Ilona Aalto

USING TIME SERIES ANALYSIS TO MONITOR DEFORESTATION DYNAMICS IN MIOMBO WOODLANDS IN SOUTHERN HIGHLANDS OF TANZANIA

Master's thesis in Geography

Turku 2020

UNIVERSITY OF TURKU
Faculty of Science and Engineering
Department of Geography and Geology

AALTO, ILONA: Using time series analysis to monitor deforestation dynamics in Miombo woodlands in Southern Highlands of Tanzania

Master's thesis, 83 pages
40 ECTS, Geography
Instructors: Niina Käyhkö & Joni Koskikala
May 2020

Deforestation and forest fragmentation are threatening the Miombo woodlands in Southern Highlands of Tanzania. Miombo ecoregion is considered one of the world's most valuable wilderness areas, providing livelihood for over 150 million people throughout the region, who are directly or indirectly depending on these ecosystem services. Monitoring deforestation process using satellite images enables the identification of ongoing changes and pressures facing the region, which is crucial for the sustainable management of the area.

In this thesis the deforestation dynamics are analysed using LandTrendr (Landsat-based detection of Trends in Disturbance and Recovery) -time series algorithm. The algorithm uses temporal segmentation of spectral trajectories to extract change information from pixel time series derived from satellite images. The study is focused in miombo woodlands around a rural village of Mantadi, which is located in Tanzanian Southern Highlands. The capacity of LandTrendr algorithm to detect changes in Miombo woodland region is evaluated through the appliance of three spectral indices. The results are combined to examine the magnitude and spatial distribution of deforestation in the study area.

The results show that to detect areas under any kind of disturbance, LandTrendr performs considerably well with all three indices. In more profound change magnitude detection, clear differences between the spectral indices can be noticed especially in finding subtler, low magnitude changes. The Normalized Burn Ratio (NBR) was found to be most stable index to detect changes in miombo woodlands. Combining the results from spectral indices increased the mapping accuracy by 10 %. The results indicate that 26,5 % of the whole study area has been under very high or high magnitude disturbance and 29,5 % under low or moderate magnitude disturbance between 1987 and 2018. This study proves that the LandTrendr algorithm is suitable for tracking long-term deforestation dynamics in Miombo woodland environments.

Keywords: deforestation, remote sensing, time series analysis, forest change, Miombo woodlands, Tanzania

TURUN YLIOPISTO
Luonnontieteiden ja tekniikan tiedekunta
Maantieteen ja geologian laitos

AALTO, ILONA: Miombo-savannien metsäkadon kartoittaminen aikasarja-analyysin avulla Tansanian eteläisillä ylänköalueilla

Pro gradu -tutkielma, 83 sivua
40 op, Maantiede
Ohjaajat: Niina Käyhkö & Joni Koskikala
Toukokuu 2020

Metsäkato ja metsien pirstoutuminen uhkaavat Tansanian eteläisten ylänköalueiden miombo-savanneja. Miombo-savannit muodostavat yhden maailman tärkeimmistä erämaa-alueista, tarjoten toimeentulon yli 150 miljoonalle ihmiselle, jotka ovat tavalla tai toisella riippuvaisia alueen ekosysteemipalveluista. Metsäkatoprosessien seuranta satelliittikuvien avulla mahdollistaa tapahtuvien muutosten ja paineiden tunnistamisen, mikä on elintärkeää alueen kestävälle hallinnalle.

Tässä opinnäytetyössä metsien häviämisen dynamiikkaa analysoidaan LandTrendr (Landsat-based detection of Trends in Disturbance and Recovery) aikasarja-algoritmin avulla. Algoritmi hyödyntää spektraalisen kulkuradan ajallista segmentointia erottaakseen muutostiedot satelliittikuvien pikselikohtaisista aikasarjoista. Tutkimus kohdistuu miombo-metsäalueisiin Tansanian eteläisillä ylänköalueilla sijaitsevan Mantadi-kylän ympärillä. LandTrendr-algoritmin kykyä havaita muutoksia miombo-savanneilla arvioidaan kolmen spektraalisen indeksin avulla. Lopulta indeksien tulokset yhdistetään, jotta metsäkadon laajuutta ja alueellista jakautumista tutkimusalueella voitaisiin tutkia entistä tarkemmin.

Tulokset osoittavat, että LandTrendr havaitsee metsissä tapahtuneet muutokset merkittävän hyvin kaikilla kolmella indeksillä. Perusteellisemmassa muutoksen voimakkuuden tarkastelussa havaitaan selviä eroja eri indeksien välillä, etenkin hienovaraisempien muutosten tunnistamisessa. Yksittäisistä indekseistä NBR (The Normalized Burn Ratio) osoittautui kaikkein vakaimmaksi miombo-savanneilla tapahtuvien muutosten tunnistamisessa. Kolmen spektraalisen indeksin tulosten yhdistäminen lisäsi kartoitustarkkuutta 10 %. Tulokset osoittavat, että 26,5 % tutkimusalueen metsistä on hävinnyt tai heikentynyt voimakkaasti ja 29,5 % on kokenut pieniä tai kohtalaisia häiriöitä vuosien 1987-2018 välillä. Tutkielma osoittaa, että LandTrendr algoritmin avulla metsäkatoa voidaan kartoittaa tehokkaasti Miombo-metsäympäristöissä pitkällä aikavälillä.

Asiasanat: metsäkato, kaukokartoitus, aikasarja-analyysi, metsien muutos, Miombo-savannit, Tansania

Table of contents

1. Introduction

Deforestation, forest fragmentation and conversion of habitats resulting from human activities, such as over-exploitation of resources and unsustainable land use practices, are threats to the biodiversity and the ecosystem resilience in Tanzanian Southern Highlands, where more than 45 % of the land consists of forest and woodland ecosystems (Jew et al. 2016). More than 2 million people live in the region, and their nutrition, energy, medicines, building materials and income originate largely from the natural resources of the region (WCS 2016). Areas where wet savannas prevail are commonly called as Miombo in reference to the main Miombo woodland tree species throughout the ecoregion (Timberlake et al. 2010). Over 3.6 million km2 area covering much of Central and Southern Africa, the Miombo ecoregion is considered one of global wilderness areas which should be prioritized in conservation efforts (Mittermeier et al. 2003). The ecoregion experiences a high level of endemicity and is a habitat for several threatened species. Over 100 million rural and 50 million urban people are believed to depend directly or indirectly on the Miombo woodlands for their subsistence (Syampungani et al. 2009).

Identifying the ongoing changes and pressures facing the region is crucial for the sustainable management of the area to ensure the vitality of ecosystems that support the local biodiversity and produce multiple services for the local people. Although the region remains sparsely settled, population growth poses an ongoing threat as the growing population has accelerated the rate of agricultural expansion. This ethnically diverse population has a close relationship with the surrounding nature and traditional lifestyles (WCS 2016).

Acquiring information on land cover change in tropical South is challenging due to the constraints in accessibility and deficiencies in national data acquisition and monitoring. Nevertheless, the high deforestation rates in tropical forests have been recognized and it outpaces the capacity for regrowth especially in the Miombo ecoregion of Southern and Eastern Africa (see e.g. Brink et al. 2014; Sawe et al. 2014; Mayes et al. 2015; Jew et al. 2016).

Monitoring land cover change has been perceived as one of the most significant contributions to study global ecological and environmental change (Horning 2010; Kuenzer et al. 2016). The spatial patterns of deforestation and reforestation are occurring across multiple temporal and spatial scales and thus are exceedingly complex and most frequently related to human impacts. To identify the driver of the change and to understand mechanisms behind deforestation, detecting and characterizing change over time becomes crucial. Opening of the Landsat data archive in 2008 and the powerful image and data processing systems enables the monitoring of land cover and land use dynamics that are taking place in the landscape from regional to global scale in a temporal range starting from the first Landsat-1 images available from 1972, until the present day (Wulder et al.

2012). Change detection research is getting increasingly diverse as the topic is currently very active and growing amount of free and high-quality data allows the development of the field.

The availability of multitemporal remote sensing data with repetitive coverage allows the investigation of alternative change detection methodologies that utilize time series data instead of relying on traditional post-classification and image differencing methods. In the 2000s, new approaches have been rising, including time series segmentation that can identify trends and breakpoints in time series (Kennedy et al. 2010). These more sophisticated approaches together with availability of denser time series data increase the accuracy of detection, as well as facilitate the identification of subtler disturbance signals, like disturbances associated with forest degradation caused by multiple factors (Hughes et al. 2017; Cohen et al. 2018). Time series of continuous data enables detection of long-term changes as well as the rate and trend of these changes, which is not possible with traditional post-classification approaches that generally use only bi-temporal images (Zhu et al. 2012; DeVries et al. 2015).

The Landsat-based detection of trends in disturbance and recovery (LandTrendr) algorithm uses temporal segmentation of spectral trajectories to extract change information from pixel time series derived from satellite images (Kennedy et al. 2010). It fits temporal trajectories from spectral variations to each pixel of annual time series of Landsat images. These trajectories provide information about specific events in time series which can be interpreted as changes in land cover. These changes include abrupt forest clearing for agricultural, industrial, or residential purposes and slow degradation of forests due to the partial wood harvesting, land degradation, insects or changes in climate conditions. Compared to other approaches, this method provides more detailed information about the change occurred (Cohen et al. 2017). The LandTrendr algorithm was previously applied in Miombo woodlands in Angola by Schneibel et al. (2017), who suggested that the approach would be suitable for regions demonstrating distinct conversion processes from woodland to agriculture in similar ecosystems and socio-ecological backgrounds. Using time series instead of post-classification or image pair approaches in Miombo woodlands allows the identification of gradual, subtler changes from background noise. It also enables the recognition of actual change from seasonal fluctuations that would otherwise be perceived as land cover change.

The deforestation dynamics are studied using LandTrendr time series segmentation algorithm. The study is focused in Miombo woodlands around rural village of Mantadi, which is located in historically forested areas in Tanzanian Southern Highlands. These landscapes are being under intensive agricultural expansion and are therefore suitable for implementing the algorithm. The thesis is both methodological and contextual; The objective is equally to evaluate the capacity of the LandTrendr algorithm to detect changes in Miombo woodland region through the use of three spectral indices, and to track woodland changes typical to the study area by combining the results from these spectral

indices. This includes tracing the forest cover changes caused by human activities together with the analysis of more subtle changes in forest vegetation cover to identify the long-term deforestation dynamics taking place in the study area during 1987—2018. Spectral indices have an advantage over single reflectance channels by augmenting the spectral effect of desired phenomena, like changes in vegetation condition, and reducing unwanted effects, such as topographic and atmospheric noise. However, there are differences in the sensitivity of these indices to detect certain types of disturbance. Therefore, the hypothesis of this study is that the accuracy of vegetation change mapping can be increased by combining results from several spectral indices in LandTrendr analysis.

Research questions:

1. What is the capacity of the LandTrendr algorithm to detect disturbances in Miombo woodlands in Southern Highlands of Tanzania?
    a. What are the strengths and weaknesses of selected spectral indices for Landtrendr-based disturbance mapping in Miombo woodlands?
    b. How does combining results from the three spectral indices affect the change mapping accuracy?
2. How has the Miombo woodland cover changed in the study area during 1987-2018?
    a. What is the scale and magnitude of deforestation and forest degradation?

2. Theoretical and methodological framework

2.1. Deforestation and forest degradation

Deforestation is an ambiguous concept as the processes involved are strongly related to the spatial and temporal environment. Depending on the interpretation, deforestation may refer to various stages in forest condition from complete deforestation to gradual long-term degradation that can be defined in several ways (Metz 2009; Sasaki & Putz 2009). In general, the concept of human-induced deforestation is used to describe the clearance of forests for other types of land use, the degradation of forests by felling or harvesting forest products, or the change of forest area following the interaction of all these activities. The ecological effects of deforestation include changes in soil composition, erosion, changes in microclimates and biodiversity loss. In addition, the ecological balance of the ecosystem is diminished, and the energy stored in humus and nutrients is reduced or eliminated, despite of the use of artificial nutrients.

Deforestation and forest degradation are often treated as separate phenomena, which confounds the conception of the ongoing forest processes and dynamics and complicates the research (Sasaki & Putz 2009). The ambiguity in defining deforestation is illustrated by the diversity of definitions between different global actors like United Nations Food and Agriculture Organization (FAO), the International Tropical Timber Agreement (ITTO), the United Nations Environment Program (UNEP) and the Intergovernmental Panel on Climate Change (IPCC), which all have slightly inconsistent definition for the concept of forest, degradation of forests and deforestation (Schoene et al. 2007: 5).

Studying deforestation and forest degradation requires setting a coherent definition for the concept of forest. The forest definition usually is set according to variables like the coverage of tree stands, crown cover and tree height per certain area. Internationally forest has been defined in the context of the definition from the Clean Development Mechanism (CDM) under the Kyoto Protocol to the UNFCCC. According to the definition of CDM, “forest” is an area of at least 0.5–1.0 hectares with a tree cover of 10–30% (UNFCCC 2002: 58). A tree is defined as a plant that has a capacity to grow 2-5 meters tall. However, this definition is problematic as the member states can apply it to their own needs (Sasaki & Putz 2009: 2). In addition, the definition does not distinguish between natural forests and forest plantations or does not consider forest degradation that is taking place within this coarse definition of forest. FAO’s Forest Resources Assessments (FRA) report defines the concept more precisely than Kyoto Protocol. According to the report, 'forest' is an area of more than 0.5 hectares, where the height of trees or their growth capacity is more than five meters and the crown cover is more than 10% (FAO 2015a). Definitions based on numerical values are not unambiguous, as forest that fill these requirements may differ significantly in structure and composition due to different environmental factors (Sasaki & Putz 2009). For clarification of the definition, the FAO has classified

forests into four categories: natural forests (untouched), modified natural forests (native species, some human traces), semi-natural forests (assisted natural regeneration and plantation of indigenous species) and plantation forests (FAO 2015a). Primary and secondary forests are both included. Current land use affects the definition, and forest classification is not valid in areas where the land use is mainly agricultural or urban even if the land cover requirements were met, leaving agroforests out from the forest definition (FAO 2000).

FAO defines deforestation according to same standards. According to this widely used definition, deforestation is a situation where the canopy cover decreases below 10 % for at least ten years or where the land use is changed permanently to something else (FAO 2000). The definition covers both natural and anthropogenic reasons. Unfortunately, there is no common definition for forest degradation and UNEP, FAO, IPCC and ITTO all have different definitions for this phenomenon (Schoene et al. 2007). According to FAO (2000), forest degradation is a gradual process which has significant but slowly progressing effects on social, ecological and cultural functions of the forests.

The FRA report show that, from 1790 to 1980, world's agricultural land increased fivefold, while forests area decreased by 20 percent (FAO 2015a). According to these statistics a large-scale deforestation occurred between 1700 and 1980 especially in Europe and North America, and since 1980 the trend has been reversed: some forest areas cleared of farming have been reforested, either by taking over or re-planting surrounding forests. In addition, urbanization since the 1950s has reduced the pressure on forests and even allowed for the afforestation of new areas. However, deforestation continues to accelerate in tropical areas of global South (FAO 2015a). It is also noteworthy that in different contexts deforestation can be regarded as either a negative or a positive process: in the global north, tropical deforestation is often perceived as a global threat which accelerates the climate change, but a farmer in global South sees the transformation of forests as a lifeline for themselves and others, just as many in the North many decades ago (Metz 2009: 39).

2.2. Underlying and proximate causes of deforestation

Human caused deforestation is not a new phenomenon. Forests have been cleared either intentionally or accidentally since man learned to use fire. The degree of deforestation increased dramatically when the hunting and collecting culture was moved to farming, and again with the industrial revolution (Metz 2009). The first consequence of population and economic growth is often land use change. The invention of farming allowed the accumulation of surplus and thus the growth of the population and the development of technology, which in turn accelerated deforestation under the growing needs of agriculture. From an ecological point of view, these changes have two general

trends: ecological simplification and redemption of natural capital (Metz 2009: 41). Crop and livestock breeding produce simplified ecosystems and reduce biodiversity.

According to Geist & Lambin (2001) the reasons behind tropical deforestation can be characterised to five driving forces. These forces are demand for land, access to land, control over land, land use intensity and encouragement of individual decision making. These factors are linked to population growth (demand for land), technology (land use intensity), and political and cultural environment (decision making and access to land). Factors driving deforestation can be viewed on different scales: global, regional and local (Metz 2009).

Proximate causes of tropical deforestation are usually originating from social factors and decisions made at local level and can be divided to three main categories: infrastructure expansion, wood extraction and agricultural expansion (Geist & Lambin 2001). The biophysical environment such as soil quality, topography and rainfall, affect the forest vulnerability locally. The value that people give to the various natural resources, such as wood or minerals, will make some areas more attractive to exploitation (Brauman et al. 2007). General perception is that agriculture is automatically the greatest driving force behind deforestation, but this is not always the case as the impact of agriculture on the forest loss depends greatly on the type of agriculture practiced. Smallholders favoring traditional parcel cultivation have different effect on land than for example shifting cultivation or a stable field crop. The impact of individual local farmers is generally small, however overall the impact of smallholders on deforestation is significant when the number of farmers exceeds the carrying capacity of the forest, and new areas needs be continuously cleared for cultivation and the remaining fields will not be able to recover sufficiently (Johnson et al. 1989).

Logging is another major proximate cause of forestation. Selective harvesting is targeted only to individual tree species, while in open logs forests are cut down on a large scale. Wood felling could be carried out sustainably so that only trees that have exceeded a certain growth stage are felled, and the surrounding nature is not damaged in the timber processing, but in practice it can be challenging to monitor and control (Salo et al. 2014: 238). Third most common cause of deforestation of tropical forests is infrastructure expansion. The construction of road infrastructure is the backbone of all these activities, as the new accessibility it provides enables extensive use of the natural resources. Additionally, activities like mining are common reason behind the deforestation of tropical forests, as the establishment of mines requires extensive felling of a forest area (Metz 2009: 46). All these forms of land use are linked to the underlying factors at national, regional and global levels (Turner et al. 2007, Geist & Lambin 2001). It is crucial to separate the proximate causes from underlying factors to be able to identify the actual drivers behind deforestation. Often the local subsistence farmers are accused for deforestation, although their options for acquiring a livelihood are restricted by prevailing political and economic situation (Turner et al. 2007).

The major underlying causes of deforestation can usually be found in state practices (Geist & Lambin 2001). The historical formation of the state, the possible colonization and the distribution of the settlement have shaped the usage of forests in diverse ways. The state can directly influence the use of forests by, for example, granting land ownership, nationalizing forests or developing development policies by granting harvesting permits and distributing property rights. Even more significant impacts are caused by indirect ways in which national governments can support the utilization of forests (Abbink et al. 2011; Andersson et al. 2014). These include development policies such as the development of transport infrastructure, promotion of export-oriented agriculture and taxes and subsidies for forest economy. Under these pressures are often three groups competing for forest rights: indigenous peoples, migrant farmers and wealthy outsiders planning to harness the land for large-scale use. In most cases, governments have supported the latter group without regard to legal violations, ensuring the best direct income for the state (Metz 2009: 49). Since the 1980s, NGOs and international pressure have driven many governments to recognize the land tenure rights of indigenous peoples.

The preservation of forests is also greatly affected by the state's ability to manage its territory (Abbink et al. 2011, Metz 2009). Many of the current countries of the global South have been arbitrarily created as a result of colonial occupation without taking into account the ecological or social differences of the regions, and as a result, decisions on the use of natural resources have caused conflicts and are difficult to control. Major economic and ethnic disparities among the population, the distribution of power to the elites and the lack of democracy are likely to lead to corruption and nepotism, which in turn hinders the sustainable use of forest areas. Regional policy decisions are also influenced by external pressures created by global markets (Shandra et al. 2011).

On a global scale, deforestation is influenced by the prevailing world order, economic theories, and international awareness of the state of the environment (Metz 2009: 48). The International Monetary Fund (IMF) has exerted pressure on indebted countries to adopt a neoliberal model in their policies, which means open access for external investors to their own markets, tariff cuts and the abolition of state-owned enterprises (Shandra et al. 2011). In the neoliberal system, the value of forests is determined based on their economic productivity, and vital ecosystem services or the value of biodiversity conservation are not recognized. This and uncontrolled exploitation of natural resources is a major threat to tropical forests globally. At the same time, however, concerns about the environment have begun to grow and the external effects of the market on vital ecosystems (climate regulation, pollination, diversity-based resilience) are recognized (Metz 2009; Shandra et al. 2011). As a result, the international community has sought to create various forest strategies that would recognize the value of ecosystem services and provide compensation for their protection. The future will show whether this is enough to protect tropical forest areas from spreading deforestation.

Forest cover decline is an important environmental issue in the tropical areas of Africa as the forest resources form a significant part of the national economies as well as the domestic resources available to local people (Geist & Lambin 2001). In contemporary Africa, attitudes towards environment are manifold (McBeath & Rosenberg 2006: 26). Even though environment is considered a resource for maintaining livelihoods, many African countries struggle to preserve it. Deeply held beliefs that relate humans to the nature prevail in most African cultures. This makes the environment spiritually and socially valuable, in contrast to the dominant, contemporary western notions of value. In various African cultures, communities live in close relationship with environment with a strong reliance to the local fauna and flora, which provide vital resources for the communities. After independence, many governments have embraced and continued the dominant approach to the environment, with development regimes and private property on its center, despite the traditional approaches of communal ownership and the sacred nature of certain areas (McBeath & Rosenberg 2006: 30). Geist & Lambin (2001) analysed the factors driving deforestation in Africa from 19 case studies and found out that growing population densities, migration, changes in markets, commercialization, agrotechnological change and changes in wood sector were the most significant underlaying factors behind deforestation. In addition to the mere population numbers, socio-economic and demographic factors have gained increasingly attention in deforestation studies (Geist et al. 2006). These include the composition and distribution of population, migration, urbanization, education, and household economics. From all the demographic factors, especially migration to sparsely populated natural areas has been recognized to cause deforestation inevitably.

2.3. Miombo woodlands

The African savannas occur across the continent and cover several soil types within an annual rainfall range of 200–1800 mm (Hutley & Setterfield 2008). Savanna is a woodland-grassland ecosystem characterized by the coexistence of both herbaceous plants and trees (Hutley & Setterfield 2008). Savannas occur largely in the seasonal tropics, covering approximately 20 % of the world's land surface. As a large, wide-spread biome, savannas have a significant role in global carbon, water and nutrient cycles. In many regions, savanna ecosystems have existed for millions of years. Therefore, the biodiversity in savanna is high and consists from both tree and herbaceous components and associated fauna.

One of the most widespread and common savanna types is Miombo, which spans to 10 countries and covers approximately 3.6 million square kilometers in Africa and 2.7 million square kilometers across Central and Southern Africa. The term Miombo is derived in reference to the main Miombo woodland tree species throughout the ecoregion (Timberlake et al. 2010). In Tanzania, more than 45 % of the land consists of forest and woodland ecosystems (Jew et al. 2016), and half of the

southwest Tanzania is known to represent Miombo woodland vegetation. According to current perception, Miombo has developed with the combination of climate, herbivory and both natural and man-made fires (Nduwamungu & Bloesch 2009). The Miombo ecoregion of Southern and Central Africa is considered one of global wilderness areas which should be prioritized in conservation efforts as the ecoregion experiences a high level of endemicity and is a habitat for several threatened species (Mittermeier et al. 2003). Miombo woodlands support the livelihoods of over 150 million people, who are directly or indirectly dependent on it for food, shelter, energy, medicines as well as spiritual and environmental needs (Campbell et al. 2007).

Miombo is usually divided to wetter and drier Miombo, according to annual rainfall limit of 1000 mm (Abdallah & Monela 2007). Dry Miombo woodland is common in central Tanzania and southern parts of Malawi, Zambia and Mozambique, whereas wet Miombo occur in southwest Tanzania, central Malawi, northern Zambia and eastern Angola. Woodlands in Miombo ecoregion are generally deciduous or semi-deciduous with highly variable canopy cover, often depending on the available moisture (Frost 1996). Other vegetation types that can be met in the ecoregion are evergreen forests, grasslands, wetlands and shrublands (Timberlake and Chidumayo 2011). The distribution of different vegetation types depends on soil conditions, rainfall, and level of human disturbance throughout the history.

The Miombo is characterized by a discontinuous canopy of 10–12 m tall deciduous tree species and an herbaceous layer of tall grasses (Hutley & Setterfield 2008). Most common tree species are Brachystegia, Isoberlinia, and Julbernardia, and grasses include mainly Andropogon species. During the dry season streams are treeless flood beds (Niemelä 2011: 54). Tree stems are high and dense in hills and areas of permeable soils. Miombo is the core region for slash and burn agriculture in Africa, and the fire has left its mark everywhere. Fire is a natural element in Miombo woodlands and Miombo is typically very resistant to fire due to the physical adaptation of trees, but recurrent burning weakens the soil condition and increase the growth of uapaca trees and succulent plants.

The structure and boundaries of Miombo ecosystems are highly dynamic and determined by a combination of environmental factors, like soil condition, availability of water and occurrence of disturbance, especially fire. Miombo landscapes are spatially diverse and its functional and physiognomic properties vary across the ecoregion, which makes the definitions broad and overlapping with other vegetation types like open savannas and deciduous forests (Malmer & Nyberg 2008). The structural variability of Miombo woodlands makes it difficult to create a strict definition for this ecosystem, and Miombo woodlands are often referred as savannas, woodlands, wooded grasslands, grasslands, shrublands and open forests. Here Miombo woodlands are described in the context of savannas.

Savanna ecosystem is a mosaic of microhabitats, where availability of sunlight, nutrients and water are inconstant (Niemelä 2011, p. 33). The existence of dry season is a key component, either as single extended period or several shorter dry periods (Hutley & Setterfield 2008). There is a clear difference between the dry season and the rainy season in savanna ecosystems. Drought is thought to be the root cause of the development of savannas, but in fact the annual variation in rainfall plays a more important role in defining vegetation type than the absolute amount of rain (Niemelä 2011: 28). Miombo woodlands look like forests during the rainy seasons, but a closer look reveals that the canopy cover is around 40 %, trees are equal length and bushes, shrubs and grasses are growing in between. The trees drop their leaves during dry season. Grassy woodlands or open savanna woodlands are scarcer, with 20-40 % canopy coverage. In wooded grasslands the canopy cover is about 20 % and most of the field layer is open. Tree density of savannas varies according to topography and reflects the variation of soil and moisture conditions, but it is not the key factor in savanna ecology. In shrub savannas or bushed grasslands vegetation consist of grasses and meadows with frequent but scarce shrub cover. Wooded grasslands are characterized by perennial and other grasses, with nearly 100 % vegetation coverage. Some isolated trees might occur, but their canopy cover is less than 2 %. Seasonal availability of water significantly influences vegetation productivity, which in turn determines the range and timing of available resources for local fauna.

