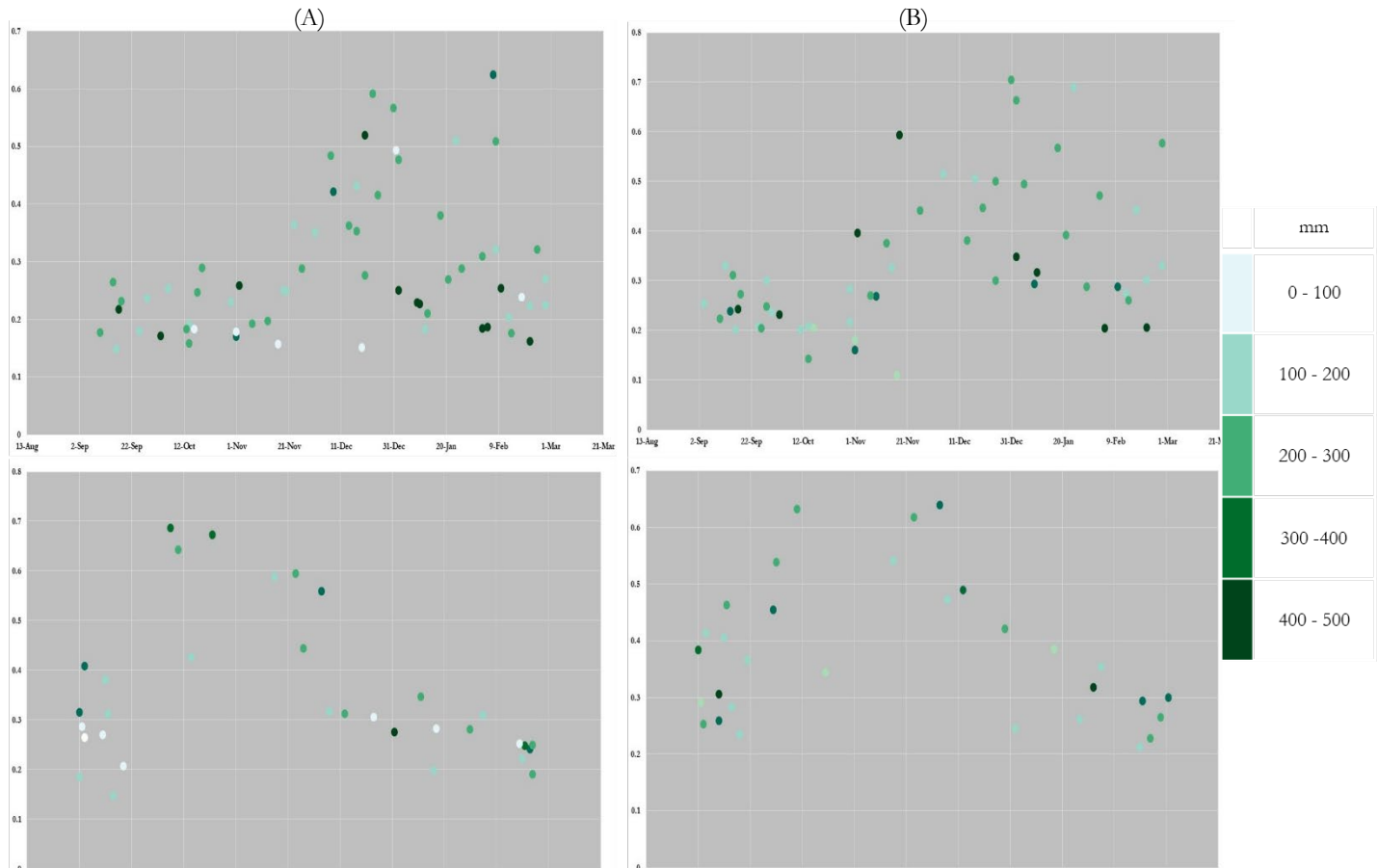


Figure 18: shows the combination of all 1999-2017 Landsat-derived NDVI values combined in a single synthetic year for two points. The colour of each observation relates to the amount of seasonal precipitation in each specific year.

4.3. Retrieved patterns in relation to vegetation composition on ground

Table 5 shows the results of the ANOVA analysis performed in excel to check for the differences of the retrieved phenological parameters and seasonal greenness between the vegetation groups for each season. The null hypothesis states that all the means of the four different groups are not different at the significance level of 0.05. This null hypothesis was rejected if the calculated F-value was greater than F-Critical or when the P-value is smaller than the level of significance used. For the long rains, only the ANOVA analysis done for the cumulative NDVI rejected the null hypothesis which meant that at least two of the vegetation community groups had different means. For the short rains, a significant difference was found for SOS₅₀, cumNDVI, maxNDVI, and AMP2 from the ANOVA analysis.

Table 5: ANOVA analysis of the vegetation community within Kapiti Farm. Results shaded in grey are those that rejected null hypothesis

		F-Value	P-value	F critical		F-value	P-value	F-critical
SOS ₅₀	Long rains	0.96373	0.429164	3.098391	Short rains	4.0859	0.03257	3.490294
EOS ₅₀		2.73219	0.070862	3.098391		0.6646	0.58962	3.490294
LGS ₅₀		1.73683	0.191715	3.098391		3.0680	0.06898	3.490294
cumNDVI		4.06754	0.020813	3.098391		5.6107	0.01221	3.490294
maxNDVI		0.34143	0.795617	3.098391		3.6953	0.04300	3.490294
AMP1		1.88213	0.165136	3.098391		1.8516	0.19447	3.490294
AMP2		0.56601	0.643752	3.098391		3.9823	0.03732	3.490294

Table 6 shows the results of the Tukey HSD analysis performed for all the parameters that rejected the null hypothesis tested in ANOVA. If the absolute mean difference between the pairs is greater than the calculated HSD value, then, it is concluded that there is a significant difference between the mean of the mention pairs. Among all the parameters that rejected the null hypothesis in both short and long rain season, it can be concluded that areas with denser tree cover (B) are the most different from all other means. Nevertheless, for the long rains, the only analysis that showed a significant difference is between the cumNDVI of the denser tree covered areas and the pure grassland areas. For the short rains, there is a significant difference between the SOS₅₀, cumNDVI, maxNDVI and AMP2 of denser tree covered (B) areas and the grasslands (A), cumNDVI and maxNDVI of grasslands (B) and diverse shrubs (C), cumNDVI of grasslands (B) and acacia (D), AMP2 of grasslands (B) and diverse shrubs, AMP2 of denser tree cover (B) and acacia shrubs (D) and maxNDVI and AMP2 of diverse shrubs (C) and acacia shrubs (C).

