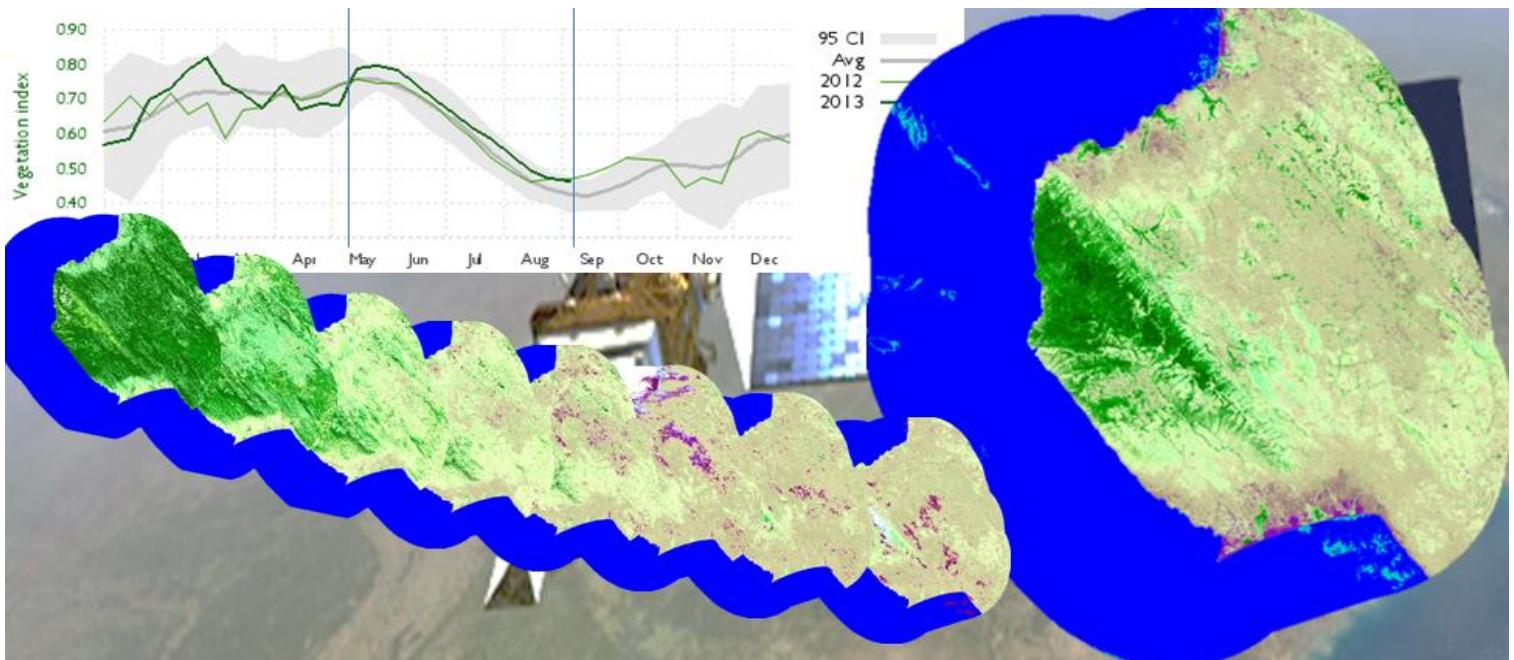


JRC TECHNICAL REPORTS

Phenology-based land cover classification using Landsat 8 time series

Edoardo Simonetti,
Dario Simonetti,
Damiano Preatoni

2014



Report EUR 26841 EN

European Commission
Joint Research Centre
Institute for Environment and Sustainability

Contact information

Simonetti Dario

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 260, 21027 Ispra (VA), Italy

E-mail: dario.simonetti@jrc.ec.europa.eu

Tel.: +39 0332 78 3871

<http://ies.jrc.ec.europa.eu/>

<http://www.jrc.ec.europa.eu/>

This publication is a Technical Report by the Joint Research Centre of the European Commission.

Legal Notice

This publication is a Technical Report by the Joint Research Centre, the European Commission's in-house science service.

It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

JRC91912

EUR 26841 EN

ISBN 978-92-79-40844-1

ISSN 1831-9424

doi: 10.2788/15561

Luxembourg: Publications Office of the European Union, 2014

© European Union, 2014

Reproduction is authorised provided the source is acknowledged.

ABSTRACT

This article describes the methodology and results of a new JRC phenology-based classification algorithm able to generate accurate land cover maps in a fully automatic manner from Landsat 8 (L8) remote sensed data available since 12th April 2013 at no charge throughout the USGS website. A preliminary study aiming to bypass the single date classification inaccuracy (mainly due to seasonality) using long term MODIS time series as a “driver” to fill gaps between high resolution data, has been carried out. The high global acquisition frequency (~16 days) and distribution policy are making Landsat 8 product extremely suitable for near real time land cover mapping and monitoring.

Five national parks in east Africa have been selected as study areas (Mahale Mountains, Mana Pools, West Lunga, Gorongosa, Tsimanampetsotsa); they are covering diverse eco-regions and vegetation types, from evergreen to deciduous. A buffer of 20 km around each park has been considered as well.

Selected single date images were first preprocessed in order to convert raw DN values to top of atmosphere (TOA) reflectance and minimizes spectral differences caused by different acquisition time, sun elevation, sun-earth distance, and after processed by the algorithm to generate a thematic raster map with land cover classes. Is worth noting that the single date classification accuracy is closely related to the acquisition date of the image, the status of the vegetation and weather conditions such as cloud and shadows often present in tropical regions; here the need of developing a phenology based algorithm that considers the vegetation evolution and generates a more accurate land cover map including evergreen and deciduous discrimination on the basis of “frequency” rules.

Land cover maps have been created for all parks and an exhaustive accuracy assessment has been carried out on Mahale Mountains and Tsimanampetsotsa. The combined overall accuracy of 82.8% demonstrates the high potentiality of this method and makes it usable at either local or regional scale.

1 Introduction

1.1 Background

Recent studies in earth science have revealed the important role of terrestrial ecosystems in sustaining the global environment. Global vegetation, covering three-fourths of the earth's land surface, has been identified as one of the key components of the climate system due to its key role in geosphere-biosphere-atmosphere interactions. The biogeochemical processes of vegetation, which involve land - atmosphere exchanges of energy, mass and momentum, are influenced by and in turn influence the climate system [1] [2].

Land surface properties such as the land cover type, leaf area index (LAI), and fraction of incident photosynthetically active radiation (0.4 - 0.7 μm) absorbed by the vegetation canopy (FPAR), are used as essential inputs in many hydrological, ecological, and climate models [2]. They are key parameters that describe the functioning of vegetation and are required for modeling vegetation productivity [3] and land surface climatology [4].

Recent studies have revealed the possibility of using remote sensing, the most effective means of collecting these information at a global extent on a regular basis, to characterize vegetation properties, and much knowledge has been gained about the role of vegetation in environmental and climate change [2] [5] [6].

Studies reporting the use of multi-temporal image data for classification often include relatively few dates, possibly due to a lack of cloud-free image availability, cost and processing requirements. A basic multi-temporal approach is the use of leaf-on leaf-off images, which provides greater vegetation phenology information that is available with single images [7]. Seasonal images have also been used in land cover classification with some success [7] [8] [9].

With the increased availability of regular multi-temporal observations, the use of time-series data and time-series derived phenological indicators is becoming more popular for regional and national scale vegetation mapping. The approach often includes the use of Vegetation Indices (VI) to monitor dynamic changes of specific vegetation growing cycles as shown in Figures 1 and 2.

The combination of spectral- and temporal-based classification techniques was also demonstrated to produce a global land cover map in operational mode for the Glob-Cover initiative [10].



Figure 1 Phenological photographs of Fujihokuroku larch forest: vegetation growing cycles.
Source: Center for Global Environmental Research, National Institute for Environmental Studies, Japan.

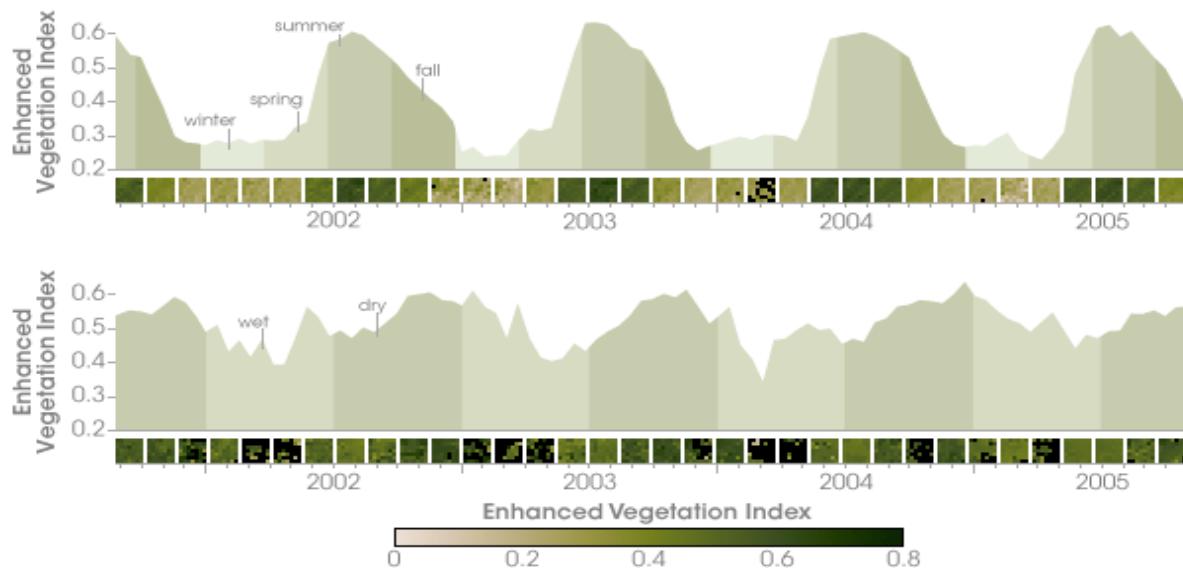


Figure 2 Seasonal differences in vegetation of two different forest type. Satellite measurements collected show seasonal changes in plant growth. The value of the enhanced vegetation index (a measurement of greenness) rises during the dry season, and falls during the wet season.
Source: <http://earthobservatory.nasa.gov/Features/AmazonEVI>.

Vegetation phenology represents a potentially significant source of land cover information. Since most of the land mass of the world is covered by vegetation, taking into account phenology when performing land cover classification may yield more accurate maps [11]. For these reasons remote sensing phenology, the use of satellites to track phenological events, is useful for assessing crop conditions, drought severity, and wildfire risk as well as tracking invasive species, infectious diseases, and insect pests. Because phenological events are sensitive to climate variation, these data also represent a powerful tool for documenting phenological trends over time and detecting the impacts of climate change on ecosystems at multiple scales [12].

1.2 Introduction to Remote Sensing

Remote sensing is a very broad field of studies. Some of the important applications of remote sensing technology are related to:

- global change detection and monitoring (global warming, deforestation, flooding, atmospheric ozone depletion, biomass);
- meteorology (atmosphere dynamics, weather prediction);
- mapping (topography, land use, leaf area index);
- forest and agriculture (vegetation condition, yield prediction);
- environmental assessment and monitoring (hazardous waste, soil erosion).

Remote sensing systems have been described in many aspects by numerous authors [13] [14] [15], and, particularly those deployed on satellites, provide a repetitive and consistent view of the earth that is invaluable to monitoring the earth system [9].

The modern era of earth remote sensing began with the first Landsat Multispectral Scanner System (MSS) in 1972, which provided for the first time a consistent set of high-resolution earth images.

The moderate imaging spectroradiometer system (MODIS), launched in 1999 onboard the Terra satellite, provides images in numerous spectral bands over a range 0.4 to 14 μm improving the quality of information that can be gathered about the earth's surface.

1.3 Remote Sensing of Vegetation

Approximately 70% of the Earth's land surface is covered with vegetation. Knowledge about variation in species and community distribution patterns, change in vegetation phenological cycles, and natural modifications in plant physiology and morphology provide invaluable insight into climatic, geological and physiographic characteristics of an area [16].

By using remote sensing data, vegetation can be distinguished from most other (mainly inorganic) materials by its nature of notable absorption in the red and blue segments of the visible spectrum, its higher green reflectance and, especially, its strong reflectance in the near-IR. As shown in Figure 3, different types of vegetation show distinctive variability from one another owing to such parameters as leaf shape and size, overall plant shape, leaf water content, and associated background (e.g., soil types and density of vegetative cover within the scene).

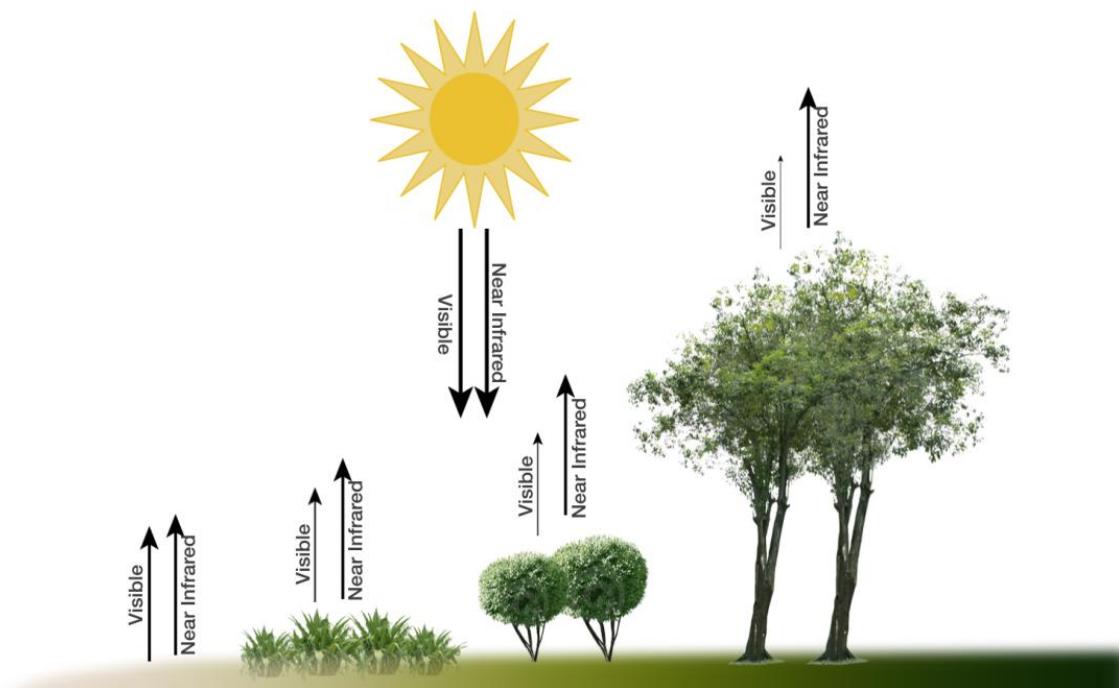


Figure 3 More vegetated areas absorb more visible light and reflect more near-infrared light back into space. Satellites can detect these relative differences, which are used to create vegetation indices.
Source; NASA/NOAA.

1.4 Leaf Reflectance

The reflectance from a leaf is determined by the leaf structure as well as the biochemical constituents of the leaf. As Figure 4 indicates, a typical leaf consists of several different layers with diverse optical characteristics. In the visible region of the spectrum the chlorophyll content controls the optical properties of the leaves. The chlorophyll absorbs the sunlight that makes photosynthesis possible. It is most absorptive in the blue and red regions. Here, as much as 70 to 90% of the incident radiation is removed. In the green region, the absorption is lower, which allows a large portion of the green light to be reflected.

In the near infrared spectrum, leaf reflectance is controlled by the structure of the spongy mesophyll tissue [17]. In this region, the healthy green leaf is characterized by high reflectance (40 – 60%), high transmittance (40 – 60%) through the leaves onto underlying leaves, and relatively low absorptance (5 – 10%).

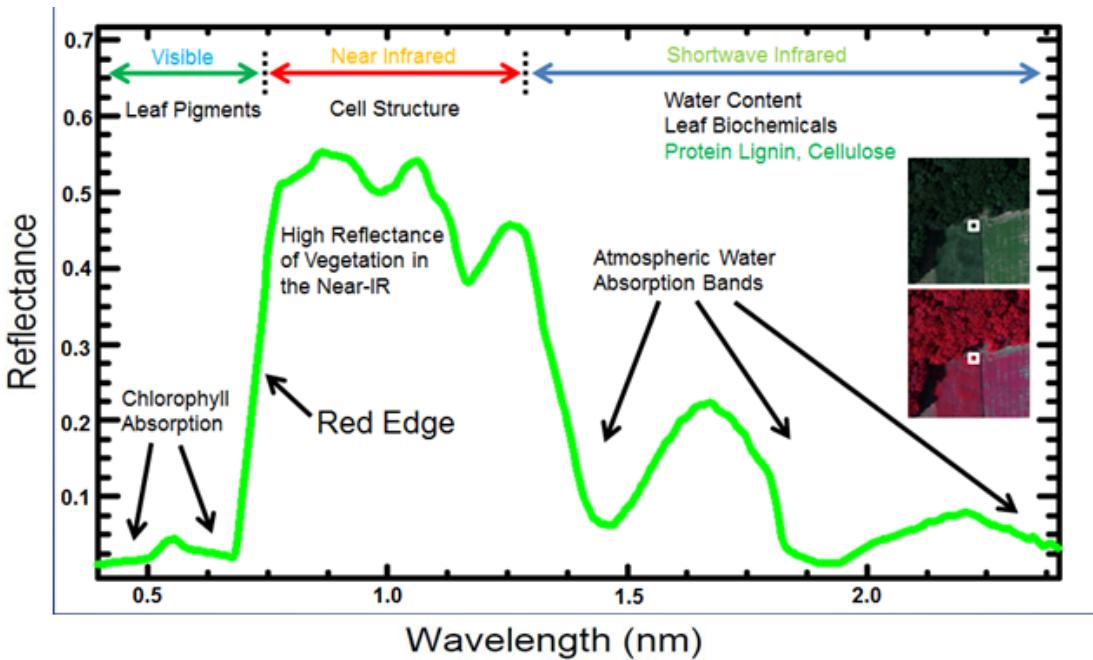


Figure 4 The vegetation spectrum typically has high absorbance in the red and blue wavelengths, high reflectance in the green wavelength, strongly reflects in the near infrared (NIR) wavelength, and displays strong absorption features in wavelengths where atmospheric water is present.