The density of tree vegetation is determined by three variables: amount of rainfall, soil's water retention capacity and depth and size of lateritic crust, which affects the water storage and diversion under the surface. Tinley (1982) revealed in his research the crucial role of the lateritic crust on the location of forest stands and grasslands in woodland environment. Trees are growing in areas where the moisture-retaining layer of soil is so thick that there is enough moisture until the beginning of the next rainy season. In dry savannah areas, increased soil moisture favors tree growth, but increased humidity in wet savannah areas increases the growth of grasses, as soil drainage prevents tree growth. In these areas, trees are spreading during successive dry periods. In mature woodlands woody plants make up over 95 % of plant biomass, but the diversity of tree species is relatively low, although the overall species richness is high (Frost 1996). Grassy depressions called "dambos" are interspersed within the woodlands and during the rainy season broad areas of these shallow grasslands area under water. These wetlands have an important role in cultivation and livestock grazing (Abdallah & Monela 2007).

Savannas are very dynamic: tree cover expands and declines with time and species composition transform in result of changes in competition, climate, and other environmental factors (Niemelä 2011: 30). The competition between trees and grasslands often turns to the benefit of trees, and savannas tend to become woodier over time. However, this trend does not last forever, as the competition for water, fires, climate conditions and herbivores are restricting the tree's expansion (Sankaran et al. 2005). Tree seedlings are particularly sensitive to fire and drought, and several

successive fires lead to the emergence of a grass- or shrublands in all savannas. On the other hand, the large number of herbivores weakens the field layer and thus reduces the destructive effect of the fires. As the rainfall decreases, the savannas turn into thornbushes and ultimately into semi-deserts characterized by thorns, succulents, and annual grasses. As the aridity increases, the vegetation cover decreases. Rainforests are easily transformed into savannas in frontier areas, and when they do, recovery takes place very slowly. Local monitoring of these conversion areas has revealed that it can take up to 30—40 years for rainforest vegetation recover after it has been turned into a savanna, even when protected from forest fires (Abbadie et al. 2006).

The contemporary classification for wet and dry savannas was developed by Huntley & Walker (1982). According to this theory, savannas can be divided in two categories according to the humidity and soil nutrition conditions. There are key differences between these savannah types. Many tree species are specialized to either of this vegetation type and there are clear variations between the flora and fauna. Arid, eutrophic savannas are located in areas where annual rainfall is under 650 mm (+-134mm, depending on the distance from the equator) and there are at least 7 dry season months, usually in two separate seasons. Dry savannas are characterized with younger soils that have accumulated in valleys and volcanic lands. Soil's ph. value is neutral or slightly acidic and either calcareous, eutrophic or volcanic. Water availability is the minimum growth factor and typical for dry savannas, tree cover is densest around the riverbanks. This type of savannah occurs mostly near to equator in Eastern Africa, for example in Somalia, Kenya and Northern Tanzania and they are bordering several great river valleys and desert areas like Sahara in the North and Namibi-Kalahari Desert in the South (Niemelä 2011: 33).

In moist dystrophic savannas the annual rainfall is generally 650—1800 mm annually and the dry season lasts at least three months. Miombo woodlands represent this type of wet savannas. Usually there is one intense rain season and one very dry season per year. Wet savannas are common farther from the equator and especially in highlands, and they are covering most of the Southern Africa inland from Kongo in the North, Tanzania and Mozambique in the East until Namibia and Angola in the West (Abdallad & Monela 2007). The soils in wet savannas are acidic and dystrophic or mesotrophic. These soils are low in nutrients as they are often located in highlands which have been washed out of during millions of years of the history of African continental shelf. Usually in the borderlands higher terrains are wet savannah and lower slopes dry savannah. Tree species richness is higher in wet savannas and large thick leaves fall gradually during the dry season. The fire resistance is also higher than in dry savannas due to the smaller amount of biomass on the ground. The differences between these savanna environments has modified the prevailing cultures during thousands of years, whereas dry savannas have given birth to nomadic cultures and in wet savannas life is based on slash-and-burn agriculture.

The dispute over the origin of savannas centralizes around fire (Niemelä 2011: 35). In Savanna ecosystem, fire is a regular phenomenon that shapes the landscape alongside other natural factors. Flash-and-burn agriculture has been part of the culture in Eastern Africa for thousands of years. Fire has a stabilizing role in savannah ecosystem (Hutley & Setterfield 2008). Fire molds and maintains the vegetation of the savanna and is an essential element of the ecosystem. It has been estimated that in Africa about 1.3 million km2 of savanna and grassland ecosystems burn annually (FAO 2015b). The consequences of fires depend on their frequency, timing, and intensity. Most fires occur during dry season (Tarimo et al. 2015: 10). It has been estimated that in Tanzanian wet savanna region up to 50.6 % of woodlands are annually affected by fire (Tarimo et al. 2015).

Fires are most intensive in areas of wet savannas, where naturally ignited fires are more frequent due to the higher amount of biomass and intensive thunders at the beginning of rainy season. In arid savannas fires are light and occasional and the effect on the vegetation is low (Frost 1987). The intensity of the fire is influenced by the amount of combustible plant material, water content and wind conditions (Ryan & Williams 2011: 58). Fire has an uneven damage to the nature, and it spreads out in mosaic pattern. Terrain and vegetation cover variations form places that are often burned, but also non-flammable refuges (Niemelä 2011: 36). Areas of dry trees and wood litter burn for a long time and might give birth to local spots where the undergrowth is totally dead. Depending on the wind, fire can be intense fast burn, or a slow burn of a calm weather. In a fast burn, shrubs and trees may suffer, as the hottest thermal profile section is higher. In slow burn the hottest section is lower and moves slowly, causing more damage to the undergrowth.

During their existence savannas have been balancing between different constraining forces. The availability of water and nutrients regulates the potential flourishment of vegetation, whereas fire and the number of herbivores determines whether this natural condition of vegetation is reached (Niemelä 2011: 32). The overall ecological balance in the region determines which of these factors are crucial in certain areas. Third major factor is human. Savannas, especially in tropical regions, are under increasing development pressure, which is why the sustainable management of savannas is becoming increasingly important (Hutley & Setterfield 2008).

2.4. Processes behind forest change in Miombo woodlands

Deforestation has been an ongoing process in Miombo woodlands during the human history, but the sparse settlement and low human pressure have left the large areas of the region relatively intact (FAO 2015b). Low soil fertility, tropical diseases and the lack of infrastructure have historically saved large areas of Miombo from permanent agriculture and settlement, conserving the woodlands until present day. However, the population growth rates in Miombo regions are high, and infrastructure

improvements are carried out, which accelerates the rate of agricultural expansion and now intensive farming is threatening the remaining areas (WWF, Niemelä 2011). Processes behind deforestation and woodland degradation in tropical woodlands are illustrated in Figure 1. Common human activities behind these processes include, but are not limited to, farming activities (also slash-and-burn), infrastructure such as roads and buildings, pastoralism and collecting firewood for charcoal market. Cultivation of cash crops like tobacco is increasingly taking space from subsistence farming and accelerating the clearance of woodlands.

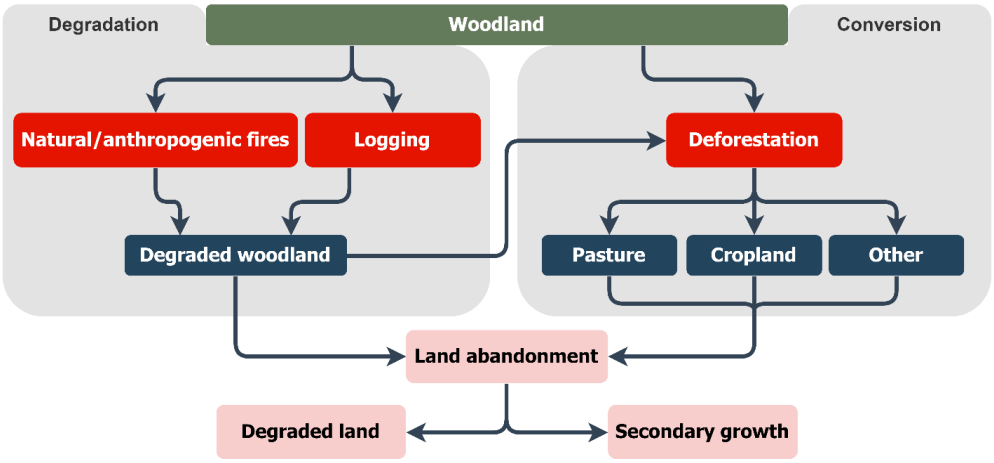


Figure 1. Processes behind deforestation and woodland degradation in tropical woodlands (modified from Souza 2006).

According to Jew et al. (2016), collecting firewood is the most common and most damaging form of forest practice in tropical Africa. Environmental impacts are profound in both forest and savanna ecosystems. In dry savannas, the felling of trees can, together with the overgrazing, destroy the entire vegetation cover and transform it into grassland or scrubland. Wood is largely used as firewood or charcoal in cooking, but also for curing tobacco (Silayo et al. 2008). Abdallah et al. (2007) have evaluated that in western Tanzania for every ton of dried tobacco, one hectare of Miombo woodland is cleared.

Fuelwood harvesting and charcoal production have been recognized as a major deforestation factor also in Southern Highlands of Tanzania (Jew et al. 2016). Local people are highly dependent on the fuelwood and charcoal that are used as energy sources and harvested from woodlands and forests as 90 % of domestic energy needs are met by charcoal and firewood (Abdallah et al. 2007, Malimbwi et al. 2004). Rural population rarely have access to electricity and high tariffs also pushes the people in urban areas to lean on charcoal and firewood. Extensive areas of woodland are cleared to meet the demand for charcoal in urban areas. Clearance often begins along the main roads and continues to spread with the developing infrastructure (WWF 2007). In addition to the forest clearance, charcoal

making is causing accidental fires, which contribute to deforestation and forest degradation (Silayo et al. 2008). Recent estimates of Tanzanian forest and woodland loss range between 372,000 and 580,000 ha per year (MNRT 2015; URT 2017). With sub-Saharan Africa being one of the only regions in the world where high population growth is predicted for the coming decades, the pressure on the Miombo ecoregion will also grow significantly (Syampungani et al. 2009). Population growth and subsequent rise in population density is the greatest single factor for the environmental change and deforestation in the Miombo region.

The Miombo region has experienced a long shift from small scale sedentary farming towards a more intensive use of the land (Jew et al. 2016). This has become possible though several socio-economic, technological, and political changes and improvements. Ideally, small area of Miombo is cut and cultivated for 1-2 years, and later left to recover for 15-20 years (Niemelä 2011: 240). In this way people can leave in harmony with nature, as part of the natural life cycle of Miombo ecosystem. Long term human impact can damage the forest structure and species composition profoundly (Cambell et al. 2007).

More recent threat for Eastern African savannas is cash-crop cultivation (Niemelä 2011: 244; Jew et al. 2016). Coffee, tobacco, tea, sisal, and cardamom are the most important cash crops in Southern Highlands of Tanzania. Growing tobacco requires large-scale burning and drying the leaves consumes a lot of firewood. In Iringa region, broad areas of Miombo have already vanished because of tobacco cultivation and curing in Iringa region. Another growing threat for native woodlands in Tanzania is forest plantations, although the effect of plantations on Miombo are twofold (Silayo et al. 2008). Miombo is likely to be more affected since it is sparsely populated, and the economic value of the region is otherwise considered low due to the poor soils. Most of the planted trees are foreign species that replace the original forests. Forest plantations demolish local ecosystems and biodiversity and profoundly alter the living conditions of local people (Cambell et al. 2007). These people's livelihoods may highly depend on firewood, building materials, naturally growing edible and medical plants that the forest provides and as for plantations, the economic benefit often accumulates outside of the area. However, if managed sustainably and considering the local vulnerability, biofuel plantations can also compensate the growing need for firewood, which is posing the most acute threat for Miombo woodlands (Silayo et al. 2008). As the collection of firewood is one of the most damaging forms of forest practice throughout the ecoregion, cultivation of biofuel components could preserve broader areas of Miombo from gradual deforestation and forest degradation caused by the fuelwood collection.

Miombo is the core region for slash and burn agriculture in Africa, and the fire has left its mark everywhere (Tarimo et al. 2015; Jew et al. 2016). Infrequently repeated fires are more destructive than annual minor fires. Even though fire is often naturally ignited in savanna ecosystem, human

has increased the incidence of fires in all savanna regions. Grasses are deliberately ignited to create new growth in the plains, to improve grazing and to reduce unwanted organisms like insects or snakes and to open the areas for hunting (Malmer & Nyberg 2008; Niemelä 2011). Paths and trails are burned to remove the thicket, and sometimes fire ignites from its remiss usage. However, the role of human activity in the savanna fires is not unambiguous (Hutley & Setterfield 2008: 3148). The impact of human-caused fires is relative to background rates of natural ignitions. Human alters the fire dynamics also by changing their timing and altering the fuel abundance and structure (Bowman el al. 2018). Frequently recurring fires have the advantage that dead plant material does not accumulate in the ground; thus, the fire does not grow dangerously. The man has consequently increased the frequency of fires but reduced their intensity (Niemelä 2011: 36). On the other hand, annual burning of savannas prevents tree seedling development and gradually transforms forest savannas into woodlands and woodlands into grasslands. Lightnings ignite grassfires more frequently during the beginning of dry season, when the vegetation is arid, and thunders are common. This is when fires damage the tree vegetation most. People burn the terrain most at the end of the rainy season, when fire damages more of the undergrowth than the trees. The undergrowth is fire sensitive during this time since the grasses have not yet settled in dormant state and humidity leads the heat into to the buds under the soil.

Majority of the fires do not damage the vegetation significantly (Bowman el al. 2018). Savannah vegetation has several characters that protects it from the thermal damage and some of them require contact with fire at some part of their life cycle to reproduce. The effect of burning is momentarily advantageous for the soil nutrition, but in the long run the effect is depleting. Long-term effects of recurring fires depend on the soil content, but normally in poor soil conditions the vegetation begins to suffer from a lack of phosphorus.

2.5. Remote sensing and change detection

Earth observation through satellite images is widely used to monitor forest cover changes across the world. Remote sensing is defined as the science of obtaining information of a research subject through the data that is acquired by device that is not in contact with the subject under investigation (Lu et al. 2004; Purkis & Klemas 2011). Remote sensing can be active or passive (Adams & Gillespie 2006: 28). In active remote sensing, the sensor emits electromagnetic radiation which interacts with objects or material it encounters. For example, Radar and Lidar systems rely on active remote sensing. The passive remote sensing system monitors the radiation which originates from the Sun but is reflected or emitted from the surface of the objects or material that is under the observation. Remote sensing enables the data collection from inaccessible areas and can be used to monitor and predict natural and human caused effects in the fields of agriculture, natural resource management,

urbanization, natural hazards, fluvial environments and so on. The methodological strengths of remote sensing are its objectivity, repeatability, and spatial explicitness (Griffiths & Hostert 2015: 308).

In remote sensing, change detection is based on the principle that landscape features have unique spectral signatures, and changes in these signatures reflect landscape changes (Adams & Gillespie 2006; Verbesselt et al. 2010). In other words, changes in land cover cause alterations in radiance values, but these changes can only be monitored with respect to radiance changes affected by other factors. Within the framework of this thesis, change detection refers to the process of identifying transitions or dissimilarities in the state of a phenomena by observing it at different times and the ability to quantify these temporal effects.

Land use/land cover change plays a crucial role in global environmental change. Numerous studies have been dealing with land cover change and forest degradation (see e.g. Banskota et al. 2014; Cohen et al. 2018; Hislop et al. 2018). Change detection research is getting increasingly diverse as the topic is currently very active and growing amount of free and high-quality data allows the development of the field. Change detection approaches are useful in extensive amount of applications, such as studying vegetation phenology, land cover changes, forest degradation, cropland biomass and other environmental phenomena. Satellite remote sensing has been used for detecting changes in land cover since the first Landsat images became available in 1970's (Lu et al. 2004). It is essential to distinguish between the terms "land cover" and "land use". Land cover refers to the physical features of the land, while land use indicates the human impact, e.g. how the land is used by people. When detecting change, it is important to notice that disturbances can result in both land cover and land use changes and others may result in a change in land cover but not land use, like fire causing deforestation in planted forest.

A variety of factors influence the spectral values recorded by a sensor like differences in illumination caused by the Sun angle, atmospheric conditions, topography, seasonal vegetation phenology and soil moisture (Lu et al. 2004; Purkis & Klemas 2011). In addition, problems and differences between the sensors which detect the radiations lead to variation in how these reflectance values are recorded. Areas of high altitudinal variation might cause some geometric distortion. Different radiometric and geometric calibration methods are available for correcting the changes that are caused by these technological challenges in satellite sensors. Thus, separating true spectral change from noise is a primary challenge in any change detection analysis (Purkis & Klemas 2010; Verbesselt et al. 2010).

The natural effects causing variation of the reflectance values can be controlled with several approaches and methods. These include reducing the effect of Sun angle and phenology changes by selecting images that are acquired same time every year, or reducing the atmospheric impacts

by identifying the pixel values that represent cloud, haze or shadows and masking these out from the analysis. Alternative approach is to use statistical approaches in selecting the “best pixel” by aggregating pixel values by certain rules or by using spectral indices derived from original reflectance values (Purkis & Klemas 2011). The idea behind the use of spectral indices is to transform multi-spectral satellite data into individual components (Adams & Gillespie 2006). The combination of often just two or three wavebands can significantly simplify the interpretation of otherwise very complex spectra by reducing the number of variables present in multispectral measurements down to one unique component. Image enhancement can be done applying mathematical functions and transformations to multispectral imagery.

Previous research literature has demonstrated that post-classification comparison, image differencing and principal component analysis are the most adopted methods for change detection (Lu et al. 2004; Gómez et al. 2016). For just monitoring change without interest in the nature or direction of it, simple image differencing and -rationing methods can be used. The post-classification techniques, which are the most traditional and widely used method in change detection, create a change matrix where rations of different land cover types can be compared. In the 2000’s new approaches have been rising, partially due to the increasing availability of satellite data. These approaches include time series segmentation, that can identify trends and breakpoints in time series, stacked generalization, spectral mixture analysis and artificial neural networks (Kennedy et al. 2010; Gómez et al 2016; Healey et al. 2018). This shift is mostly resulting from open data policies, such as releasing complete Landsat archive for open usage in 2008. This has led to substantial development in change detection methods based on Landsat time series, which allows higher accuracies in forest change timing and resolution (Wulder et al 2012). Increasing computer processing capabilities together with freely available Landsat data from last four decades has facilitated the development of advanced time-series techniques for analysing forest change. The section of the most suitable change detection method can be a difficult task as there is no single approach is optimal and applicable to all cases, but the method must be selected in relation to the nature of problem (Lu 2004).

2.6. Time series analysis as an approach for deforestation research

Time series can be described as a continuous or discrete series of data, where values are collections of observations that are taken sequentially in time (Chatfield 2016: 1—6). The subject of interest in time series analysis is the temporal variation within these sequences that is dependent on the time variable and reveals time series effects like trends or correlations as well as cyclic, transitory and random behavior. In land remote sensing science, time series analysis is generally understood as temporal monitoring of land surface dynamics through sequential observations in time that can

capture the trajectory of dependent variables (Jamali et al. 2015; Kuenzer et al. 2016; Chatfield 2016). These variables can be for example geophysical-, index-, or thematic variables. Time series approach as a methodological framework allows tracking changes in forest cover with high temporal resolution which enables the detection and quantification of gradual canopy changes in long term (Griffiths & Hostert 2015: 308). Change in ecosystems can be characterized in three categories: seasonal change, which is affected by annual variations in rainfall and temperature or the distribution of land cover types with different plant phenology; gradual change such as forest degradation or interannual climate variability; and abrupt change, caused by severe disturbances such as deforestation, fires, agriculture and construction of infrastructure. The ability of any change detection approach to identify changes depends on its capacity to separate actual change from variability at seasonal scale (Purkis & Klemas 2010; Verbesselt et al. 2010).

Different methods are available for characterising forest dynamics through time series analysis, but the use of spectral indices is widespread (Hislop et al. 2018). According to Adams & Gillespie (2006: 254), geophysical models like index variables provide the best linking between image and ground as they contain insight to the criteria which is used to monitor certain land cover or vegetation type. We especially need physical models to interpret spectral data for the non-visible wavelengths, to be able to extract the valuable information in areas outside our visual abilities. Geophysical variables are various reflectance values derived from raw digital numbers, whereas index variables, which represent physical models, are dimensionless, such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Burn Ratio (NBR) or Leaf Area Index (LAI). Each of these variables can be analysed in long term time series to derive for example minimum and maximum, mean, standard deviation, variability and trend values from daily to inter-annual timespan. Thematic variables in time series analysis are derived from original values usually through classification or regression approaches and stacked as time series to retrieve statistical parameters. Most remote sensing based phenological studies apply curve fitting techniques to timeseries of spectral indices that are sensitive to vegetation dynamics (DeVries et al. 2015; Roy et al. 2015; Cohen et al. 2018). Landsat is commonly used in regional level change detection studies due to its temporal and spectral resolution, availability, and sensibility to forest change (Wulder et al. 2012).

Forest disturbance monitoring based on the Landsat time series has been significantly improved in the 2010s, and new algorithms utilizing a well-calibrated image archive have emerged (Roy et al. 2014). These algorithms often utilize all available high-quality images to obtain the best annual or seasonal pixel composites, which are then used as an input in time series algorithm. Prior to data opening, Landsat time series analysis was limited to areas with technological and economic deficiencies to access the data, and even then, most applications relied on bi-temporal approaches such as image differencing and post-classification methods.

With 30-meter pixel resolution, Landsat provides an appropriate scale for phenological monitoring. Several studies have demonstrated that moderate resolution Landsat data reduces the mixing of different landscape components and increases the ability to capture smaller scale phenological variations in long-term time series and enable more specific scaling of ground-based observations (Verbesselt et al. 2010; Zhu et al. 2012; DeVries et al. 2015). Using time series instead of post-classification or image pair approaches, the change can be separated from background noise and gradual, subtler changes can be identified.

2.7. Spectral indices

Forest dynamics, like abrupt changes and longer-term trends, can be studied with several time series approaches, but the use of spectral indices is an exceedingly common practise in the field (Hislop et al. 2018). Spectral indices transform multi-spectral satellite data into a single component, so the trajectory of individual pixels can be traced through time. Different materials have complex characteristic spectra, which indicates the amount of absorption or reflectance of different wavelengths of solar radiation (Jones & Vaughan 2010: 164). This complexity can be processed to simplified measures that can be used to describe wide variety of physical phenomena on the ground. The combination of often just two or three wavebands can significantly simplify the interpretation of otherwise very complex spectra by reducing the number of variables present in multispectral measurements down to one unique component. This provides opportunities for enhancing the sensitivity of the variable by measuring differences or ratios between the selected bands (Jones & Vaughan 2010: 164). In addition, it also allows the normalization of the data, by minimizing the effects of irradiance caused by other unrelated reflectance variables such as variation in atmospheric conditions. The spectral bands are chosen based on their ability to convey information about the biophysical parameters of interest, like water content, amount of chlorophyll or light interception (Horning 2010: 108). In remote sensing, the principle of deriving spectral indices from measures at certain wavelengths is extensively adopted, particularly for studying vegetation cover (Jones & Vaughan 2010: 165). With the availability of even more complex, hyperspectral data, the principle has been progressively expanded to the development of more detailed indices that can be used for not only detection and quantification of vegetation but observing specific vegetation characteristics. Example applications include environmental health and natural resource monitoring, biodiversity assessment, carbon management and agricultural efficiency evaluations.

Vegetation indices (VI) are combinations of spectral channels, blended in such way that they strengthen the spectral impact of green vegetation, measuring the canopy “greenness” that can be used to quantify vegetation volume and health (Hislop et al. 2018). Traditionally this measurement combines the leaf reflectance signal in the near infrared region (NIR) with the chlorophyll-absorbing

red spectral region (RED) to provide a constant and vigorous rate of per-pixel average canopy photosynthetic capacity (Njoku 2013). There are several examples in which the NIR, red and other bands may be blended to measure vegetation, resulting in variety of vegetation index equations including linear and optimized band combinations, band ratios and normalized differences (Tucker 1979).

Two areas of electromagnetic spectrum are strongly absorbed by the chlorophyll pigment: blue (450nm) and red (670) light (Purkis & Klemas 2010: 64—65). The magnitude of absorption is directly related to radiant energy utilized by the plant. Near-infrared region (700-1200nm) acts inversely, displaying a pronounced peak in reflectance. Most vegetation indices are based on the sharp increase in vegetation related reflectance that occurs around 700 nm, which is also known as the red-edge, because the transition from low to reflectance in the red to the high in the near-IR is very abrupt. This shift where the reflectance of near-IR is several times as large as the visible band is characteristic specially for green vegetation as other natural surfaces show relatively slow variations of reflectance in these wavelengths. Over the years, many vegetation indices have been developed to exploit this phenomenon and most vegetation indices utilize the red and near-infrared bands that capture this shift (Jones & Vaughan 2010: 164—168; Purkis & Klemas 2010: 64—65). To truly represent the vegetation conditions on the ground, satellite derived vegetation index should always, when possible, be calculated from surface reflectance values, in which the atmospheric, radiometric and illumination effects have been corrected.

The Normalized Difference Vegetation Index (NDVI) is one of the most widely used vegetation indices, as it has been in use since 1970s (Rouse et al. 1974) and it is frequently used in Landsat derived time series analysis (Verbesselt et al. 2010; Hislop et al. 2018). NDVI is a measure of photosynthetic biomass, which correlates well with ecological variables, such as the fraction of green vegetation cover (Hislop et al. 2018). NDVI is generated by dividing the difference between reflectance in the infrared and the red with the sum of near infrared and red reflectance (Purkis & Klemas 2010: 65).

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{RED})/(\mathrm{NIR} + \mathrm{RED})$$

Division by the sum of reflectance from these bands reduces the effect of non-uniform illumination and thus enhances the comparability of resulting vegetation index values across the satellite image. The range of values obtained by NDVI is between -1 and +1. Positive values correlate with the vegetated pixels, and the higher the index, the greater the chlorophyll content of the target. Negative values indicate higher reflectance in the visible wavelength region than in the infrared region and is often due to snow, clouds, bare soils or rock. The value of the NDVI varies depending on the season,

hydric conditions, and the local climate (Purkis & Klemas 2010: 66). The index gives a relative indication of the vegetation rather than an absolute measurement.