Table 6: Tukey HSD analysis carried for the parameters that rejected the ANOVA null hypothesis. A = grasslands, B = Denser tree cover, C = Diverse shrubs and D = Acacia shrubs. Results shaded in grey are those that rejected the Tukey HSD test

		HSD Value	AB	AC	AD	BC	BD	CD
Long Rains	cumNDVI	0.321	0.376	0.078	0.117	0.298	0.259	0.039
Short rains	SOS ₅₀	8.7	16.5	8.5	3.5	8	13	5
	cumNDVI	0.241	0.495	0.208	0.106	0.286	0.389	0.102
	maxNDVI	0.056	0.065	0.020	0.037	0.086	0.028	0.057
	AMP2	0.044	0.056	0.046	0.001	0.010	0.057	0.047

5. DISCUSSION

This study showed that it was possible to estimate vegetation phenology in a complex savanna ecosystem at a finer spatial resolution ($\sim 30\text{m}$) for a single season or a multi-annual average. Very little was found in the literature regarding the possibility of retrieving vegetation phenology for a single season by fusing Landsat and Sentinel-2 images. The closely related study is research carried out by Vrieling et al. (2017a) where they fused RapidEye and SPOT5 sensors to estimate vegetation phenology for a single season. Recently Vrieling et al. (2017b) showed the possibility of even using Sentinel-2 images alone to study the vegetation phenology, but this is only possible for regions near the overlapping orbits which have more Sentinel-2 images (~ 5 images) in a month compared to other regions of the world. In most cases, studies have been fusing sensors having a short revisit time (e.g. MODIS) with finer spatial resolution sensors (e.g. Landsat) to increase the number of cloud-free observations to enable vegetation phenology retrievals for a single season (Gao et al., 2017; Hazaymeh & Hassan, 2017; Walker et al., 2014). On the other hand, use of a Landsat ‘synthetic year’ to estimate multi-annual average vegetation phenology, had until recently only existed for temperate climates regions having a relatively homogeneous vegetation structure unlike the savannas (Melaas et al., 2016, 2013; Nijland et al., 2016; Vrieling et al., 2017a). While persistent cloud cover is an issue during acquisition of satellite images all over the world, the study area is located in the tropics near the equator which have 75% of the total images acquired in a single year showing heavy clouds contamination (Hashim et al., 2014). This made it challenging to estimate vegetation phenology, especially for a single season. Other than that, the study area has two short seasons (~ 3 months each), which makes it harder to accurately describe the vegetation dynamics for within these short periods of time only a few observations are left to capture the quick growth and decay phases. The frequent droughts in the area, which can sometimes lead to a failing season (i.e., NDVI values remains low) together with the inter-annual variations in vegetation greenness challenged the use of combined years of Landsat into a ‘synthetic year’, for until now, this method had only been applied in regions that have a single season that is relatively stable between single years. In general, persistent cloud cover, high inter-annual variation in vegetation greenness, two short seasons and frequent droughts were the major challenges encountered in this study.

Despite these difficulties, retrievals for a single season tested between September 2015 and September 2017 were possible. The long rains of 2017 had the highest percentage (53.7%) of successful pixels retrievals which is still lower than that attained (77.5%) by Vrieling et al. (2017b). A possible explanation of these differences might be because the location of Vrieling et al. (2017) study has a much longer season with far more gradual green-up and decay phase and it is located around the overlapping orbits which have more images hence more observations. Nonetheless, from the successful retrievals, the SOS and EOS for the short rains was within November and January while for the long rains (2017) was April and July respectively which seem to be consistent with other studies (Broadhead et al., 2003; Oscar et al., 2015; Vrieling et al., 2016) as well as the known seasonality of the area. Although the retrievals showed a consistent spatial pattern, analysis of the twenty-four sample locations for the LGS of the short rains showed large differences between the 2015/2016 (84 ± 11 days) season and the 2016/2017 (45 ± 7 days) season. These differences can be partially explained by the total seasonal precipitation where the short rains of 2015/2016 had 258mm and the short rains of 2016/2017 had 136.5mm. These differences in total seasonal precipitation could also partially explain the differences in cumNDVI in the same season, confirming the arguments made by Hawinkel et al. (2016) that variability of precipitation influence the amount of total seasonal vegetation greenness.

For the multi-annual average phenological results from the Landsat ‘synthetic year’, 99.1% pixels had successful retrieval during the long rains and 64.22% during the short rains. On average the start and end of the season for the long rains was within April and June while for the short rains it was within November

and January which agreed with the single-season results of this study as well as the regions known seasonality. Comparisons between aggregated multi-annual average phenology from Landsat with those derived from eMODIS showed an RMSD of 14 days (SOS₅₀) and 12 days (EOS₅₀) during the long rains while during the short rains had 9 days (SOS₅₀) and 13 days (EOS₅₀). The MSD values of similar comparisons were 13 days (SOS₅₀), 8 days (EOS₅₀) for long rains and -6 days (SOS₅₀), 5 days (EOS₅₀) during the short rains. However, these deviations differ from similar comparisons done by Melaas et al. (2016) who attained an RMSD of ± 5 days. There could be several possibilities as to why the RMSD values are different. For instance, Melaas et al. (2016) used the Enhanced Vegetation Index (EVI) calculated from MODIS data (MDC43A4) to estimate the vegetation phenology while in this study we used NDVI values calculated from eMODIS data. Other than that, Melaas et al. (2016) also employed some data normalization, where they first made average years of EVI trajectories observed within a window of three years successively from the long-term Landsat EVI series before generating the ‘synthetic year.’ This procedure was not carried out in this study because some years had very few observations and performing this normalization would have somehow resulted in misleading average years. Another possible explanation could be the heterogeneous vegetation state of the savannas, compared to the relatively homogenous forest structure in the case of Melaas et al. (2016) study. In search of other ways that may in future improve the generation of a ‘synthetic year’ that would result to less scattered NDVI values, grouping of the NDVI profiles was attempted based on the amount of total seasonal precipitation in each season. The hypothesis here was that years of Landsat NDVI within the same range of seasonal precipitation would align along a similar temporal trajectory. Surprisingly, seasonal precipitation alone was found not to be a good predictor of the amount of greenness experienced in a certain location. It is still difficult to explain these results in the present study.