Source: www.markelowitz.com/Hyperspectral.html

Notice that a healthy green leaf reflectance and transmittance spectra throughout the visible and near-infrared spectrum are almost mirror images of one another [18].

Differences in reflective properties of plant species are more pronounced here than in the visible region [13].

1.5 Remote Sensing Phenology

Remote sensing phenology studies use data gathered by satellite sensors that measure wavelengths of light absorbed and reflected by green plants.

Many sensors carried aboard satellites measure red and near-infrared light waves reflected by land surfaces. Using mathematical formulas (algorithms), scientists transform raw satellite data about these light waves into vegetation indices, an indicator that describe the greenness, the relative density and health of vegetation for each picture element, or pixel, in a satellite image. Although there are several vegetation indices, one of the most widely used is the Normalized Difference Vegetation Index (NDVI), calculated from the visible and near-infrared light reflected by vegetation (Figure 5).

Calculations of NDVI for a given pixel always result in a number that ranges from minus one to plus one; however, no green leave gives a value close to zero. A zero means no vegetation and close to +1 indicates the highest possible density of green leaves [19].

Areas of barren rock, sand, or snow usually show very low NDVI values, for example, 0.1 or less. Sparse vegetation such as shrubs and grasslands or senescent crops may result in moderate NDVI values, approximately 0.2 to 0.5.

High NDVI values, approximately 0.6 to 0.9, correspond to dense vegetation such as that found in temperate and tropical forests or crops at their peak growth stage.

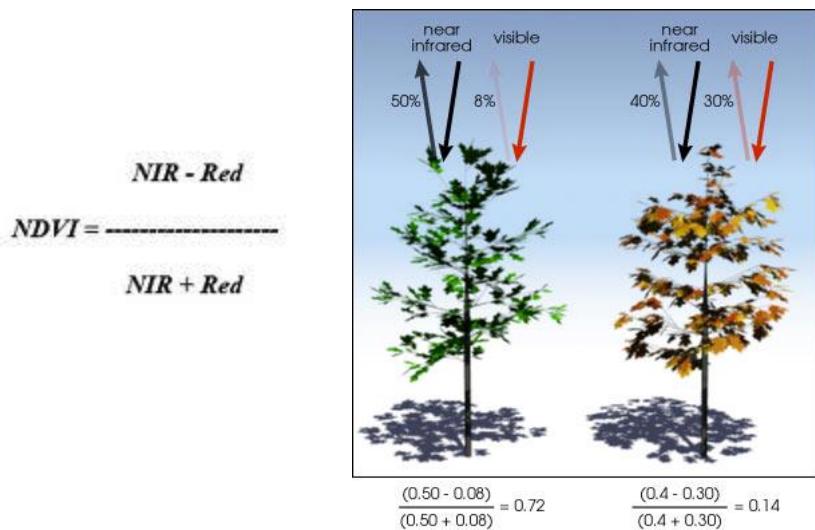


Figure 5 Basic principle of NDVI. (A) The healthy vegetation absorbs more Red light than the unhealthy or stressed vegetation shown in (B). Source: <http://earthobservatory.nasa.gov>.

By transforming raw satellite data into NDVI values, researchers can create images and other products that give a rough measure of vegetation type, amount, and condition on land surfaces around the world. NDVI is especially useful for continental to global-scale vegetation monitoring because it can compensate for changing illumination conditions, surface slope, and viewing angle.

NDVI values can be averaged over time to establish "normal" growing conditions in a region for a given time of year as illustrated in Figure 6.

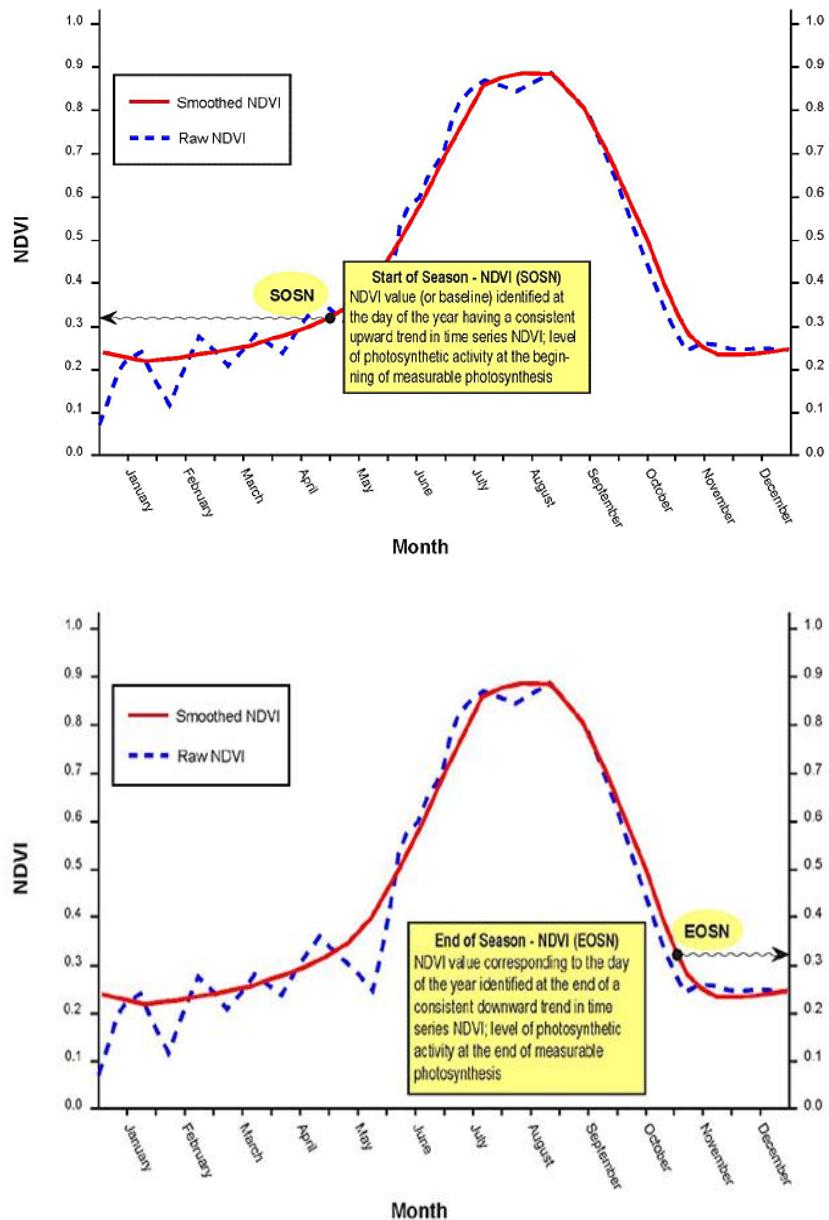


Figure 6 Time-series NDVI temporal curve that summarizes the various stages that green vegetation undergoes during a complete growing season. Such curves can be analyzed to extract key phenological variables, such as the start of the growing season (SOS) and end of the season (EOS).

Source: http://phenology.cr.usgs.gov/methods_metrics.php.

1.6 Background on Image Classification

Image classification is the process used to produce thematic maps, which shows the spatial distribution of identifiable earth surface features and provides an informative description of a given area. As a result, the image is partitioned into some non-intersecting regions, such that each region is homogeneous and the union of two adjacent regions is heterogeneous.

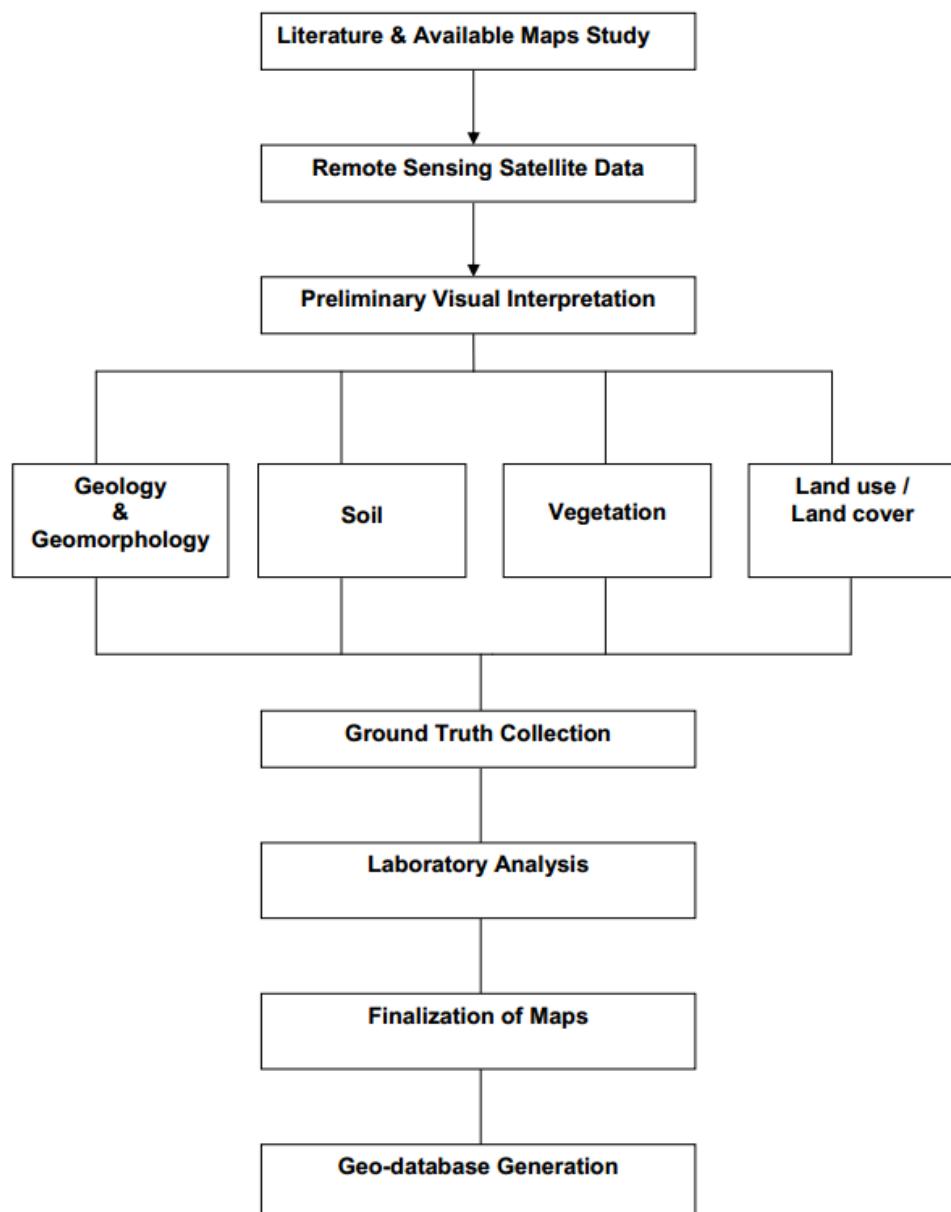


Figure 7 Image classification steps to produce thematic maps.

The use of remote sensing data for vegetation study can be seen in two ways:

The traditional approach is called image-centered. Its primary purpose is in determining spatial relationships among objects and features on the ground (pixel or object-based classification). In fact, the goal of image-centered analyses is creation of a map [20]. Various degrees of success in image classification have been achieved with different sources of images at different scales, from species mapping at stand scale to land cover mapping at regional, continental, and global scale [21] [22].

The second approach applied in remote sensing of vegetation is called data-centered. Here, the information inferred from remote sensing data itself is of interest, rather than the spatial relationship of features on the ground. The most common approach is to make use of the reflectance measurement from remote sensing data for deriving biophysical parameters, either by applying empirical or process-based models. For example, phenological cycles of vegetation can be determined by analyzing the temporal variation of NDVI from time series of satellite images [23]. The results and products of data-centered analysis should also be presented in the context of spatial maps in order to fully understand the spatial distribution and behavior of the biophysical parameters. In recent years, increasing interest in global change and in long-term monitoring of the human effects on environment has led to the use of remote sensing data at global scale [23].

In this sense, the two approaches, image-centered and data-centered, converge.

2 Data and Methodology

2.1 MODIS Data: Preliminary Study

Moderate Resolution Imaging Spectroradiometer (MODIS) is a key instrument aboard the Terra (EOS AM) and Aqua (EOS PM) satellites. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths.

The first MODIS Flight Instrument, (integrated on the Terra (EOS AM-1) spacecraft successfully launched on December 18, 1999) and the second MODIS flight instrument (integrated on the Aqua (EOS PM-1) spacecraft launched on May 4, 2002) provide high radiometric sensitivity (12 bit) in 36 spectral bands ranging in wavelength from 0.4 μm to 14.4 μm . Two bands are imaged at a nominal resolution of 250 m, five bands at 500 m, and the remaining 29 bands at 1 km. One of the unique features of the MODIS instrument is its Direct Broadcast capability: in addition to storing data for later download at designated intervals, MODIS immediately broadcasts the raw data it collects from all 36 spectral bands. The Terra MODIS instrument was one of the first satellites to constantly broadcast data for anyone with the right equipment and software to download, free of charge [24].

The responses, global maps of several land surface characteristics, including surface reflectance, albedo, land surface temperature, and vegetation indices are custom tailored to the individual needs of the user community and provide exceptionally low out-of-band response, offering an unprecedented look at terrestrial, atmospheric, and ocean phenomenology for a wide and diverse community of users throughout the world [25].

In this study we tested the possibility of using MODIS NDVI 10 days time-series together with single date high resolution imagery to overcome the single date classification inaccuracy mainly due to phenological vegetation cycles. Is worth noting that at the beginning of this research study (beginning of 2013), Landsat 8 was not yet launched and the main source of free satellite data was offered by Landsat 4/5/7. Limitations imposed by the low acquisition frequency of the high resolution imagery together with dissemination policy, satellite problems (Slc-off) and weather condition (cloud/haze coverage), might sometimes bring only few images per year on specific areas for Landsat path 172, row 064, covering Mahale Mountains National Park in Tanzania [26].

A preliminary study aiming to bypass the single date classification inaccuracy (mainly due to seasonality) using long term MODIS time series as a “driver” to fill gaps between high resolution data, has been carried out. The high temporal resolution (4 acquisitions per day using Terra and Aqua satellites and 10 days NDVI synthesis products) might compensate the low spatial resolution of about 250 m. The long term NDVI time series allows a better understanding of the vegetation phenology over time bypassing the lack of high resolution imagery.

For this study, MODIS NDVI v5 datasets (MOD13Q1 and MYD13Q1) were acquired from the NASA MRTweb tool interface [27].

The Mahale Mountains study area is covered by two adjacent tiles, namely h20v09 and h21v09, shown in Figure 8, for a total of 138 individual files from beginning of January 2012 to end of July 2013.