NDVI is responsive to changes in vegetation condition and has been demonstrated to accurately detect forest disturbances (Hislop et al. 2018). However, there are also some disadvantages related to the saturation effect in higher vegetation densities, nonlinear performance of ratios and sensitivity to soil reflectance in areas of sparse vegetation. NDVI is also reactive to grasses and other non-forest vegetation, which might colonize the site after a disturbance and give a false signal of forest recovery. Consequently, new variants of the NDVI have been developed to minimize the external influences and to improve the overall performance. Soil Adjusted Vegetation Index (SAVI) has been developed by Huete (1988) to correct the soil effect from NDVI formula by incorporating an adjustment parameter (L) in the NDVI equation:

$$\mathrm{SAVI} = \frac{(1 + \mathrm{L})(\mathrm{NIR} - \mathrm{RED})}{\mathrm{L} + \mathrm{NIR} + \mathrm{RED}}$$

Other versions of SAVI include the generalized SAVI (GESAVI), modified SAVI (MSAVI) and transformed SAVI (TSAVI) (Jiang et al. 2007).

Another external distorting factor for vegetation index performance is the atmospheric influence. The Enhanced Vegetation Index (EVI), which apply the blue spectral band to correct the aerosol impact in the red band, has been optimized for resistance to atmospheric effects. Furthermore, the optimized use of the NIR band reduces the sensitivity to soil reflectance and signal saturation problems in densely vegetated areas through the greater optical penetration (Huete et al. 2002).

$$EVI = \frac{2.5(NIR - RED)}{(NIR + 6.0 \times RED - 7.5 \times BLUE + 1.0)}$$

Another common approach is using short-wave infrared (SWIR) band combinations. Indices using short-wave infrared (SWIR) bands are sensitive to alterations in forest moisture, structure and vegetation density and thus are frequently used in Landsat time-series (Epting et al. 2005; Kennedy et al. 2012; Hislop et al. 2018). The Normalized Burn Ratio (NBR) was developed by Key & Benson (2006) to identify burned areas from satellite images. The index is generated similarly to NDVI, except that it uses near-infrared (NIR) and shortwave-infrared (SWIR 2.08—2.35) portions of the electromagnetic spectrum. This is a powerful combination of bands to use for vegetation change

studies as vegetation reflects strongly in the NIR region and very little in the SWIR. NBR is especially sensitive to reflectance changes caused by fire as a fire scar contains scarred woody vegetation and reflects more strongly in SWIR region of electromagnetic spectrum (Fornacca et al. 2018). Several authors have found out that NBR correlates highly with forest ecosystem condition and that the index can accurately track changes in forest ecosystems [Kennedy et al. 2012, Cohen et al. 2018, Epting et al 2005, White et al. 2017).

$$NRB = (NIR - SWIR)/(NIR + SWIR)$$

The Normalized Difference Moisture Index (NDMI) is very similar to NBR, but it uses the shorter SWIR band instead of the longer one, which has shown to be just as sensitive to the moisture content of soils and vegetation, even though it is less often used for burned area recognition.

Another technique that can be used to measure the amount and status of green vegetation in multispectral imagery is principal components analysis (PCA). PCA is a vector space transformation for multivariate data that is commonly used in multivariate statistics to decrease the redundancy of different channels (Horning 2010: 111). It transforms several correlated variables into a smaller number of uncorrelated variables called principal components. In remote sensing, these multiple variables are bands that are converted into a smaller set of principle components that describe the variance in the image. The Tasselled Cap transformation is a standardized PCA approach, first presented for Landsat Multispectral Scanner (MSS) data by Kauth and Thomas in 1976 and later for Landsat TM data by Crist and Cicone in 1984 (Crist et al. 1984). The Tasselled Cap transformation (TCT) is a physical model that was originally designed for agricultural landscapes (Adams & Gillespie 2006: 270). When applying Tasselled Cap transformation coefficients, it is possible to compare component images to identify changes in land cover over time when applied to time series (Horning 2010: 111).

Three main components from Tasselled Cap transformation are Tasselled Cap Brightness (TCB), which in simplified terms represents the overall brightness of all bands, Tasselled Cap Greenness (TCG) which a contrast between the visible and near-infrared bands, similarly to NDVI, and Tasselled Cap Wetness (TCW) which is the contrast of the visible and near-infrared with the SWIR bands and is useful for tracking vegetation dynamics based on the moisture content of the vegetation (Crist et al 1984). Tasselled cap images are especially helpful in discriminating areas of exposed substrate from those of closed canopy vegetation. Greenness and wetness are spectral attributes which could not be mapped on the ground, but they can be used to identify certain elements from satellite images and to monitor changes that are taking place in them. Various time-series studies have proven

success with these TC components (Epting et al. 2005; Kennedy et al. 2010; Senf et al. 2015; White et al. 2017).

Although the range of different spectral indices is broad, in Landsat time series the use of NDVI in forest dynamic studies remains (Hislop et al. 2018). Both NDVI and NBR can accurately capture fire disturbance in a pixel-based time-series approach in sclerophyll forests of southeast Asia. However, there are some differences that appear when taking a closer look to the specific characters of these indices. NDVI, which captures the chlorophyll content of the forest canopy, returns to pre-fire levels soon (in 3—5 years) after the fire event, even though the vegetation has not fully recovered. On the other hand, indices that are more sensitive to the moisture and structure, such as NBR and TCW have much longer recovery rates (8—10 years), indicating more trustworthy results for the full recovery of the forest vegetation after fire disturbance. Stress is more challenging to interpret than other disturbances and it is usually developing gradually. It is also easily mixed with natural variability in vegetation index values, for instance a normal drought response might result in an erroneous stress detection. Another challenging disturbance that cannot be accurately detected from satellite images without additional information is insect/disease outbreaks. As the Landsat timeseries-based approaches are becoming more popular, understanding the limitations and advantages of these indices becomes crucial.

3. Study area

The study focuses on landscapes around rural town called Mantadi in Tanzanian Southern Highlands. The area of Southern Highlands refers to the region encompassing the four sub-regions of Mbeya, Iringa, Njombe, Rukwa and Ruvuma (Figure 2). The climate in Southern Highlands is characterized by long annual rainy seasons, starting at the end of October, and continuing until May, which makes the region one of the most agroecologically productive zones in Tanzania. Dry season is from June to October, during which several forest and grass fires are taking place (URT 2012).

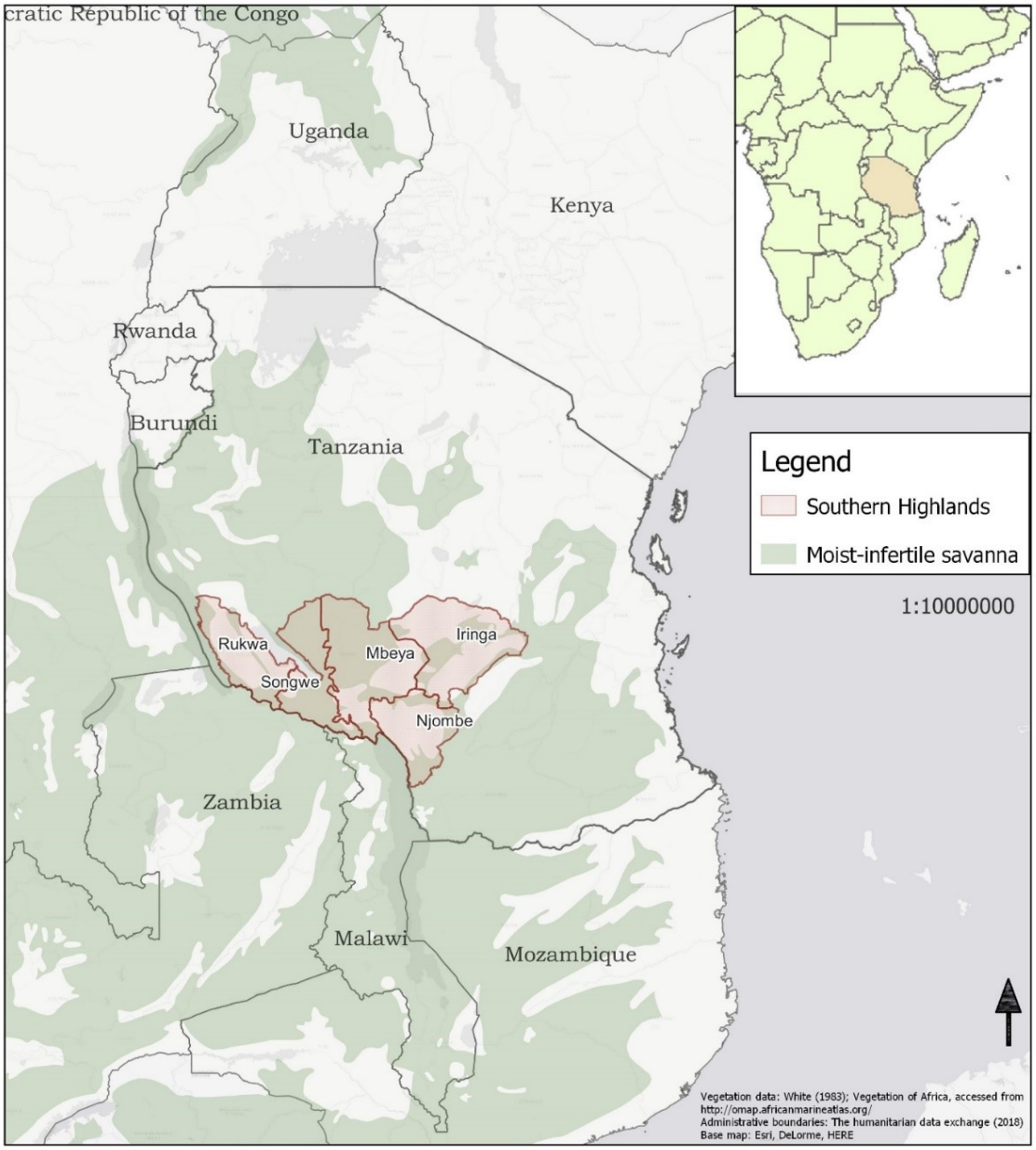


Figure 2. Southern Highlands of Tanzania and the prevalence of moist-infertile savanna (Miombo) vegetation according to White (1983).

Mantadi (Isangawana) is located in Mbeya region, a town which has spread its agricultural zone on the both sides of the B6 Sirari-Mbeya road towards North (Figure 3). Despite of the small size of this town, it has a significant role in the accessibility of lands around it. The agricultural zone around Mantadi has been spreading gradually further to the Miombo woodlands in the North, clearing the forest and turning them into fields and pastures. These landscapes were chosen as they have been experiencing exceptionally high deforestation rates according to Global Forest Watch (Hansen et al. 2013), National Forest Resources Monitoring and Assessment of Tanzania Mainland -report (NAFORMA, MNRT 2015) and Tanzania's forest reference emission level report (URT 2017), and therefore were suitable for studying deforestation.

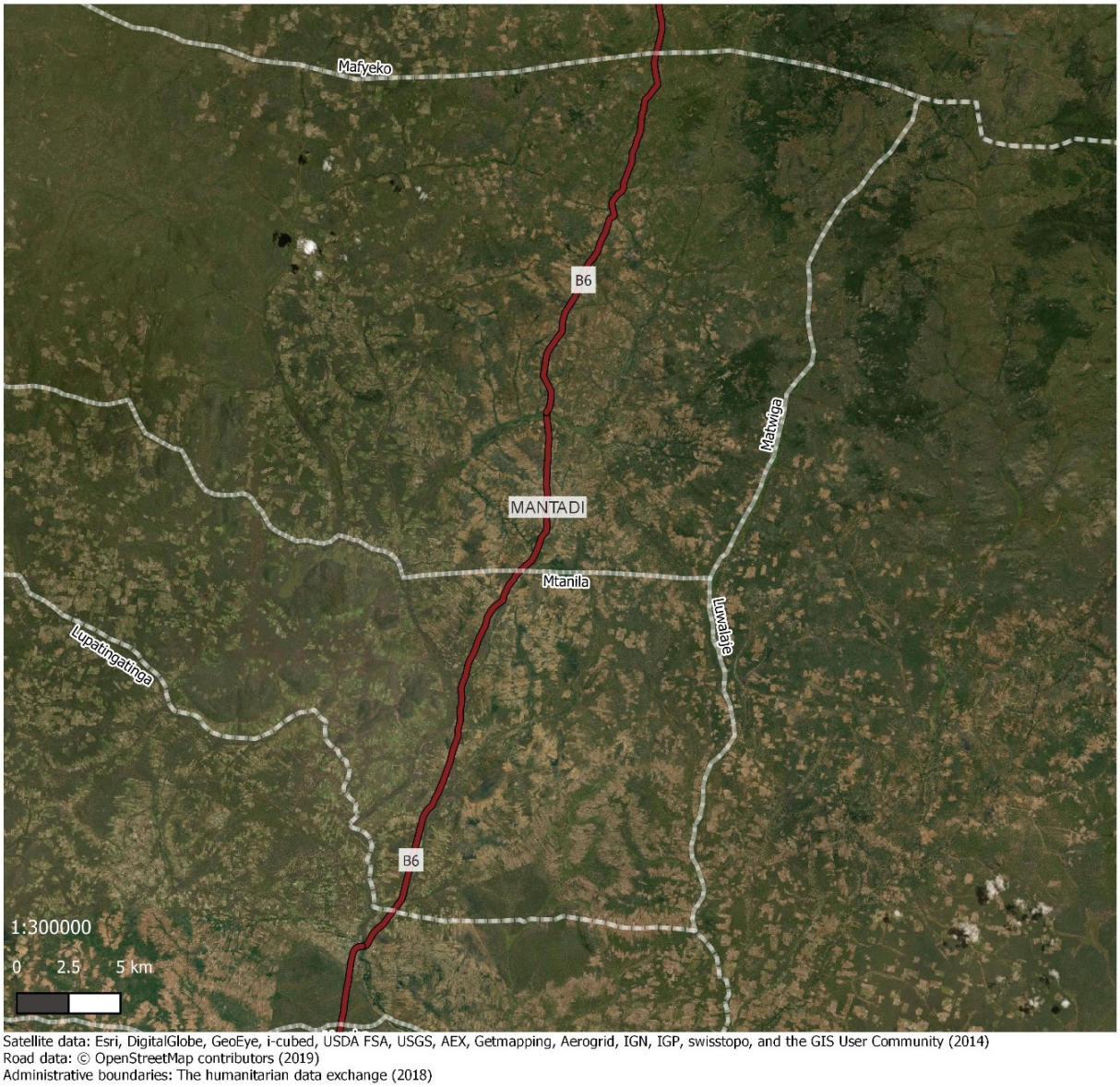


Figure 3. The study area in satellite image in 2014. Vegetation in the area is mainly Miombo woodland and the deforestation rates have escalated since 2000 (Hansen et al. 2013; MNRT 2015; URT 2017).

The study area is located in Chunya district in the North-Western part of Mbeya Region (Figure 4). The district is among the seven districts of Mbeya region and it lies between 7 and 9 Latitudes and 32 and 34 Longitudes. Chunya is the largest district in Mbeya region and more than 75 % of the land is arable land (URT 1997a). Despite of large area of arable land, it was estimated that only 2 % of the land is simultaneously under cultivation (URT 1997b). This indicates that even though reserved as arable land, a large portion of this area was covered by natural vegetation before the 21st century.

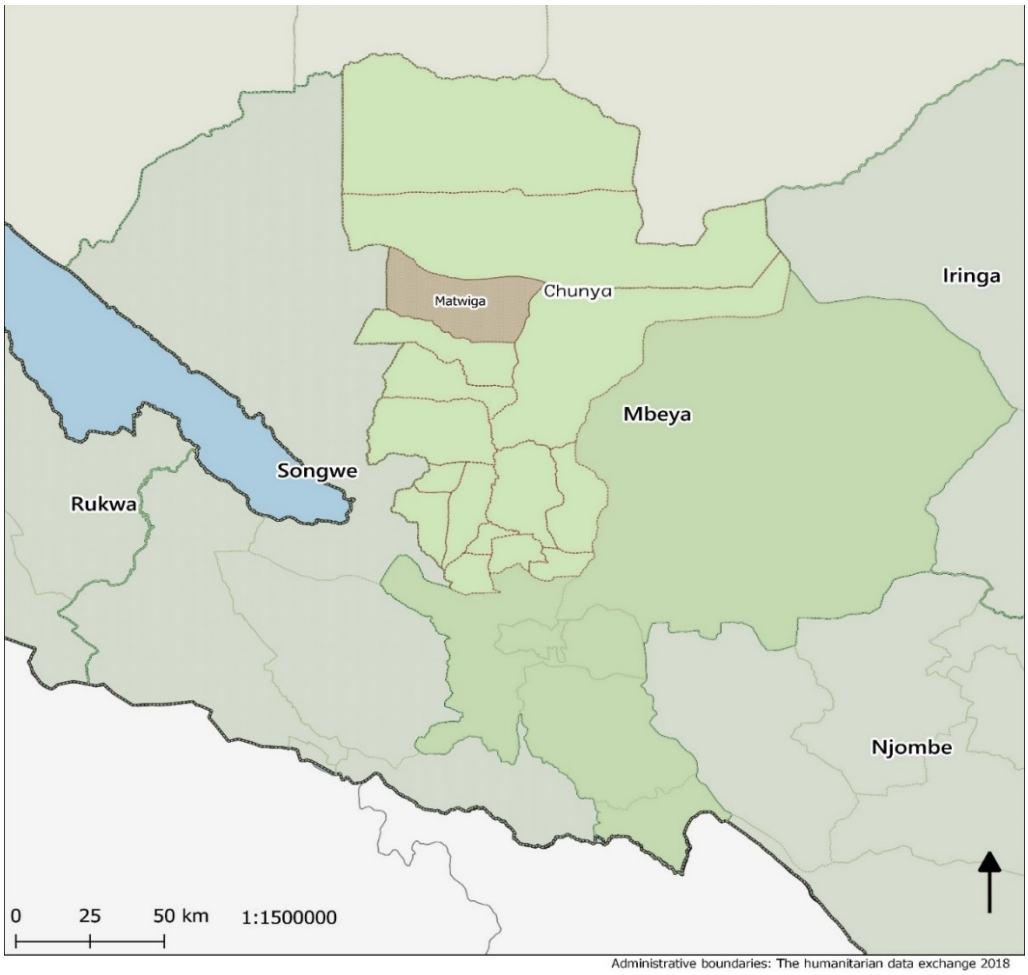


Figure 4. Administrative boundaries of Mbeya region. Town of Mantadi is located in the Matwiga ward of Chunya district.

Chunya district has been classified into three agro-economic zones and the study area is located in Miombo woodland zone. The Miombo woodland zone is characterized by, in addition to Miombo woodlands, thick forests, scattered trees, bush and thickets. The altitude in Miombo woodland zone ranges between 1400-1800 meter above the sea level and the mean annual rainfall is 750mm, with 4—6 wet months annually. Landscape is predominantly flat with shallow and sandy soils.

These soils provide good base for growing tobacco, sorghum, maize, fingermillet, cassava, sweet potatoes and groundnuts and have also potential in gold mining.

The vegetation varies from thick forests to Miombo woodlands, bushlands, and thickets with scattered trees. Average annual rainfall is 933 ± 36 mm (n = 28 years) and most of it falls between October and May (Jew et al. 2017). Miombo woodlands are the most predominant natural vegetation type and represents dry Miombo according to the 1000 mm limit, even though the classification in this area is not clear as the annual precipitation is very close to the limit.

As village-level statistics are not available, socio-economic data of the study area is based on the ward and district level statistics. Study area belongs to the Matwiga Ward but the town around which the deforestation zone has grown is officially called Isangawana, although in most regional maps it has been marked as Mantadi or Mantandi. In Matwiga ward the average household size according to the 2012 census was 5,1 and total population 8939 (NBS 2013). Across all villages, farming is the dominant occupation. Other activities include livestock keeping and harvesting forest products. At the end of the 20th century, establishment of new settlements was a common phenomenon in the area, initiated by the herdsmen and cotton and tobacco growers from Arusha, Shinyanga, Singida and Tabora regions (URT 1997b; Jew et al. 2017). Normally the migration from the southern regions is seasonal. Farmers clear forested landscapes for tobacco and cotton fields whereas herdsmen move into wooded grassland areas. Both migration groups are altering the natural environment of these landscapes. The population distribution patterns are formed according to climatic conditions and soil fertility, particularly in the highlands. The availability of rivers or water bodies are also attracting settlements. Village-level Participatory Forest Management committees oversee five Reserves, and the District Forestry Department govern three Forest Reserves within the division. However, Jew et al. (2017) found out that the reserves are poorly managed due to insufficient funding and limited personnel and transport capacity. This indicates that access to woodland is mainly unrestricted across both protected and unprotected areas.

The economy in Chunya district is predominantly subsistence and informal, which complicates the GDP calculations. It has been estimated that 60 % of the population are employed in the agricultural sector, where 22 % small-scale farming and 10 % fishing. Other small-scale industries and trades occupy less than 1 % of the population. Determination of GDP per district or capita has not been possible due to insufficient data. The economic potential of Chunya district is in large extent drained out with the cash crops and not directly contributing to its development. This is also why the nutritional status of the district is below average. (URT 1997b).

The major sources of energy for both domestic and industrial use are firewood, charcoal, hydro/thermal electricity, and oil products. Until 1994 the population of the whole district was depended on firewood, charcoal, and petroleum and these are still the most common sources of energy since rural electrification has not taken place. Other energy sourced are mainly used in the urban areas. In 1997 about 80—90 % of the energy was from firewood or charcoal and it was already then stated that if the total consumption of firewood remains the same, Chunya district would become a desert and the need of alternative sources for energy is serious (URT 1997b).

In 1997, farming and animal husbandry covered about 50—95 percent of the district GDP where other sectors, beekeeping, fishing, and forest activities were estimated to cover 10—20 percent of the district GDP (URT 1997b). Tobacco is still the leading cash crop, followed by cotton, sorghum, and paddy. Main food crops include beans, maize, cassava, sweet potatoes, and millet. All the farming is small-scale, even though the abundant arable lands and good climatic conditions would permit large-scale farming as well. Poor road network might be the main reason for the lack of large-scale cultivation of these crops. (MNRT 2015, Jew et al. 2016).

According to the Naforma report (MNRT 2015), Mbeya counts among the regions with largest volumes of commercial woodland but does not give more specific numbers on a regional level. Commercial woodland is defined by the type of trees that can be used commercially, but the definition does not consider the land area's legal status or accessibility. Commercial trees are found in production forests, protection forests and wildlife protected areas. Timber, poles, and charcoal are main products in the forest sector. It was estimated that the district annual wood demand is almost double of the available supply, which was leading to deforestation already before 21st century (URT 1997b). Between 1985 and 1995 the forest product harvesting multiplied six-fold in Chunya district. Felling of trees has led to soil erosion, loss of fertility and drying of the water catchment areas, which has a negative effect to the agricultural productivity. Agriculture, overgrazing and gold mining have seriously reduced the biodiversity in the Chunya district (URT 1997b).

4. Materials and methods

4.1. Study design

The study was built on three main phases that included several minor work steps (Figure 5). The first phase included the preparation and collection of reference data through several platforms and data sources. The LandTrendr algorithm was run first time already in this phase to ensure the selection of meaningful set of reference points. The reference samples were selected based on the preliminary LandTrendr results to ensure that samples representing forest changes and stable pixels were included. To collect the reference data through time travelling across the satellite images, Open Foris Collect -survey had to be prepared. Survey platform was constructed to contain the parameters that were to be filled during the actual data collection. Finally, the Collect Earth application was utilized to access the high-resolution satellite imagery simultaneously with the historical Landsat satellite composites. By exploring all the available data, visual interpretation of the change events in reference plots was conducted and filled to the survey.

In the next phase LandTrendr algorithm was parametrised separately for each vegetation index by utilizing the reference values from example pixels with very high or low vegetation, modified to match each index's relative values. After the parametrisation LandTrendr was run, and the raster layers showing the change magnitude and rate in each disturbed pixel were downloaded to ArcGIS software for further analysis. These raster layers were compared and analysed using raster statistics and reclassified to four quartile classes according to the magnitude of change in each pixel. Finally, the classified change maps were combined to create a new map composite representing the optimized results from all three index-based change raster layers.

The final phase involved the evaluation and accuracy assessment of these forest change maps. Error matrices were calculated for each index-based change map, combination of these maps and for the map showing the change rate. The change statistics and accuracy statistics between different change maps and the map combination were compared, and results visualized in figures and tables, where the results could be evaluated.

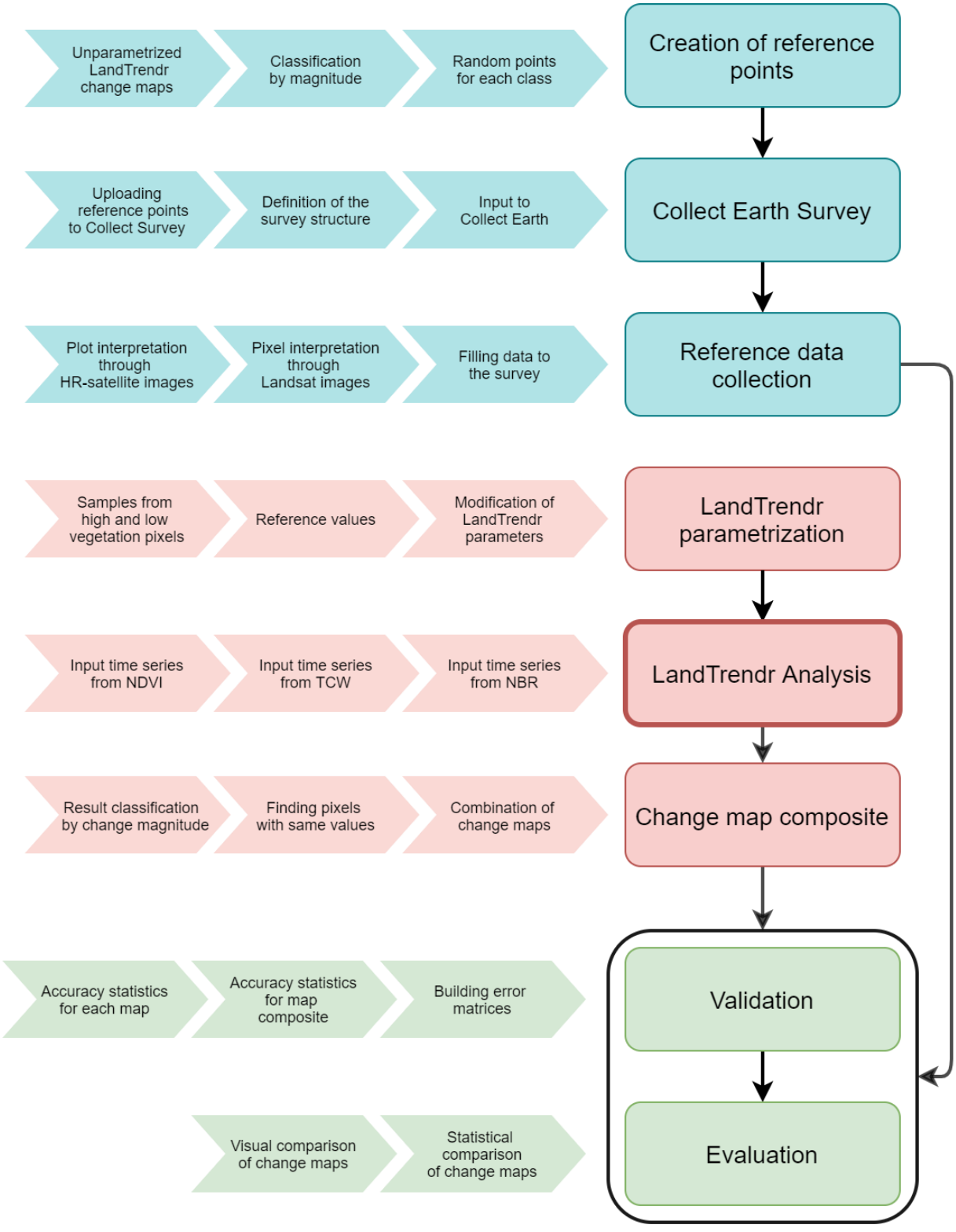


Figure 5. Flowchart illustrating the methodological steps in the research process.