Investigations whether the retrieved multi-annual average retrievals (SOS/EOS) would show a significant difference between the different vegetation communities at this scale, showed no evidence during the long rains season. However, a significant difference from the Tukey HSD test was noted between the start of the season of the Acacia shrublands and denser tree cover areas during the short rains. Studies (Boke-Olén, 2017; Higgins et al., 2011) that have shown the existence of this temporal separation were done at a field-camera observation scale and that could partly explain the results of this study which was done at a 30m resolution scale. Other interesting findings were that at this scale (30m) it was possible to find a significant difference of the Tukey HSD test between the cumNDVI of the grasslands and denser tree cover areas during both long and short seasons as well as the maxNDVI and AMP2 during the short rains seasons. It is therefore likely that such temporal separations exist.

The combinations of the findings of this study show that even for a complex ecosystem, in a region that has two short seasons, inter-annual vegetation greenness variation, frequent droughts and persistent cloud cover, estimation of vegetation phenology is still possible using freely available finer spatial resolution data even though it is not so frequent. Despite these promising results, further work is required to test the validity of these results. While ground-based vegetation phenology studies are still limited in Africa (Adole et al., 2016), there is hope for this study area with the recent (October 2017), installation of three ‘tracking’ cameras : (1) Acacia shrublands, (2) grasslands and (3) areas with denser tree covers (Appendix 4). These cameras are set to capture images at an interval of thirty minutes between 0800hrs and 1730hrs and one additional image at 0000hrs creating more chances of getting quality images (Appendix 5).

Future studies focusing on single-season retrievals can consider adding more images from the newly launched Sentinel-2B mission which will increase the chances of having more cloud-free observations and therefore raise a possibility that could allow phenology retrievals at a 10m resolution. For those focusing on multi-annual average phenology, finding better ways of having less scattered NDVI trajectories is key. A suggestion could be removing years exhibiting a ‘strange’ NDVI temporal pattern (e.g., years with very low

NDVI values probably due to drought) before constructing the ‘synthetic year.’ Nonetheless, there is still abundant room for further progress in this topic.

6. CONCLUSION

This study set out to investigate if it was possible to use Landsat and Sentinel-2 images to estimate vegetation phenology for a rangeland site in Kenya in a single season or as a multi-annual average. This study has shown that for a single season (fusing Landsat and Sentinel-2) it was possible to have successful retrieval of ~50% of the total pixels in the area during the short rains and long rains of 2017. For the multi-annual average (‘synthetic year’ of Landsat), comparisons with frequent sensors (eMODIS) showed a deviation that is considered high for this region with fast green-up and decay phase. One significant finding from both approaches was that the start- and end- of the season in both seasons agreed with each other as well as the known seasonality in the region and other studies carried out here. Although this study focuses on estimating vegetation phenology, it was found out that for the single-season retrievals of cumNDVI and maxNDVI were partly related to the amount of seasonal precipitation. No clear evidence was found of different SOS and EOS dates for different vegetation types. This could be because 30m pixels with a relatively high occurrence of trees and shrubs still contained a large amount of grasses that may have an important influence on the NDVI trajectory. This study appears to be the first to utilize high spatial resolution data to study the vegetation phenology of the savannas therefore laying the groundwork for future research.