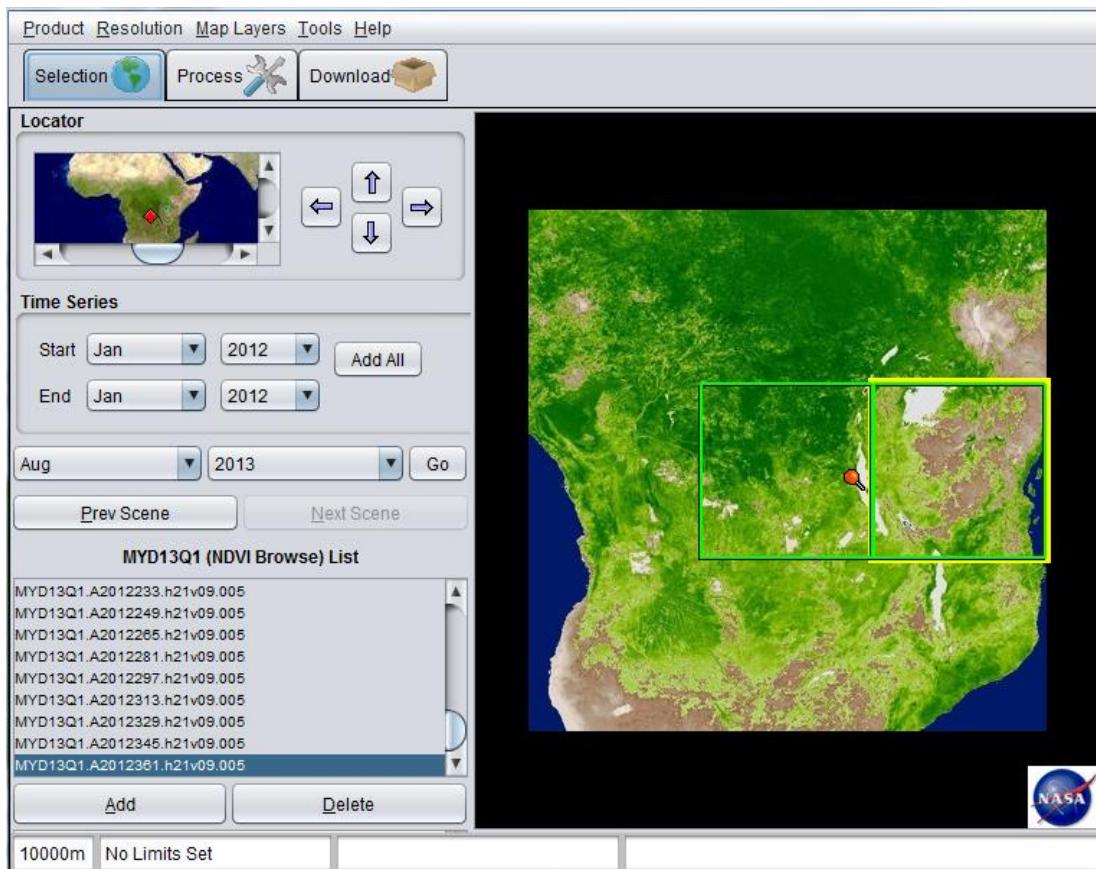


Figure 8 MODIS NDVI v5 datasets (MOD13Q1 and MYD13Q1) acquired from the NASA MRTweb tool interface. Mahale Mountains study area is covered by two adjacent tiles, namely h20v09 and h21v09 for a total of 138 individual files from beginning of January 2012 to end of July 2013.

Before submitting the order we specified a WGS84 lat/lion reference system instead of the original sinusoidal projection and a mosaicked GeoTiff output instead of the original HDF tiling format. No spatial clip has been done at this stage in order to download the whole area for further analysis. The download phase is performed batching the “wget” Unix command instead of manually clicking the 69 links provided by the MRTWeb tool (download tab).

Clip to Mahale Mountains National Park boundaries and layer stacking are done using Geospatial Data Abstraction Library (GDAL) libraries [28]:

```
- gdal_merge.py -o MODIS_16d_NDVI_2012_2013.tif -separate *.tif -co  
"COMPRESS=LZW"  
- gdalwarp MODIS_16d_NDVI_2012_2013.tif 7521_MODIS_16d_NDVI_2012_2013.tif -  
cutline 7521.shp -crop_to_cutline -co "COMPRESS=LZW"
```

2.1.1 Data Smoothing

The reflected light waves that satellite sensors detect coming from vegetation on the Earth's surface can be altered or blocked by a variety of phenomena, including aerosols and clouds in the atmosphere as well as changing illumination patterns and the angle at which the satellite views the ground at any given time. These phenomena introduce "noise" into raw satellite data. To address this problem, raw data are processed using techniques that filter out noise and produce a clearer, more representative data set. Filtering techniques are diverse. Compositing/merging maximum NDVI values acquired over (typically) 7-, 8-, 10-, 14-, or 16-day intervals increases data quality. But residual effects of sub-pixel clouds, prolonged cloudiness, and other negative elements require further processing in the form of data smoothing. Data smoothing facilitates time-series analysis by reducing aberrant, noise-induced peaks and valleys that appear when NDVI values are plotted graphically to reveal vegetation changes over time as shown in Figure 10 [29]. For this research study an in-house smooth filter has been implemented in IDL language [30] following the schema in Figure 9.

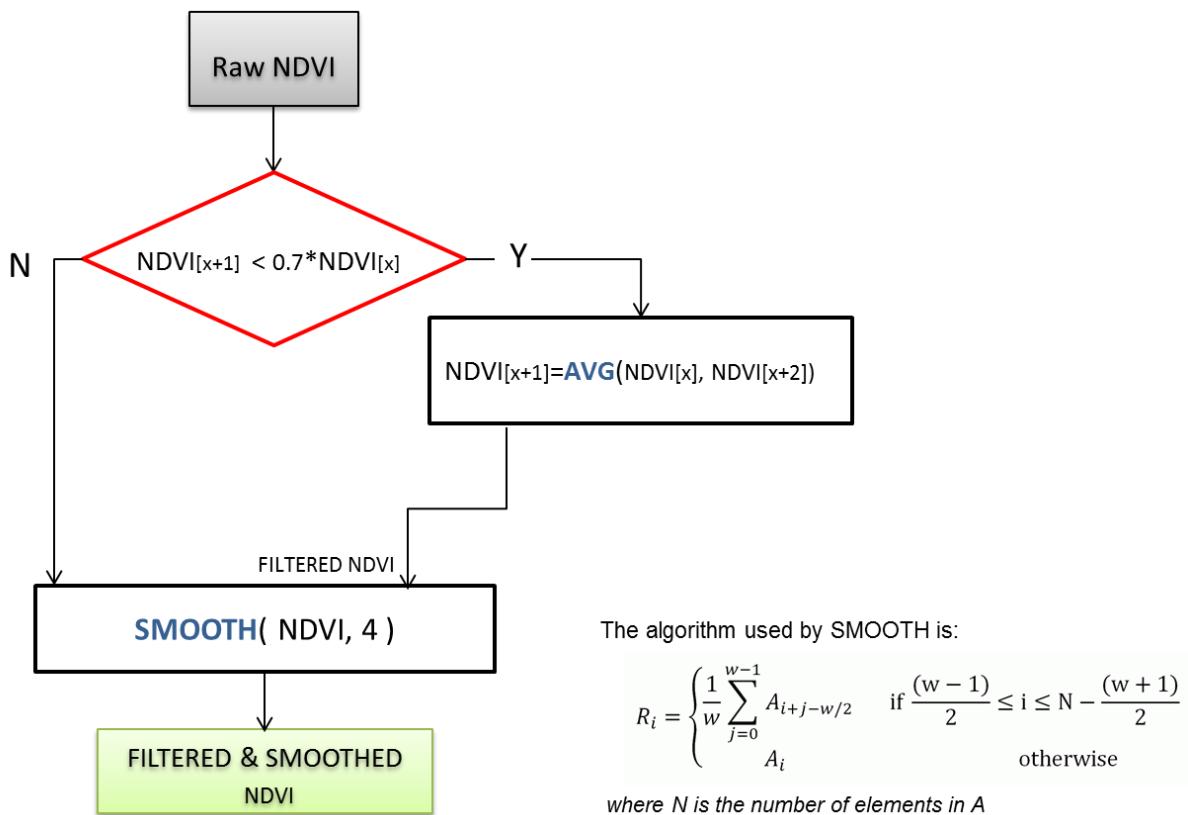


Figure 9 Scheme used for implementation of smooth filters using IDL language.

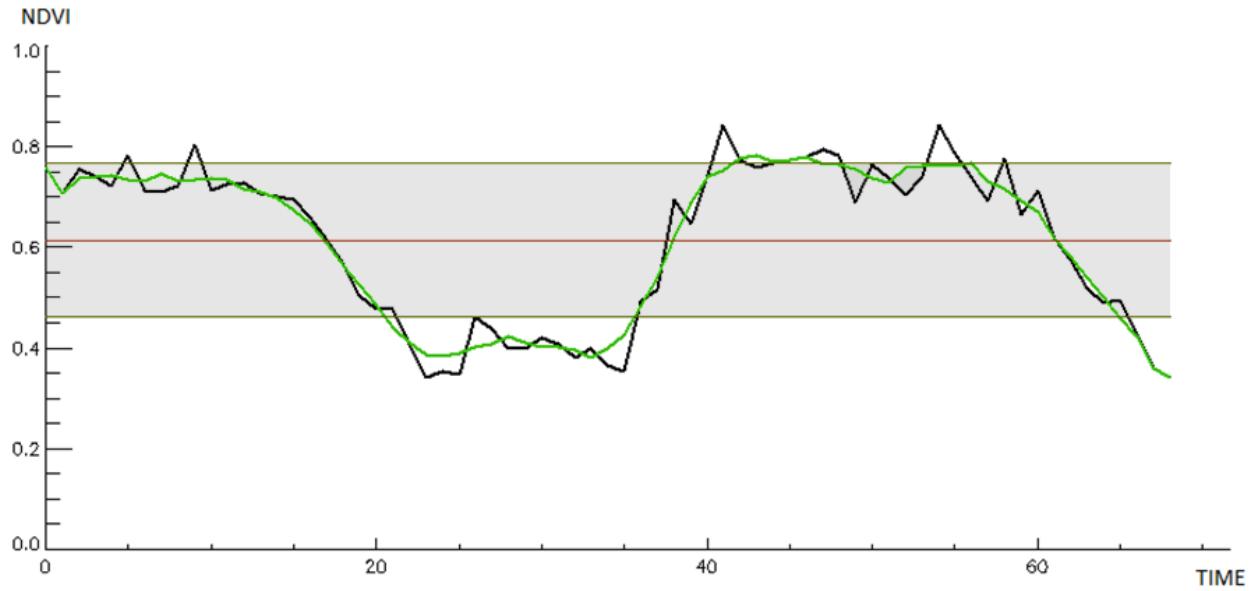


Figure 10 The black line represents the raw NDVI, the green line the filtered and smoothed NDVI, red line indicates average NDVI and the olive lines indicate ± 1 standard deviation.

2.1.2 Use of NDVI Index

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

Calculated for each year or growing season, these metrics are the basis for diverse research and monitoring application [31].

Despite several techniques of data fusing [32] have been proposed in literature, we realize that the spatial resolution of the MODIS data is not enough to detect and keep small (within the 50 m range) but important land cover types especially in fragmented landscapes. The accuracy of the land cover map derived from remote sensing data depends on the spatial resolution of data. Especially, in fragmented landscape, most pixels at 1 km and 250 m resolution are a mixture of several land cover types.

Previous results from the MODIS validation team have suggested that the MODIS land cover product is realistic, and that the algorithm performs well at the global scale. At site scale, comparing to ground truth maps derived from different sources, fragmentation in landscapes is a fundamental problem encountered in the use of MODIS products [33].

In order to detect smaller areas which can be classified in a particular land cover class like small field plots or small vegetation openings is necessary a better resolution, from 3 to 80 meters.

While working on this topic the USGS launched the latest Landsat satellite (Landsat 8) and few months later (April 2013) any acquired image was made available at no charge. Thanks to the 16 days revisiting time and dissemination policy (all acquired images are distributed in a near-real time), few months after we could build a short NDVI time series at 30 m resolution of the study area as shown in Figure 11. The ideal situation of having high spatial and temporal resolution imagery for land cover mapping purposes was now becoming reality and we decided to invest more on this new product.

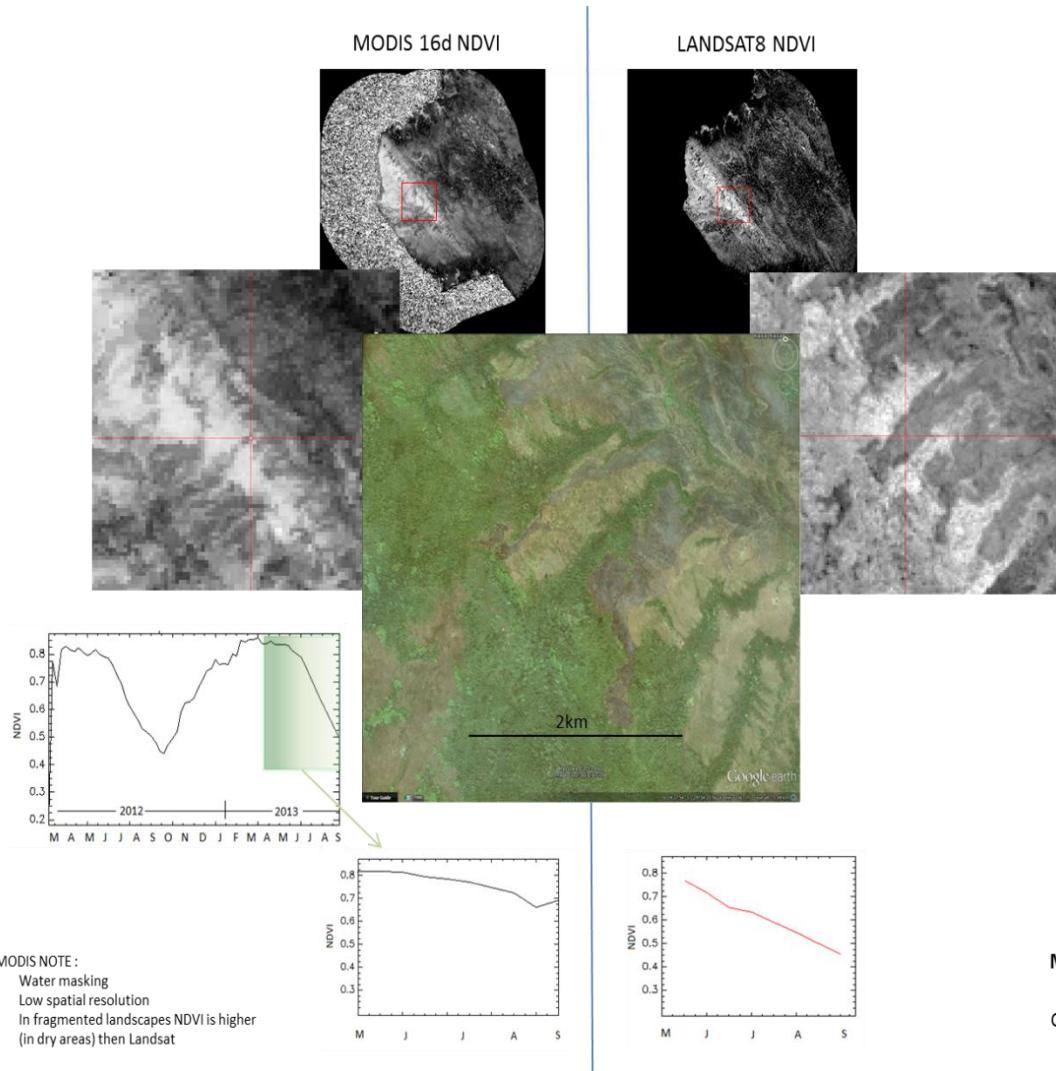


Figure 11 Difference between MODIS 16 days NDVI and Landsat 8 NDVI.

2.2 LANDSAT Data

Landsat represents the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data [34].

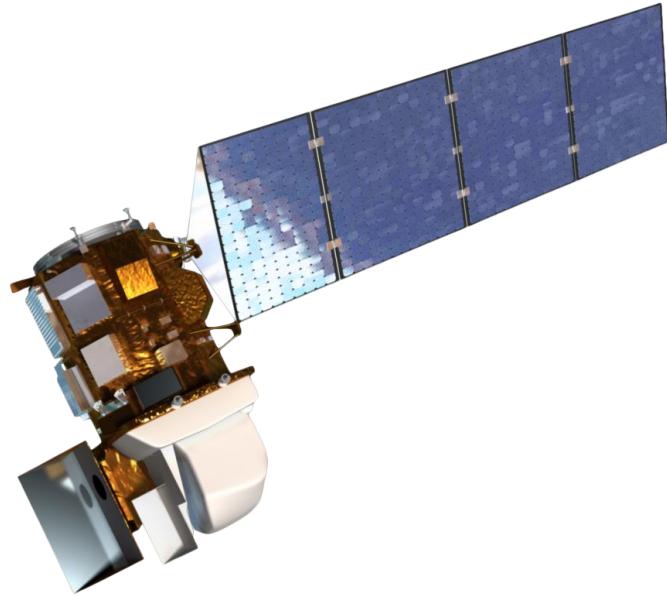


Figure 12 Landsat 8 satellite. Source: <http://landsat.gsfc.nasa.gov>.

In 1965, NASA began methodical investigations of Earth remote sensing using instruments mounted on planes and few years later, in 1972 Landsat 1 was launched, heralding a new age of satellite remote sensing of land from space. For over 40 years, the Landsat program has collected spectral information from Earth's surface, from seven improved Landsat satellites, creating a historical archive unmatched in quality, detail, coverage, and length. On May 30, 2013, data from the Landsat 8 satellite became available. As with previous partnerships, this mission continues the acquisition of high-quality data that meet both NASA and USGS scientific and operational requirements for observing land use and land change [34].