4.2. Satellite data

The Landtrendr time series analysis was conducted based on Landsat Tier 1 products from Landsat 4, 5 and 8, covering the time period from 1987 until 2018 (Table 1). Landsat Tier 1 products include Level 1 Precision Terrain processed data, which are inter-calibrated across different Landsat sensors, radiometrically calibrated and geolocated consistently across the full collection for all the sensors. They are considered suitable for time-series processing analysis (USGS 2019). This data is available freely in Google Earth Engine catalogue and thus was selected for the research as it is high quality ready-to-use product, which saves time in preprocessing and allows focusing on the research objectives.

Time series was limited only to images available from dry season at the beginning of June until the end of September in 1987—2018. First three years (1984—1986) that would have been available in Landsat 4 and 5 were left outside of the analysis due to lack of data and high amount of clouds. All the Landsat satellite images have 30-meter spatial resolution and 16-day revisit time. Landsat 4 and 5 had similar functionalities and both carried Multi Spectral Scanner (MSS) and Thematic Mapper (TM) sensors, from which the TM-products were used as they provide 7 spectral bands and 30-meter spatial resolution (USGS 2018). In Landsat 7 the Thematic Mapper was updated to Enhanced Thematic Mapper (ETM+). In May 2003, the scan line corrector (SLC) of Landsat 7 failed, and the satellite images taken after that have systematic data gaps but can still be partly utilized. In Landsat 8 the sensor was updated to Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS), together containing eleven spectral bands and improved radiometric precision (USGS 2018). All the images were accessed from Google Earth Engine data catalogue (Gorelick et al. 2017).

Table 1. The Landsat data that was used as an input in LandTrendr algorithm. Data accessed through Google Earth Engine.

NAME	DESCRIPTION	DATASET AVAILABILITY	BANDS	RESOLUTION
LANDSAT 4 TM	Landsat 4 TM Collection 1 Tier 1 DN values, representing scaled, calibrated at-sensor radiance.	1982-August – 1993-December	1-7	30 meters
LANDSAT 5 TM	Landsat 5 TM Collection 1 Tier 1 DN values, representing scaled, calibrated at-sensor radiance.	1984-March – 2012-May	1-7	30 meters
LANDSAT 7 ETM+	Landsat 7 Collection 1 Tier 1 DN values, representing scaled, calibrated at-sensor radiance.	1999-January – present	1-8	30 meters
LANDSAT 8 OLI	Landsat 8 Collection 1 Tier 1 DN values, representing scaled, calibrated at-sensor radiance.	2013-April – present	1-11	30 meters

Several satellite data sources and products were exploited during the visual interpretation process for collecting the reference data (Table 2). Atmospherically and radiometrically corrected surface reflectance products from Landsat 4, 5, 7 and 8 were used, as well as the NDVI and EVI products derived from them. Where Landsat data were not available due to quality problems or data gaps, MODIS terra NDVI and EVI products were applied from the 21st century onwards. These images have lower spatial resolution (250 meters), but with 1—2-day temporal resolution MODIS provides more continuous image stream, enabling the identification of change events from areas and dates where Landsat data was not adequate. Additionally, high resolution satellite imagery available from Google Earth during varying time scales was explored. These were the most accurate images for the interpretation, and they were primarily used whenever available. The availability was highly depending on the location, resulting in different interpretation processes in every plot; sometimes the Google Earth imagery was enough for disturbance detection, but in most cases careful interpretation demanded additional imagery from other sources. All the data was accessed through Google Earth Engine data catalogue and Google Earth software.

Table 2. All the satellite data that was utilized for the visual interpretation of reference plots. Data accessed through Google Earth and Google Earth Engine.

NAME	DESCRIPTION	AVAILABILITY	RESOLUTION
USGS LANDSAT 4, 5, 7 AND 8 SURFACE REFLECTANCE	Atmospherically corrected Landsat-derived surface reflectance products	1982-present	30 meters
LANDSAT 8-DAY EVI COMPOSITES FROM LANDSAT 4, 5, 7 AND 8	Enhanced Vegetation Index (EVI)-composites are made from Tier 1 orthorectified scenes, using the computed top-of-atmosphere (TOA) reflectance.	1982-present	30 meters
LANDSAT 8-DAY NDVI COMPOSITES FROM LANDSAT 4, 5, 7 AND 8	Normalized Difference Vegetation Index (EVI)-composites are made from Tier 1 orthorectified scenes, using the computed top-of-atmosphere (TOA) reflectance.	1982-present	30 meters
MOD13Q1.006 TERRA VEGETATION INDICES 16-DAY GLOBAL 250M	The MODIS NDVI and EVI products are computed from atmospherically corrected bi-directional surface reflectances that have been masked for water, clouds, heavy aerosols, and cloud shadows.	2000-present	250 meters
GOOGLE EARTH HIGH RESOLUTION AERIAL AND SATELLITE IMAGES	High resolution imagery for viewing purposes, provided by Maxar Technologies, DigitalGlobe, Google and CNES/Airbus	2002-2019, depending on the area	1-5 meters

4.3. Reference data collection

Validating the LandTrendr disturbance maps and the land cover maps required collecting data through visual photo interpretation methods. In remote sensing, the visual interpretation refers to the human recognition of two-dimensional satellite images based on an interpretation made from visual elements such as hue, pattern, shape, texture and shadow (Campell 2008: 122). Without the field data it became crucial to understand the geographical context of the study area and to clarify the objectives of the photo interpretation. I prepared myself by orienting to the environment through literature and by time-traveling the area through Google Earth.

The traditional method for validation of LandTrendr disturbance maps is the TimeSync application, which was developed as companion to LandTrendr (Cohen et al. 2010). In TimeSync validation process the characteristics of spectral trajectories for sample plots are interpreted in combination with visual interpretation of the image to describe change processes, thereby segmenting a plot's spectral trajectory. This includes placing vertices in the spectral curve which correspond to changes in the landscapes. However, the usage of up to date TimeSync web application demanded a contact with the Oregon State University Environmental Monitoring, Analysis and Process Recognition Lab, which unfortunately did not react to the service request in time. The final solution was to develop a customized validation framework by combining the use of the Open Foris tools and the time series of several Landsat imageries and the spectral indices derived from them in Google Earth Engine. The Open Foris Initiative was established in 2009 by the FAO and its objective is to produce and distribute open source web-based software to facilitate monitoring and management of forest areas in cooperation with Google (Bey et al. 2016), and therefore was considered suitable for the purposes of this study.

The actual pixel-based reference data collection was done by using Collect Earth -tools within Google Earth platform. Collect Earth combines the strengths of different satellite data, by enabling their use side by side, complementing each other (Bey et al. 2016). The software provides access to several satellite catalogs that can be viewed using Google Earth, Bing Maps, or Google Earth Engine. The Collect Earth enabled the accessing of high-resolution satellite images from dry season in 2000-2018, depending on the image coverage in the study area, as well as the Landsat images from the whole time period (1984—2018) and MODIS images between 2000 and 2018.

The idea of customized reference and validation data collection scheme was to combine the information derived from the traditional Collect Earth tools through visual interpretation of satellite images and the information from time-series analysis that could be built for each sampling plot representing an individual Landsat pixel. This was possible by utilizing the part of original LandTrendr script, which builds the time-series composite from all the available Landsat images. The script could be modified in Google Earth Engine to apply for one sampling pixel at a time with defined year and

date range, to generate annual surface reflectance composites and spectral indices, from which the time series trajectory was produced for visual interpretation (Kennedy et al. 2018). The script includes all the preprocessing steps that were presented in the LandTrendr description.

4.3.1. Sampling scheme

First, an appropriate sampling grid was assigned to the research area using ArcGIS-software. To ensure versatile but random coverage for reference data plots, the Landtrendr algorithm was run with default parameter settings for TCW, NBR and NDVI index values. TCW, NDVI and NBR indices were used to confirm the widest possible coverage of different types of change events. If points were to select in totally random scheme, without preliminary change analysis, statistically great number of plots would have been from areas where no change had been occurring, which would have decreased the number of plots where the change events could be investigated. Now ¾ of plots were selected from areas where change events were more likely, but not evident, and ¼ from areas where algorithm did not predict change, to take into comparison the reference plots that were not identified by any index applied in LandTrendr.

To obtain a well-represented set of samples, stratified random sampling method is recommended, to collect a determined number of samples from each class (Lillesand et al 2008: 588). Samples were selected using a stratified random sampling method based on the results from LandTrendr algorithm. Study area was extracted by creating 20*20 km square buffer radius around the town of Mantadi. In remote sensing, the collection of minimum sample of 0,5 % that is recommended in statistics is often unrealistic, due to the large number of pixels representing the population. Three disturbance maps were downloaded from Google Earth Engine: one for each spectral index. Lillesand et al. (2008: 289) suggest that the sample size should be calculated in relation to proportional size of each class from total land cover. Thus, the disturbed pixels where divided in four quartiles (25 %, 50 %, 75 %) according to the magnitude of change values to get a reference data that represents changes as diversely as possible. Spatial autocorrelation, which can be resulting from clustering of samples, can be avoided by setting an appropriate distance between the sampling points (Lunetta & Lyon 2000: 6). Spatial autocorrelation was avoided by setting the sample interval in 500 meters. 15 random points were created separately for each change class using ArcGIS-software and finally combined into same layer. This resulted in 15 samples per quartile class per index and in total of 180 random sample points from the predicted change areas. In addition, 60 random points were selected outside of these change maps to validate the areas with no change. The final reference sample set included 240 samples from all around the study area. In addition, coordinates and altitude data were added to the grid for each point using ArcGIS-software and all

the unnecessary fields were removed. The final file was converted to csv format and loaded into the browser-based Collect survey designer.

4.3.2. Collect -survey designer

The sampling survey for Collect Earth tool was created in Open Foris Collect-tool. The Survey Designer platform allows the definition of the frame used to collect data, the so-called questionnaire, which is filled in at each sampling block according to the interpreter (Bey et al. 2016). A customized data entry form was created with questions about the land cover and its change. The survey included land cover classes based on Naforma (MNRT 2015), which was then supplemented with more detailed information derived from the actual data. Eight different main land cover classes were coded into the questionnaire, including subclasses that are typical to heterogenous landscapes (Table 3). Land cover classes that were not characteristic to the area were excluded from the survey. In addition to the land cover classes, questionnaire was filled with the year of the interpretation, the assessment of historical land cover, estimation of year of change, magnitude of change (from 1-low to 4-high), change type as gradual or abrupt and the estimation of canopy cover. Change events like burning incidences, human disturbances, and certainty of interpretation were also written down in the survey whenever possible. The size of the sample plot was adjusted to Landsat image's pixel size 30m*30m.

Table 3. Naforma Land Cover classification (MNRT 2015). Classes that are highlighted in green were included in survey.

No.	***Level 1 Class***	***Level 2 Class***
1	Forest	Humid montane
		Lowland
		Mangrove
		Plantation
2	Woodland	Closed
		Open
		With scattered cultivation
3	Bushland	Thicket
		Bushland
		With scattered cultivation
4	Grassland	Grassland
5	Cultivated land	Wooded crops
		Grain and other crops
6	Open land	Bare soil
		Rock outcrop
		Ice cap/snow
7	Water features	Ocean
		Inland water
		Wetlands
8	Built-up areas	Settlements
		Mining

4.3.3. Collect Earth analysis

The method of collecting reference data based on visual interpretation of satellite images is highly dependent on the skills of the interpreter. To identify land cover classes, visual interpretation elements are often set to reduce subjectivity and to keep the result consistent **(**Lunetta & Lyon 2000: 7). In this study the emphasis was more in recognizing change events from high- and medium-resolution satellite images and related time-series than to classify land cover, so systematic interpretation elements were not used in the approach. Land cover classification was a secondary objective, and the land cover class was assessed in the least accurate classification scale. Change was interpreted based on several different resolution images available in Google Earth and Google Earth Engine. Related literature emphasizes the importance of high-quality reference data: reference data should have higher spatial resolution than target data, time gaps should be minimal, and the land cover categories used should be identifiable (Lunetta & Lyon 2000: 7). However, when the research is targeting to rural areas in Global South, reference data had to be determined by availability of data rather than guidelines set in the literature.

The Collect-survey that was built on previous phase was uploaded to Collect Earth, where visual interpretation of the land cover plots was performed via Google Earth and Google Earth Engine. The software automatically organizes the collected data into databases that can be managed later in the Collect database window. Each sample cell in the Google Earth was assigned with corresponding land cover class value from the most recent high-resolution satellite image and the possible changes in land cover were identified and their magnitude estimated a four-step scale (1—4). I had no previous knowledge about the algorithm classification result for each validation point during the interpretation, avoiding bias. Class values were determined based on the majority rule, choosing the dominant land use class in each cell. Once the data for each point was filled in and sent, the results could be analysed through Calc and Saiku Analysis Tools, which are part of the Open Foris toolbox, or downloaded for other uses. In this study, final reference data set was downloaded as csv-file and opened in ArcGIS 10.5.

4.4. LandTrendr vegetation change analysis

4.4.1. LandTrendr algorithm

The Landsat-based detection of trends in disturbance and recovery (LandTrendr) is a semi-automatic algorithm developed by Kennedy et al. (2010). The algorithm uses temporal segmentation of spectral trajectories to extract change information from Landsat time series. LandTrendr algorithm fits temporal trajectories from spectral variations to each pixel of annual image time series of Landsat images. Annual spectral values of a certain plot create a response curve, which is usually defined as spectral trajectory (Figure 6). These trajectories provide information about the specific events in time series which can be interpreted as a forest cover change. Vertices are turning points in spectral trajectories and segments are linear trends between two vertices that represent the change process occurring between vertices, which can be interpreted as stable, disturbance or regrowth in the vegetation change analysis. Direction of the vector indicates the nature of change i.e. if the values are increasing or decreasing over time. The angle and the length of the vector provides information about the magnitude and rate of the change. The LandTrendr time series algorithm produces seven per-pixel attributes that are calculated from the time series segmentation: the starting year of the change, the end year of the change, magnitude of change, the value before the change, the value after the change, duration of the change, and the change rate (Kennedy et al. 2018). In this study, the magnitude of change and the change rate were selected as the attributes under interest.

The input values for LandTrendr analysis are usually spectral bands or spectral indices derived from them, but other metrics that represent yearly behavior are possible too (Kennedy et al. 2010). Image data is reduced to a single band or spectral index so that there is only one value for every year. Temporal trajectory of data values is tracked by using linear segments that form a sequence, in which segments are bounded by vertices, which can be interpreted as breakpoints. After maximum number of segments is identified, linear segments are chronologically fit to the spectral metric values. At every iteration step fit statistics are calculated and compared against a threshold, which is defined by user (Cohen et al. 2018). The model is evaluated by measuring the residual between the fitted segmentation and the actual measurement.

The pre-processing steps that are required before segmentation are divided in three categories in LandTrendr script and are profoundly explained in the original article (Kennedy et al. 2010). The first one is the traditional satellite image pre-processing, which includes geometric rectification, radiometric calibration, cloud, haze and shadow masking and calculation of surface reflectance values out of top of atmosphere values. Image quality and calibration are especially important when evaluating spectral differences in time series of images. Each pixel should match the same footprint

on the ground, and the channels of images that are being compared should be correctly registered together to diminish the radiometric effects (Adams & Gillespie 2006: 29). Main challenges in comparing images in time series are variability in brightness levels and atmospheric effects in each scene. In addition, when using imagery from several satellite detectors, differences between the detector responses has to be taken into account.

Secondly, the pixels representing the annual value must be identified (Cohen et al. 2018). There are several approaches for this phase: pixels can be selected according to derived metrics across multiple dates, like maximum or median value. Another good option is to use best-pixel approach, where the best pixels representing a single year are identified using different criteria like the closeness to date of interest or cloud scores. Thirdly, the selected pixels are stacked into an image mosaic and arranged for the trajectory segmentation.

After completing the pre-processing steps, pixel-based time series of data values are entered to the actual LandTrend algorithm (Kennedy et al. 2010). Inside the algorithm the data is further processed in two steps. First step is the despiking of the algorithm to remove anomalous vertices with user-defined parameters. Secondly, when pixel values are missing or inappropriate due to data gaps caused by the cloud cover or other distraction, pixels are flagged as missing values and thus ignored in the vertex identification. Once the segments are constructed these missing values will be filled using existing values from before or after the gap. Finally, a maximum number of segments is defined for trajectory construction and, if necessary, the segments with smallest angles are removed to simplify the initial temporal trajectory. Geometry calculations on these line segments enables user to distinguish information about distinct spectral turning points.

Compared to other approaches, this method provides the most detailed information about the change occurred (Banskota et al. 2014). However, defining appropriate thresholds to capture the phenomena accurately in relation to the variation caused by the confusing factors mentioned earlier is challenging. Another practical challenge is the validation of the historical changes, as the ground truth data of these areas might not be available or even exist.

The LandTrendr algorithm has been applied in a variety of applications (see e.g. Fragal et al. 2016; Schneibel et al. 2017; Cohen et al. 2018; Yang et al. 2018), but hasn’t reached the wider remote sensing community before for three main reasons: Firstly, building a time series image stacks has previously required a significant time investment due to the data management and pre-processing steps that are required for the code to run properly. Secondly, the processing language in which the code was written has been an obstacle to many, due to its proprietary nature and high learning curve. Finally, the extensive amount of data (series of satellite images where each pixel represents its own temporal signature) and the algorithm itself require great amount of computing efficiency and time

(Kennedy et al. 2018). Application to Google Earth Engine has set aside many of these barriers that prevented the utilization of the algorithm before.

As a cloud-based, geoprocessing platform with extensive catalogue of multitemporal satellite imagery together with built in powerful application programming interface (API), Google Earth Engine provides an excellent capacity for implementing the LandTrend algorithm (Kennedy et al. 2018). Assessing all the historical Landsat data in analysis-ready surface reflectance products, as well as the associated cloud, shadow, haze, and aerosol masks, significantly saves effort on pre-processing steps allowing the user to focus in change analysis. Thus, the Google Earth Engine platform for LandTrendr allows the focus on the main temporal segmentation algorithm by simplifying the extensive pre-processing steps. According to Kennedy et al. (2018), when compared with the original LandTrendr code (LT-IDL) it was stated that the practical impact of any modifications taking place while translating the code into Earth Engine were minimal, and the code can be safely applied by the broader user community.

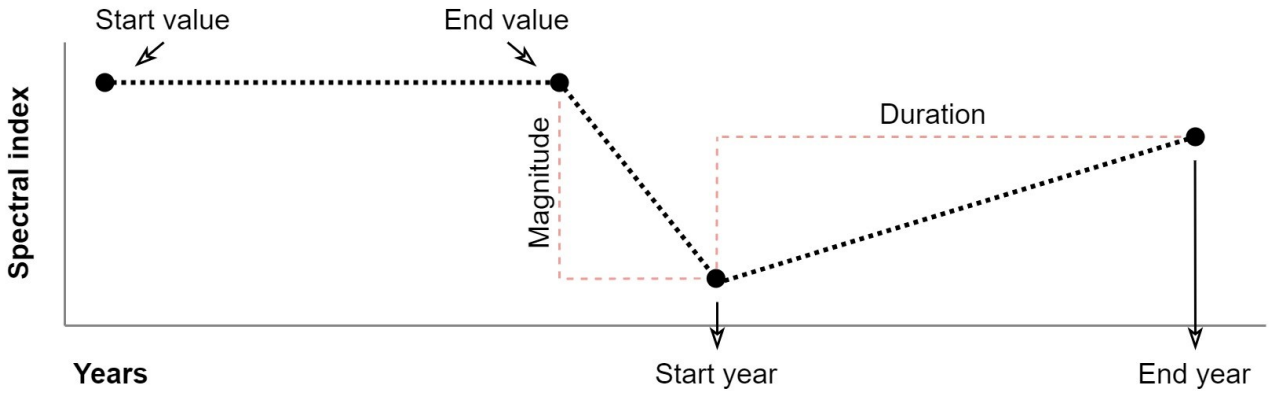


Figure 6. LandTrendr segmentation attributes (modified from Kennedy et al. 2018).

4.4.2. Optimising LandTrendr for the study area

Several approaches use land cover mask to limit the coverage of the analysis only to plots with forest/woodland vegetation. After a deep consideration, this approach was not applied in this study as classification of the heterogenic woodland landscape covering the whole study area would have resulted in further uncertainties about the validity of this classification. In addition, study area was selected to this study based on its typical vegetation cover, which represents mostly Miombo woodland and agricultural area. Decrease of vegetation cover in this area most presumably represents deforestation, as the LandTrendr algorithm can distinguish the decline in permanent

vegetation cover from the inter-annual variations in croplands and fields. However, the analysis was directed to forested areas by setting a threshold for preliminary value for each pixel to be included in the analysis. This approach is explained in detail in chapter 4.4.3.

One of the most important considerations in running the LandTrendr algorithm is the selection of suitable spectral band or index. Cohen et al. (2018) tested the capacity of different spectral indices to detect different change types when implemented to LandTrendr algorithm. These changes included abrupt forest clearing for agricultural, industrial, or residential purposes, slow degradation of forest due to the partial wood harvesting, land degradation, insects, or changes in climate conditions. According to their findings, TCW has the best performance when objective is to detect various types of changes. Moreover, using TC components is widely investigated and various time-series approaches have shown success with TCW or TCG in studying deforestation due to its ability to track overall moisture content (Kennedy et al. 2010; Frazier et al. 2015; White et al. 2017), so it was chosen to be applied in this study as well. Same research indicated that NBR can also accurately detect several type of forest disturbances, not only fire, and had high scores in overall performance. In LandTrendr forest disturbance mapping conducted by Kennedy et al. (2012), NBR was the most sensitive index to capture of disturbance events but was also more vulnerable to noise and therefore had slightly lower overall performance than TCW. NBR is widely used in Landsat time-series and applied in LandTrendr algorithm and thus has proven its power at characterizing forest dynamics in different forest areas (White et al. 2017; Schneibel 2017; Cohen et al 2018). Consequently, NBR was chosen to the study due to its overall performance and special sensitivity to detect fire disturbances.

NDVI was selected as it is still regularly used for forest disturbance detection and frequently in LandTrendr-based analysis (see e.g. Fragal 2016; Yang 2018), despite its reputation as “old-fashioned” vegetation index. NDVI was also thought to complement the selection of TCW and NBR as NDVI is based on reflection values from near-infrared channel, whereas TCW relies on shortwave-infrared reflection channel and NBR to combination of both. Where NBR is more sensitive to moisture content of vegetation, NDVI is sensitive to more subtle changes in forest health, such as disturbances which will not necessarily decrease the canopy cover. Another rationale for selecting these indices was that Kennedy et al. (2010) had profoundly examined these indices and therefore provided recommendations for parametrisation.

Without modification of the change parameters, LandTrendr identifies disturbances that are covering almost the whole study area. This could be caused by actual disturbances, since most of the area has had some of incidences of human disturbance, fire, drought, or insects during the 31-year time period. The occasional lower values in time series might also be a result of atmospheric or radiometric effects that distort the actual surface reflectance. Thus, it became crucial to find suitable

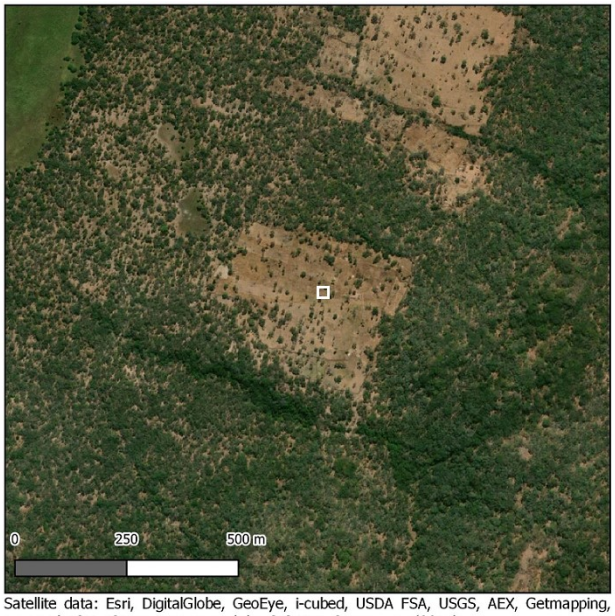
segmentation parameters as well as the most accurate change parameters that would minimize the occurrence of these errors without too much delimiting out the actual change events.