LIST OF REFERENCES

- Adole, T., Dash, J., & Atkinson, P. M. (2016). A systematic review of vegetation phenology in Africa. *Ecological Informatics*, 34, 117–128. <https://doi.org/10.1016/j.ecoinf.2016.05.004>
- Ahl, D. E., Gower, S. T., Burrows, S. N., Shabanov, N. V., Myneni, R. B., & Knyazikhin, Y. (2006). Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. *Remote Sensing of Environment*, 104(1), 88–95. <https://doi.org/10.1016/J.RSE.2006.05.003>
- Baudena, M., Dekker, S. C., Van Bodegom, P. M., Cuesta, B., Higgins, S. I., Lehsten, V., ... Brovkin, V. (2015). Forests, savannas, and grasslands: bridging the knowledge gap between ecology and Dynamic Global Vegetation Models. *Biogeosciences*, 12, 1833–1848. <https://doi.org/10.5194/bg-12-1833-2015>
- Bekure, de L. P. N., Grandin, B. E., & Bekure, S. (1991). Introduction to the Kenyan rangelands and Kajiado district. In de L. P. N. Bekure, B. E. Grandin, & S. Bekure (Eds.), *Maasai herding: An analysis of the livestock production system of Maasai pastoralists in eastern Kajiado District, Kenya* (p. 172). ILCA (International Livestock Centre for Africa), Addis Ababa, Ethiopia. Retrieved from <http://www.fao.org/Wairdocs/ILRI/x5552E/x5552e04.htm>
- Blom, A., Lan, G., & Adil, M. (2015). *Sub-Saharan African Science, Technology, Engineering, and Mathematics Research: A Decade of Development*. The World Bank. <https://doi.org/10.1596/978-1-4648-0700-8>
- Boke-Olén, N. (2017). *Global Savanna Phenology Integrating Earth Observation, Ecosystem Modeling, and PhenoCams (Doctoral Dissertation)*. Faculty of Science Department of Physical Geography and Ecosystem Science. Retrieved from https://lucris.lub.lu.se/ws/files/21760230/Niklas_Boke_Olen_webb.pdf
- Boke-Olén, N., Lehsten, V., Ardö, J., Beringer, J., Eklundh, L., Holst, T., ... Tagesson, T. (2016). Estimating and Analyzing Savanna Phenology with a Lagged Time Series Model, 15. <https://doi.org/10.1371/journal.pone.0154615>
- Broadhead, J. S., Black, C. R., & Ong, C. K. (2003). Tree leafing phenology and crop productivity in semi-arid agroforestry systems in Kenya. *Agroforestry Systems*, 58, 137–148. Retrieved from <https://ezproxy.utwente.nl:3351/content/pdf/10.1023%2FA%3A1026091921043.pdf>
- Cardoso Da Silva, J. M., & Bates, J. M. (2002). Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot The Cerrado, which includes both forest and savanna habitats, is the second largest South American biome, and among the most threatened on the continent. *BioScience*, 52(3), 225–234. [https://doi.org/10.1641/0006-3568\(2002\)052\[0225:bpacit\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0225:bpacit]2.0.co;2)
- Cipriotti, P. A., Flombaum, P., Sala, O. E., & Aguiar, M. R. (2008). Does drought control emergence and survival of grass seedlings in semi-arid rangelands? *Journal of Arid Environments*, 72(3), 162–174. <https://doi.org/10.1016/j.jaridenv.2007.06.012>
- Conservation of Kenya. (2010). Athi Kapiti Conservation Initiative - Conservation for Kenya. Retrieved December 18, 2017, from <https://akci.wordpress.com/>
- Corinna, R., Jeffrey E., H., Siva R., S., Cary, F., & Jayne, B. (2011). A Simple Graphical Approach to Quantitative Monitoring of Rangelands. *Society for Range Management*, 8. Retrieved from <https://mail.google.com/mail/u/0/#search/a.vrieling%40utwente.nl/15eb89ffdfcdccf?projector=1>
- de Beurs, K. M., & Henebry, G. M. (2010). Spatio-Temporal Statistical Methods for Modelling Land Surface Phenology. In *Phenological Research* (pp. 177–208). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-3335-2_9
- Department of the Interior U.S. Geological Survey. (2017). *Landsat 4-7 surface reflectance (LEDAPS) product*. Retrieved from https://landsat.usgs.gov/sites/default/files/documents/ledaps_product_guide.pdf
- Elmendorf, S. C., Jones, K. D., Cook, B. I., Diez, J. M., Enquist, C. A. F., Hufft, R. A., ... Weltzin, J. F. (2016). The plant phenology monitoring design for The National Ecological Observatory Network. *Ecosphere*, 7(4), e01303. <https://doi.org/10.1002/ecs2.1303>
- Frantz, D., Stellmes, M., Roder, A., Udelhoven, T., Mader, S., & Hill, J. (2016). Improving the Spatial Resolution of Land Surface Phenology by Fusing Medium- and Coarse-Resolution Inputs. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7), 4153–4164. <https://doi.org/10.1109/TGRS.2016.2537929>
- Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., ... Prueger, J. H. (2017). Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sensing*

- of *Environment*, 188, 9–25. <https://doi.org/10.1016/J.RSE.2016.11.004>
- GEOGLAM RAPP. (2017). Kenya | Kapiti Farm -ILRI site. Retrieved July 31, 2017, from <http://www.geo-rapp.org/pilot-activities-2/kenya/>
- Giglio, L., Van, G. R., Werf, D., Randerson, J. T., Collatz, G. J., Kasibhatla, P., & Van Der Werf, G. R. (2006). Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics European Geosciences Union Atmos. Chem. Phys. Atmospheric Chemistry and Physics*, 6(6), 957–974. Retrieved from www.atmos-chem-phys.net/6/957/2006/
- Githeko, A., Semmazzi, F., Senouci, M., Niang, I., Nyong, A., Vogel, C., ... Hanson, C. (2007). Africa. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. In *IPCC* (pp. 433–467). Cambridge UK: Cambridge University Press., Retrieved from <https://www.ipcc.ch/pdf/assessment-report/ar4/wg2/ar4-wg2-chapter9.pdf>
- Glen, S. (2016). Tukey Test / Tukey Procedure / Honest Significant Difference. Retrieved December 13, 2017, from <http://www.statisticshowto.com/tukey-test-honest-significant-difference/>
- Government of Kenya (GOK). (1993). *National development plan for the period 1994 to 1996*. Office of the Vice President and Ministry of Planning and National Development. Nairobi, Kenya: Government Printers. Retrieved from <https://searchworks.stanford.edu/view/2838042>
- Grace, J., Jose, J. S., Meir, P., Miranda, H. S., & Montes, R. A. (2006). Productivity and carbon fluxes of tropical savannas. *Journal of Biogeography*, 33(3), 387–400. <https://doi.org/10.1111/j.1365-2699.2005.01448.x>
- Gräler, B., Pebesma, E., & Heuvelink, G. (2016). Spatio-Temporal Interpolation using gstat. Retrieved from <https://cran.r-project.org/web/packages/gstat/vignettes/spatio-temporal-kriging.pdf>
- Hashim, M., Pour, A. B., & Onn, C. H. (2014). Optimizing cloud removal from satellite remotely sensed data for monitoring vegetation dynamics in humid tropical climate. *IOP Conference Series: Earth and Environmental Science*, 18(1), 12010. <https://doi.org/10.1088/1755-1315/18/1/012010>
- Hawinkel, P., Thiery, W., Lhermitte, S., Swinnen, E., Verbist, B., Van Orshoven, J., & Muys, B. (2016). Vegetation response to precipitation variability in East Africa controlled by biogeographical factors. *Journal of Geophysical Research: Biogeosciences*, 121(9), 2422–2444. <https://doi.org/10.1002/2016JG003436>
- Hazaymeh, K., & Hassan, Q. K. (2017). A remote sensing-based agricultural drought indicator and its implementation over a semi-arid region, Jordan. *Journal of Arid Land*, 9(3), 319–330. <https://doi.org/10.1007/s40333-017-0014-6>
- Herrick, J. E., Urama, K. C., Karl, J. W., Boos, J., Johnson, M.-V. V., Shepherd, K. D., ... West, L. T. (2013). The global Land-Potential Knowledge System (LandPKS): Supporting evidence-based, site-specific land use and management through cloud computing, mobile applications, and crowdsourcing. *Journal of Soil and Water Conservation*, 68. <https://doi.org/10.2489/jswc.68.1.5A>
- Higgins, S. I., Delgado-Cartay, M. D., February, E. C., & Combrink, H. J. (2011). Is there a temporal niche separation in the leaf phenology of savanna trees and grasses? *Journal of Biogeography*, 38(11), 2165–2175. <https://doi.org/10.1111/j.1365-2699.2011.02549.x>
- Hoffman, T., & Vogel, C. (2008). Climate Change Impacts on African Rangelands. *BioOne Research Evolved*, 30, 12–17. [https://doi.org/https://doi.org/10.2111/1551-501X\(2008\)30\[12:CCIOAR\]2.0.CO;2](https://doi.org/https://doi.org/10.2111/1551-501X(2008)30[12:CCIOAR]2.0.CO;2)
- House, J. I., Archer, S., Breshears, D. D., & Scholes, R. J. (2003). Conundrums in mixed woody-herbaceous plant systems. *Journal of Biogeography*, 30(11), 1763–1777. <https://doi.org/10.1046/j.1365-2699.2003.00873.x>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. ., Gao, X., & Ferreira, L. . (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- IPCC. (2007). *Climate change adaptation: According to the Fourth Assessment Report Panel on Climate Change (IPCC)*. The Global Environment Facility. Retrieved from <http://www.unccd.int/Lists/SiteDocumentLibrary/Publications/SLM-english-2.pdf>
- Jeltsch, F., Milton, S. J., Dean, W. R. J., & van Rooyen, N. (1996). Tree Spacing and Coexistence in Semiarid Savannas. *The Journal of Ecology*, 84(4), 583. <https://doi.org/10.2307/2261480>
- King, E. G. (2008). Facilitative effects of *Aloe secundiflora* shrubs in degraded semi-arid rangelands in Kenya. *Journal of Arid Environments*, 72(4), 358–369. <https://doi.org/10.1016/j.jaridenv.2007.08.001>
- Laurance, W. F., Alonso, A., Lee, M., & Campbell, P. (2006). Challenges for forest conservation in Gabon,