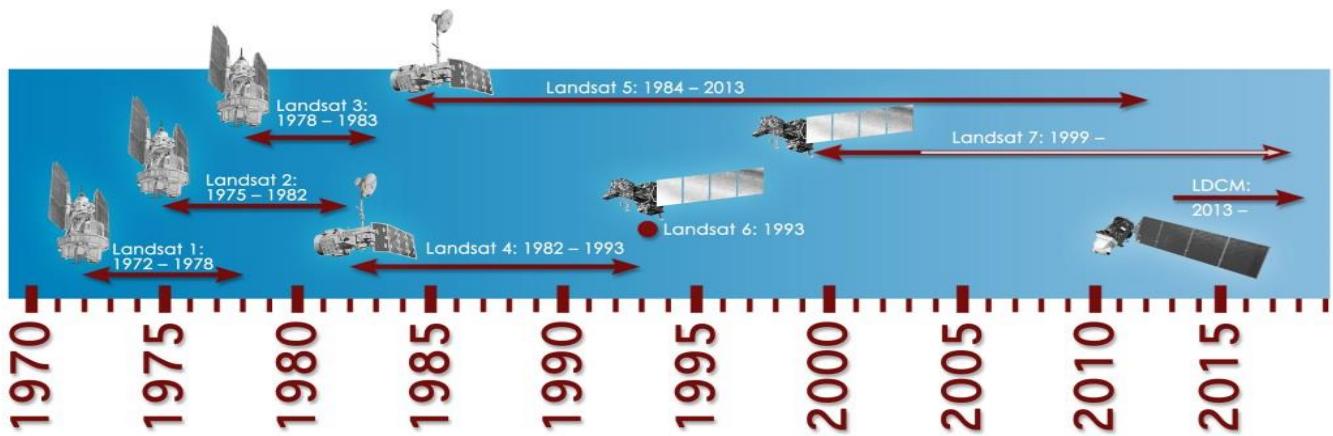


Figure 13 Landsat Timeline from Landsat 1 to Landsat 8. Source: <http://landsat.gsfc.nasa.gov>.

Landsat 8 carries two improved instruments, the Observational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), that together observe the same wavelengths of light as earlier Landsat satellites and measure different ranges of frequencies along the electromagnetic spectrum. Each range is called a band, and Landsat 8 adds two new bands, 1 and 9. Additionally, the single thermal infrared band sensed by previous Landsat instruments is split into two thermal bands to help improve sensitivity to surface temperature. Landsat 8 also improves the radiometric quality of the imagery, for example by increasing the number of bits used to represent each pixel value in an image and using an improved range of 4096 potential grey levels in an image respect to only 256 grey levels in previous 8-bit instruments [35]. The following Figure 14 shows the wavelengths from both Landsat 7 and Landsat 8 and the specifications of the 11 bands.

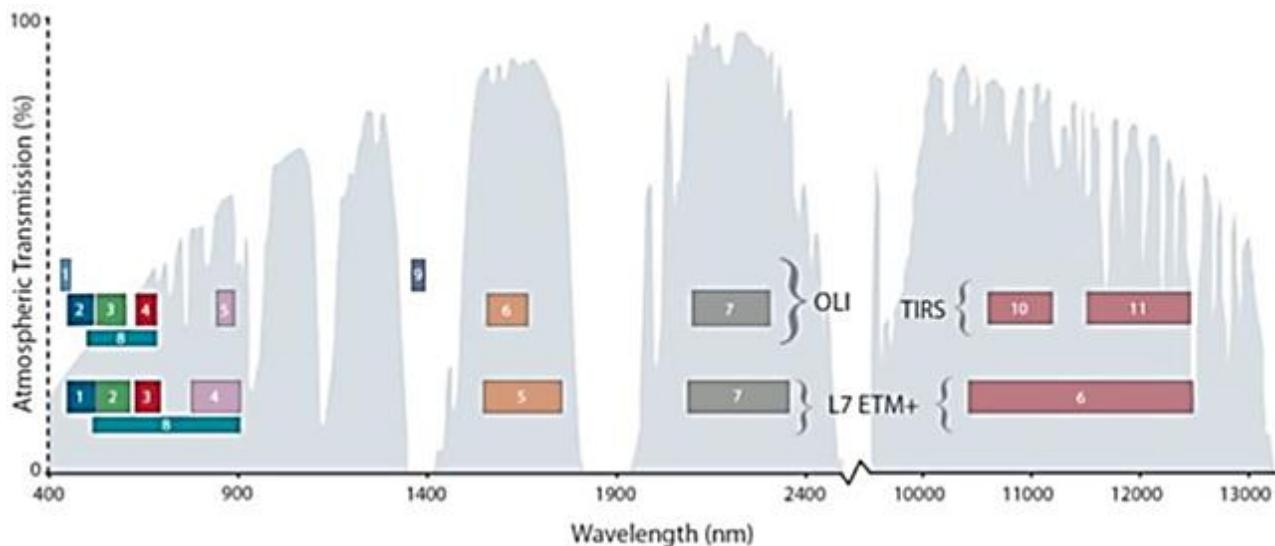


Figure 14 Bandpass wavelength for Landsat 8 OLI and TIRS sensor, compared to Landsat 7 ETM+ sensor. Landsat 7 (bottom row) and Landsat 8 (top row). Source :<http://landsat.gsfc.nasa.gov>.

The 30 m spatial resolution and 185 km swath of Landsat imagery fills an important scientific niche because the orbit swaths are wide enough for global coverage every season of the year, yet the images are detailed enough to characterize human-scale processes such as urban growth, agricultural irrigation, and deforestation.

Land imaging offers the critical and irreplaceable capability to observe land use and land use change across those scales [36] and provides repetitive and synoptic observations of the Earth otherwise unavailable to researchers and managers who work across wide geographical areas and applications.

Worldwide, millions of people are helped by Landsat-data-based decisions that impact food and water management, human health, agriculture, climate, energy, fire, natural disasters, urban growth, ecosystems and biodiversity, forest management, land use and land cover [37].

Landsat 8 satellite images the entire Earth every 16 days in an 8-day offset from Landsat 7. Data collected by the instruments onboard the satellite is available to download from the LandsatLook Viewer within 24 hours of reception [36].

The USGS LandsatLook Viewer [38] uses ArcGIS Server technology and allows comprehensive searching and downloading of full-resolution Landsat images.

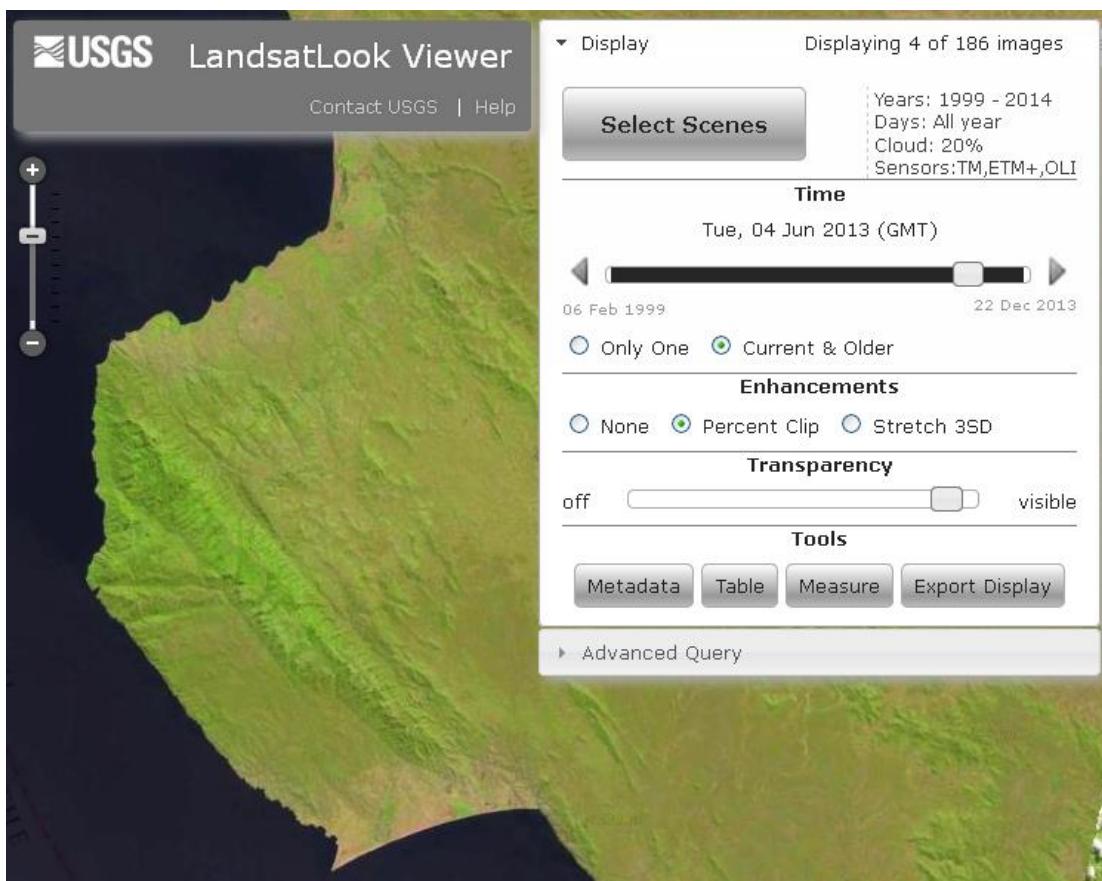


Figure 15 Image of Mahale Mountains National Park. Scene id: LC81720652013155LGN00, Sensor OLI, Acquisition Date: 04/06/2013. Source: <http://landsatlook.usgs.gov>.

2.2.1 Use of New LANDSAT 8

Recently, research communities had the possibility to download Landsat 8 imagery acquired since 12th April 2013 at no charge throughout the USGS website.

The high global acquisition frequency (~16 days) and distribution policy (all images are available almost at near real time) are making this product extremely suitable for land cover mapping and monitoring as well as for land cover/use change detection. For this study we did not consider imagery that were not freely available or covered by specific agreement/copyrights.

Image below shows the availability of satellite imagery from Landsat sensors (TM, ETM+, OLI/TIRS) spanning from April to September over path 172, row 064 (Mahale Mountains National Park, Tanzania) from 1998 to 2013; in 2003 the Landsat 7 faced a partial sensor failure visible in the remote sensed image as black horizontal lines limiting the study to the central part only. Despite several techniques of gap filling have been implemented, the launch of the new Landsat 8 satellite offers a valid source of information for researchers and has the highest revisiting frequency, data availability and quality among its predecessors.

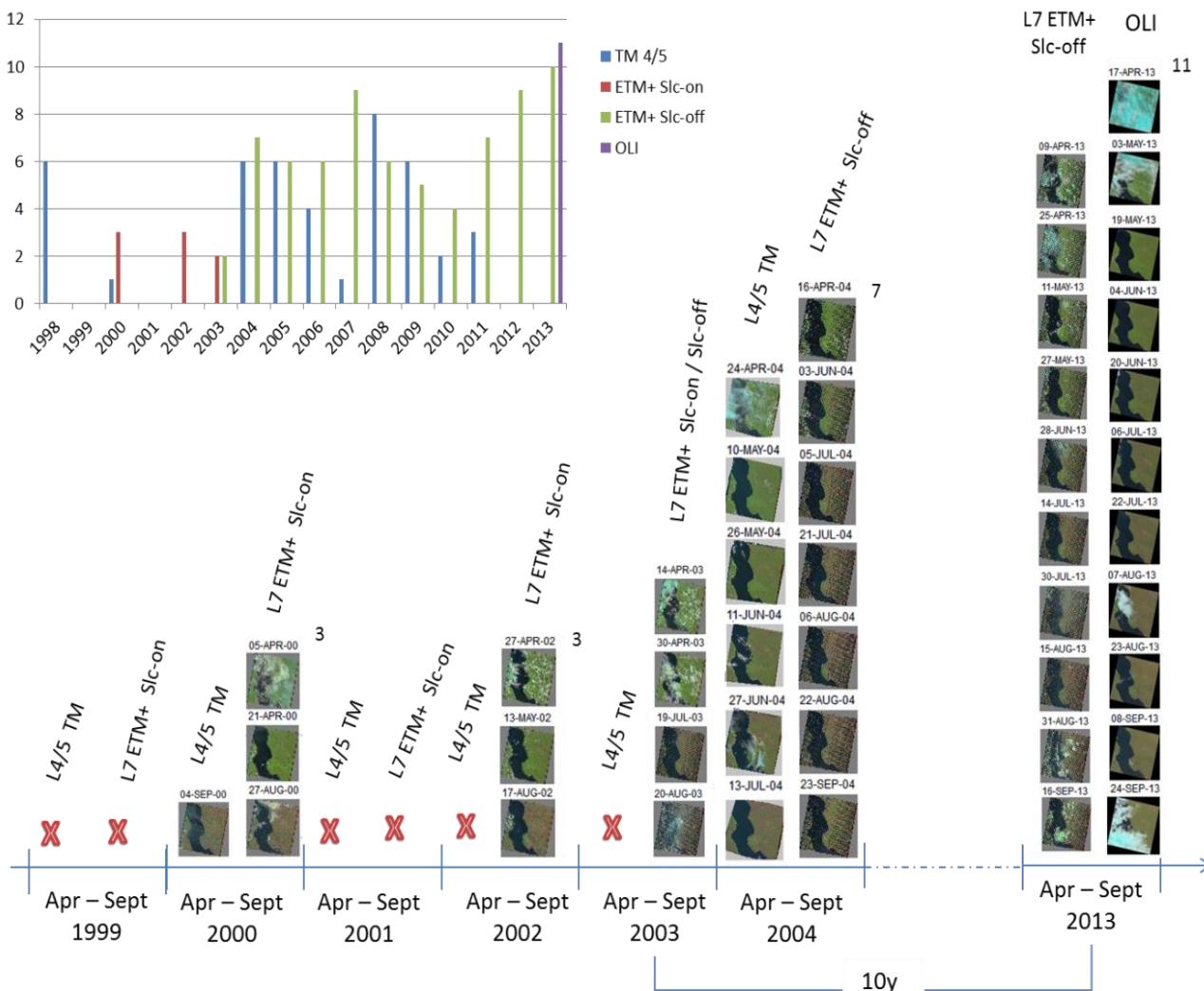


Figure 16 Availability of satellite imagery from Landsat sensors (TM, ETM+, OLI/TIRS) spanning from April to September over path 172, row 064 (Mahale Mountains National Park, Tanzania) from 1998 to 2013;

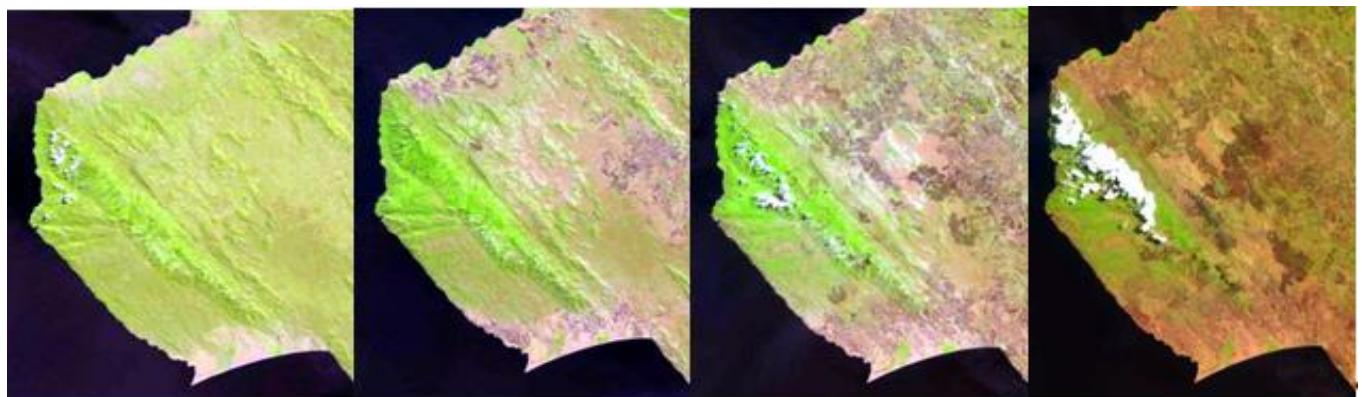
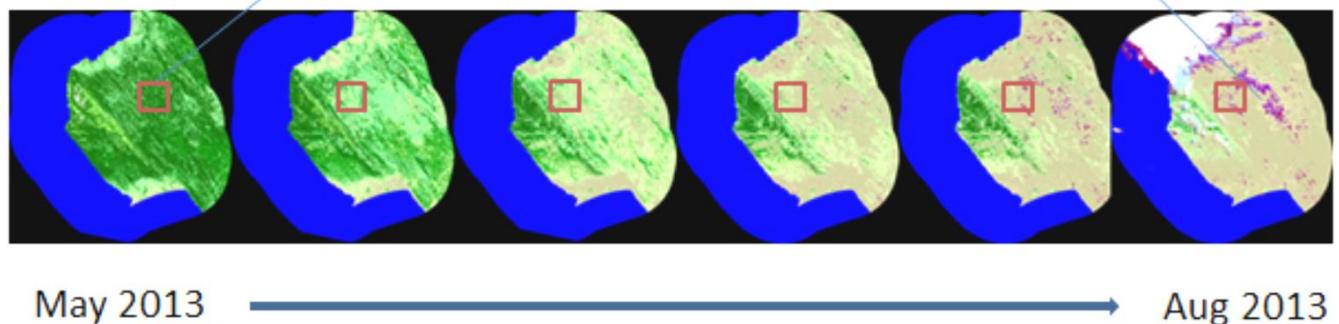
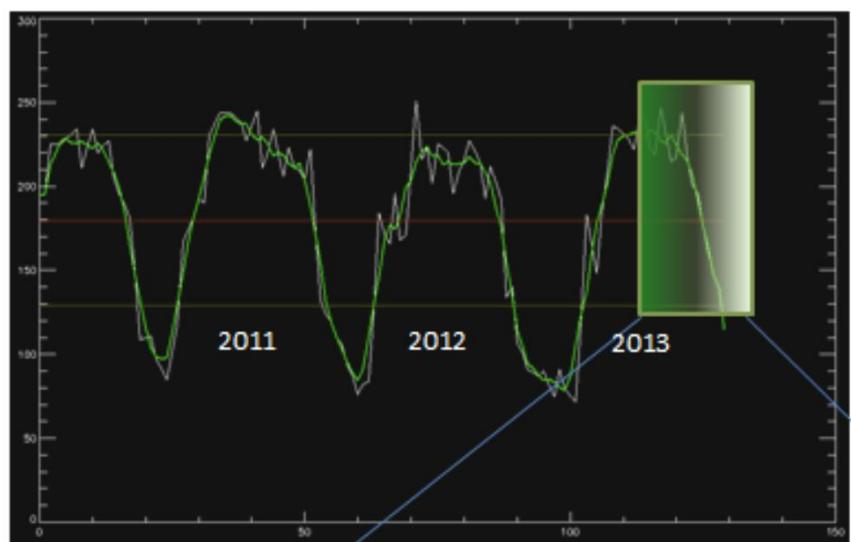


Figure 17 Availability of Landsat 8 and MODIS NDVI profile from May to August (Mahale Mountains National Park, Tanzania).