Detecting deforestation from time series analysis can be straightforward when the trend thresholds are determined accordingly (Cohen et al 2018). To calibrate the algorithm, the use of accurate reference data is crucial. By allowing any arbitrary segmentation there is a risk of over- and under-fitting in Landtrendr trajectory segmentation. Therefore, a series of control parameters and filtering steps has been included in the algorithm, which are designed to reduce overfitting while still capturing the desired features of the trajectories (Kennedy et al. 2010). The LandTrendr function takes in 9 arguments: the annual image collection and 8 control parameters which adjust the spectral-temporal segmentation process. These parameters, and the implications of adjusting them, have been fully described in original LandTrendr paper (Kennedy et al. 2010). The parameters were mostly adjusted according to recommendations from Kennedy et al. (2010) and Cohen et al. (2018) and tested with small sample set of 30 previously well identified change events. These included samples from turning the woodland into agricultural fields (12), fire events (5), general degradation (5), as well as tree felling (8) and clear-cut events (2) (Table 4). According to the reference data collected in previous study phase, these are the most common change types in the study area. While other change types exist and parameter optimization could itself be another research topic, in this study the parameter adjustment was mostly based on the literature and the sample set was utilised only to make minor adjustments. Figure 7 represents the example of a plot that was cleared for agriculture in 2010 and the index-specific LandTrendr segmentation trajectories from this plot during 1987—2018 with final parameter settings. Each of these indices can identify the abrupt clearance for agriculture that took place in 2010.

Table 4. The test sample set from different disturbance types for brief testing of LandTrendr functionality.

	Fire	**Agriculture**	**Tree felling**	**Clear cut**	**Degradation**
Count	5	10	8	2	5

To minimize external effects, it is essential to select seasonally continuous imagery. It is often recommended to use imagery representing the peak growing season conditions for vegetation change mapping (Verbesselt et al. 2010). However, the growing season of the study area happens during the rainy season, which has considerable effects to the cloudiness of the imagery. To reduce the effects of clouds seasonal variations, images were selected annually from dry season between June and September.



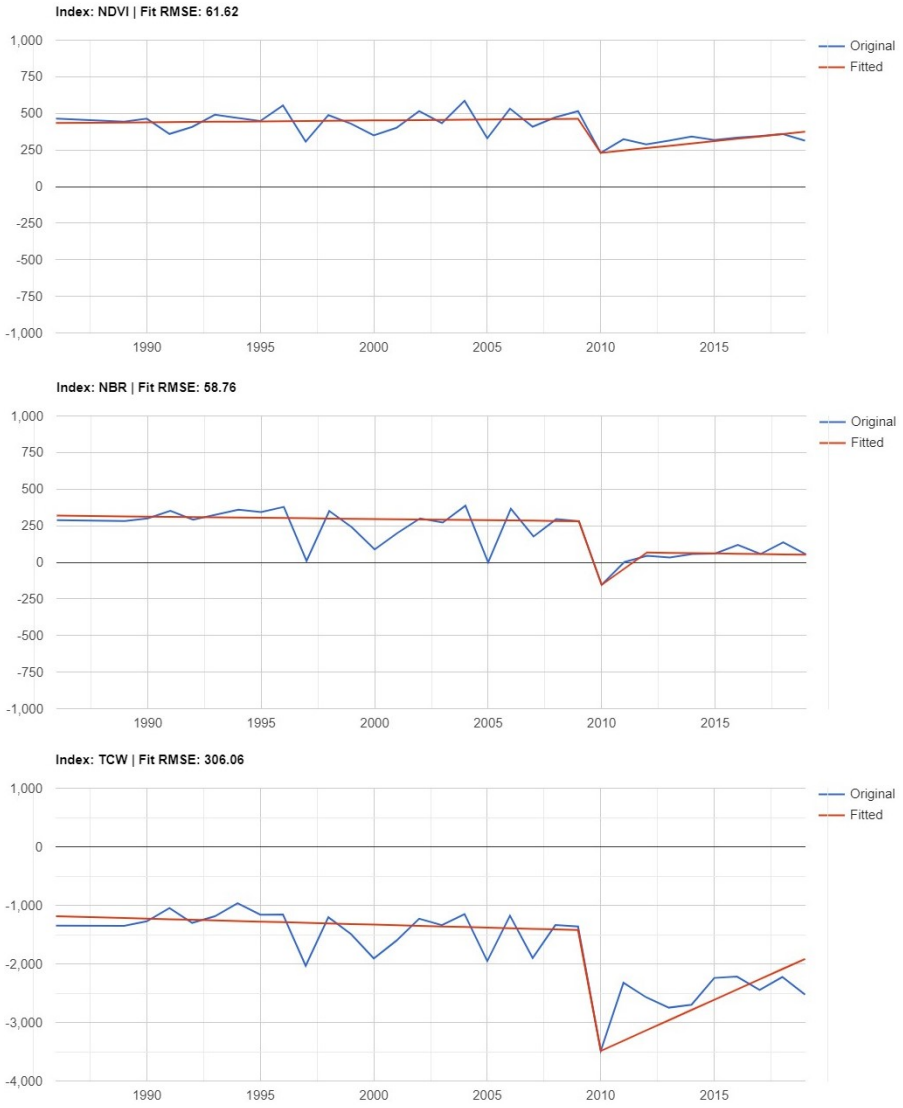


Figure 7. LandTrendr segmentation of NDVI, NBR and TCW time series for woodland that has turned into agricultural field in 2010. Time series plots created in Google Earth Engine.

4.4.3. Parameter settings for time series segmentation and change detection

Effects of changing the LandTrendr segmentation parameters were studied in depth by Kennedy et al. (2010) and Cohen et al. (2018), where the impacts of segmentation parameters on LandTrendr fitting was evaluated and compared between different Landsat bands and vegetation indices derived from them. Full list and descriptions of default parameters can be explored in Table 5. Selection of potential vertices through both angel-based criteria and regression is crucial, which is why the *vertex count overshoot* parameter is set to 3 by default (Kennedy et al. 2010). This approach detects vertices as a separate phase from the actual fitting. Parameter settings that favour capture of disturbance helps to decrease the false negatives, in the cost of higher number of false positives, making it less beneficial for overall matching of trajectory shape, which reduces the capability of capturing stable periods, or timing of the recovery. Most disturbance is captured with parameter settings that allow higher number of segments (max_segments), allow fast recovery (recovery threshold) and permit the robustness of fit by setting the p-value higher (pval). Parameters were set to moderate disturbance sensitivity according the recommendations of Kennedy et al. (2010) and Cohen et al. (2018). The algorithm was run with the segmentation parameters listed in Table 6.

Table 5. LandTrendr parameters (Kennedy et al. 2018).

Parameter	Type	Default	Definition
maxSegments	Integer		Maximum number of segments to be fitted on the time series
spikeThreshold	Float	0.9	Threshold for dampening the spikes (1.0 means no dampening)
vertexCountOvershoot	Integer	3	The initial model can overshoot the maxSegments + 1 verticles by this amount. Later, it will be pruned down to maxSegments + 1
preventOneYearRecovery	Boolean	false	Prevent segments that represent one year recoveries
recoveryThreshold	Float	0.25	If a segment has a recovery rate faster than 1/recoveryThreshold (in years), then the segment is disallowed
pvalThreshold	Float	0.1	If the p-value of the fitted model exceeds this threshold, then the current mode is discarded and another one is fitted using the Levenberg-Marquardt optimizer
bestModelProportion	Float	1.25	Takes the model with most verticles that has a p-value that is at most this proportion away from the model with lowest p-value
minObservationsNeeded	Integer	6	Min observations needed to perform output fitting
timeSeries	ImageCollection		Collection from which to extract trends (it's assumed that each image in the collection represents one year. The first band is used to find breakpoints, and all subsequent bands are fitted using those breakpoints

Table 6. Segmentation parameter settings for LandTrendr algorithm

maxSegments	6
spikeThreshold	0.75
vertexCountOvershoot	3
preventOneYearRecovery	true
recoveryThreshold	0.25
pvalThreshold	0.05
bestModelProportion	0.75
minObservationsNeeded	6

In order to reduce commission errors throughout the disturbance detection, the time series trajectories were filtered according to the relative percent of canopy cover change magnitude following the example of Griffiths & Hostert (2015), who applied this method in LandTrendr analysis. Filtering was performed by linearly scaling the values of three indices from areas of no canopy cover to 100 % canopy cover. This was accomplished by taking in comparison again 10 plots representing bare soil or minor undergrowth, 10 plots representing 50 % canopy cover and 10 plots with 100 % cover from dispersed plots in the study area for each spectral index and calculating average values from the index values. It should be noted that this is a crude approximation that enable the calculation of relative disturbance magnitude and facilitates the distinction of actual change values from external noise in spectral trajectory (Griffiths & Hostert 2015: 312). These estimations were then used in filtering of the pixels that were taken into the LandTrendr time series segmentation so that only areas with pixel values higher than those that represent 30 % tree cover, or values representing similar vegetation conditions, were taken into the analysis (Table 7).

It is expected that thresholding the disturbance detection by magnitude of change could improve accuracy by reducing false positives, but it also predefines the type of allowable change by limiting the change event only to those with more significant magnitude. As international agreements are including forest degradation to the discussion about climate change (Kissinger et al. 2012), the importance of detecting low magnitude disturbances with remote sensing methods is underlined (McDowell et al. 2015). However, when targeting to the lower magnitude disturbances in time series analysis, high omission rates are likely. According to previous studies on LandTrendr functionality (see e.g. Kennedy et al. 2010; Fragal et al. 2016; Yang et al. 2018; Cohen et al. 2018), false positives are more common source of inaccuracy in LandTrendr analysis than false negatives. To diminish the number of false positives, change magnitude was filtered to cover only changes equal to 10 % or higher decrease in canopy cover based on the same scaling of canopy cover that was used in determination of prevalues for the change algorithm (Table 7).

With this threshold value, the number of false positives were expected to decline without losing the information about subtler forest degradation trends, as 10 % trend variation could well be originated from external noise in time series or natural variation in vegetation condition (Healey et al. 2018).

Table 7. Change thresholds to use in LandTrendr parameters.

Index	Average value for 0 % canopy cover	Average value for 100 % canopy cover	Average value for 30 % canopy cover	Relative magnitude value corresponding to 10 % canopy cover change
NBR	-80	520	100	60
NDVI	110	590	254	48
TCW	-3500	-800	-2690	270

The change parameters were same for each spectral index, except prevalue and change magnitude filter (Table 8). Prevalue was moderated for each index separately according to the value corresponding to scaled 30 % canopy cover. Change magnitude filter was defined by using the same scale but setting the minimum magnitude value to values that correspond to 10 % decrease in canopy cover reflectance calculated separately for NBR, NDVI and TCW.

Table 8. Change mapping parameters for LandTrendr

delta	'loss',
sort	'greatest',
year	{checked:true, start:1987, end:2018},
month	{checked:true, start:6, end:9},
magnitude	{checked:true, value:1, operator:'>', dsnr:true},
duration	{checked:false}
prevalue	{checked:true, value: x, operator:'>'},
mmu	{checked:false}

Testing the algorithm with several different parameter settings brought in light some weaknesses that were also noted in literature. The segmentation approach identifies change in totally relative sense, which gives a disproportionate number of emphasis on the first and last years of the trajectory (Cohen et al. 2018). Originally the algorithm was run with the longest possible time scale, from 1984 until 2018. However, it was found that there were fewer high-quality Landsat images available from

years 1984—1989, which resulted in gap years and errors in analysis. It was noticed that the algorithm repeatedly and erroneously recognised the year 1985 as the year of change. This was fixed by taking the year 1984 and 1985 in the calculations of the algorithm but starting the change analysis only in year 1987.

4.5. Land cover change maps

With the parameter settings defined earlier, the Landtrendr algorithm was run separately for NBR, NDVI and TCW indices. The output layers were downloaded and opened in ArcGIS software for further processing.

The raster layers were reclassified in four quartile classes, based on change magnitude value. This generalization had to be done to be able to compare the analysis results between different indices and the reference data. Cell statistics and histograms were calculated from three raster layers representing the index-based change magnitude values from LandTrendr analysis. Null values indicate pixels, where index values had remained approximately same during the 31-year time series period. The classified raster layers were then analysed further by tracing pixels which belonged in same magnitude category by applying CombinatorialOr and CombinatorialAnd -tools in ArcGIS 10.5. The question in mind was could these three maps be combined in a way that takes advantage of the different ways how each spectral index identify the change, to provide overall more accurate disturbance map? Numerous sophisticated methods exist for fusing several change maps or the original time series segmentation (Healey et al. 2018). However, in this thesis the focus has been in testing the functionality of LandTrendr in Miombo environments and compare the usability of different spectral indices. Thus, the combination of the change maps that was performed from a heuristic premise, as the more advanced approach would necessitate a broader knowledge of statistical methods and overall a deeper study on the issue, which would not be sensible for the purpose of this paper.

First, those pixels that were included into the same magnitude category were found and categorized as high probability change pixels, assuming that if multiple indices detect the change in the same manner, it is more likely that the predicted change matches the observed change on the ground. As an example, pixels that were classified in magnitude category 4 by all three indices, were categorized as highest probability change pixels. High probability was also labelled for pixels where at least two indices had the same magnitude value, or all the three indices had found change but had disagreement on the category. If several magnitude categories were found in the same pixel, average value was calculated for those pixels. Moderate probability class was labelled to pixels

where two indices detected change of some sort and the average value was calculated for those pixels as well. Highest uncertainty was with pixels, where only one index out of three detected change, and these pixels were excluded from the final map. Similar, yet much more advanced ensemble approach was developed by Healey et al. (2018), in which annual disturbance map series derived from different time series algorithms were stacked to create a single disturbance map.

One of the output bands from the LandTrendr algorithm includes values that represent rate of the spectral change occurred. This rate is calculated as magnitude of the negative trajectory segment divided by duration of this segment in years. Using this band, changes were categorized as gradual and abrupt change. The range between the minimum and maximum change values is different for each index, so categorizing had to be done separately for all the indices. The change thresholds created earlier in LandTrendr parametrisation were used for this purpose. Values representing relative magnitude value corresponding to 10 % vegetation cover change were applied to the calculation. All the pixels where change magnitude was relatively matched to 10 % or less canopy cover decrease per year were counted as gradual change, and higher than 10 % annual change rate pixels were classified as abrupt change according to the rate value. For NBR the rate threshold was then 60, for NDVI 48 and TCW 270 (Table 7 earlier). The rate values from original index-based output layers were generalized based on majority rule after classification to gradual and abrupt change. Finally, the values were filtered only to cover the pixels that were earlier labelled as disturbed in high or moderate probability.

4.6. Accuracy assessment

The basic idea in accuracy assessment approaches is to provide information on the credibility of the thematic map for the map user (Jones & Vaughan 2010: 267). The visually interpreted disturbance in reference data is compared against the results produced by the automatic algorithm. Any automatic classification from remote sensing data contains some errors from various steps in satellite image processing, calibration or change detection itself, not to mention errors that result from the lack or poor quality of data.

In order to assess the change detection accuracy, the outputs of the algorithm were compared against the ground reference information. To calculate accuracy statistics, the magnitude values recorded by LandTrendr were recalculated by relative magnitude in four quartiles to match the reference data (1. low magnitude, 2. moderate 3. high and 4. very high change magnitude class) following the example of Kennedy et al. (2012).

The error matrix table was built manually in Microsoft Excel and cross-tabulation was employed to calculate user's accuracy, producer's accuracy and overall accuracy for all the change maps where pixels were classified into four change magnitude classes and to one "no change" class.

Overall accuracy reveals the total proportion of reference sites that were mapped correctly, with 100% accuracy being a perfect classification where all reference sites were classified correctly (Jones & Vaughan 2010: 267). It is defined as a total of correctly classified samples divided by the total number of samples. Overall accuracy is easy to calculate, but it only provides basic accuracy information for the map user and producer. The diagonal elements in error matrices represent the areas that were correctly classified. Overall accuracy was calculated by dividing the number of correctly classified sites by the total number of reference plots.

Classification error occurs when a pixel is not classified into a right category (Lunetta & Lyon 2000: 3–4). When pixel is incorrectly included in the evaluated category, the error is preferred as Error of commission. On the other hand, if the pixel is left out of the category being evaluated, the classification error is called as Error of omission. Thus, an error of omission in one category will be considered as an error in commission in another. These error types are used to further calculate user's and producer's accuracies. User's accuracy is the accuracy from the map user's point of view. It shows how often the mapped class represent the actual situation on the ground. This is referred to as reliability of the map. User's accuracy reveals the so called "false positives", where pixels are incorrectly classified as a recognized class when they should have been classified as something else. The User's Accuracy is calculated based on Commission Error, where

User's Accuracy = 1 - Commission Error

Producer's accuracy is the map accuracy from the map producer's point of view. In this case it indicates the probability in which the real change events on the ground are correctly shown on the classified map. In other words, it is the number of accurately classified reference sites divided by the total number of reference sites for that class. The Producer's Accuracy is calculated based on Commission Error, where

Producer's Accuracy = 1—Omission Error

The user and producer accuracy provide necessary information for evaluation of the analysis and classification results. These accuracy percentages are not typically similar but indicate different aspects of the classification (Jones & Vaughan 2010: 267—269).

5. Results

5.1. Change detection by spectral indices

5.1.1. Index susceptibility for detecting negative trends in time series

Despite the modification of the normal distribution of change magnitude values (by cutting the values representing 10 % of canopy loss), every histogram illustrating the distribution of change magnitude values are skewed to the right (Figures 8—10). This indicates that the data has some large values that drive the mean upward but do not significantly affect the median, which is the middle number in a given sequence of numbers i.e. the value in the center of the distribution (Nummenmaa et al. 2014). In TCW derived change magnitude map, remote values are more common than maps derived from NBR and NDVI, suggesting instability in spectral values derived from TCW-based analysis. However, these values were not considered as outliers as it is possible that LandTrendr produces extreme change magnitude values as a summary from a long time series, which might include high variance in spectral values throughout the change trajectory (Kennedy et al. 2018). Statistics also reveal the similar distribution of change magnitude values derived from LandTrendr analysis based on NDVI and NBR indices. Equal shapes of histograms suggest that NDVI and NBR are detecting change in similar manner when studying distribution of values in this context. However, in NBR-based analysis, more pixels have been labelled as changed, demonstrating the higher sensitivity to change detection (Table 9). NDVI recorded disturbance in 46 % and NBR 56 % of all the pixels in the study area.

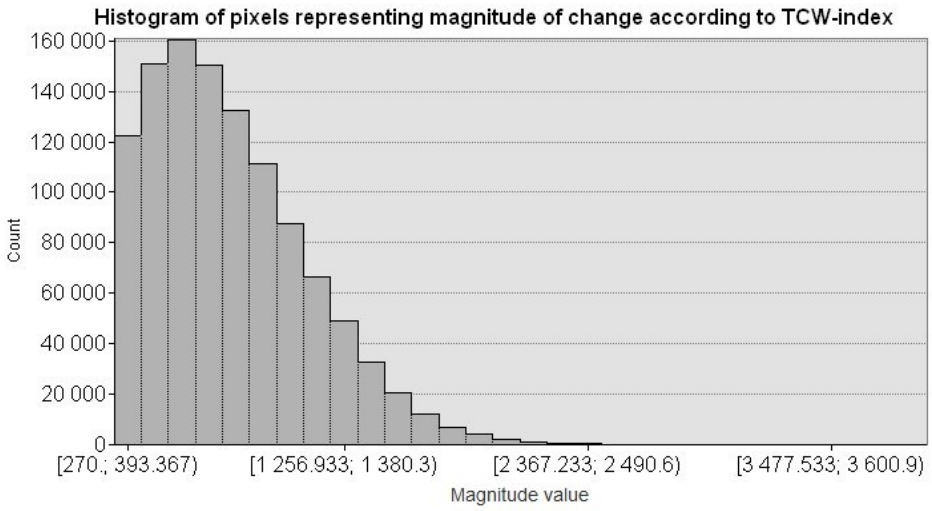


Figure 8. The distribution of change magnitude values when applying TCW to LandTrendr analysis in the study area. Values start from 270, which represents 10 % decrease in canopy cover.

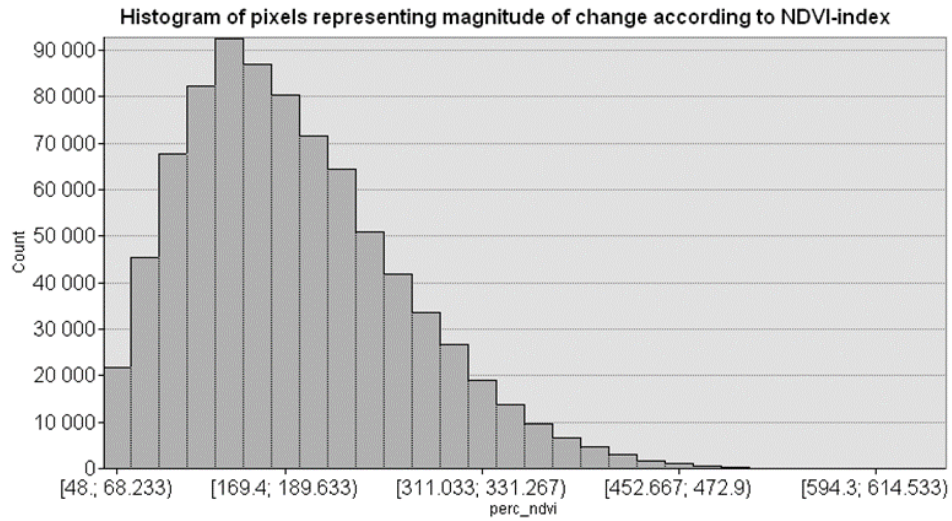


Figure 9. The distribution of change magnitude values when applying NDVI to LandTrendr analysis in the study area. Values start from 48, which represents the 10 % decrease in canopy cover.

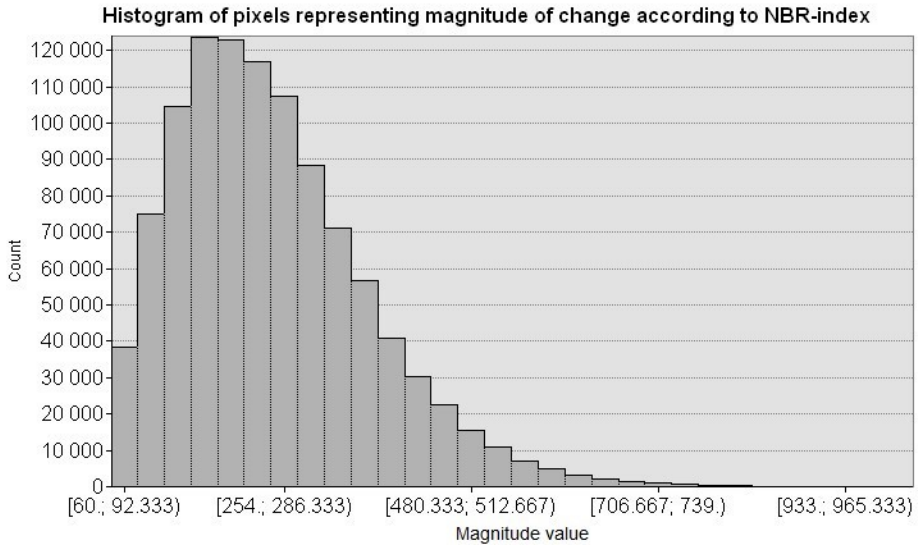


Figure 10. The distribution of change magnitude values when applying NBR to LandTrendr analysis in the study area. Values start from 60, which represents the 10 % decrease in canopy cover.

Table 9. Statistics of individual spectral index-based disturbance layers.

Layer	***MIN***	***MAX***	***MEAN***	***MEDIAN***	***SDT***	***SUM***	***% of disturbed pixels***
NBR	60	1030	255,17	236	116,11	1 074 007	56 %
NDVI	48	655	185,36	172	78,29	837 595	46 %
TCW	270	3971	800,04	709	358,12	1 185 431	66 %

TCW-based histogram of change magnitude values shows more independent shape and greater amount of changed pixels (66 %, Table 9), even though proportionally large amount of lower change magnitude values has been left outside of the analysis. TCW also produces greater scale in change magnitude values as the difference between no vegetation values and 100 % vegetation is significantly higher than in NBR and NDVI, which have more related distribution. The number of disturbed pixels is highest in TCW-derived disturbance map, followed by NBR (56 %) and NDVI (46 %).

The correlation between change magnitude raster layers can be studied in tables 10 and 11. The calculation was done for raster layers including “no change” areas (nulls) and for layers where nulls were not included into the calculation. Positive correlation between the indices can be found, however it is only moderate between NBR and NDVI. TCW and NBR correlate little less, and only a weak correlation is found between NDVI and TCW. Including the “no change” areas increased the correlation between all the indices, showing more consistent result in areas where no change has occurred. In both cases, NDVI and NBR correlated more strongly with each other than with TCW. This again indicates similar capacity for change detection between these two indices.

Table 10. The correlation matrix without 0 values (only disturbed pixels).

Layer	*NBR*	*NDVI*	*TCW*
NBR	1,00	0,45	0,39
NDVI	0,45	1,00	0,26
TCW	0,39	0,26	1,00

Table 11. Correlation matrix including 0 values.

Layer	*NBR*	*NDVI*	*TCW*
NBR	1,00	0,57	0,45
NDVI	0,57	1,00	0,42
TCW	0,45	0,42	1,00

Difference between largest and smallest value is demonstrated in the value range map (Figure 11). The map describes the variation of the magnitude values between different indices. Magnitude values have been reclassified to quartiles. Fewest variation can be found in areas that were outside of the human influence (top-right corner) and areas that were already under cultivation and

infrastructure (in the middle). Some of the greatest variation areas form long green/yellow areas in elongated patterns. These are wetland areas around rivers, where especially moisture conditions vary greatly depending on the climate conditions, causing confusion in indices that rely on SWIR-bands (TCW and NBR). Some dark blue/black areas in bottom left are resulting from clear cuts that have been taking place during 2000's and they can be identified by all three indices, causing only little variation between the results from the different spectral indices.