- Central Africa. *Futures*, 38(4), 454–470. <https://doi.org/10.1016/J.FUTURES.2005.07.012>
- Leeuw, P. N. de, Grandin, B. E., & Bekure, S. (1991). Chapter 2: Introduction to the Kenyan rangelands and Kajiado district. In *Maasai herding - An analysis of the livestock production system of Maasai pastoralists in eastern Kajiado District, Kenya* (p. 172). Addis Ababa, Ethiopia: ILCA (International Livestock Centre for Africa). Retrieved from <http://www.fao.org/Wairdocs/ILRI/x5552E/x5552e04.htm>
- Ma, X., Huete, A., Yu, Q., Restrepo Coupe, N., Davies, K., Broich, M., ... Eamus, D. (2013). Spatial patterns and temporal dynamics in savanna vegetation phenology across the North Australian Tropical Transect. *Remote Sensing of Environment*, 139, 97–115. <https://doi.org/10.1016/j.rse.2013.07.030>
- Macharia, P. N., & Ekaya, W. N. (2005). The Impact of Rangeland Condition and Trend to the Grazing Resources of a Semi-arid Environment in Kenya. *J. Hum. Ecol.*, 17(2), 143–147. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.6777&rep=rep1&type=pdf>
- Maitima, J. M., Olson, J. M., Alagarswamy, G., Andresen, J., Campbell, D., Conway, D., ... Yanda, P. (2003). Implications of climate change on rangeland productivity in Kenya CLIP policy brief. *Journal of Mammalogy Journal of Mammalogy Quaternary Research Republic of Kenya Republic of Kenya National Development Plan*, 84(5), 456–470. Retrieved from [http://clip.msu.edu/Policy Briefs 2008/Rangelands_policybrief.pdf](http://clip.msu.edu/Policy%20Briefs%202008/Rangelands_policybrief.pdf)
- Markwardt, C. B. (2009). Non-linear Least Squares Fitting in IDL with MPFIT, XXX, 1–4. <https://doi.org/citeulike-article-id:4067445>
- Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., ... Lim, T.-K. (2006). A Landsat Surface Reflectance Dataset for North America, 1990–2000. *IEEE Geoscience and Remote Sensing Letters*, 3(1), 68–72. <https://doi.org/10.1109/LGRS.2005.857030>
- Mathu, W. (2009). United Nations Environment programme terminal evaluation of UNEP/GEF project gf/3010-05-10 (4857) – dry land livestock wild life environment interface project (DLWEIP). Retrieved from <https://wedocs.unep.org/rest/bitstreams/10708/retrieve>
- Melaas, E. K., Friedl, M. A., & Zhu, Z. (2013). Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. *Remote Sensing of Environment*, 132, 176–185. <https://doi.org/10.1016/j.rse.2013.01.011>
- Melaas, E. K., Sulla-Menashe, D., Gray, J. M., Black, T. A., Morin, T. H., Richardson, A. D., & Friedl, M. A. (2016). Multisite analysis of land surface phenology in North American temperate and boreal deciduous forests from Landsat. *Remote Sensing of Environment*, 186, 452–464. <https://doi.org/10.1016/j.rse.2016.09.014>
- Meroni, M., Fasbender, D., Kayitakire, F., Pini, G., Rembold, F., Urbano, F., & Verstraete, M. (2013). Regional drought monitoring using phenologically tuned biomass production estimates from SPOTVEGETATION FAPAR. In *2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* (pp. 495–499). IEEE. <https://doi.org/10.1109/Argo-Geoinformatics.2013.6621970>
- Meroni, M., Verstraete, M. M., Rembold, F., Urbano, F., & Kayitakire, F. (2014). A phenology-based method to derive biomass production anomalies for food security monitoring in the Horn of Africa. *International Journal of Remote Sensing*, 35(7), 2472–2492. <https://doi.org/10.1080/01431161.2014.883090>
- Mganga, K. Z., Musimba, K. . N., Nyariki, M. D., Nyangito, M. M., Ekaya, N. W., Muiru, M. W., & Mwang'ombe, W. A. (2011). Different land use types in the semi-arid rangelands of Kenya influence soil properties. *Journal of Soil Science and Environmental Management*, 2, 370–374. Retrieved from https://www.researchgate.net/publication/236646411_Different_land_use_types_in_the_semi-arid_rangelands_of_Kenya_influence_soil_properties
- Ming, L., Zhengfang, W., Lijie, Q., & Xiangjun, M. (2011). Extracting Vegetation Phenology Metrics in Changbai Mountains Using an Improved Logistic Model. *Chinese Geographical Science*, 21(213), 304–311. <https://doi.org/10.1007/s11769-011-0471-3>
- Moore, C. E., Beringer, J., Evans, B., Hutley, L. B., & Tapper, N. J. (2017). Tree–grass phenology information improves light use efficiency modelling of gross primary productivity for an Australian tropical savanna. *Biogeosciences*, 14, 111–129. <https://doi.org/10.5194/bg-14-111-2017>
- Nijland, W., Bolton, D. K., Coops, N. C., & Stenhouse, G. (2016). Imaging phenology; scaling from camera plots to landscapes. *Remote Sensing of Environment*, 177, 13–20. <https://doi.org/10.1016/j.rse.2016.02.018>
- O'Farrell, P. J., Donaldson, J. S., & Hoffman, M. T. (2007). The influence of ecosystem goods and services