2.3 Study Areas and Satellite Imagery

Landsat 8 sensor has been identified as the main source of remote sensed data for our study, free of charge and having a spectral resolution high enough (30 m) to map regional fragmented landscapes. Since data are available from April onwards, our study is focusing on East Africa, a semi dry region characterized by a peak of vegetation growth around that month. Satellite series will be able to catch the vegetation evolution from green to dry.

The selection procedure was focused on:

- semi dry regions where the seasonal behavior of the vegetation is strongly reflected in the single date automatic classification.
- Areas not affected by fast land cover changes (anthropological changes) that might be detected in the study time frame (April-September); consequently we restrict the scope over protected areas and 20km buffer zone with a limited human density/influence.
- Availability of cloud free Landsat 8 imagery. Scenes totally covered by clouds were not acquired.

Imagery from NASA's Landsat 8 covering each study areas are listed on Table 1.

Area	Path/Row	Date 1	Date 2	Date 3	Date 4	Date 5	Date 6	Date 7	Date 8
Mahale M.	172/064	19/05	04/06	20/06	06/07	22/07	07/08	23/08	08/09
West Lunga	174/069	15/04	01/05	17/05	18/06	04/07	20/07	05/08	06/09
Gorongosa	167/063	30/04	16/05	17/06	03/07	19/07	04/08		
Mana Pools	171/071	26/04	12/05	28/05	13/06	29/06	15/07	16/08	
Tsimanam.	160/077	13/04	29/04	15/05	31/05	02/07	03/08	04/09	

Table 1 In first column, path/row of Landsat 8 images covering each study area (Mahale Mountains, West Lunga, Gorongosa, Mana Pools and Tsimanampetsotsa National Parks). Number of Path and Row in second column.

Selected study areas are presented in Figure 18.

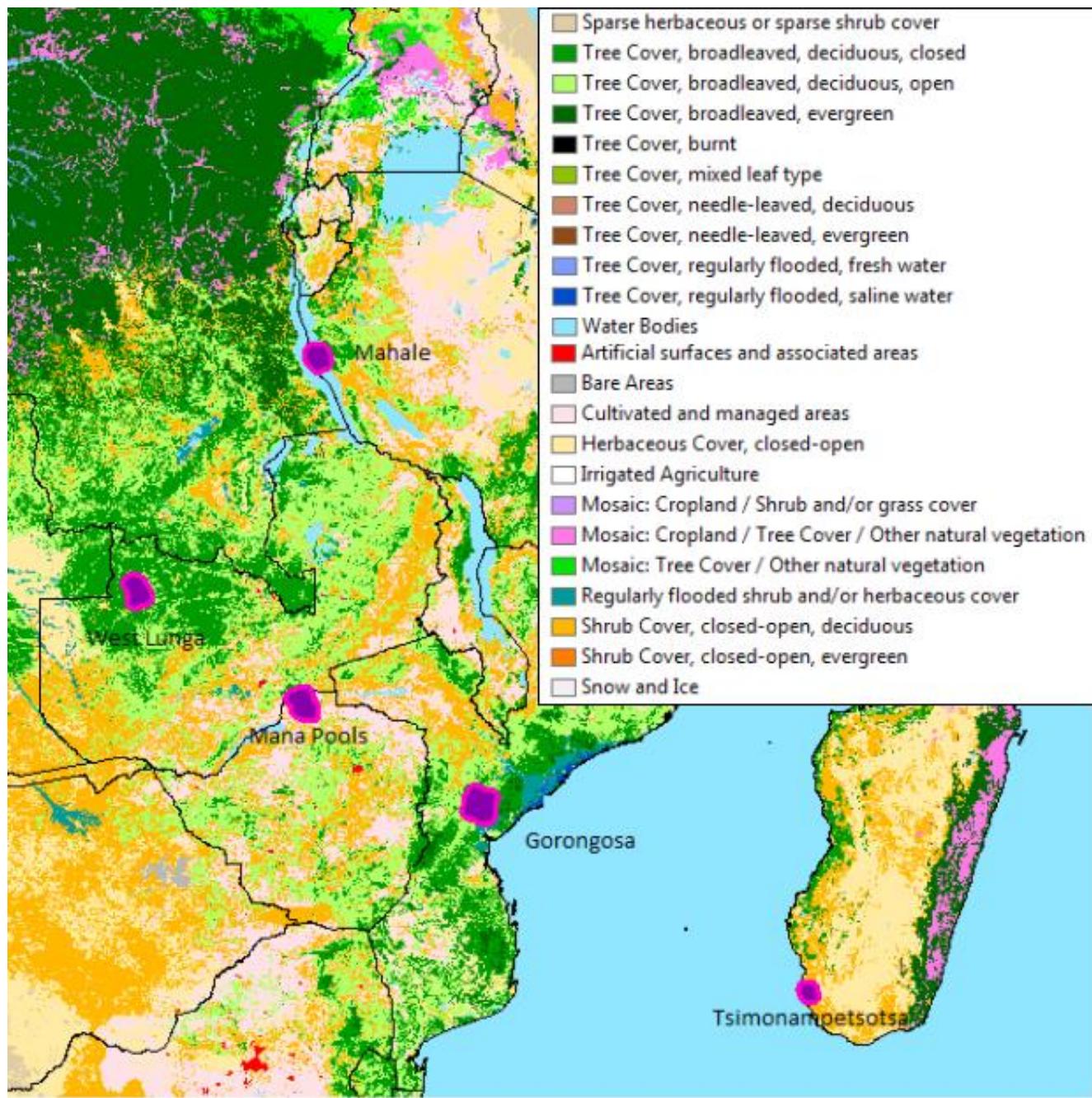
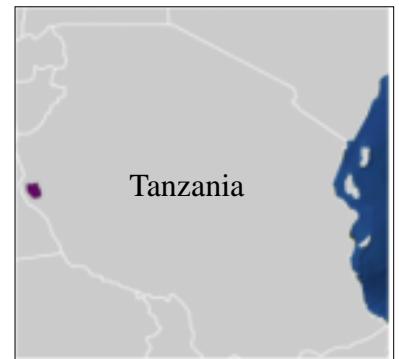


Figure 18 Map of Africa based on GLC 2000 [39], showing National Parks considered in this study.

Specific information on protected area is provided by The World Database on Protected Areas which is held at the UNEP-World Conservation Monitoring Centre (UNEP-WCMC) and is managed in partnership with IUCN - The World Conservation Union [40].

2.3.1 Mahale Mountains National Park

Mahale Mountains National Park, in Tanzania, was gazetted in 1985, covers an area of 1613 km² and is located about 128 km south of Kigoma town on the eastern shore of Lake Tanganyika. The western boundary of the park protects an adjacent 1.6 km wide strip of Lake Tanganyika's waters.



The terrain is mostly rugged and hilly, and is dominated by the Mahale Mountains chain that runs through the heart of the park. The tallest peaks reach as high as 2460 meters. The effect of the topography on precipitation is immediately apparent in the scene, as dense, green forests exist on the western side, while browner, more arid landscapes appear on the other as shown in Figure 20.

Miombo woodland covers about three quarters of the park. This is criss-crossed by narrow strips of riverine forest that grow along some watercourses [41].

Figure 19 Mahale Mountains National park position in Tanzania.



Figure 20 Landsat8 from 20-06-2013
RGB: band 753.

Where the mountain chain converges with the lake, there is a broad blanket of lowland forest, known locally as 'Kasoge', which extends up to 1300 m a.s.l. The Kasoge forest is an enclave of the central-African, tropical, semi-deciduous forest type. Above 2300 m the forest gives way to montane grassland [42]. This peculiar habitat mosaic results from the combined presence of Lake Tanganyika and the Mahale Mountains, which affect climatic conditions on a very localized scale. There are two seasons in the park. Generally, the dry season starts in mid-May and ends in mid-October.

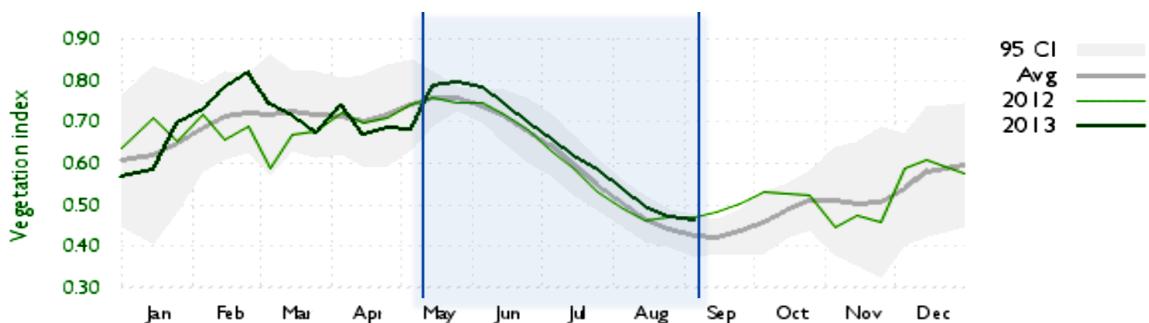


Figure 21 NDVI Trends for Mahale Mountains National Park. Blue vertical lines define the period of the year considered in the study. The dark gray line is the average for each dekad based on the available timeseries, and the light grey areas indicate the 95% confidence limits around this average. Source: African Protected areas.

There are no roads or other infrastructures within the park boundaries, and the only way in and out of the park is via boat on the lake. An important unusual things about Mahale vegetation is the wide array of habitat types that it contains. The park is a mosaic of overlapping rainforest, woodland, bamboo forest, montane forest and grasslands [43].

The most recent mapping study is the one proposed in Mahale Mountains National Park General Management Plan (2006-2016) [44] developed through a participatory planning process involving a cross section of Mahale Mountains National Park stakeholders, under the coordination of a Core Planning Team comprising representatives from Tanzania National Parks Headquarters, Mahale Mountains National Park managers, the Frankfurt Zoological Society and Conservation Development Centre.

According to TANRIC map, shown in Figure 22, there are 9 categories of vegetation present in MMNP. This includes evergreen or Kasoge forest, large areas of Miombo woodland, montane forest and high altitude grassland. This classification roughly coincides with the results of a study by the Japan International Cooperation Agency that simplified MMNP's vegetation classification into five broad types.

It describes the Miombo woodland as the dominant vegetation type covering up to 75% of the park's area. As altitude increases this is replaced by mountain forest and high altitude grassland and mountain bushland. In the northwest of the park there is an isolated patch of acacia woodland.

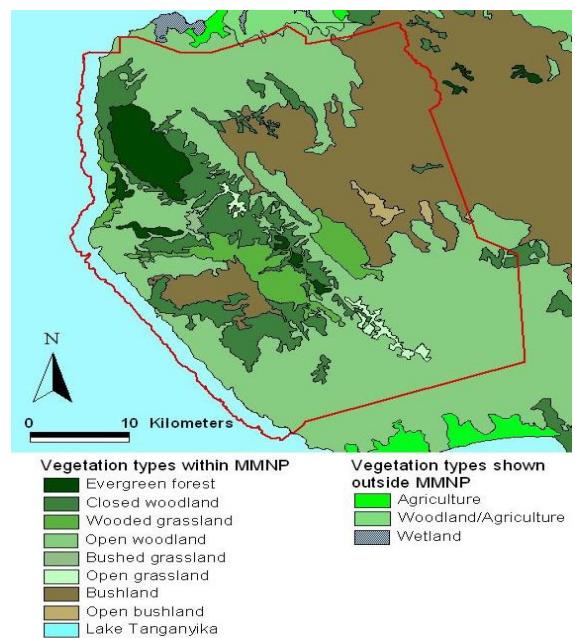


Figure 22 TANRIC Mahale Mountain National Park Land-Cover map provided by Frankfurt Zoological Society and Conservation Development Centre [44].

More information about biodiversity of Mahale Mountains National Park can be found in “Biodiversity of Mahale Mountains National Park” report [43].

2.3.2 Mana Pools National Park

Mana pools is a wildlife conservation area in northern Zimbabwe constituting a National Park (Figure 23). It is a region of the lower Zambezi River in Zimbabwe where the flood plain turns into a broad expanse of lakes after each rainy season. As the lakes gradually dry up and recede, the region attracts many large animals in search of water, making it one of Africa's most renowned game-viewing regions.

Mana means ‘four’ in Shona, in reference to the four large permanent pools formed by the meanderings of the middle Zambezi. These 2500 km² of river frontage, islands, sandbanks and pools, flanked by forests of mahogany, wild figs, ebonies and baobabs, is one of the least developed National Parks in Southern Africa. It was saved from a hydro-electric scheme in the early eighties which would have seen the flooding of this subsequent World Heritage site. It has the country’s biggest concentration of hippopotamuses and crocodiles and large dry season mammal populations of elephant and buffalo [45] [46].



Figure 23 Mana Pools National Park position in Zimbabwe.



Figure 24 Landsat8 from 12-05-2013
RGB: band 753.



Figure 25 NDVI Trends for Mana Pools National Park. Blue vertical lines define the period of the year considered in the study. The dark gray line is the average for each dekad based on the available timeseries, and the light grey areas indicate the 95% confidence limits around this average. Source: African Protected areas.

There was no comprehensive baseline data on vegetation in Zimbabwe and no GIS-based land cover map of Mana Pools National Park was created.

Landform classification within these regions is based on the Soil and Terrain Database of Zimbabwe, compiled by the Soil Survey and Cartography Section of the Chemistry and Soil Research Institute with technical assistance of ISRIC [47].

One of the major activities of “The Impact of Desertification on Food Security in Southern Africa: a Case Study in Zimbabwe”, was to compile a Soil and Terrain Database (SOTER) for Zimbabwe [48].

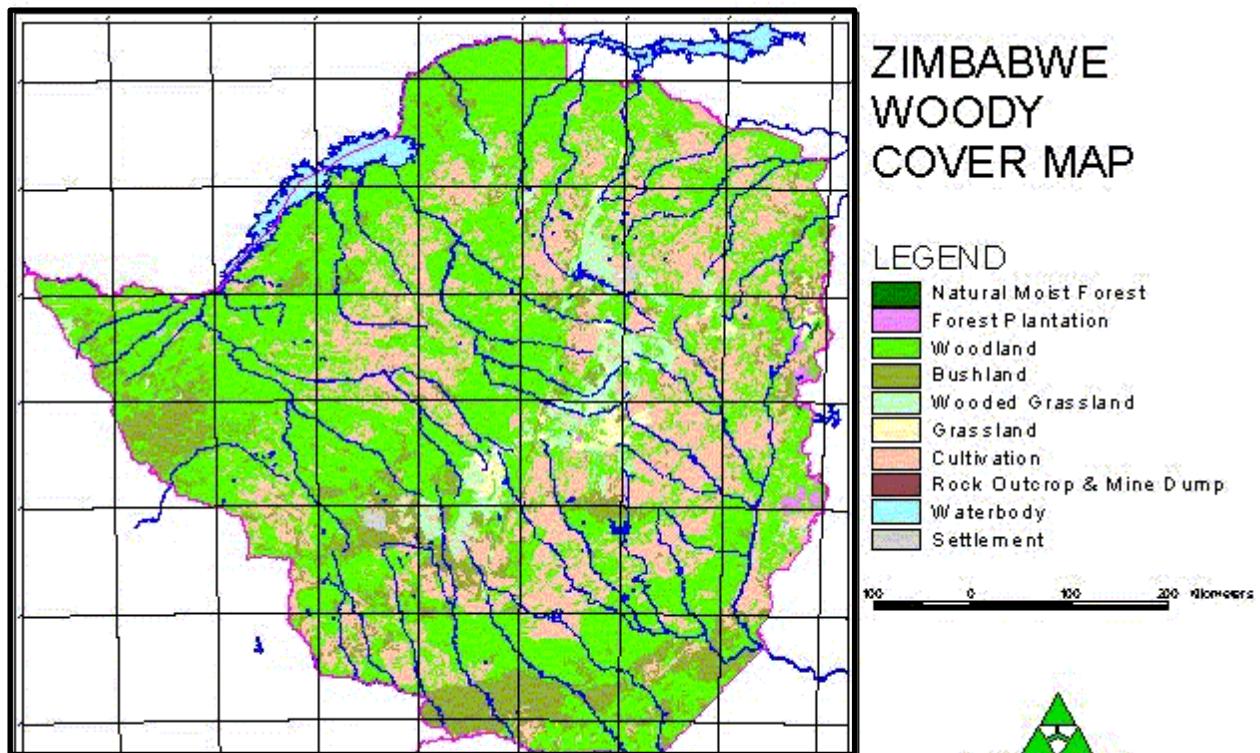


Figure 26 SOTER- derived maps of Zimbabwe: Impact of desertification on food security [48].