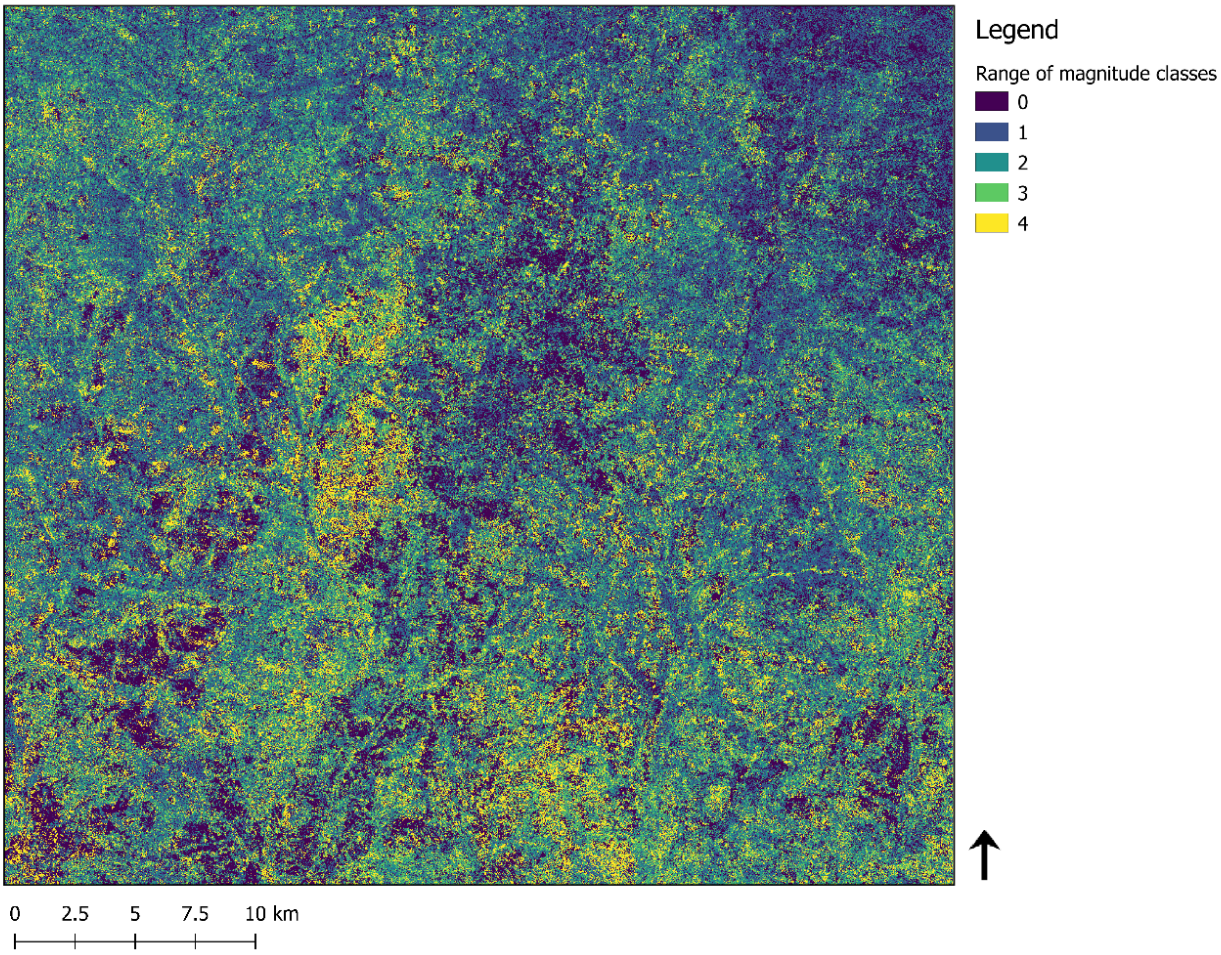


Figure 11. Range of magnitude values from the change raster layers derived from NDVI, NBR and TCW based LandTrendr analysis. Pixels have been reclassified to 4 quartiles according to their change magnitude value and the range map describes the range between the classified values in these maps. Null values representing "no change" are also included.

5.1.2. Identification of disturbance events

From 240 reference samples, where 180 were selected from areas where the algorithm predicted change and 60 from areas where there was no sign of a change, in total 142 pixels were classified as changed and 98 as stable pixels. From all pixels categorized as changed, clear change (magnitude classes 3 and 4) was detected in 77 pixels and in 65 pixels subtler changes (magnitude classes 1 and 2) was noticed at some point of the time series. In total 38 disturbed plots showed stable recovery after the change. In 12 sampling plots the vegetation cover was increasing, but they were classified as stable to keep the focus in negative change trajectory trends. The rest 86 plots did not show sign of changes or the changes were considered insignificant or temporal. Only in five samples the original land cover was other than forest or woodland.

Almost half (44 %) of the identified change events was clearance for agriculture or other land use (Figure 12). The second greatest disturbance was fire (26 %), followed by tree felling (18 %). 2 % of the identified disturbance was resulting from wetland variation. For 10 % of disturbed plots the change event was classified as other, as a result of uncertainties in interpretation.

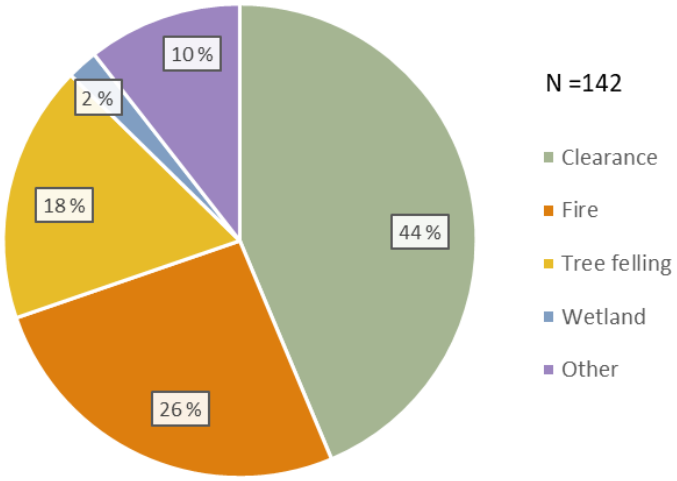


Figure 12. Relative proportion of identified disturbance events based on reference data.

81 % of significantly changed pixels was interpreted as human induced clearance for cultivation or settlement. 2 % was resulting from wetland variation (Figure 13). The rest 17 % of high magnitude change plots were severely disturbed by one or several fire events or gradually degraded due to natural or other effects. From 65 plots in lower change magnitude categories (1 and 2), most of the disturbance events were caused by fire (45 %) or tree felling (32 %). In the remaining plots the disturbance event was classified as wetland variation or other.

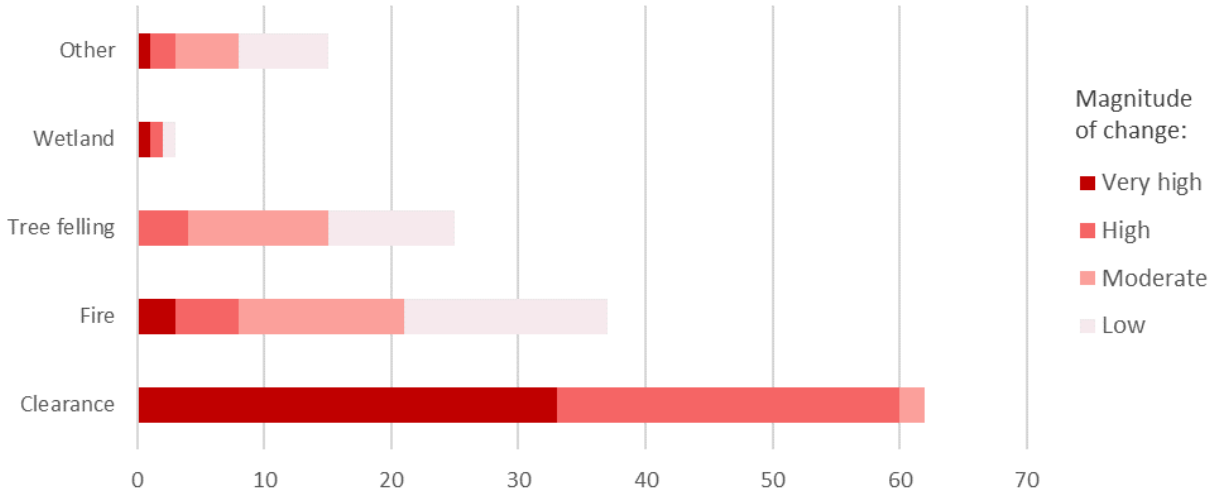


Figure 13. Change magnitudes in relation to identified disturbance events.

Of all plots that were identified as disturbed in visual interpretation process, TCW detected change in 97 %, NBR in 93 % and NDVI in 85 %. All indices had agreement on change in 71 % of the reference plots, although there were differences in change magnitude category. Where indices were not in agreement, they were complementing each other. From 77 severely disturbed plots only one could not be recognized by any index and in 60 % of plots all indices had agreement in change magnitude category. In four samples TCW missed the change but NBR and/or NDVI identified it. Subtler changes took place in 65 reference plots (magnitude classes 1 and 2). Within these classes, high amount of variance was found between the observed magnitude and predicted value produced by LandTrendr. Surprisingly all three indices produced change maps with similar accuracies when applied in LandTrendr analysis, although closer look revealed differences in the ways in which these indices recognized disturbances in all magnitude scales. It can be noticed that the higher the change category, the higher the probability for all three indices to identify change.

The comparison of error matrices reveals that TCW was detecting change most often, but it also produced high number of false positives (Tables 12—13, Figure 14). The overall accuracy of TCW based LandTrendr analysis, when using the quartile categories was 61 %. Both producer's and User's accuracies were significantly greater in high magnitude change classes.

Table 12. Error matrix for TCW derived disturbance map and all the change classes.

Classified change map (TCW)	**Reference data**							
	Class 1	*Class 2*	*Class 3*	*Class 4*	*No change*	***Total***	*User's accuracy*	*Commission error*
Class 1	13	9	2	0	9	33	0,39	0,61
Class 2	11	11	5	2	9	38	0,29	0,71
Class 3	6	6	25	7	5	49	0,51	0,49
Class 4	3	3	6	29	7	48	0,60	0,40
No change	1	2	1	0	68	72	0,94	0,06
Total	34	31	39	38	98	240	0	
Producer's accuracy	0,38	0,35	0,64	0,76	0,69	0	0,61	
Omission error	0,62	0,65	0,36	0,24	0,31			0,39

Table 13. Error matrix for changed/not-changed pixels in TCW derived disturbance map.

TCW	**Reference data**			
	Changed	Not changed	Total	User's accuracy
Changed	138	30	168	0,82
Not changed	4	68	72	0,94
Total	142	98	240	
Producer's accuracy	0,97	0,69		0,86

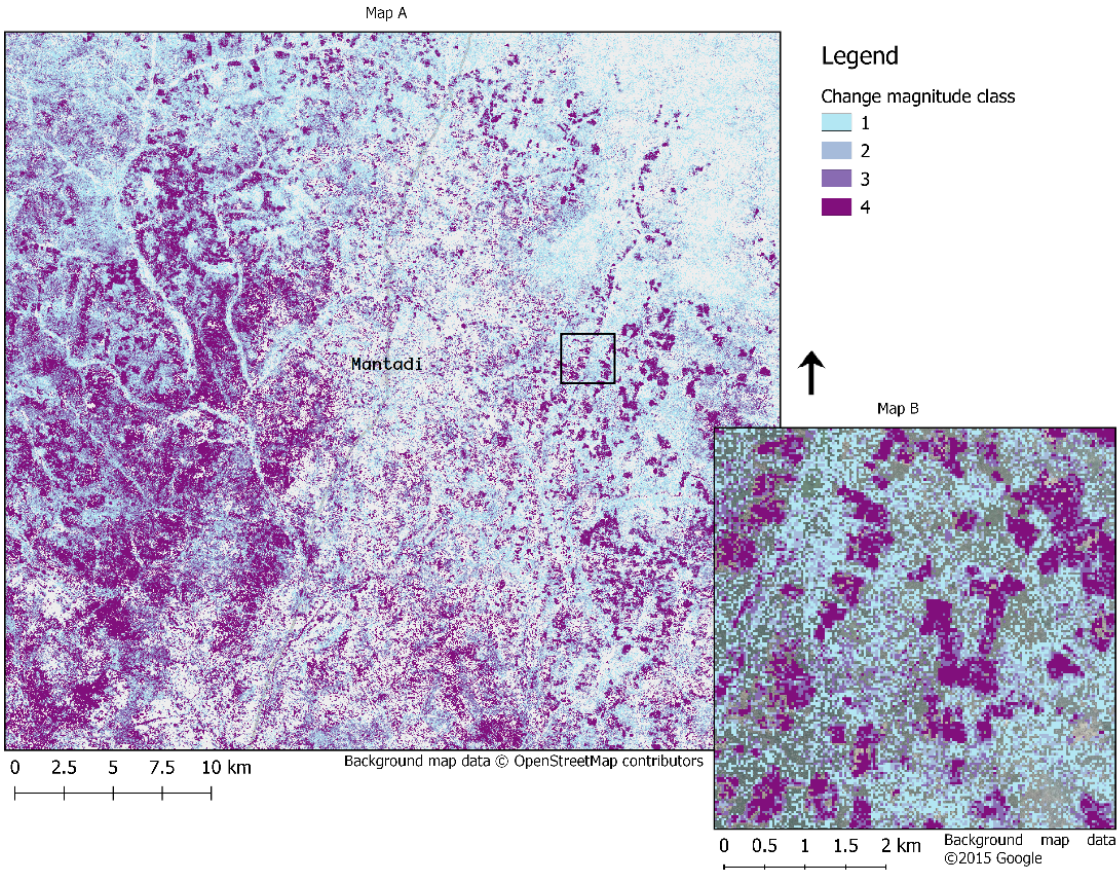


Figure 14. Disturbance map from LandTrendr analysis based on TCW. Map A represents the whole study area, whereas Map B reveals differences in closer look.

The overall accuracy of NDVI based disturbance map was 58 % (Table 14). Both producer's (83 %) and user's (79 %) accuracies were high in "No change" class. However, compared to other disturbance maps, in NDVI derived map commission error of "no change" class was higher (21 %), indicating that NDVI was not able to detect change as efficiently as NBR and TCW, in which the commission errors were 11 % and 6 %.

Overall accuracy was decreased due to the dispersion of disturbed pixels through all the magnitude classes. Especially in the first two magnitude classes there was a lot of confusion in the classification. Generally, errors of commission and omission were decreasing while the magnitude of change was increasing, demonstrating again that the most accurate results was achieved when the change magnitude was high.

In NDVI based disturbance map (Figure 15) the number of pixels identified as disturbed was lowest, with 38 % of reference pixels classified as no change. In comparison with NBR the percentage was 36,6 % and with TCW 30,8 %. This indicates that TCW is more sensitive to detect all kinds of disturbances, but also more prone to false positives, as the 31 % omission error in "no change" class

reveals. In comparison, in change maps derived from NDVI and NBR spectral indices, omission error for stable pixels was only 17 %. Overall NBR and NDVI produced more similar results compared to TCW, in which the number of accurately classified pixels were greater in high magnitude spectral change classes (3 and 4), but lower change magnitude had bigger error rates.

Table 14. Error matrix for NDVI derived disturbance map and all the change classes.

Classified change map (NDVI)	**Reference data**							
	Class 1	*Class 2*	*Class 3*	*Class 4*	*No change*	***Total***	*User's accuracy*	*Commission error*
Class 1	12	8	4	2	8	34	0,35	0,65
Class 2	10	12	9	6	5	42	0,29	0,71
Class 3	2	4	14	9	1	30	0,47	0,53
Class 4	0	2	8	18	1	29	0,62	0,38
No change	10	5	4	3	83	105	0,79	0,21
Total	34	31	39	38	98	240	0	
Producer's accuracy	0,35	0,39	0,36	0,47	0,85	0	0,58	
Omission error	0,65	0,61	0,64	0,53	0,15			0,42

Table 15. Error matrix for changed/not-changed pixels in NDVI deriver disturbance map.

NDVI	**Reference data**			
	Changed	Not changed	Total	User's accuracy
Changed	120	15	135	0,89
Not changed	22	83	105	0,79
Total	142	98	240	
Producer's accuracy	0,85	0,85		0,85

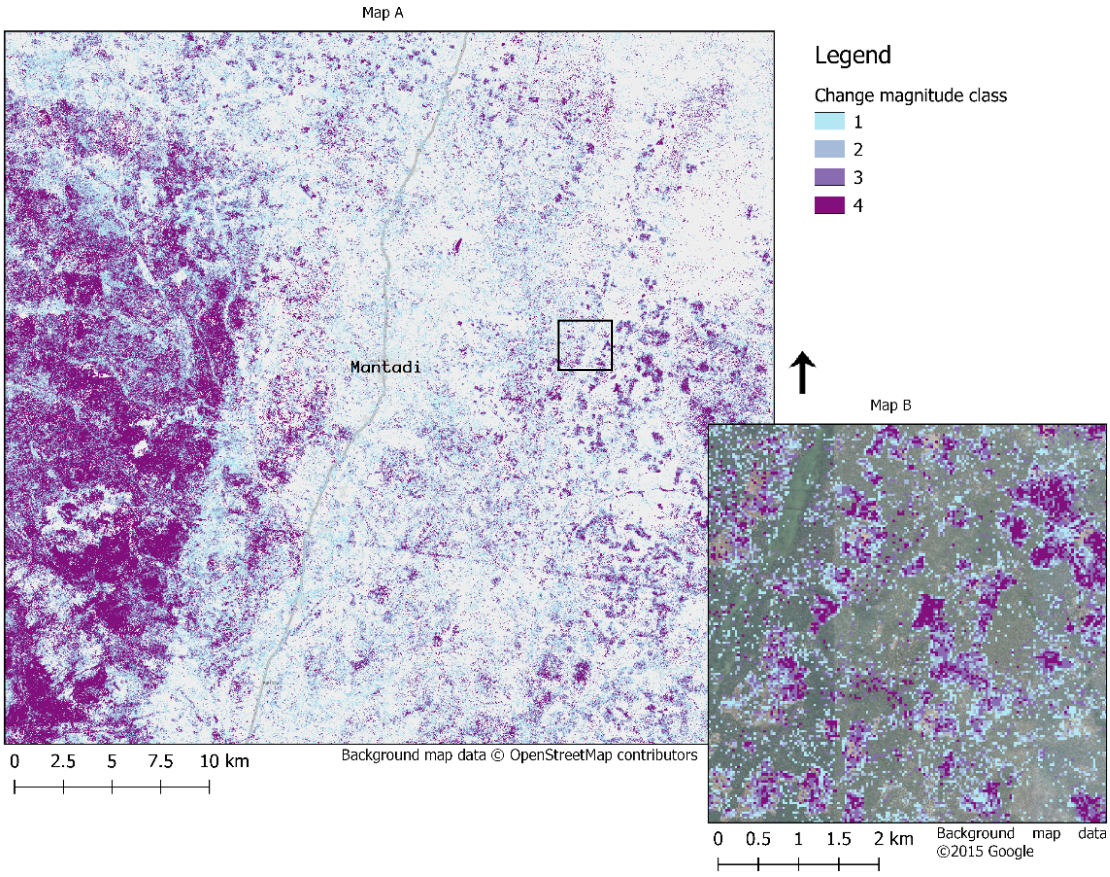


Figure 15. Disturbance map from LandTrendr analysis based on NDVI. Map A represents the whole study area, whereas Map B reveals differences in closer look.

The best spectral index to apply in LandTrendr analysis was found to be NBR, although the overall accuracy was only slightly higher (62 %) compared to TCW (61 %) (Table 16). When comparing the error matrices with only “changed” on “not change” categories, NBR manages to widen the gap with other indices (Table 17).

Despite of the minor difference to the overall accuracy of TCW based analysis, comparing the class-specific accuracies reveals NBR index’s more stable capacity to correctly detect all kinds of changes, and it was not as sensitive to false change signals than TCW.

Table 16. Error matrix for NBR derived disturbance map and all the change classes.

Classified change map (NBR)	**Reference data**							
	Class 1	*Class 2*	*Class 3*	*Class 4*	*No change*	***Total***	*User's accuracy*	*Commission error*
Class 1	15	11	5	0	11	42	0,36	0,64
Class 2	9	13	8	3	2	35	0,37	0,63
Class 3	3	3	18	11	3	38	0,47	0,53
Class 4	2	2	7	22	1	34	0,65	0,35
No change	5	2	1	2	81	91	0,89	0,11
Total	34	31	39	38	98	240		
Producer's accuracy	0,44	0,42	0,46	0,58	0,83		0,62	
Omission error	0,56	0,58	0,54	0,42	0,17			0,38

Table 17. Error matrix for changed/not-changed pixels in NBR derive disturbance map.

NBR	**Reference data**			
	Changed	Not changed	Total	User's accuracy
Changed	132	17	149	0,89
Not changed	10	81	91	0,89
Total	142	98	240	
Producer's accuracy	0,93	0,83		0,89

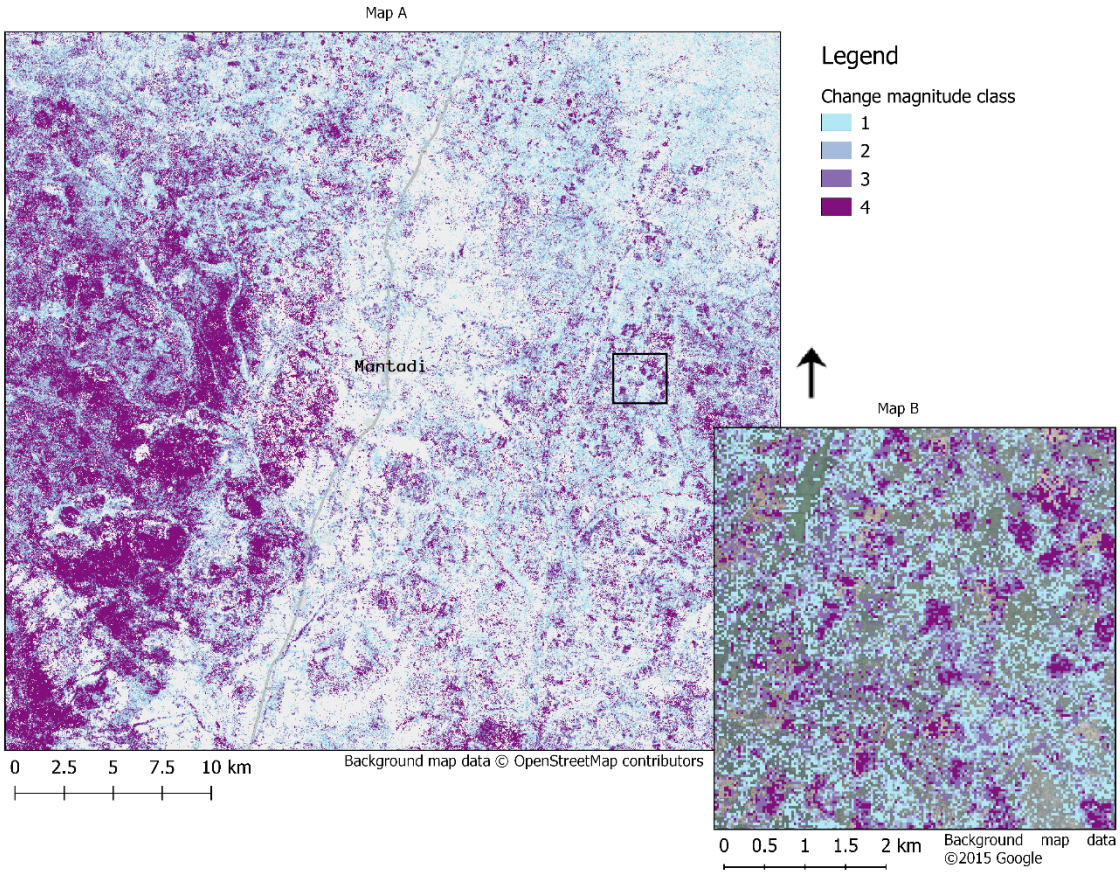


Figure 16. Disturbance map based on NBR. Map A represents the whole study area, whereas Map B reveals differences in closer look.

Comparing change classes of the reference data separately, it can be noted that highest accuracy between observed and predicted disturbance was in very high magnitude change, which has the highest producer's accuracy of all the change classes in all indices.

When applying NBR to LandTrendr analysis, 48 % of the identified disturbance pixels were consistent with the magnitude category that was defined in reference plots. Applying TCW, 54 % of the identified disturbance plots we in the same change magnitude category as reference plots were, out of which in the highest magnitude category omission error was only 24 %, compared to the lowest magnitude category's error of 62 %. In NDVI derived disturbance map only 39 % of disturbed pixels were in the right magnitude category. These results do not necessarily reflect the actual correlation between observed and predicted disturbance but are also affected by the dispersion of different change types in reference sample data. Great number of pixels were labelled as stable during the reference data collection. However, many of these no change-plots were misinterpreted as changed by some of the indices. These observations are called as "false positives", indicating the algorithm gives positives values in areas where they should not occur. It became clear, that even though TCW performed strongly in recognizing all types of change, its

sensitivity to variation in moisture conditions retaliated as high number of false positives. Producer's accuracy for no change areas was only 69 % for TCW, 83 % for NBR 85 % for NDVI, showing that NDVI was least sensitive to false signals.

5.2. Regional distribution and intensity of the forest change between 1987—2018

The combination of the three index-based change raster layers, that were produced by the LandTrendr algorithm, was used to calculate final area of deforestation and forest degradation on the study area (Figures 17—18). Considering both high and moderate probability, 56 % of the total study area has been under disturbance during 1987—2018. 26,5 % of the whole study area has been under high or very high magnitude forest change (classes 3 and 4) and 29,5 % low or moderate magnitude forest change (Table 18).