- on livestock management practices on the Bokkeveld plateau, South Africa. *Agriculture, Ecosystems & Environment*, 122(3), 312–324. <https://doi.org/10.1016/J.AGEE.2007.01.025>
- Oscar Kipchirchir, K., Robinson Ngugi, K., Stephen Mwangi, M., George Njomo, K., Raphael, W., & George, K. (2015). Water Stress Tolerance of Six Rangeland Grasses in the Kenyan Semi-arid Rangelands. *American Journal of Agriculture and Forestry*, 3(5), 222–229. <https://doi.org/10.11648/j.ajaf.20150305.18>
- Pebesema, E., & Graeler, B. (2017). Package “gstat” - Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation. CRAN. Retrieved from <https://cran.r-project.org/web/packages/gstat/gstat.pdf>
- Reed, B. C., Schwartz, M. D., & Xiao, X. (2009). Remote Sensing Phenology. In *Phenology of Ecosystem Processes* (pp. 231–246). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-0026-5_10
- Richardson, A. D., Klosterman, S., & Toomey, M. (2013). Near-Surface Sensor-Derived Phenology. In *Phenology: An Integrative Environmental Science* (pp. 413–430). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6925-0_22
- Sandhage-Hofmann, A., Kotz E, E., Van Delden, L., Dominiak, M., Fouch E, H. J., Van Der Westhuizen, H. C., ... Amelung, W. (2015). Rangeland management effects on soil properties in the savanna biome, South Africa: A case study along grazing gradients in communal and commercial farms. *Journal of Arid Environments*, 120, 14–25. <https://doi.org/10.1016/j.jaridenv.2015.04.004>
- Sankaran, M., Hanan, N. P., Scholes, R. J., Ratnam, J., Augustine, D. J., Cade, B. S., ... Coughenour, M. B. (2005). Determinants of woody cover in African savannas, 438. <https://doi.org/10.1038/nature04070>
- Sankaran, M., Ratnam, J., & Hanan, N. P. (2004). Tree-grass coexistence in savannas revisited - insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters*, 7(6), 480–490. <https://doi.org/10.1111/j.1461-0248.2004.00596.x>
- Scholes, R. J., & Walker, L. A. (1993). *An African Savanna: Synthesis of the Nylsvley Study*. Cambridge University Press. Retrieved from [https://books.google.nl/books?hl=en&lr=&id=FtXeA8IxmWQC&oi=fnd&pg=PP1&dq=Scholes,+J.+%26+Walker,+B.H.+\(1993\)+An+Africa+savanna&ots=TPYFm_Nb5E&sig=DVTt1VNHpkZc9m7fyVydKUXvH0M#v=onepage&q&f=false](https://books.google.nl/books?hl=en&lr=&id=FtXeA8IxmWQC&oi=fnd&pg=PP1&dq=Scholes,+J.+%26+Walker,+B.H.+(1993)+An+Africa+savanna&ots=TPYFm_Nb5E&sig=DVTt1VNHpkZc9m7fyVydKUXvH0M#v=onepage&q&f=false)
- Scholes, R. J., & Archer, S. R. (1997). Tree-grass interactions in savannas. *Annual Review of Ecology and Systematics*, 28(1), 517–544. <https://doi.org/10.1146/annurev.ecolsys.28.1.517>
- Studer, S., Stöckli, R., Appenzeller, C., & Vidale, P. L. (2007). A comparative study of satellite and ground-based phenology. *International Journal of Biometeorology*, 51(5), 405–414. <https://doi.org/10.1007/s00484-006-0080-5>
- Swets, D. L., Reed, B. C., Rowland, J. D., & Marko, S. E. (1999). A Weighted Least-Squares Approach to Temporal NDVI Smoothing. *American Society of Photogrammetric Remote Sensing*, 526–536. Retrieved from https://phenology.cr.usgs.gov/pubs/ASPRS_Swets_et_al_Smoothing.pdf
- Vrieling, A., Meroni, M., Darvishzadeh, R., Skidmore, A. K., Wang, T., Zurita-Milla, R., ... Paganini, M. (2017). Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sensing of Environment*, (Under review).
- Vrieling, A., Meroni, M., Mude, A. G., Chantararat, S., Ummenhofer, C. C., & de Bie, K. (2016). Early assessment of seasonal forage availability for mitigating the impact of drought on East African pastoralists. *Remote Sensing of Environment*, 174, 44–55. <https://doi.org/10.1016/j.rse.2015.12.003>
- Vrieling, A., Skidmore, A. K., Wang, T., Meroni, M., Ens, B. J., Oosterbeek, K., ... Paganini, M. (2017). Spatially detailed retrievals of spring phenology from single-season high-resolution image time series. *International Journal of Applied Earth Observation and Geoinformation*, 59, 19–30. <https://doi.org/10.1016/j.jag.2017.02.021>
- Walker, J. J., De Beurs, K. M., & Wynne, R. H. (2014). Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data. <https://doi.org/10.1016/j.rse.2014.01.007>
- Werner, P. A. (2009). *Savanna Ecology and Management: Australian Perspectives and Intercontinental Comparisons*. John Wiley & Sons. Retrieved from https://books.google.nl/books?id=sb_sDXasggIC
- White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, 11(2), 217–234. <https://doi.org/10.1029/97GB00330>