2.3.3 West Lunga National Park

West Lunga is one of Zambia's parks which is very wild and little visited. In 2008 the African Parks Network began managing West Lunga National Park. It was originally gazetted as a game reserve in the late 1940s, mainly to preserve its population of yellow-backed duiker. Then elephants were also abundant, along with a multitude of antelope species including Angolan (giant) sable and Lichtenstein's hartebeest. However, in the last few decades it's been used very little, except as an area for hunting and fishing by the local communities, and so has been off the map for most visitors.

West Lunga National Park covers 1684 km² of forests, open grasslands and swamps and it is the only Park in Zambia that is predominantly covered with forest (Figure 28).

It is bounded by the Kabompo River to the east and south and by the West Lunga River to the west.

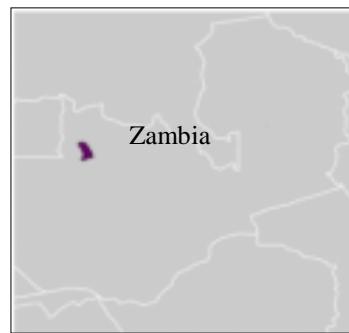


Figure 27 West Lunga National Park position in Zambia.

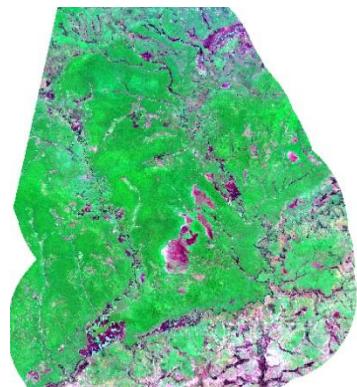


Figure 28 Landsat8 from 05-08-2013
RGB: band 753.

These are distinctive, dry, evergreen forests which are regarded by botanists as forming the largest area of tropical evergreen forest in Africa outside of the equatorial zone. Although the rainfall is relatively high at around 1000mm per annum, the soils are very sandy and well-drained, which means that, apart from the rivers, there are no permanent surface water bodies. The dryness, the nutrient-poor soils, and the density of the vegetation are the reasons that this area is so sparsely populated [49]

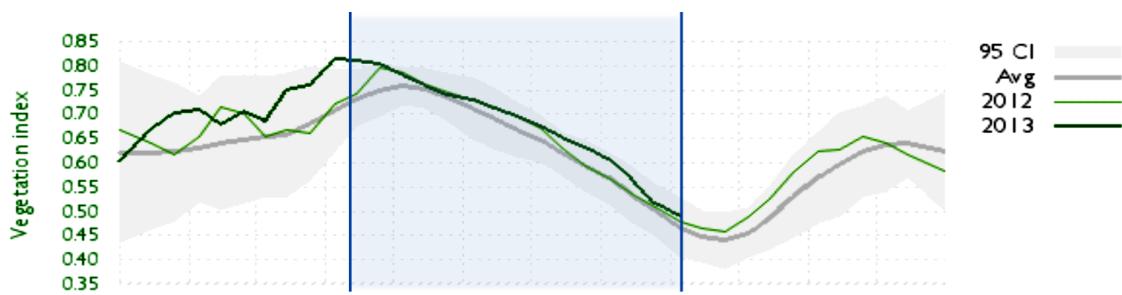


Figure 29 NDVI Trends for West Lunga National Park. Blue vertical lines define the period of the year considered in the study. The dark gray line is the average for each dekad based on the available timeseries, and the light grey areas indicate the 95% confidence limits around this average. Source: African Protected areas.

To the best of our knowledge there are no up to date high resolution vegetation maps for this site.

2.3.4 Gorongosa National Park

Gorongosa National Park is situated at 19°S and 34°W, within the Sofala province in the central part of Mozambique as shown in Figure 30. The morphology of the central Mozambican landscape is rising inland, from a broad shallow continental shelf and a low coast line, to the Great Eastern Escarpment on the border between Mozambique and Zimbabwe [50]. The park is situated within the Urema Rift and covers an area of approximately 3770 km².

The park together with the 8 surrounding areas comprises four morphological units, which are from east to west; the Cheringoma Platform, the Rift valley floor, the Bárue Platform and the Gorongosa Mountain [50].

Seasonal flooding and waterlogging of the valley, which is composed of a mosaic of different soil types, creates a variety of distinct ecosystems. Grasslands are dotted with patches of acacia trees, savannah, dry forest on sands and seasonally rain-filled pans and termite hill thickets. The plateaus contain Miombo and montane forests and a spectacular rain forest at the base of a series of limestone gorges.

This combination of unique features at one time supported some of the densest wildlife populations in all of Africa, including charismatic carnivores, herbivores and over 500 bird species. But large mammal numbers were reduced by as much as 95% and ecosystems stressed during Mozambique's long civil conflict at the end of the 20th century [51] [52]. Park ecologist, Dr. Kenneth Tinley, in 1977 [50] have identified three main vegetation types supporting the Gorongosa ecosystem's wealth of wildlife. Seventy-six percent is savanna — combinations of grasses and woody species that favors well-drained soils.



Figure 30 Gorongosa National park position in Mozambique.

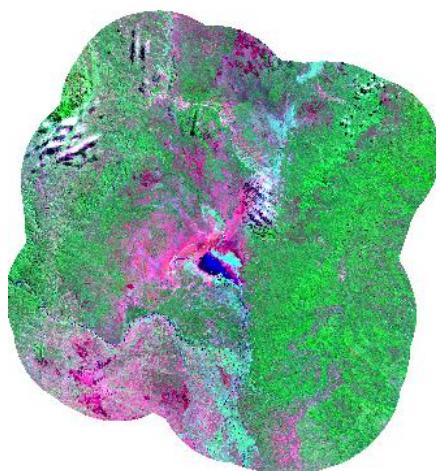


Figure 31 Landsat8 from 17-06-2013
RGB: band 753.

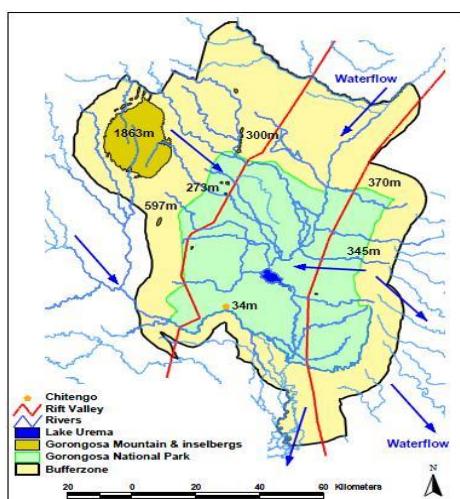


Figure 2 Map of Gorongosa National Park (green) and bufferzone(yellow).Source <http://www.gorongosa.org/>

Fourteen percent is woodlands – several kinds of forest and thickets. The rest is grasslands subjected to harsh seasonal conditions that prevent trees from growing. All three types are found throughout the system, with many different sub-types and varieties [53].

Mount Gorongosa has rainforests, montane grasslands, riverine forests along its rivers, and forests and savanna woodlands at lower elevations. Both plateaus are covered with a kind of closed-canopy savanna, widespread in southern Africa, called Miombo. Mozambique's tropical savanna climate, with an annual cycle of wet and dry seasons, has added yet another factor to the complex equation: constant change in soil moisture that varies with elevation.

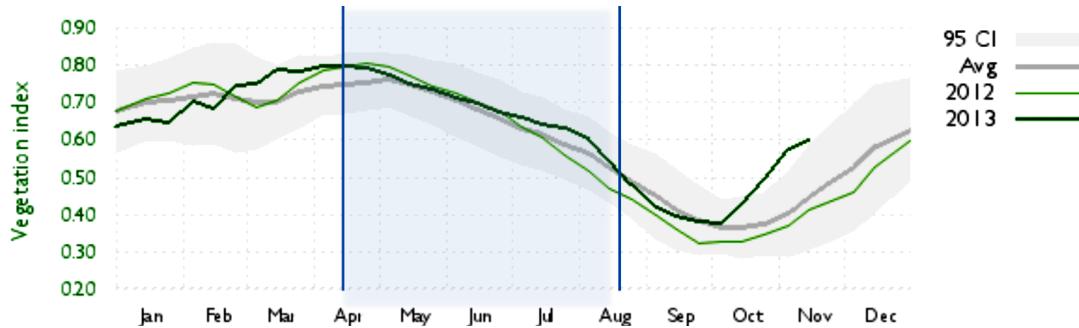


Figure 33 NDVI Trends for Gorongosa National Park. Blue vertical lines define the period of the year considered in the study. The dark gray line is the average for each dekad based on the available timeseries, and the light grey areas indicate the 95% confidence limits around this average. Source: African Protected areas.

There are three major sources of map information available on the vegetation of the Gorongosa National Park and its buffer zone [54].

Cenacarta land cover map, (Carta de Uso e Cobertura da Terra) land cover at a 1:250,000 scale for the whole of Mozambique. The individual units have been derived through visual identification and manual digitizing of units from LANDSAT satellite imagery. This map has a limited spatial resolution and do not include detailed vegetation information for this study area.

Cunliffe & Lynam [55] produced a map using automated classification of LANDSAT 7 imagery. Some of the drawbacks of this map are the little evidence of a linkage between the identified units and Tinley's and extremely limited ground truthing (also identified by the authors).

Carr Foundation landscapes map for Gorongosa National Park research center.

Regional, landscape and (selective) community map in digital format was created according to the different units described by Tinley [50] implementing the existing map information on the vegetation (CENACARTA land cover maps and Cunliffe & Lynam). Satellite images was processed using automated techniques (supervised and unsupervised classification) to produce homogeneous units and with Aerial and ground control points used for the survey and validation [54] [52].

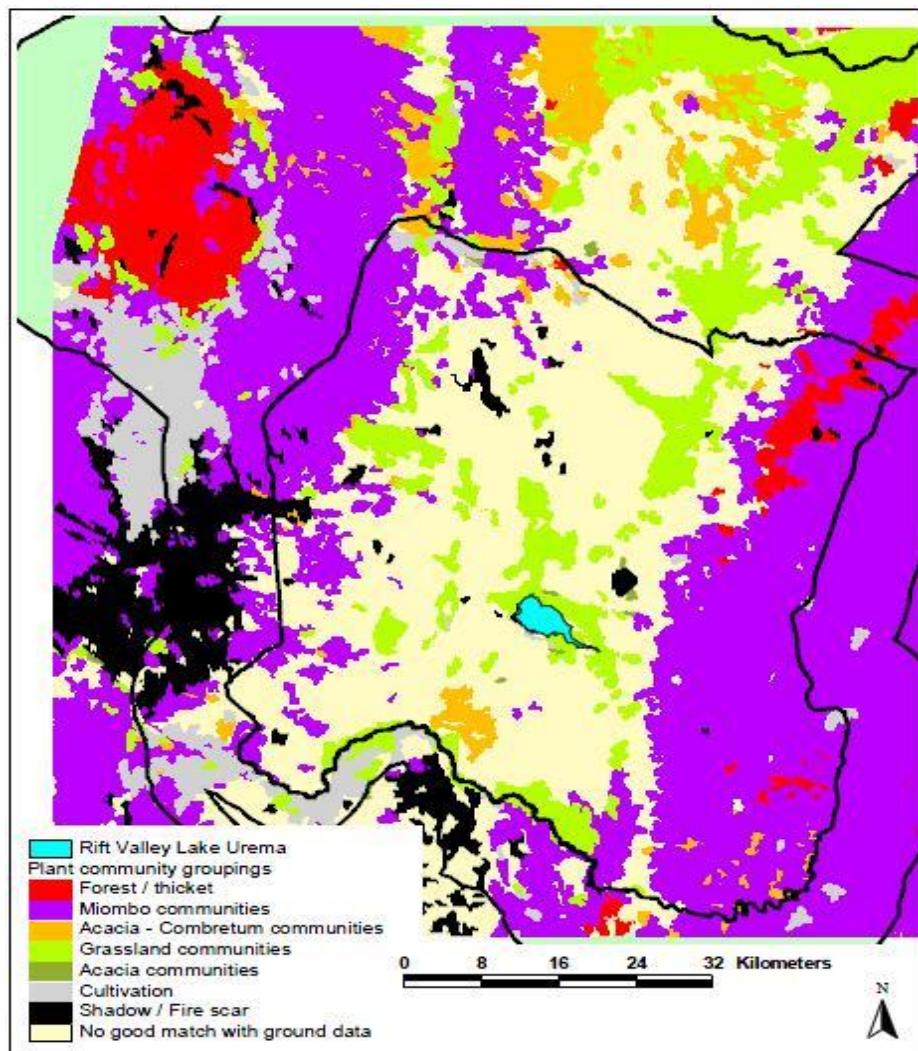


Figure 34 Carr Foundation landscape map of Gorongosa National Park. Source <http://www.gorongosa.org>, Gorongosa landscapes.

Related readings: *Vegetation survey of Mount Gorongosa*, Tom Müller, 2012 Biodiversity Foundation for Africa Zimbabwe [56].

2.3.5 Tsimanampetsotsa National Park

Tsimanampetsotsa National Park, named after its spectacular lake and main park highlight, is located in the southwest of Madagascar, not far from the coast and approximately 90 km south of Tulear and comprises 430 km² of spiny forest and wetlands (Figure 35). This area was already protected in 1927 due to its biological meaning (90% of the flora and fauna is endemic) mainly as habitat for water birds, and became a National Park in 1966.

These water supplies are basic both for humans and for animals and plants, since Tsimanampetsotsa (which means “lake without dolphins” in Malagasy) lies in the most arid zone of the island, which receives only around 300 mm rain each year.

Apart from the lake itself, there are two main landscapes inside the park: a calcareous plateau covered with dense xerophilous thickets (a locally endemic spiny forest), huge baobabs and banyan trees, and the sand dunes along the coast covered by a grassy blanket. The park also has numerous caves and sinkholes, which are the result of an underground stream [57].

To the best of the author’s knowledge there are no up to date high resolution vegetation maps for this site.

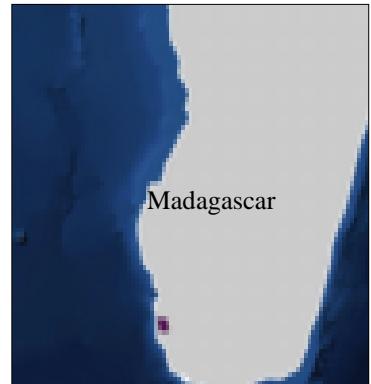


Figure 35 Tsimanampetsotsa National park position in Madagascar.

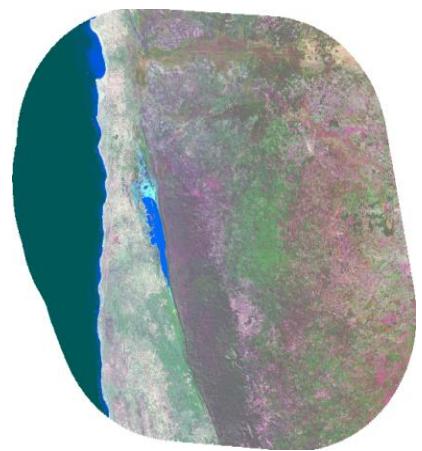


Figure 36 Landsat8 from 12-05-2013
RGB: band 753.

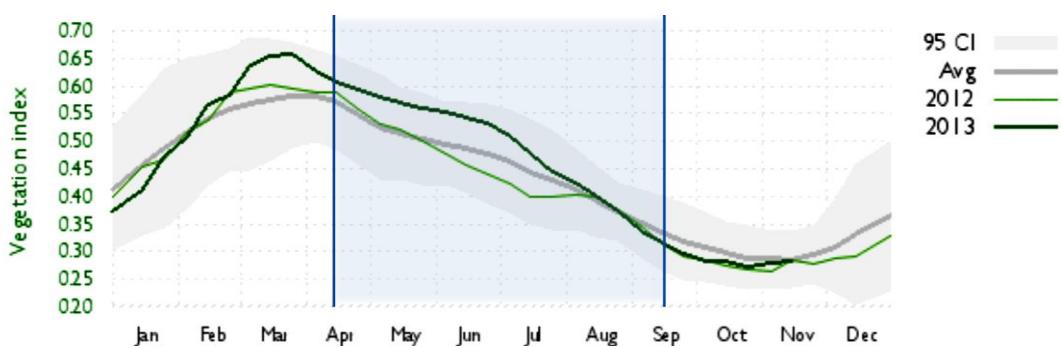


Figure 37 NDVI Trends for Tsimanampetsotsa National Park. Blue vertical lines define the period of the year considered in the study. The dark gray line is the average for each dekad based on the available timeseries, and the light grey areas indicate the 95% confidence limits around this average. Source: African Protected areas.

2.4 Image Preprocessing

Selected images were preprocessed in order to convert raw DN values to top of atmosphere (TOA) reflectance, a physical unit that allows putting data from different sensor/platform onto a common radiometric scale using conversion parameters retrieved from the metadata. Working with TOA reflectance also minimizes spectral differences between images caused by different acquisition time, sun elevation and sun-earth distance. In this study only Landsat 8 band that correspond to Landsat TM/ETM bands were used, therefore band 1, 9, were not processed as well as TIR and PAN bands.