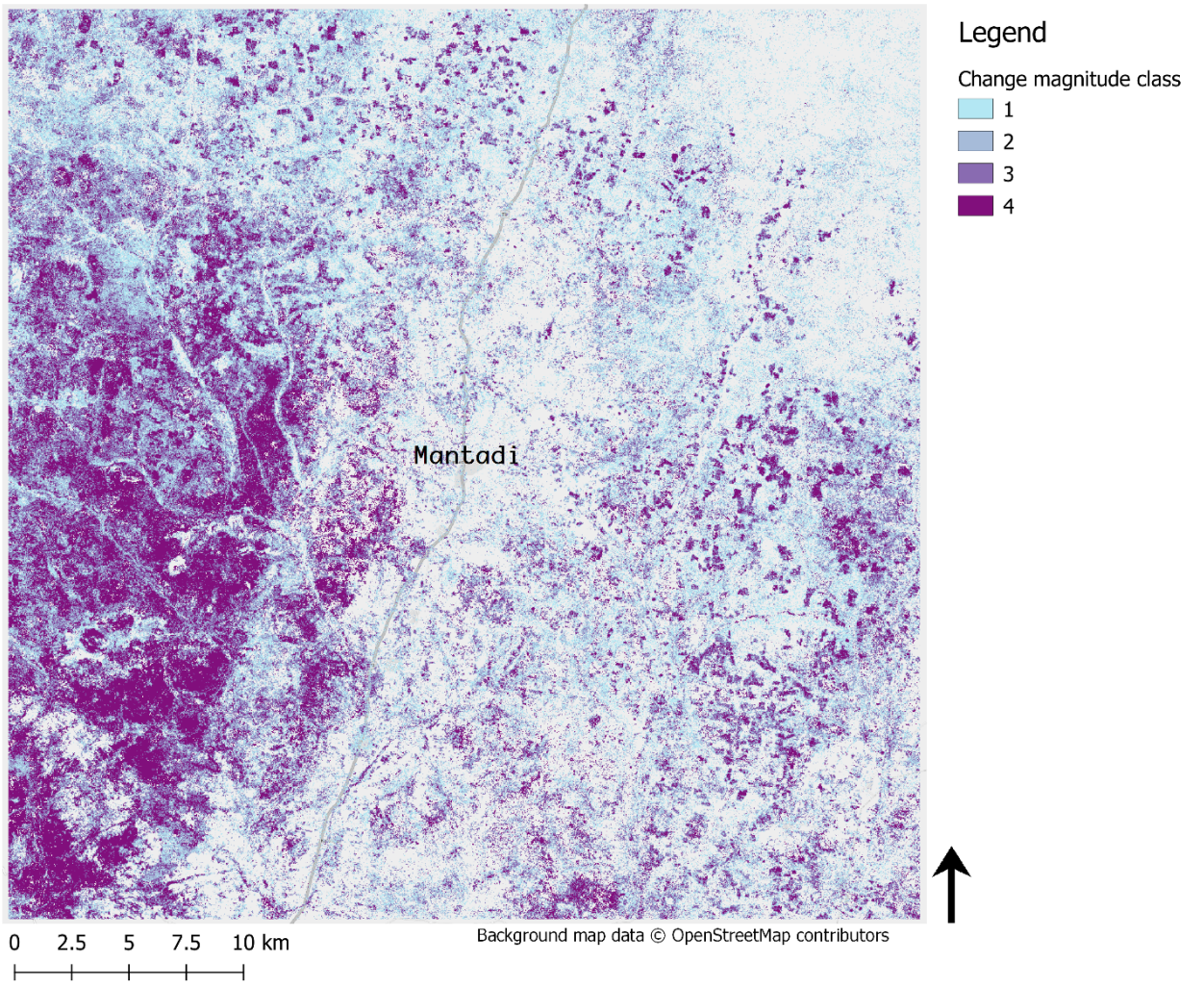


Figure 17. Final change map from surroundings of Mantadi village centre and the Sirari-Mbeya road. Change magnitude is classified per quartiles from 1 to 4, where 1 is low magnitude of change and 4 high magnitude of change.

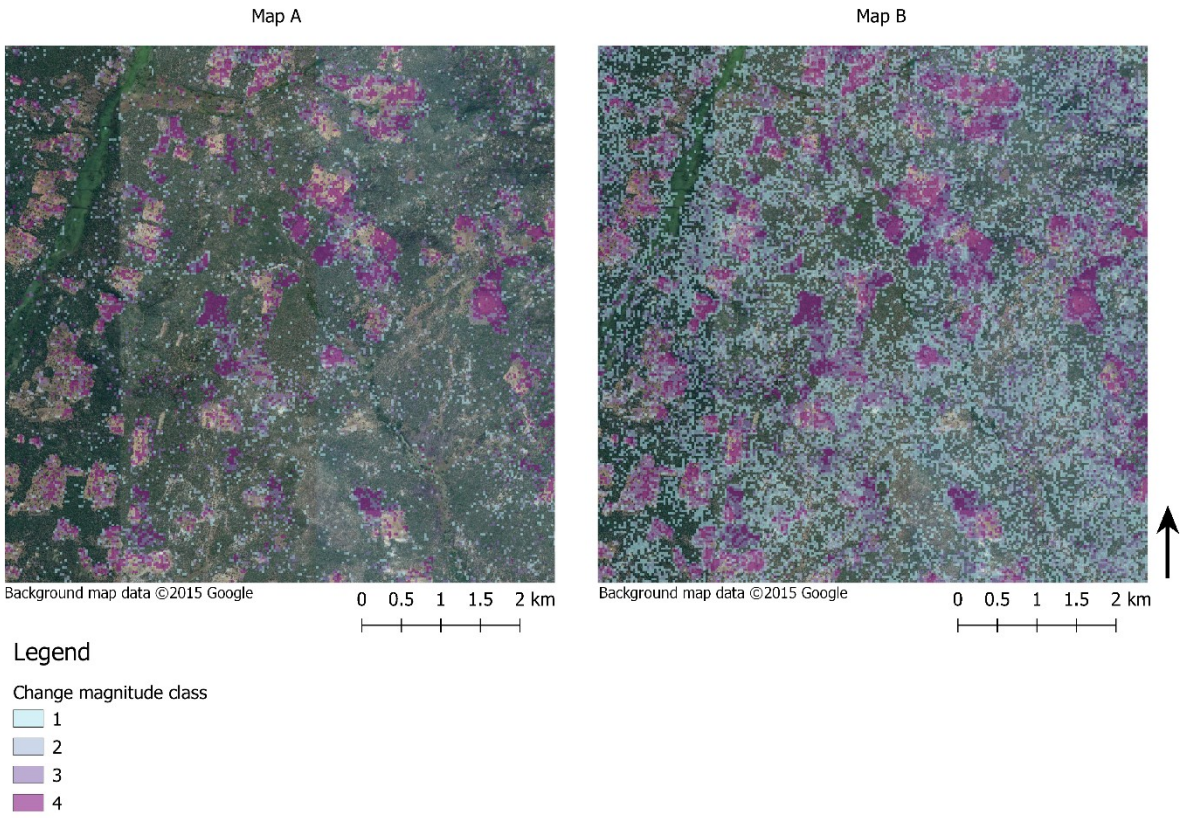


Figure 18. Map A represents pixels that have been labelled as high probability change, indicating that two or more indices had same magnitude category values. Moderate probability change pixels were added in Map B, including also pixels where at least two indices identified change in any magnitude category, thus the value of change magnitude class represents the average of these spectral index-based change maps. Opacity has been reduced to 60 % to demonstrate change in relation to the satellite image in the background.

These results show that the woodlands in the study area have been under serious conversion and degradation. Not all the degradation is caused by human as significant amount of lower magnitude disturbances are resulting from natural fires that are occurring regularly in the study area. However, to be classified as changed by the LandTrendr analysis, these change events are required to last for at least two years, as the parameters were set to ignore the temporary negative change signals. The possible recovery trends were not considered in this study leaving us without knowledge about the areas that may have been recovered later after the change event.

Table 18. Proportion of change magnitude classes in total disturbed area.

Magnitude class	**Area km2**	**%**
1 Low	203	12,6
2 Moderate	272	16,9
3 High	250	15,5
4 Very high	178	11,0
Total	902	56,0

The intensity of the disturbance events can be studied in Figure 19. In 36 % of the disturbed pixels in final map, change rate was less than 10 % annually. More intense change rate was found in 64 % of the pixels, indicating that high proportion of disturbance has occurred relatively fast, suggesting that abrupt clear-cuts or intensive burning have been taking place. The change rate map confirms this conclusion as it can clearly be reflected that agricultural fields, which are creating patchy red pattern all around the map, have high change rates. The areas in between represent gradual, lower intensity change and these areas occur almost everywhere in the map around the red patches. On the upper-right corner of the map a thicker forested area prevails, which does not yet belong to the agricultural zone and is only experiencing subtler change rates. On the left side of the map, the whole area has been cleared from woodland during the 31-year study interval. Closely around the village of Mantadi and the main road, most of the woodlands had already been cleared in the beginning of the observations and are mainly experiencing gradual negative vegetation change.

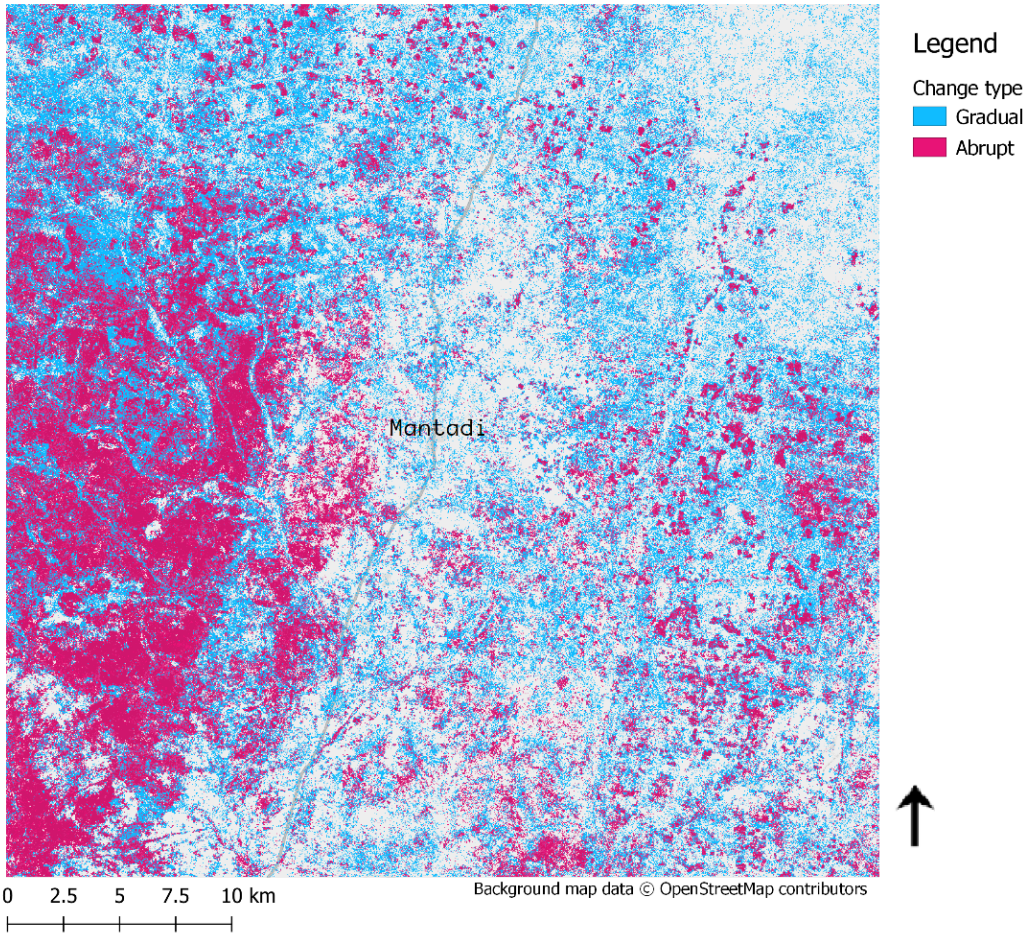


Figure 19. Annual change rate in the study area. Gradual change refers to less than 10 % decrease in canopy cover annually, whereas abrupt change refers to more intense change rate, over 10 % annually.

The accuracy of change rate map was generally high, with 85 % overall accuracy (Table 19). Producer's accuracy of abrupt change events was slightly higher than gradual change events. Commission errors for both gradual and abrupt change events were minimal and, in both cases, very similar. The change rate mapping was not evaluated in index-specific manner, but nevertheless the results show that the LandTrendr can accurately detect the change rates in addition to change magnitudes.

Table 19. Error matrix for the change rate values of final combined disturbance map.

Classified change map	Reference data				
	Gradual	Abrupt	Total	User's accuracy	Commission error
Gradual	51	9	60	0,85	0,15
Abrupt	12	66	79	0,84	0,16
Total	63	75	138		
Producer's accuracy	0,81	0,88		0,85	
Omission error	0,19	0,12			0,15

5.3. The effect of multi-index approach on accuracy of LandTrendr change mapping

The overall accuracy of the combined change map representing the regional distribution and magnitude of change was 72 % showing that compared to individually best performed NBR, the accuracy increased by 10 % (Table 20). All the change categories got less omission and commission errors after blending these maps by leaving out the change events that were not recognised by at least two indices. The misclassification was nevertheless mostly due to confusion in change magnitude category.

Accuracy metrics were also calculated for generalized categories “Changed” and “Not changed”, as was done earlier for individual index-based error matrices (Table 21). This led to higher overall accuracy of 89 % indicating that if not separately classified by the change magnitude, this approach can detect change accordingly in very high accuracy. However, the same accuracy was reached for generalised categories when applying NBR in LandTrendr analysis. Therefore, this approach did not increase the overall accuracy for simple change detection. It is also noteworthy that in assessing the accuracy of the generalized categories, the weighting in "change" category is one third larger than “no change” category, as more reference pixels are included in this category.

User’s Accuracy of “No change” category was 95 %, indicating that when this category was to be found in map, it is very unlikely that ground truth would suggest it to be something else. In comparison, producer’s accuracy for the “No change” category was 77 %. This means that 95 % percent of the areas identified as “No change” in the classified map are actually stable on the ground, but only 77 % of the reference stable areas have been correctly identified as “No change” by LandTrendr algorithm, demonstrating that the algorithm’s tendency to detect false positives could not be totally avoided by combining the results from several change maps. Closer look reveals that lower magnitude change classes have been included in this category. However, these “false positives” were 100 % in magnitude category of 1—2, indicating that, if taking only into consideration severe disturbance, the producer’s accuracy for the “No change” class would have been 100 %.

Table 20. Error matrix for the final combined disturbance map and all the change classes.

Combined change map	**Reference data**							
	Class 1	*Class 2*	*Class 3*	*Class 4*	*No change*	*Total*	*User's accuracy*	*Commission error*
Class 1	21	6	1	0	15	43	0,49	0,51
Class 2	7	19	3	3	8	40	0,48	0,52
Class 3	3	5	28	5	0	41	0,68	0,32
Class 4	1	0	6	30	0	37	0,81	0,19
No change	2	1	1	0	75	79	0,95	0,05
Total	34	31	39	38	98	240		
Producer's accuracy	0,62	0,61	0,72	0,79	0,77		0,72	
Omission error	0,38	0,39	0,28	0,21	0,23			0,28

Table 21. Error matrix for changed/not changed pixels in final combined disturbance map.

Classified change map	**Reference data**			
	Changed	Not changed	Total	User's accuracy
Changed	138	23	161	0,86
Not changed	4	75	79	0,95
Total	142	98	240	0
Producer's accuracy	0,97	0,77	0	0,89

6. Discussion

6.1. Change detection with LandTrendr algorithm and spectral indices

All three spectral indices have proven to be efficient in recognizing significant disturbances or land cover changes in Landsat time series with LandTrendr algorithm. High magnitude change was most likely recognized accurately by all the indices and major differences appear only when identifying subtler changes. TCW is oversensitive for all kinds of changes in vegetation or soil humidity conditions, and thus identifies actual changes efficiently. However, when applied in LandTrendr analysis in Miombo woodlands of Angola (Schneibel et al. 2017), TCW was found to correlate with the amount of rainfall, which might explain its tendency to produce high amounts of false positives. NBR and NDVI especially are less sensitive for subtler changes in vegetation conditions, and less often give false positives in the cost of ignoring actual changes and increasing the number of false negatives.

NBR is the most stable of three indices, producing comparable number of false positives and negatives and has therefore highest overall accuracy. Nearly equal accuracies were reached by Schneibel et al. 2017, when implementing NBR in LandTrendr algorithm to study Miombo woodland conversion to agriculture in Angola. According to their findings, NBR was found to reliably detect clearcuttings prior to cultivation in Miombo woodland systems and suggest the use of NBR especially for these purposes. Other indices like NDVI or Tasselled Cap Transformations might in their opinion be more appropriate for characterization of other vegetation dynamics. This might be explained with to the higher water content and water stress resilience of Miombo woodlands. Kennedy et al. (2012) found out that when implementing to LandTrendr algorithm, NDVI performed more poorly than TCW and NBR, which indicates that shortwave-infrared region is more useful in tracking forest disturbances. NBR was also more vulnerable to noise, which made TCW most suitable for LandTrendr algorithm fitting, according to their study.

Combining the strengths of these indices, especially the lower magnitude disturbances can be tracked even more accurately, as this study has demonstrated. Compared to individually best performed NBR, the accuracy increased by 10 % when the results were combined, and all the change categories got less omission and commission errors. The misclassification was nevertheless mostly due to confusion in change magnitude category. Categorizing magnitude in 4 classes might provide more insight to the analysis but it also could be a source of error, as the classification of reference samples in four categories was in some cases challenging. Lower magnitude class accuracy does not necessarily indicate the false interpretation of the algorithm but the lack of reference data to rely on in the visual interpretation process when the algorithm does the evaluation

in more sophisticated way. While the classification accuracy was generalized only to categories “Changed” and “Not changed”, the overall accuracy of change detection was exceptionally high for the final disturbance map.

When two or three different spectral indices instead of one record a negative trend in vegetation cover, the change event becomes more evident. Utilizing more than just one spectral index stabilizes the outcome, increasing the accuracy of the change detection. Cohen et al. (2018) highlighted that the outcome of time series approach can significantly alter depending on the selected method and approach. Thus, matching together results from analysis with different parameter settings, even when performed in simplified manner, decreases the variation of change detection results caused by the special characteristics of different indices and emphasizes the detection of actual change regardless of these specifications. Healey et al. (2018) managed to reduce omission errors effectively by taking the advantage of the strengths of different algorithms using secondary classification that employs stacking of output maps. They also demonstrated that using the LandTrendr alone could provide comparable, if not improved results in multispectral ensemble.

The algorithm’s definition of change in any time period depends on the preceding and latter spectral properties of time series and might be distorted in the beginning or the end of the trajectory, as the segmentation is not stabilising itself in relation to earlier or later observations. The relativity allows the capture of subtle but consistent trends, though it can also produce several small-magnitude false positives (Cohen et al. 2018). Most of the errors in each change map were associated with low magnitude spectral change. Therefore, the accuracy of LandTrendr derived change maps could be higher if a certain proportion of the lowest spectral change would be removed from the consideration. Some mapping applications do not require the full range of disturbance magnitude, and greater the magnitude, higher the mapping accuracy. By trimming the magnitude range, disturbance map can therefore be tailored to the required application. In this approach only the values representing first 10 % decrease in vegetation cover were trimmed to avoid false positives caused by external noise. Increasing the percentage would probably lead to higher accuracies by removing further uncertainties that are related in the lowest magnitude categories.

Google's software has significantly altered the remote sensing field by streamlining the acquisition and processing of images, and by developing easy-to-use software that makes Earth observation available to everyone (Bey et al. 2016). In future, the processing power of Google Earth Engine or other platforms, and numerous spectral indices and time series approaches could be exploited by ensemble approaches that was introduced briefly by Healey et al. (2017). Using secondary classification, the power of these indices and time series approaches could be combined even more efficiently to find best approaches for various vegetation zones and types.

6.2. Deforestation dynamics in Miombo woodlands of the study area

This study demonstrated that significant amount Miombo woodlands have been under deforestation and forest degradation processes between 1987—2019 in the study area. Half of these change events were occurring in high intensity and magnitude, indicating a fast-phased clearance of forests. According to Jew et al. (2016), Miombo woodlands in the study area affected by both deforestation through the clearance for agriculture and woodland degradation through the collection of timber for building material, products, and the most importantly, charcoal and fuelwood. The visual interpretation of the deforested plots supports this conclusion, indicating that the proximate cause for deforestation and woodland degradation in and around the village of Mantadi is smallholder tree felling and farming. There were no signs of intensive agriculture, but progressive spreading of smallholder farmland towards the unoccupied landscapes increases the deforestation zone annually in circular pattern around the main road and the village. A common form of land use in Mantadi region is slash-and-burn agriculture (Jew et al. 2016). The forest is cut down, after which the trees are burned to transfer nutrients to the soil. The land can be cultivated and burned again for a few harvests, after which the land is left to reforest, and the farmer moves to new areas. In small-scale, the forest is reforested in time but when the population pressure increases, burning often becomes recurrent, causing the soil to deteriorate which weakens the ability of the forest to regenerate (Kalaba et al. 2013). This phenomenon was identified through visual interpretation process. The deforestation process recognized in several reference plots started with natural or intentionally ignited fires and led to establishment of farmland. After few years of cultivation, farmland was abandoned and areas further away are again burnt and cultivated. However, the abandoned agricultural areas did not show major recovery trends according to the reference data, indicating soil deterioration. These dynamics could have been better captured if reforestation trends were analysed as part of the study. Within this thesis, only the negative vegetation trends taking place in the landscape could be confirmed.

Lower magnitude rates were associated with less intensive land cover change events such as tree felling and fire. Fire is a natural element in Miombo woodlands, and the consequences of fires depend on their frequency, timing, and intensity (Niemelä 2011: 36). The parameters in LandTrendr algorithm were set so that change events lasting less than one year were not considered as land cover change. Hence these lower magnitude disturbances are results of longer land cover processes like recurring fires, harvesting and degradation due to other reasons.

6.3. Critical assessment of the methods

Despite the obvious strengths and great utility, there are shortcoming and potential sources of error in the research methods used in this study. Any automated analysis from remote sensing data contains some errors from various steps in satellite image processing, calibration, and the change detection itself (Jones & Vaughan 2010: 267). In addition, there are almost certainly distortions that result from the lack or poor quality of data. The Landsat data had three data gaps during 1987-1994 and in some years the number of cloudy pixels were limiting the usability of the images as these pixels had to be masked out. The season under observation lasted four months (from June to September), which might result in some seasonal variation between annual images causing negative trend in change trajectory, which then is interpreted as disturbance. Even if the images were acquired on a same date, annual variations in climate such as precipitation timing and rainfall would affect the reflectance values, especially in indices that rely strongly on SWIR-bands, which are detecting changes in moisture conditions. However, the LandTrendr segmentation does not consider changes that last less than one year in, thus avoiding most of the annual distortive fluctuations (Kennedy et al. 2010).

There are also shortcomings in Collect Earth. Sampling by point is not a spatially comprehensive method, as it does not produce accurate site-specific information and regional coverage depends on the density of the generated grid (Bey et al. 2016). In addition, selecting the appropriate sampling plan and density for the area requires familiarization with the geographical features of the area to maintain high validity and ensure that the generalizations made based on the samples can be relied upon. The geographical coverage of high-resolution satellite images is still limited, and the irregular temporal interval of images restricts the full utilization of the software for research purposes, especially in areas where cloud cover is common, and changes occur quickly. The availability of high-quality satellite images varies from region to region, and occasionally there are instances where a cloud cover or other disturbance has affected all available satellite data for a single sampling block. With the Landsat satellite image archive, geographic and temporal coverage can be achieved, but the resolution of the images is rather low (30m), making the detection and scheduling of small changes challenging.

Detecting change through visual interpretation is a challenging task and is highly dependent on the skills of the interpreter. The interpretation process involves understanding the spectral characteristics of indices being used and how they reflect the changes on the ground. Without the possibility to visit the landscapes in the field, characterisation of the vegetation and the change dynamics relied greatly on the available literature and the accessibility and quality of the satellite images. The accuracy of change maps could be also more reliably estimated if the reference material were larger. However, the analysis and identification of disturbance simultaneously from multiple

satellite data sources required extensive time investment, that had to be limited appropriately to match the objectives of this thesis.

It should also be noted that full potential of the LandTrendr algorithm could not be harnessed for this study approach nor could all the features of the algorithm be considered in the parameterization process. LandTrendr also considers the starting and ending points, and therefore the duration of the change events, and it would be fascinating to consider the timing factor as part of the results. However, confirming of the exact year of the change event through visual interpretation requires more sophisticated approach for the reference sampling and validation system, and therefore the time factor was left out of the analysis. With the currently unavailable TimeSync validation platform, the year of change could be more reliably identified and validated.

Profound evaluation on how different settings effect to the algorithm functionality and results in certain types of vegetation dynamics, such as changes in Miombo woodlands, would be a broad research topic of its own. Therefore, this study relied strongly in literature in parametrisation process, and confirmed the functionality only with a small amount of test samples. More research is required in the future to properly parametrise LandTrendr precisely to be suitable for studying disturbance and recovery dynamics in Miombo woodlands with different vegetation indices. However, this study served as a preliminary preview for future research and presented the potential of this powerful algorithm and the benefits that could be acquired by combining the results from different indices in Landsat based time series analysis.

The methodologies used in this research were novel even in the field of remote sensing and thus required deep self-familiarization and independent work from the author. Some of greatest challenges in this study were related to the author's inexperience with mathematical and statistical modelling, therefore creating restrictions on the application of the method. However, with the help of other recently published international research articles and the latest implementation of the algorithm to Google Earth Engine by Kennedy et al. (2018) and the Oregon State University team, the study could be conducted without experience in time series analysis by implementing scrips in cloud computing platform. Due to the great amount of resources placed into understanding and implementation of the code and to constructing an applied reference data collection scheme in the same platform combined with Collect Earth, the data processing from the algorithm results received less attention. Some promising stacked ensemble approaches have been applied lately for Landsat based time series algorithm results, to get the most out of the algorithm and spectral indices (Healey et al. 2018). This approach would have been fascinating to apply in this study as well but would have required more familiarization with the topic. Therefore, the results of the LandTrendr analysis were combined in a very practical manner using the skills that the author already possessed in geoinformatics and that could be evaluated in a very straightforward manner.

6.4. Miombo woodlands under threat

Miombo is lost gradually due to the population pressure (Niemelä 2011: 264). Erosion and soil depletion are one of the most serious problems everywhere in Africa. Logging, grazing, and burning weaken the soil little by little, causing erosion and soil depletion. Smallholder clear-cuts for agricultural purposes and wood extraction for energy are the greatest threat to Miombo woodland in Tanzania (Abdallah & Monela 2007; Cabral et al. 2011). Slash and burn agriculture is not sustainable, when there is not enough time for the cultivated area to recover. For local people, agriculture and utilization of woodland products provides both food and income and are vital to their livelihoods when they are dependent on self-subsistence. Mobile lifestyle turns into permanent settlements and the pressure towards landscapes grows even further. Intensive agriculture poses a new threat for Miombo woodlands (Niemelä 2011; Jew 2016). Turning these forests to agricultural lands would damage the self-sufficiency of local peoples and reduce the biodiversity of these ecologically valuable landscapes. Furthermore, illegal logging threatens forests even inside the conservation areas in many East African countries, including Tanzania (Jew et al. 2016). Most of the Miombo woodlands remain outside of conservation areas and are affected by human disturbance (Timberlake & Chidumayo 2011).