- Whitecross, M. A., Witkowski, E. T. F., & Archibald, S. (2017). Savanna tree-grass interactions: A phenological investigation of green-up in relation to water availability over three seasons. *South African Journal of Botany*, 108, 29–40. <https://doi.org/10.1016/j.sajb.2016.09.003>
- Yamagiwa, J., Kanyunyi Basabose, A., & Kaleme, K. P. (2008). Phenology of fruits consumed by a sympatric population of gorillas and chimpanzees in Kahuzi Biega National Park, Democratic Republic of Congo. *African Study Monographs*, 39, 3–22. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.513.7146&rep=rep1&type=pdf>
- Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., & Liu, Z. (2005). Monitoring the response of vegetation phenology to precipitation in Africa by coupling MODIS and TRMM instruments. *Journal of Geophysical Research*, 110(D12), D12103. <https://doi.org/10.1029/2004JD005263>
- Zhang, X., Tan, B., & Yu, Y. (2014). Interannual variations and trends in global land surface phenology derived from enhanced vegetation index during 1982–2010. *International Journal of Biometeorology*, 58(4), 547–564. <https://doi.org/10.1007/s00484-014-0802-z>

7. APPENDICES

7.1. Appendix 1

This is an example of a LandPKS protocol field form. There were some adjustments compared with the original form. Some of the adjustment was the addition of photo-protocol and visual estimate cover %.

Plot#: 18 Date: 04/10/17

Time_start: 14:02

Time_end: 14:26

Lat: 290470 Lon: 9816156

Error GPS: 4 Altitude: 5654 ft

Names: Stella & Maria

Slope:

0.2% 35% 6.10% 11.15% 16.30% 35.60% 61.100%

WEST

NORTH

START + HERE

EAST

HLNOS

Visual estimate %cover (+ picture)

	Trees	Shrub	Grass
NW	0	20	40
NE	0	10	40
SE	0	70	40
SW	0	60	30

No canopy - + X - Soil

No basal - line

Density of plants > 1m in 10x10m

	Trees	Shrubs
Number of plants	0	15

Animal perp.

Photo-protocol

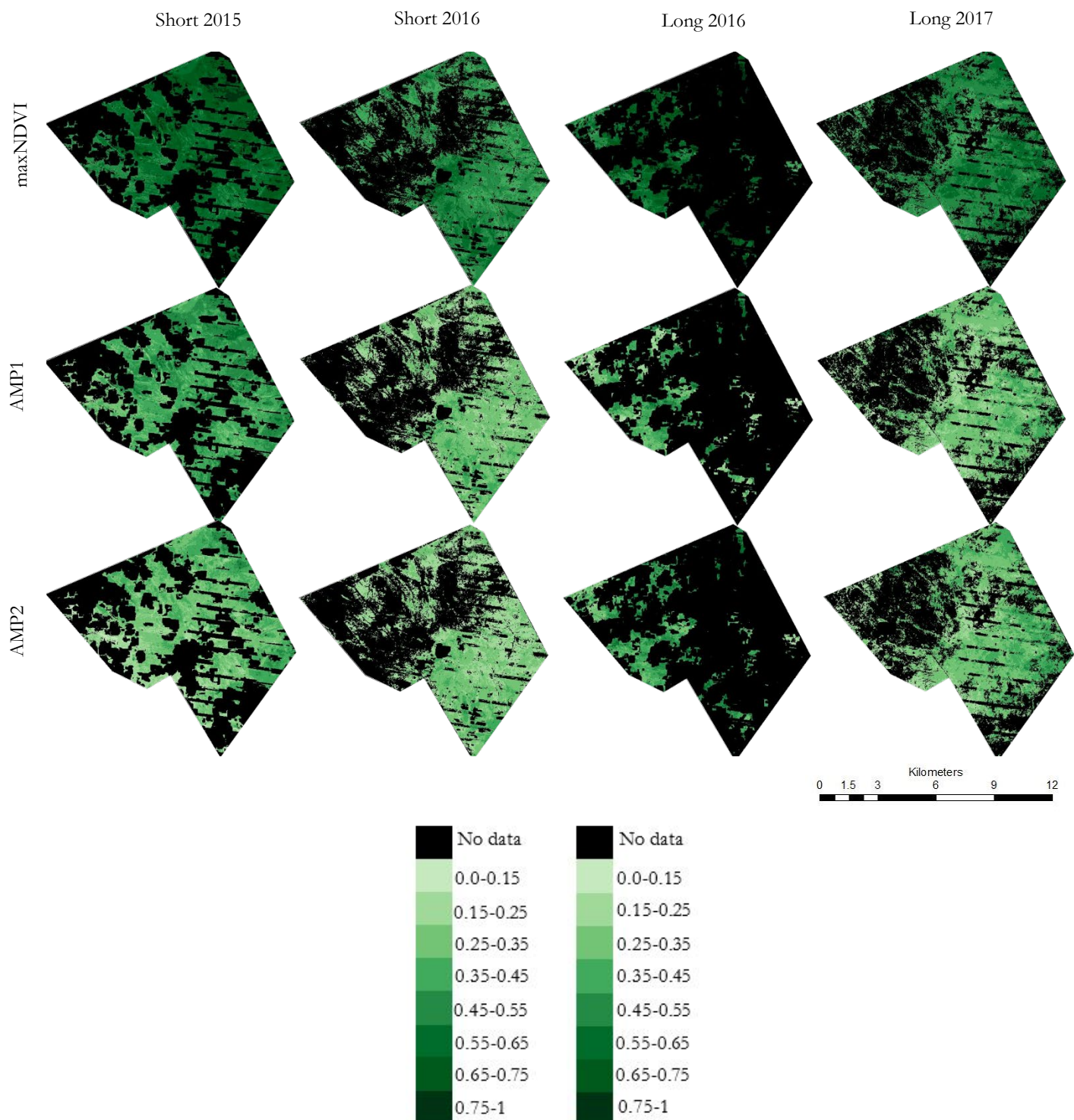
- 4 photos breast-height NW,NE,SE,SW: f
- 1 photo overview plot (upslope): f
- Detail photos vegetation: 1- Acacia, 2- Shrub, 3- grass
- 4- shrub 5- Shrub 5
- POLE: ~5m Center, 20m N, 20m E, 20m S, 20m W: f