The following formula describes the conversion steps and parameters involved:

$$\rho_\lambda = \pi * L_\lambda * d^2 / ESUN_\lambda * \cos \theta_{SZ}$$

where

ρ_λ = TOA reflectance for band λ

L_λ = Radiance for band $\lambda = M_\lambda * Q_{cal} + A_\lambda$

M_ρ = Band-specific multiplicative rescaling factor

A_ρ = Band-specific additive rescaling factor

Q_{cal} = Quantized and calibrated standard product pixel values (DN in 16bit)

$d = (1-0.01672*COS(0.01745*(0.9856*(JulianDayImage - 4)))) []$

θ_{SZ} = Local solar zenith angle

bESUN = [2067.000,1893.000,1603.000,972.6,245.0,79.72] [<http://www.gisagmaps.com/landsat-8-atco/>]

Processed bands were then stack together into a floating point [0-1] GeoTif image to be classified using an automated classification algorithm developed at Joint Research Centre-Ispra, as proposed hereafter.

2.5 Single Date Classification

The algorithm developed by the Joint Research Centre, as proposed in *Fast and Robust Topographic Correction Method* [58], originally designed to work with Landsat TM/ETM+ imagery has been used to process collected Landsat 8 datasets per our study areas. The algorithm receives top of atmosphere reflectance (float [0-1]) and generates a thematic raster map with land cover classes as summarized on Table 3 and an additional cloud/shadow mask based on spectral properties and 3D modeling involving sun azimuth and cloud/shadow direction together with morphological filters to fill holes and buffering the edge, usually characterized by thin clouds or haze that causes classification confusions [59]. It does not require any human interaction since it is based on an internal automatic decision tree based on spectral signatures collected at global scale.

1 st step categorization based on NDVI	2 nd step
[-1,0] CAT1	<pre> if Snowshape = 1 then SI else if Watershape and B >= 0.078 and 0.04 <=G <=0.12 and max(SWIR1,SWIR2) <= 0.04 then DWAT else if R >= max(NIR,SWIR1,SWIR2) and 0.04 <=R <= 0.19 and B > 0.078 and max(SWIR1,SWIR2) < 0.04 then SWAT else if B > 0.94 and G > 0.94 and R > 0.94 and NIR > 0.94 then CL else if Wetness > 5 then DWAT else SS </pre>
[0,0.45] CAT2	<pre> if Snowshape = 1 then SI else if B > 0.94 and G > 0.94 and R > 0.94 and NIR > 0.94 then CL else if Other_watershape and B > 0.078 and max(SWIR1 , SWIR2) < 0.058 then SWAT else if Cloudshape or Cloudshape1 or Cloudshape2 then CL else if B > G > R and NIR > 0.254 and B > 0.165 and NDVI < 0.40) then CL else if B > G) and B > 0.27and G > 0.21 and ABS(R-G) <= 0.1 and NIR > 0.35 then CL else if 0.13 > B > G > R < 0.05 and (B-NIR) < -0.04 then SV else if Wetness > 5 then DWAT else if 0.13 > B > G > R < 0.05 and B-NIR < 0.04 then SV else if 0.14 > B > 0.10 > G > R < 0.06 and NIR < 0.14 and (NIR-B) < 0.02 then SS else if B > G > R and NIR >= 0.06 and ABS(NIR-G) < 0.01 and B-NIR >= 0.01 then SS else if NDVI <= 0.09 and NIR < 0.4 and G <= R <= NIR then OLD else if NDVI <= 0.20 and NIR > 0.3 and B <= G <= R <= NIR then OLD if NDVI >= 0.35and B >= G and ABS(R-G) < 0.04 then SPV else if NDVI >= 0.20 and ABS(R-G) < 0.05 then OLL else OLL </pre>
[0.45,1] CAT3	<pre> if NDVI < 0.5 and NIR >= 0.15 then SPV else if NDVI < 0.5 and NIR < 0.15 then SV else if NDVI < 0.55 and B <= NIR <= 0.15 then SV else if NDVI < 0.55 and B <= NIR then GRS else if NDVI < 0.55 and B > NIR then SV else if NDVI < 0.65 and NIR >= 0.22 then TCL else if NDVI < 0.65 and NIR >= 0.165 then TCD else if NDVI < 0.65 and NIR < 0.165 then TCD else if NDVI < 0.78 and NIR < 0.30 then TCD else if NDVI >= 0.78 SHR else TCL </pre>

Where:

B: blue, **G:** green, **R:** red, **NIR:** band 4, **SWIR1:** band 5; **SWIR2:** band 7;
Watershape = ((B-G) > -0.2) >= G >= R >= NIR >= SWIR1
Other_watershape= B >= G >= R <= NIR < R*1.3 < 0.12 > SWIR1 > SWIR2 and 0.039 < NIR < G
NDSI= (G - SWIR1) / (G + SWIR1)
Wetness = 66.96*B + 53.55*G + 23.61*R + 16.72*NIR – 194.53*SWIR1 – 137.19*SWIR2
Brightsoil= (B < 0.27 and Growing15) or (B < 0.27 and Growing14 and SWIR1-SWIR2 > 0.038)
Saturation= max(G,R,NIR)-min(G,R,NIR)/max(G,R,NIR)
Snowshape= min(B,G,R,NIR) > 0.30 and NDSI > 0.65
Cloudshape= min(B,G,R)> 0.17 and max(G,B,R,NIR) > 0.30 and (NIR/R) >= 1.3 and (NIR/G) >= 1.3 and (NIR/SWIR1) >= 0.95 and SWIR1 > min(B,G,R) and NDSI < 0.65
Cloudshape1= max(B,G,R,NIR,SWIR1,SWIR2) > 0.47 and min(B,G,R,NIR) > 0.37 and Snowshape = 0 and Bright_soil = 0
Cloudshape2= min(B,G,R) > 0.21 and SWIR1 > min(B,G,R) and 0.2 <= Saturation <= 0.4 and max(G,R,NIR) >= 0.35 and Snowshape = 0 and NDSI > -0.3 and Bright_soil = 0
Growing14= B<G<R<NIR
Growing15= B<G<R<NIR<SWIR1

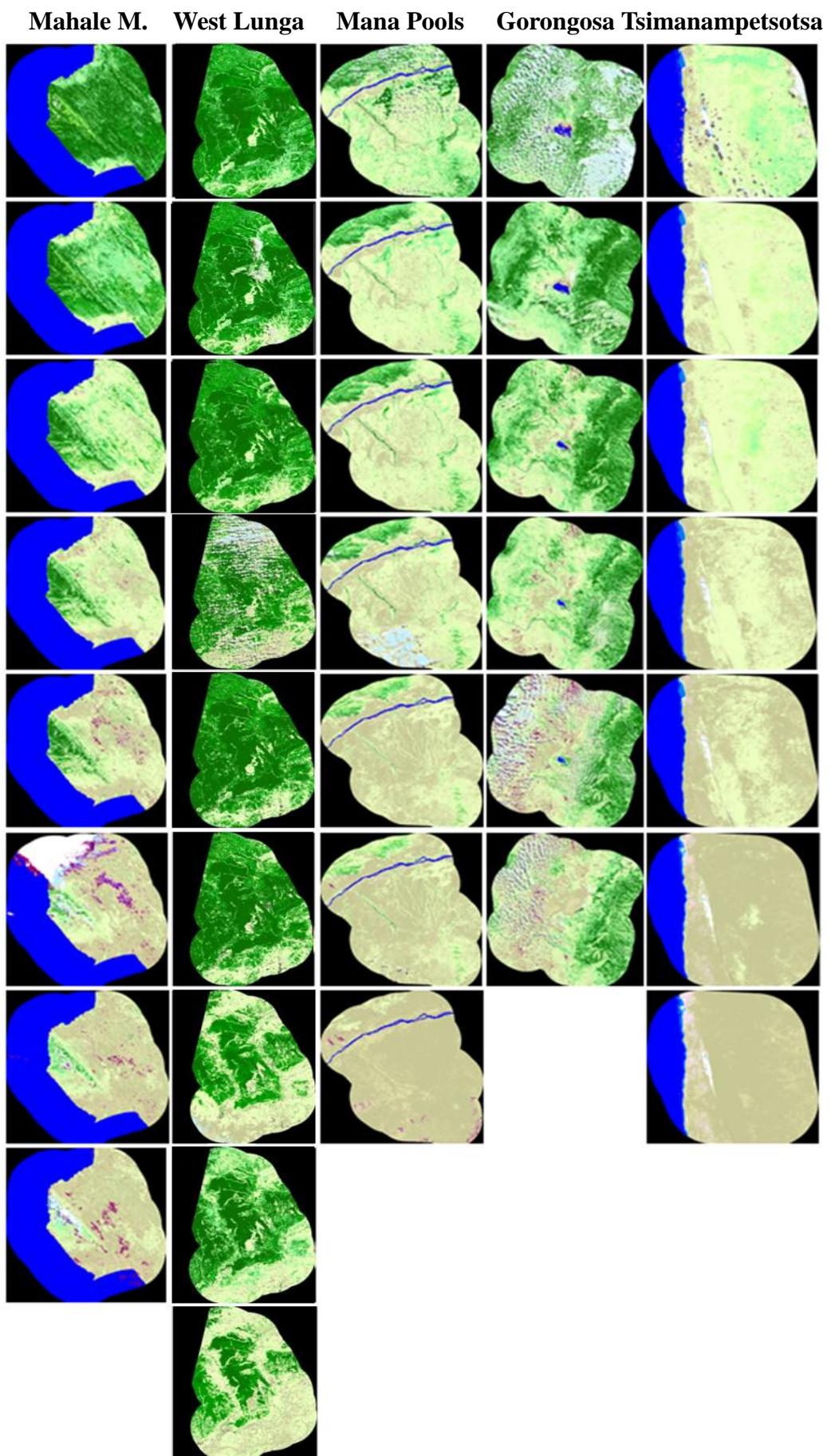
Table 2 Pseudo code of automatic decision tree used in the classification algorithm.

Description	Land cover classes	Class ID	Color
No Data	-	ND	
Water	Deep Water bodies/Rivers (DWAT/SWAT)	WAT	Blue
Clouds	Clouds	CL	
Snow	Snow / Ice	SI	Cyan
Tree Cover Dark	Dense Forest/Dense Shrub	TCD	Green
Tree Cover Light	Open Forest/Shrub	TCL	Light Green
Shrub	Dense Shrub	SHR	Dark Olive Green
Grassland	Dense Grassland/ Open Shrub	GRS	Light Olive Green
Sparse vegetation	Sparse Grassland/Sparse Shrub	SPV	Yellow
Other Land Light	Light soil/rocks/sand	OLL	Light Brown
Other Land Dark	Dark soil/rocks/sand	OLD	Dark Brown
Shadowed Vegetation	Shadowed / Low Illuminated Vegetation	SV	Dark Olive Green
Shadowed Soil	Shadowed Soil / Burnt Areas	SS	Purple

Table 3 Output associated land cover classes, ID and colour of automatic classification tool.

Hereafter the single date maps for each study areas. Top images are corresponding to the peak of the vegetation growth; the bottom ones, characterized by brownish colors, to the end of the dry season. Is worth noting that the single date classification accuracy is closely related to the acquisition date of the image, the status of the vegetation and weather conditions such as cloud and shadows often present in tropical regions; therefore deciduous forest cover will be mapped as TCD at the peak of the growth, as SHR or GRS while drying up and OLL/OLD when leafless (Table 3).

Here the need of developing a phenology based algorithm that considers the vegetation evolution and generates a more accurate land cover map.



2.6 Phenology Based Classification

The objective of this chapter is to describe the methodology used to combine the single date classification into a more reliable and detailed land cover map including evergreen and deciduous discrimination on the basis of “frequency” rules. The use of thematic products (8 bit layer with 12 classes), rather than the original multispectral bands time series, reduces the complexity of the model and speed up processing time. To better understand different land cover type dynamics and their separability, several samples have been collected within the 5 study areas and plotted together as shown on Figure 39. As expected, in April and March (peak of the green season) the spectral response of deciduous and evergreen forest, humid agriculture and shrub is almost similar at 30 m resolution and the high NDVI values are driving the classification towards dark and light tree coverage (TCD/TCL). However, during dryer months (September, October), deciduous leafless vegetation is classified as bright or dark soil (low NDVI response and high reflectivity on IR bands) reaching the highest separability compared to evergreen vegetation.

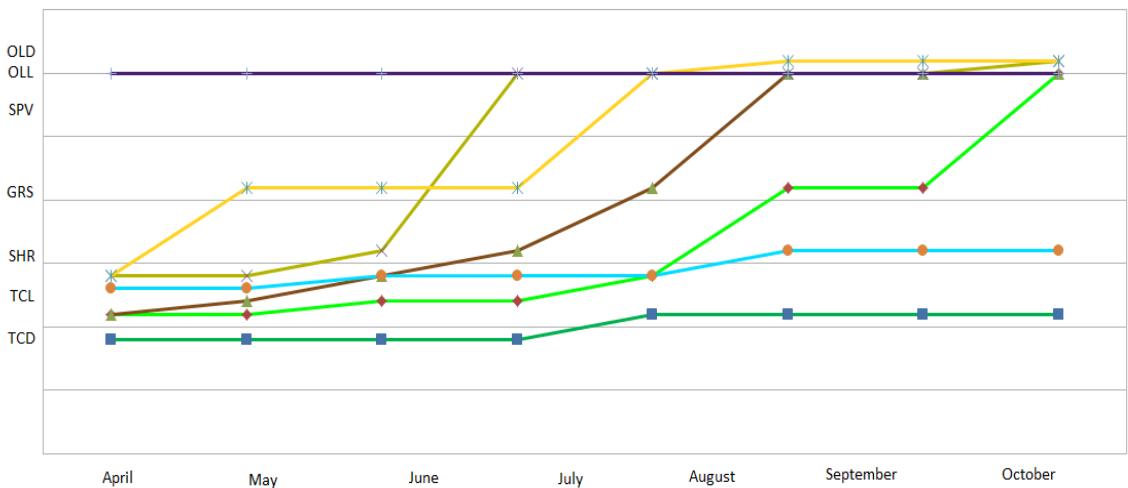


Figure 39 Time profile of different land cover classes selected among the five study areas. Horizontal axe shows the time while vertical, the single date classification ID as provided by the single date classifier.

Is worth noting that our study areas are national parks with a limited anthropization and built up areas are mainly characterized by small villages (most extent is Msomsoko, 1.5 km², buffer zone south of Mahale Mountains National park). The absence of asphalt (stable spectral response over time) and the scattered distribution of houses, vegetation and sandy soil make extremely difficult the identification of unique spectral behavior for this class. We have therefore decided to not concentrate on Land Use classes (Urban Areas and Agriculture) in this research activity but leaving them in the proposed legend.



Figure 40 Msomsoko, 1.5 km², buffer zone south of Mahale Mountains N.P. Source: Google earth .

Hereafter we present the legend and the decision rules implemented in the Phenology Based Classification to derive Land Cover classes from the analysis of the single date classification categories.

Legend

According to Land Cover Classification System [60] proposed by FAO/ UNEP and in line with the legend adopted by the JRC in the framework of Land Cover/Land Use cove change detection in African protected areas, we identified the following dominant classes:

ID	Class Description	ID	Class Description
10	EVG Closed Tree Cover (>40%)	15	Tree or Shrub crops
11	EVG Open Tree Cover (40% to 15 %)	16	Arable Agriculture
13	EVG Shrub Cover	27	Humid Agriculture
14	EVG Herbaceous Cover	17	Flooded Shrub and Herbaceous
		18	Open Water
23	DEC Closed Tree Cover (>40%)	19	Urban
24	DEC Open Tree Cover (40% to 15 %)	20	Bare Areas
25	DEC Shrub Cover	21	Clouds
26	DEC Herbaceous Cover	22	Shadow

Primarily vegetated areas include vegetative cover of at least 4% for at least two months a year.

This cover may consist of the life forms Woody (Trees, Shrubs), Herbaceous (Forbs, Graminoids) or a combination of them, or consist of Lichens/Mosses.

Cover can be considered as the presence of a particular area of the surface covered by a layer of plants. A distinction is made between Closed (> 40%) and Open (between 40% and 15%).

Height is a discriminant factor for Trees (>30 – 3 m), Shrubs (5 - 0.3mt) and Herbaceous (3 - 0.03 m). Height is used to separate woody plants higher than 5 m (trees) from those < 5 m (shrubs) Cultivated areas or Arable Agriculture refers to areas where the natural vegetation has been removed or modified and replaced by other types of vegetative cover of anthropogenic origin. This vegetation is artificial and requires human activities to maintain it in the long term.

Cultivated Aquatic or Regularly Flooded (Humid Agriculture) class includes areas where an aquatic crop is purposely planted, cultivated and harvested, and which is standing in water over extensive periods during its cultivation period. In general, it is the emerging part of the plant that is fully or partly harvested.

Class Description	Photo	Google Earth	Landsat 8
Closed/Open Evergreen Forest			
Closed/Open Deciduous Forest			
Shrub Cover			
Herbaceous Cover			
Arable Agriculture			
Humid Agriculture			

Figure 41 Dominant classes in primarily vegetated areas: In columns, for each class (Closed/Open Evergreen Forest, Closed/Open Deciduous Forest, Shrub Cover, Herbaceous Cover, Arable Agriculture, Humid Agriculture) different images sources: ground or aerial photo, Google Earth and Landsat 8 Image.