Recent studies have estimated that the Miombo woodlands in Southern Africa are one of the ecoregions that will be most seriously affected by climate change (Warren et al. 2018). Increasing temperature would in the worst-case scenario threaten to abolish a quarter of the Miombo woodland species. On the other hand, Miombo woodlands have a great capacity for carbon storage to mitigate carbon emissions, presenting another rationale for conservation of these savannas (Munishi et al. 2010). Research is key to the sustainable development and nature conservation in these woodlands. African nature is still relatively poorly studied, and with knowledge the regions under highest pressure and conservation value can be recognized. According to Munishi et al. (2011), most published research papers are from areas of dry Miombo, and very few evaluates Miombo in areas where cultivation is currently taking place. Only a few studies have been carried out in areas with precipitation above 900 mm per year (Munishi et al. 2011), which emphasize the need for further research in these areas. On the positive side, there is still a lot of unique, native and high biodiversity forest in Eastern Africa, which have a significant conservation value (Niemelä 2011).

The attitudes of local people towards nature conservation are often twofold. Establishment of natural parks and wildlife related tourism can increase income for some but restrict the use of natural resources for others. Illegal logging threatens even protected forests (Timberlake & Chidumayo 2011). Threat scenarios are similar across all tropical regions: population growth and unsustainable exploitation of natural resources. As long as population growth continues, more and more forest will be replaced with agriculture and human pressure (Metz 2009). Overexploitation causes soil depletion

and erosion which lead to clearing of new areas. There is a global demand for new farmland and some of the wealthier countries and companies have taken advantage of cheap land of Africa. At the same time, the benefits from the land are transferred abroad. A large part of the world's unused land is in Africa, but the ecological value of this land is irreplaceable for the local people as well as for the conservation of biodiversity.

Even though more attention is paid to nature conservation and environmental issues, as long as the population grows in Tanzania, the human pressure to the surrounding nature will increase. Locals cannot be blamed for taking advantage of the surrounding nature if that is the only way to secure livelihood to themselves and their families. Man has left a permanent mark everywhere and the deforestation in Africa is not more of a crime than the historical deforestation that took place earlier in the North. Miombo woodland in western Tanzania will continue to degrade and the woodland area decreases as conversation to agriculture and utilisation of local people keeps increasing (Jew et al. 2016). If regulation remains insufficient it is likely that land use will not be sustainable, leading to losses in biodiversity and ecosystem services. Even though Miombo woodlands, due to the effective tree regeneration, have demonstrated a noteworthy capacity to recover after disturbance like logging, agriculture and fire, it’s unlikely that the regeneration periods will be sufficient in the future, as reaching the original mature woodland structure requires 20—30 years (Kalaba et al. 2013). It is vital that land management plans would be developed taking into consideration ecological sustainability and the needs of the local communities.

In 2017 FAO was investigating the sustainability aspects in Miombo woodlands (Gumbo et al. 2018). Key stakeholders from Miombo countries were invited to work together to learn and share knowledge from sustainable management of the woodlands. This workshop highlighted the need for regional collaboration among Miombo countries. The key conclusions from the workshops were that 1) the importance of the Miombo woodlands for food and nutrition security and poverty alleviation needs to be recognized; 2) cooperation between the Miombo countries on sustainable forest management and restoration must be emphasized; 3) a framework for harmonizing policies and guidelines on the sustainable use of forests must be established; 4) transboundary monitoring and sustainable management of the Miombo woodlands needs to be enabled 5) alternative livelihood opportunities must be identified and acquired training provided. According to the report, Miombo woodland countries are lacking the enabling policy environment and resource management that would consider the conservation needs or establish the required institutions to sustainably manage these areas. The full value of Miombo woodlands social, economic, and environmental benefits are commonly not recognized in community or national level, and regional research and literature is often missing or outdated. There are no data at country level reporting precisely the coverage of Miombo woodlands and the woodlands are often reported as savannah or simply forests.

6.5. Further steps in deforestation research

The global change is creating increasing demands to see the earth as a single complex system where human societies and forest dynamics are interconnected. The causes, regional structure and consequences of deforestation are highly interdimensional and interrelated with different social and environmental processes (Lambin et al. 2006: 7). Every change is the result of many factors, land uses and responses to social, economic, and environmental conditions under broad range of regional and temporal scales and connections between people and their land. This makes the construction of universal theories very demanding, but complex research topics such as climate change, carbon budgets and biodiversity loss require more comprehensive deforestation studies.

Through remote sensing it is possible to reach even the remotest areas where changes in forests are taking place. Compared to alternative information sources such as inventories and field measurements, remote sensing data is easily obtained and enables relatively fast analysis with lower economic costs. The Landsat program has offered an incomparable amount of temporal and spatial coverage among available satellite images, and together with the free accessibility and available cloud processing power, increasingly sophisticated approaches can be implemented with lower economic and personal investment. The development of open access algorithms like LandTrendr continues, creating more possibilities to fully understand the complex change trajectories in forest dynamics and how these are affected by the changing climate condition and human influence.

To understand the deforestation dynamics, it is not enough to just identify the changes. Miombo woodlands are part of the complicated ecological, social, and economic network that have a crucial role in providing livelihoods for rural and urban communities through nutrition, products, and energy, not to mention the spiritual and environmental needs. Research has a key role in understanding the ongoing changes and thus to find the root causes and possible solutions for the protection and sustainable management of forest landscapes that support the maintenance of the local livelihoods and the precious biodiversity. Landscape studies could provide deeper understanding of the underlying causes of deforestation (Babin 2004: 19). A landscape is the space that is formed between people and the land and its resources. The activities that are taking place in the landscape are leaving traces in spatial structures and dynamics, which can be observed and interpreted from the satellite images. Therefore, landscape studies offer a possibility to understand the dynamics behind the deforestation processes in terms of space and time. Furthermore, social and environmental sciences are important contributors to increase the understanding of the broader social and economic interactions by investigating the role that forests play in the lives of the local people that are often both causing the deforestation, and suffering from its consequences the most.

7. Conclusions

- The capacity of the LandTrendr algorithm to detect forest disturbance and deforestation in Miombo woodlands depends highly on the spectral index that has been used.
- To detect areas under any kind of disturbance, LandTrendr performs considerably well with all three indices (overall accuracy from 85—89 %).
- In more profound change magnitude detection, clear differences between the indices can be noticed especially in finding of subtler, low magnitude changes.
- NBR has highest overall accuracy and it detects disturbances most evenly in all change magnitude classes in Miombo woodlands. TCW performs almost equally well, even though they have distinct tendencies for change detection. TCW provides the most accurate detection disturbance but is compromised with a high number of false positives, which decreases the overall accuracy of the index.
- NDVI is less sensitive to detect changes through LandTrendr analysis in Miombo woodlands.
- By combining the results from three spectral indices in LandTrendr analysis, the overall accuracy of disturbance mapping was increased by 10 %, compared to the best performing index. It also stabilized results by decreasing either the number of false negatives from NDVI and NBR or the false positives from TCW.
- The results indicate that 56 % of the total study area has been under disturbance during 1987—2018.
- 26,5 % of the whole study area has been under very high or high magnitude disturbance and 29,5 % under low or moderate magnitude disturbance.
- High magnitude changes are mostly related to abrupt change events like forest clearance for agriculture and infrastructure, that lead to total deforestation of the plot. In some cases, high magnitude changes also relate to gradual, but considerable degradation like gradual tree felling or successive forest fires that lead to a prominent loss or deterioration of vegetation.
- Moderate magnitude changes are consequences from prominent forest degradation through harvest, partial cultivation, or fire events. In these areas the tree vegetation has decreased or weakened notably, but not totally deforested.
- Low magnitude change events are usually resulting from minor harvesting, fire or other factors that cannot be identified from satellite images such as climate or insects.
- Miombo woodlands in the study area are thus affected by deforestation through woodland clearance for agriculture, and woodland degradation through tree felling and natural and anthropogenic fires.

References

Abbadie, L., Gignoux, J., Roux, X. & Lepage, M. 2006. *Lamto: Structure, Functioning and Dynamics of a Savanna Ecosystem* 179. 408 p. Springer Science & business Media.

Abbink, I. J., Bruijn, M. D. & Hesseling, G. 2011. Land, law and politics in Africa: Mediating conflict and reshaping the state. Leiden; Boston: Brill, 1—12.

Abdallah, J. M., & Monela, G. G. 2007. Overview of Miombo woodlands in Tanzania. In *Proceedings of the First MITMIOMBO Project Workshop, Morogoro*, 6—12.

Adams, J. & Gillespie, A. 2006. *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach*. 336 p. Cambridge: Cambridge University Press.

Andersson, K., Evans, T., Gibson, C. C., and Wright, D. 2014. Decentralization and Deforestation: Comparing Local Forest Governance Regimes in Latin America. In: Duit, A. 2014. *State and Environment: The Comparative Study of Environmental Governance*. Cambridge, Mass: The MIT Press, 239—259.

Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M. A., Froese, R. E., & White, J. C. 2014. Forest monitoring using Landsat time series data: A review. *Canadian Journal of Remote Sensing* 40(5), 362—384

Bey, A., Sánchez-Paus Díaz, A., Maniatis, D., Marchi, G., Mollicone, D., Ricci, S. & Patriarca, C. 2016. Collect earth: Land use and land cover assessment through augmented visual interpretation. *Remote Sensing* 8(10), 807.

Bowman, D. M. J. S., Haverkamp, C., Rann, K. D. & Prior, L. D. 2018. Differential demographic filtering by surface fires: How fuel type and fuel load affect sapling mortality of an obligate seeder savanna tree. *Journal of Ecology* 106(3), 1010—1022.

Brauman, K.A., Daily G.C., Duarte T.K., Mooney H.A. 2007. The nature and value of ecosystem services: an overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32, 67—98.

Brink, A. B., Bodart, C., Brodsky, L., Defourney, P., Ernst, C., Donney, F., Lupi, A. & Tuckova, K. 2014. Anthropogenic pressure in East Africa—Monitoring 20 years of land cover changes by means of medium resolution satellite data. *International Journal of Applied Earth Observation and Geoinformation* 28, 60—69.

Cabral, A.I.R., Vasconcelos, M.J., Oom, D. & Sardinha R. 2011. Spatial dynamics and quantification of deforestation in the central-plateau woodlands of Angola (1990–2009) *Applied Geography* 31(3), 1185—1193.

Campbell, B. M., Angelsen, A., Cunningham, A., Katerere, Y., Sitoe, A., & Wunder, S. 2007. *Miombo woodlands–opportunities and barriers to sustainable forest management.* CIFOR, Bogor, Indonesia.

Chatfield, C. 2016. *The Analysis of Time Series: An Introduction.* Sixth edition. Chapman and Hall/CRC. London, 1—6.

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

Cohen, W.B., Healey, S.P., Yang, Z., Kennedy, R.E., Gorelick, N. 2018. A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sensing of Environment* 205, 131—140.

Crist, E.P., Cicone, R.C. 1984. A Physically Based Transformation of Thematic Mapper Data-The TM Tasseled Cap. *IEEE Trans. Geosci. Remote Sensing of the Environment* 22, 256—263.

DeVries, B., Verbesselt, J., Kooistra, L., & Herold, M. 2015. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sensing of Environment* 161, 107—121.

Epting, J., Verbyla, D., Sorbel, B. 2005. Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. *Remote Sensing of Environment* 96, 328—339.

FAO (Food and Agriculture Organization of the United Nations). 2015a. Global Forest Resources Assessments (FRA) 2015. 56 p., FAO, Rome.

FAO (Food and Agriculture Organization of the United Nations). 2015b. Global Forest resource Assessment (FRA) 2015 – Country Report, Tanzania.

FAO (Food and Agriculture Organization of the United Nations). 2007. State of the World's forests 2007. 140 p. FAO, Rome.

FAO (Food and Agriculture Organization of the United Nations). 2000. On Definition of Forest and Forest Change. 15 p. Forest Resource Assessment Program, Working paper 33, Rome.

Fornacca, D., Ren, G. & Xiao, W. 2018. Evaluating the Best Spectral Indices for the Detection of Burn Scars at Several Post-Fire Dates in a Mountainous Region of Northwest Yunnan, China. Remote Sensing 10(8), 1196—1212.

Fragal, E. H., Silva, T. S. F., Novo, E. M. 2016. Reconstructing historical forest cover change in the Lower Amazon floodplains using the LandTrendr algorithm. *Acta Amazonica* 46(1), 13—24.

Geist, H., W. McConnell, E. Lambin, E. Moran, D. Alves & T. Rudel. 2006. Cause and trajectories of land-use/cover change. In: Lambin, E. & H. Geist (eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts.* Springer-Verlag Berlin Heidelberg, Germany. 41—70.

Geist, H. & Lambin, E. 2001. What drives tropical deforestation: A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence. LUCC Report Series No. 4, 136 p.

Gómez, C., White, J., & Wulder, M. 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 116(C), 55—72.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202, 18—27.

Griffiths, P., & Hostert, P. 2015. Forest Cover Dynamics During Massive Ownership Changes–Annual Disturbance Mapping Using Annual Landsat Time-Series. *Remote Sensing Time Series*, Springer, Cham, 307—322.

Gumbo, D. J., Dumas-Johansen, M., Muir, G., Boerstler, F., & Zuzhang, X. 2018. *Sustainable management of Miombo woodlands: food security, nutrition and wood energy*. FAO.

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

Healey, S.P., Cohen, B., Yang, Z., Brewer, C.K., Brooks, E.B., Gorelick, N., Hernandez, A.J., Huang, C., Hughes, M.J., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroedes, T.A., Stehman, S.V., Vogelmann, J.E., Woodcocl, C.E., Yanf, L., Zhu, Z. 2018. Mapping forest change using stacked generalization: An ensemble approach. *Remote Sensing of Environment* 204, 717—728.

Hislop, S., Jones, S., Soto-berelov, M., Skidmore, A., Haywood, A., Nguyen, T. 2018. Using Landsat Spectral Indices in Time-Series to Assess Wildfire Disturbance and Recovery. *Remote Sensing* 10(3), 1–17.

Horning, N. 2010. *Remote sensing for ecology and conservation: A handbook of techniques.* Oxford: Oxford University Press.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83, 195—213.

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

Hughes, M., Kaylor, S., Hayes, D. 2017. Patch-based forest change detection from Landsat time series. *Forests* 8, 23—32.

Huntley, B. J. & Walker, B. H. 1982. *Ecology of tropical savannas*. ECOLSTUL 42, Springer. 626 p.

Hutley, L.B. and Setterfield. S.A. 2008. Savannas. Encyclopedia of Ecology, Elsevier B.V, Oxford. 3143—3154.

Jamali, S., Jönsson, P., Eklundh, L., Ardö, J., & Seaquist, J. 2015. Detecting changes in vegetation trends using time series segmentation. *Remote Sensing of Environment 156*(C), 182—195.

Jew, Eleanor & Dougill, Andrew & Sallu, Susannah & O Connell, Jerome & Benton, Tim. 2016. Miombo woodland under threat: Consequences for tree diversity and carbon storage. *Forest Ecology and Management*. 361 p.

Jiang, Z., Huete, A. R., Li, J., and Qi, J. 2007. Interpretation of the modified soil-adjusted vegetation index isolines in red-NIR reflectance space. *Journal of Applied Remote Sensing* 1, 503—512.

Johnson, Knowles, C. R. & Colchester, M. 1989. Rainforests. Land use options for Amazonia. 75 p. Oxford University Press and WWF United Kingdom, London.

Jones, H. G. & Vaughan, R. A. 2010. Remote sensing of vegetation: Principles, techniques, and applications. 353 p. Oxford: Oxford University Press. New York.

Kalaba, F., Quinn, C., Dougill, A., & Vinya, R. 2013. Floristic composition, species diversity and carbon storage in charcoal and agriculture fallows and management implications in Miombo woodlands of Zambia. *Forest Ecology and Management 304*(C), 99—109.

Kennedy, R. E., Yang, Z. & Cohen, W. B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment* 114(12), 2897—2910.

Kennedy, R. E., Yang, Z., Cohen, W. B., Pfaff, E., Braaten, J., & Nelson, P. 2012. Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment 122*, 117—133.

Kennedy, R.E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W.B., Healey, S. 2018. Implementation of the LandTrendr Algorithm on Google Earth Engine. *Remote Sensing 10*(5), 691, 1—10.

Key, C.H. & Benson, N.C. 2006. Landscape Assessment: Sampling and Analysis Methods. In FIREMON: Fire Effects Monitoring and Inventory System; *General Technical Report, RMRS-GTR-164-CD; Lutes, D.C., Ed.; US Forest Service*: Washington, DC, USA, LA1—LA51.

Kuenzer, C., Dech, S. and Wagner, W. 2016. Remote Sensing Time Series Revealing Land Surface Dynamics: Status Quo and the Pathway Ahead*. Remote Sensing and Digital Image Processing* 22, 1—24.

Lambin, E., Geist, H. & Rindfuss, R. 2006. Introduction: Local Processes with Global Impacts. In. Lambin. E.F. & Geist, H. (Eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts,* Springer, Germany, 41—70.

Lu, D., Mausel, P., Brondízio, E. & Moran, E. 2004. Change detection techniques. *International Journal of Remote Sensing* 25(12), 2365—2401.

Lunetta, R. & Lyon, J. 2000. *Remote Sensing and GIS Accuracy Assessments.* 304 p. CRC Press LLC, Boca Raton

Malimbwi, R. E., Nduwamungu, J., Misana, S., Jambiya, G. C., Monela, G. C., & Zahabu, E. 2004. Charcoal supply in Dar es Salaam city, Tanzania. *Tanzania journal of forestry and nature conservation 75*(1), 108—118.

Mayes, M. T., Mustard, J. F., & Melillo, J. M. 2015. Forest cover change in Miombo woodlands: modelling land cover of African dry tropical forests with linear spectral mixture analysis. *Remote Sensing of Environment* 165, 203—215.

Malmer, A., & Nyberg, G. 2008. Forest and water relations in Miombo woodlands: need for understanding of complex stand management. *Research and development for sustainable management of semiarid/Miombo/woodlands in East Africa. Working Papers of the Finnish Forest Research Institute 98*, 70—86.

McBeath, G. A. & Rosenberg, J. 2006. Comparative environmental politics. 182 p. Dordrecht, the Netherlands: Springer.

Metz, J. J. 2009. Deforestation. *International Encyclopedia of Human Geography.* Elsevier Ltd., 39—50.

Mittermeier, R. A., Mittermeier, C. G., Brooks, T. M., Pilgrim, J. D., Konstant, W. R., Da Fonseca, G. A. B. & Kormos, C. 2003. Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences of the United States of America* 100(18), 10309.

MNRT (Ministry of Natural Resources and Tourism) 2015. *NAFORMA - National Forest Resources Monitoring and Assessment of Tanzania Mainland*.

Munishi, P. K., Temu, R. P. C., & Soka, G. 2011. Plant communities and tree species associations in a Miombo ecosystem in the Lake Rukwa basin, Southern Tanzania: Implications for conservation. *Journal of Ecology and the Natural Environment 3*(2), 63—71.

Munishi, P. K. T., Mringi, S., Shirima, D. D. & Linda, S. K. 2010. The role of the Miombo Woodlands of the Southern Highlands of Tanzania as carbon sinks. *Journal of Ecoogy and the Natural Environment* 2(12), 261—269.

National Bureau of Statistics (NBS) 2013. *Population and Housing Census, 2012. Population Distribution by Administrative Areas.* National Bureau of Statistics, Ministry of Finance, Dar es Salaam and Office of Chief Government Statistician, President's Office, Finance, Economy and Development Planning, Zanzibar.

Nduwamungu, Jean & Bloesch, Urs & Hagedorn, P. 2008. Recent land cover and use changes in Miombo woodlands of eastern Tanzania. *Tanzanian Journal For Nature Conservation 78*(1).

Niemelä, T. 2011. *Vihreä Afrikka: Kasveja ja kasvillisuutta*. 320 p. Helsinki: Kasvimuseo, Luonnontieteellinen keskusmuseo.

Njoku, E. 2013. *Encyclopedia of remote sensing*. 939 p. New York: Springer.

Nummenmaa, L., Holopainen, M., Pulkkinen, P., & Kimpimäki, K. 2014. *Tilastollisten menetelmien perusteet* (1. ed). Helsinki: Sanoma Pro.

Purkis, S. & Klemas, V. 2011. Remote sensing and global environmental change. 384 p. Chichester; Hoboken, N.J.: Wiley-Blackwell.

Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering 1974. Monitoring vegetation systems in the Great Plains with ERTS, In 3rd ERTS Symposium, NASA SP-351 I, 309—317.

Roy, D., Wulder, M., Loveland, T., C.E, W., Allen, R., Anderson, M., Hipple, J. 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment 145(C), 154—172.

Ryan CM, Williams M. 2011. How does fire intensity and frequency affect Miombo woodland tree populations and biomass? Ecology Applications 21(1), 48—60.

Salo, M., A. Sirén & R. Kalliola 2014. *Diagnosing wild species harvest.* 479 p. Elsevier Inc, London.

Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F. & Ardo, J. 2005. Determinants of woody cover in African savannas. *Nature* 438(7069), 846—849.

Sasaki, N. & Putz, F. E. 2009. Critical need for new definitions of “forest” and “forest degradation” in global climate change agreements. *Conservation Letters 2*(5), 226—232.

Sauer, J., & Abdallah, J. 2007. Forest diversity, tobacco production and resource management in Tanzania. *Forest Policy and Economics 9*(5), 421—439.

Sawe, T. C., Munishi, P. K., & Maliondo, S. M. 2014. Woodlands degradation in the Southern Highlands Miombo of Tanzania: implications on conservation and carbon stocks. *International Journal of Biodiversity and Conservation* 6(3), 230—237.

Schneibel, A., Stellmes, M., Röder, A., Frantz, D., Kowalski, B., Haß, E., & Hill, J. 2017. Assessment of spatio-temporal changes of smallholder cultivation patterns in the Angolan Miombo belt using segmentation of Landsat time series. *Remote sensing of Environment 195* 118—129.

Schoene, D., W. Killmann, H. von Lüpke & Wilkie, M. L. 2007. Definitional issues related to reducing emissions from deforestation in developing countries. *Forests and climate change working paper* 5. Food and Agriculture Organization of the United Nations, Rome.

Senf, C.; Dirk, P.; Wulder, M.A.; Hostert, P. 2015. Characterizing spectral−temporal patterns of defoliator and bark beetle disturbances using Landsat time series. *Remote Sensing of the Environment* 170, 166—177.

Shandra, J., Shircliff, E. & London, B. 2011. The International Monetary Fund, World Bank, and structural adjustment: A cross-national analysis of forest loss. *Social science research* 40(1), 210—225.

Silayo, D. A., Katani, J. Z., Maliondo, S. M. S., & Tarimo, M. C. T. 2008. Forest plantation for biofuels to serve natural forest resources. *Research and development for sustainable management of semiarid Miombo woodlands in East Africa. Working Papers of the Finnish Forest Research Institute 98*, 115—124.

Souza, C. M. 2006. Mapping land use of tropical regions from space. *Proceedings of the National Academy of Sciences 103*(39), 14261—14262.

Syampungani, S., Chirwa, P. W., Akinnifesi, F. K., Sileshi, G. & Ajayi, O. C. 2009. The Miombo woodlands at the crossroads: Potential threats, sustainable livelihoods, policy gaps and challenges. *Natural Resources Forum*, 33(2), 150—159.

Tarimo, B., Dick, Ø. B., Gobakken, T. & Totland, Ø. 2015. Spatial distribution of temporal dynamics in anthropogenic fires in Miombo savanna woodlands of Tanzania. *Carbon balance and management 10*(1), 18.

Timberlake, J. and Chidumayo, E. N. 2011. Miombo Ecoregion Vision Report. *Occasional Publications in Biodiversity*, WWF – SARPO.

Timberlake, J., Chidumayo, E.N., & Sawadogo, L. 2010. Distribution and characteristics of African dry forests. In Chidumayo, E. N. & Gumbo, D.J. (Eds.), *The dry forests and woodlands of Africa: managing for products and services.* Earthscan, 11—42.

Tinley, K. L. 1982. The influence of soil moisture balance on ecosystem patterns in southern Africa. *Ecology of Tropical Savannas* (Ed. by B. J. Huntley & B. H. Walker), Springer, Heidelburg, 175—192.

Tucker, C. J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, 127—150.

Turner, M.G., Gardner R.H. & O'Neill R.V.O. 2001. *Landscape Ecology in Theory and Practice: Pattern and Process.* 401 p. Springer Verlag, New York.

United Nations Framework Convention on Climate Change (UNFCCC) 2002. FCCC/CP/2001/13/Add.1. 22.2.2016

United Republic of Tanzania (URT) 1997a. *Mbeya region socio-economic profile, 1997.*

United Republic of Tanzania (URT) 1997b. *Chunya district council socio-economic profile, 1997*.

United Republic of Tanzania (URT) 2012. *National sample census of agriculture 2007/2008. Volume Vk: Regional report – Mbeya Region.*

United Republic of Tanzania (URT) 2017. *Tanzania's forest reference emission level submission to the UNFCCC.*

USGS (U.S. Geological Survey) 2019. Landsat Missions. *Landsat Collection 1.* <https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-1> (Accessed 9.1.2019)

Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D. 2010. Detecting trend and seasonal changes in satellite image time series. Remote Sensing of Environment 114, 106—115.

Warren, R., Price, J., VanDerWal, J., Cornelius, S., & Sohl, H. 2018. The implications of the United Nations Paris Agreement on climate change for globally significant biodiversity areas. *Climatic change 147*(3-4), 395—409.

WCS (Wildlife Conservation Society) of Tanzania. 2016. Southern Highlands. <*https://tanzania.wcs.org/Landscapes/Southern-Highlands.aspx*> (Accessed 3.1.2018).

White, F. 2017. The vegetation of Africa: a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa.

White, J.C., Wulder, M.A., Hermosilla, T., Coops, N.C., Hobart, G.W. 2017. Remote Sensing of Environment - A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. *Remote Sensing of Environment* 194, 303—321.

Wulder Ma, Masek JG, Cohen WB, Loveland TR, Woodcock CE. 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment* 122, 2—10.

Wulder, Michael & Loveland, Thomas & Roy, David & Crawford, Christopher & Masek, Jeffrey & Woodcock, Curtis & Allen, Richard & Anderson, Martha & Belward, Alan & Cohen, Warren & Dwyer, John & Erb, Angela & Gao, Feng & Griffiths, Patrick & Helder, Dennis & Hermosilla, Txomin & Hipple, James & Hostert, Patrick & Hughes, M. & Zhu, Zhe. 2019. Status of Landsat program, science, and applications. *Remote Sensing of Environment 225*, 127—147.

Yang, Y., Erskine, P., Lechner, A., Mulligan, D., Zhang, S., & Wang, Z. 2018. Detecting the dynamics of vegetation disturbance and recovery in surface mining area via Landsat imagery and LandTrendr algorithm. *Journal of Cleaner Production* 178, 353—362.

Zhu Z, Woodcock CE, Olofsson P. 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sensing of Environment* 122, 75—91.