LEGEND

- Tree
- Shrub
- Grass
- Plant Base
- Litter
- Rock

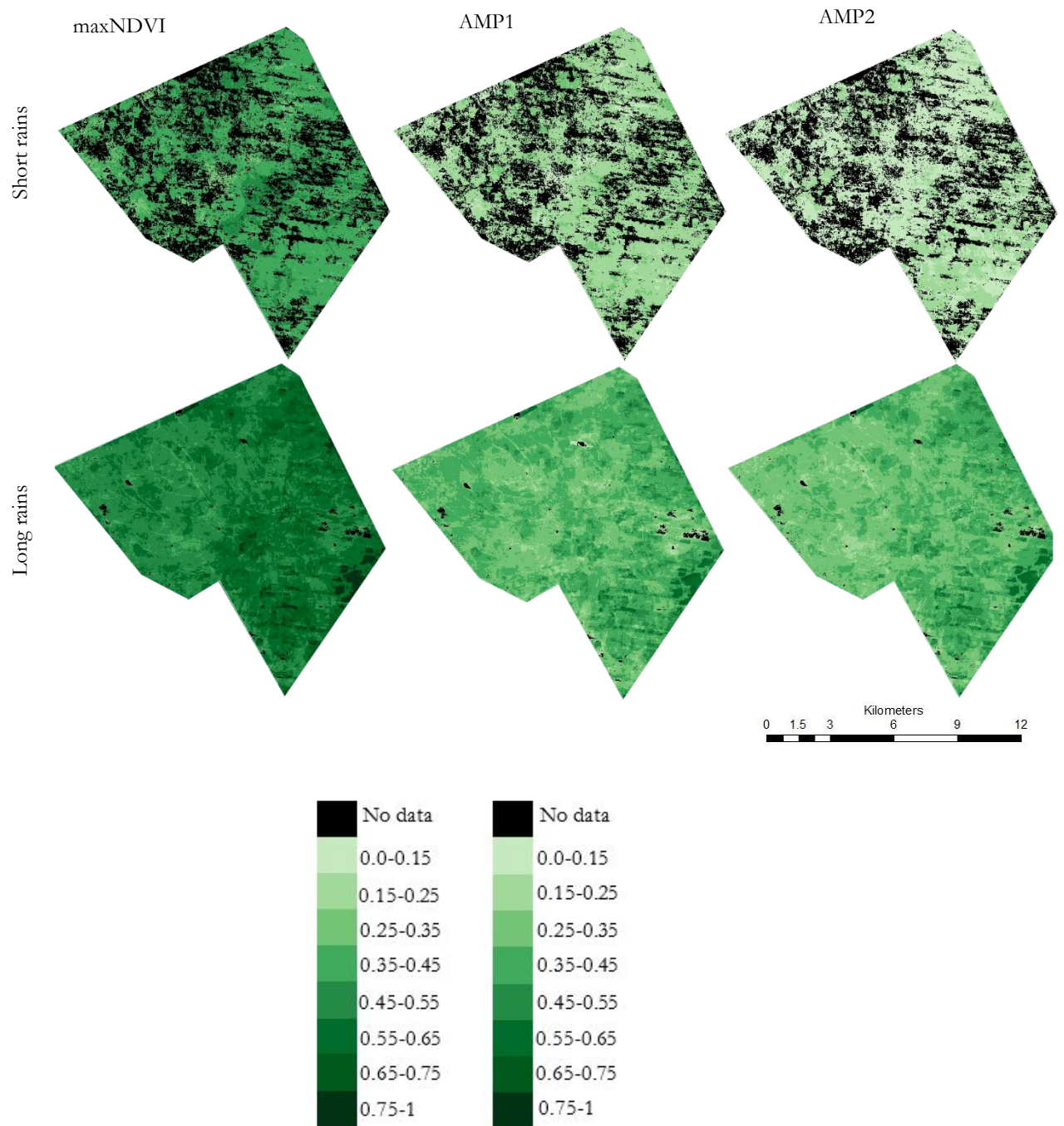
7.2. Appendix 2

Other seasonal greenness parameters; maximum NDVI, AMP1 (growth phase), AMP2 (decay phase), for the different seasons between September 2015 and September 2017 as extracted from the fused Landsat and Sentinel-2 images.



7.3. Appendix 3

Other seasonal greenness parameters from the combined years of Landsat into a synthetic Year.



7.4. Appendix 4

Images retrieved from one cameras installed in Kapiti Farm



7.5. Appendix 5

GCC time-series derived from the three cameras located in Kapiti Farm. Source: Dr. A. Vrieling