Primarily Non-Vegetated includes areas that have a total vegetative cover of less than 4% for more than 10 months of the year, or in the absence of Woody or Herbaceous life forms less than 25% cover of Lichens/Mosses.

Artificial surfaces or Urban Area describes areas that have an artificial cover as a result of human activities such as construction (cities, towns, and transportation), extraction (open mines and quarries) or waste disposal.

Bare Areas is characterized by less than 4% vegetative cover and includes rock, sands and deserts.

Natural water bodies, Snow and Ice class refers to areas that are naturally covered by water, such as lakes, rivers, snow or ice.

Class Description	Photo	Google Earth	Landsat 8
Urban Areas	 	 	
Bare Areas			
Water bodies			

Figure 42 Dominant classes in primarily non-vegetated areas: In columns, for each class (Urban Areas, Bare Areas and water bodies) different images sources: ground or aerial photo, Google Earth and Landsat 8 Images.

Data Preparation

Each single date classification and related cloud/shadow mask are stack together into a single geo-referenced file following the chronological order to ease the classification phase.

Classification

“Frequency” rule is applied pixel by pixel analyzing the land cover “Class ID” in the stack of the single date classification layers; no data (ND) and clouds/shadow (CL/SH) pixels are excluded from the statistics only if other classes are available in the stack. In this context “frequency” does not mean that the most frequent class (temporal majority) is consider to be the output land cover; this assumption might be valid for permanent water bodies, bare areas or evergreen vegetation but not for deciduous vegetation. Dry deciduous forest cannot be detected using a majority approach since only few months a year has green leaves while the remaining months is dry and the soil component is the dominant spectral response detected by satellite at 30 m resolution like Landsat 8. “Frequency” means respecting a precise pattern over the observed period (in our study is spanning from April to October) as illustrated on Table 4.

ND	CL SH	SI	WAT	SV	TCD	TCL	SHR	GRS	SPV	OLL	OLD	SS	Output	
100													No data	
	100												Cloud	
		100											Snow/Ice	
			100										Permanent Water	
	*	50	*	*	*	*	*	*	*	*	*	*	Seasonal Water	
			*	80	*	*							Evergreen Closed Forest	
			*	*	80	*	*	*	*	*	*	*	Evergreen Open Forest	
				*	*	80	*						Evergreen Shrub	
			*	20	20	20	*	*	*	*	*		Close Deciduous Forest	
				*	20	30	*	*		30		*	Open Deciduous Forest	
		*	*	*	*	*	30	30		30			Grassland	
												100	Dark Rocks / Urban	
											100		Dark Soil	
									100				Light Soil/ Sand	

Table 4 Frequency distribution tables with the percentage of observations for each class.

Permanent Water/Snow or Light/Dark Soil are easy to be detected since are characterized by the same stable class (WAT, SI, OLL, OLD) over the time;

Evergreen Closed Forest is characterized by two variables: 1) each single date classification has to be in the range [SV, TCD, and TCL]; 2) at least 80% of the time is TCD. Small variations (max 20%) are mainly due to different illumination conditions or presence of small percentage of deciduous species that might create openings (at sub-pixel level, less the 30 m) at the end of the season.

Close Deciduous Forest covers a wider class range during time; the pic of the season is well represented by dark forest class (TCD), brighter (TCL and SHR) before becoming leafless and soil class (OLL, OLD) once dry.

Open Forest (Evergreen and Deciduous) are both characterized by wider class range spanning from TCD to OLL. For better mapping those land cover types we suggest using post classification spatial filters or working with segments/objects rather than pixels. However, we decided to keep them in the legend.

Ad hoc rules can be defined to discriminate particular land cover type (e.g. flooded vegetation) or different species within a class. Currently, land uses classes like agriculture or urban are not optimally characterized (spectrally and temporally) and will be subject of further developments. However, they have been included in the final accuracy assessment as well to evaluate the confusion with other classes. Dark and light soils are then aggregated into Bare Areas.

Hereafter, in Figure 43, the final classification maps of the 5 study areas.

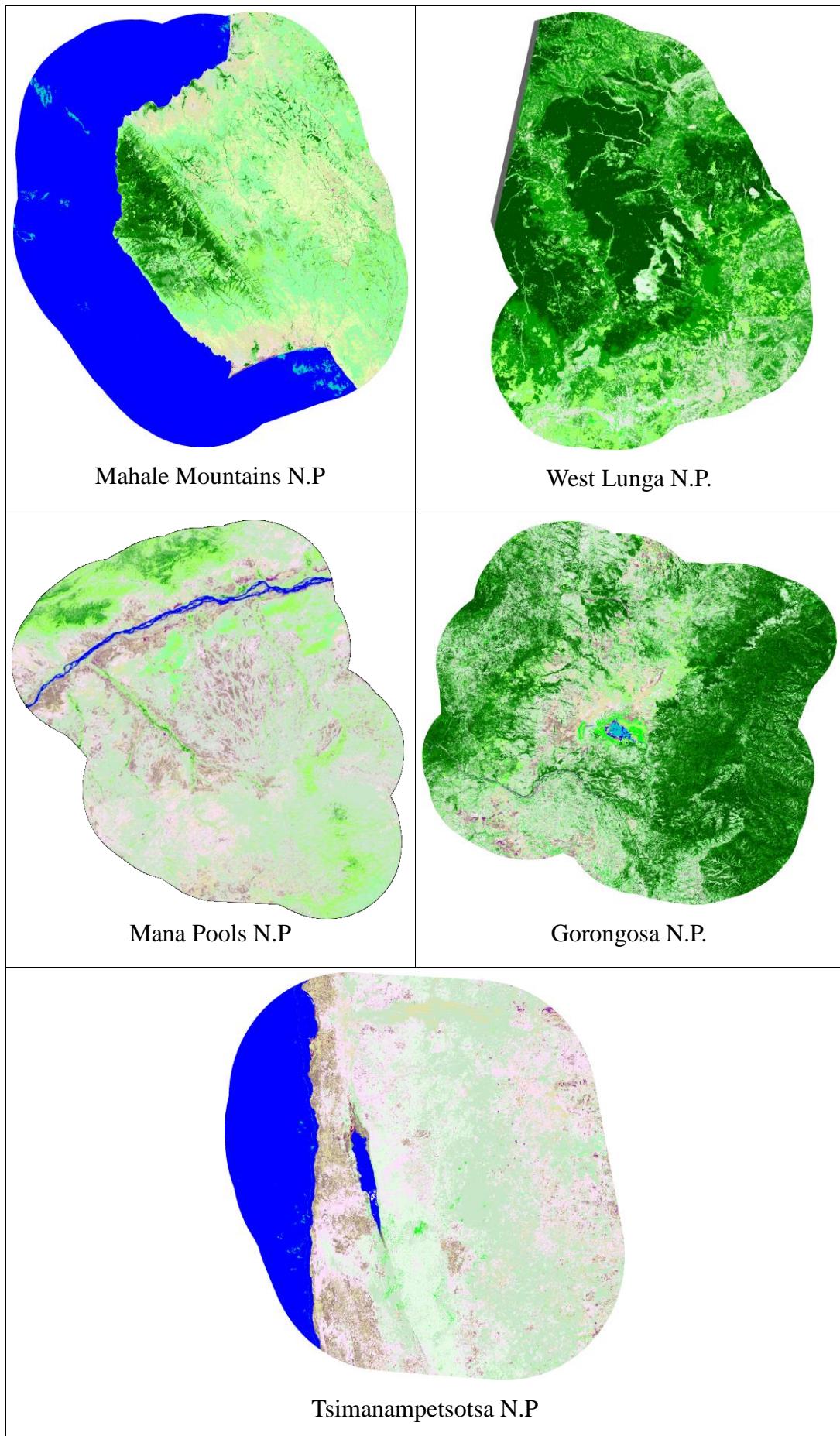


Figure 43 Final phenology based land cover maps of the study areas.

2.7 Classification Accuracy Assessment

Documenting the effectiveness of the phenology based classification is difficult given the current scarcity of ground verification data at the appropriate scale. High resolution land cover map are rarely available at country scale and even within protected areas boundaries.

Products at 1 km resolution like GLC-2000 [39] are not suitable for assessing a 30 m resolution classification of 2013 land cover maps due to its coarse resolution and reference date (2000). Recently TANAPA in collaboration with Frankfurt Zoological Society and European Union [44], published a land cover map of Mahale National Park giving a more up-to date classification but not detailed enough to be used as a reference dataset for this study area.

Figure 44 shows GLC-2000 [39], GIZ [44] our map and the L8 false color image for the above mentioned site. As visible from the L8 the fragmented landscape, the thin gallery forest and grassland patches are well mapped by our classification.

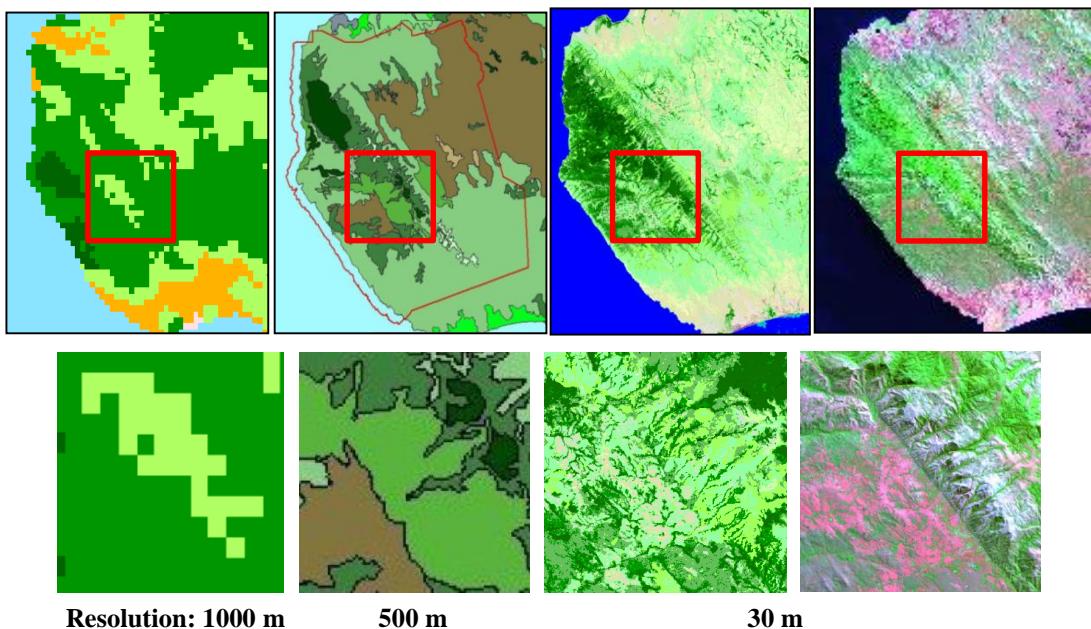


Figure 44 Different Land cover classification for Mahale Mountain National Park. From left to right GLC-2000 [39], GIZ [44] our map and the L8 false color following the rise of resolution, from 1000 m to 30 m.

Other protected sites do not have any up to date high resolution reference map therefore we decided to collect ground truth data by visual interpretation of high resolution data, available on Google Earth, using a web-based tool designed and developed for this purpose by the Joint Research Centre, Ispra. For each study area a regular grid of approximately 1000 x 1000 m has been made and each confluence point has been visually interpreted using the dominant land cover class (as shown in Figure 45) within a circle of 60 m radius. If the sample area is either too heterogeneous or no high resolution images are available, the confluence point is rejected and labeled as no data.

The tool offers the possibility of panning all samples, assigning land cover class using a drop down menu, overlay collected L8 images with desired band combination and stretch options; Google Earth time bar is provided as well to better understand the vegetation evolution over time.

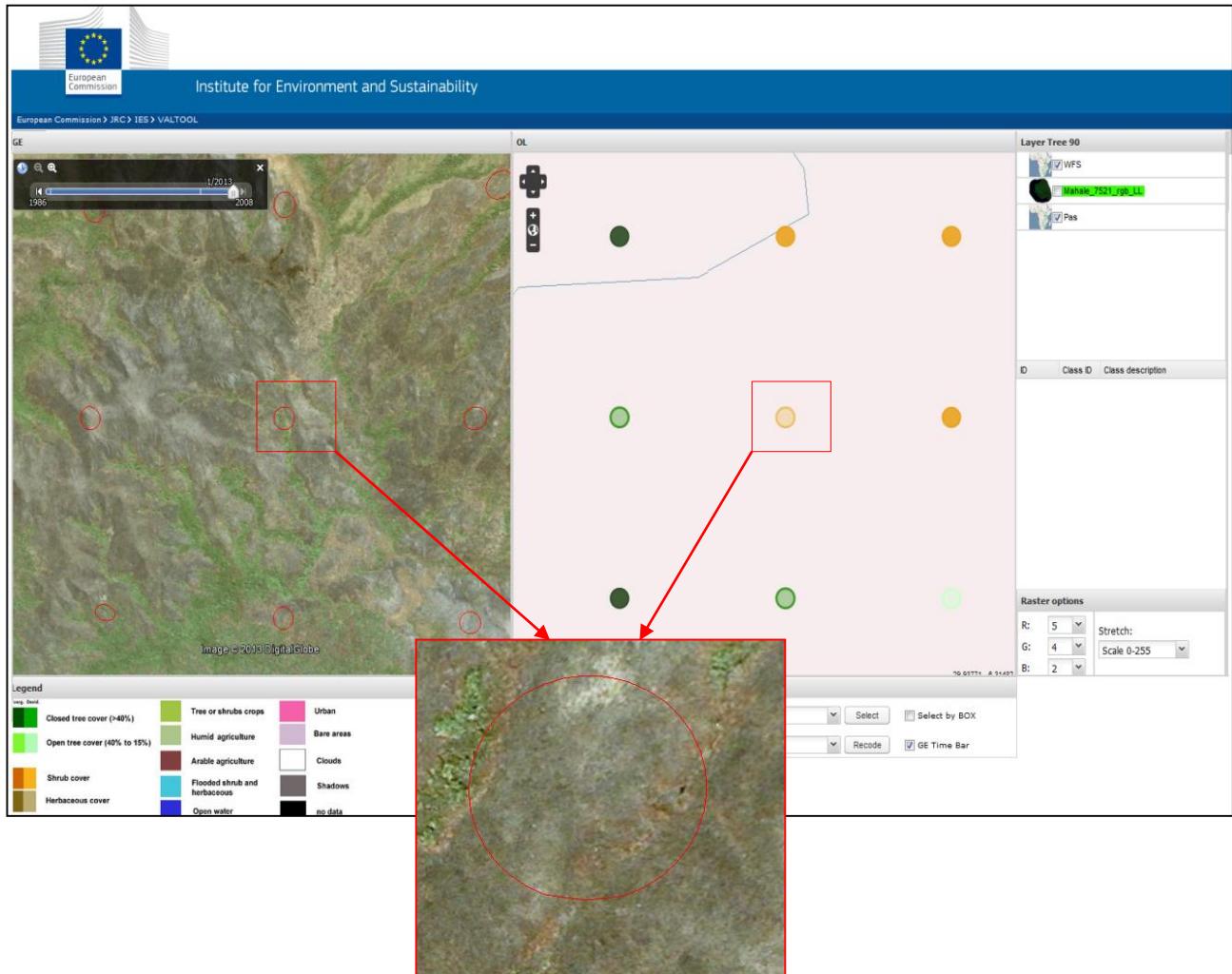


Figure 45 JRC web-based tool used for the accuracy assessment.

Collecting ground truth data is a crucial step for providing objective measure of a map quality, but can be time consuming and costly. However, the result provides us with an overall accuracy of the map and the accuracy for each class in the map. We are able to perform an accuracy assessment only on Mahale Mountains National Park and Tsimanampetsotsa National Park. These two areas have been selected because they cover, in a complementary way, all the vegetation classes spanning from evergreen forest of the first site, to the deciduous dry vegetation of the latter. The Joint Research Centre, Ispra, will continue the exercise on the remaining protected sites to provide better and robust quality estimation.

Is worth noting that the proposed classification method generates Land Cover and not land use maps and therefore we cannot access accuracy level for typical Land Use classes like Arable/Humid Agriculture, Tree/Shrub Crops and Urban Areas.

Additionally, 320 ground truth samples (~7%) of Mahale Mountains and 37 (~3%) of Tsimanampetsotsa National Park have been removed from the analysis because none of the land cover classes reached the absolute majority within the sample unit (red circle, Figure 46), or because the landscapes was too heterogeneous.



Figure 46 Examples of removed ground truth samples: none of the land cover classes reached the absolute majority within the sample unit (red circle) Source: Google Earth images.

The following tables (tables 5, 6, 7) show the accuracy assessment for the two sites cited before. Automatic classification labels are reported in columns while rows correspond to ground truth observations. Number of sample points and accuracy percentage has been reported. Clouds (21), Shadows (22) and Evergreen Herbaceous (14) classes are not included since no samples have been collected and no commission error have been observed.

Tables identified with “b” suffix are reporting the same accuracy assessment but aggregating class 10+11 (Evergreen Tree Cover Close + Open) and 23+24 (Deciduous Tree Cover Close + Open) resulting in a significantly higher accuracy (89.3%).

Confusion between open and close is mainly related to two causes:

- The resolution of the satellite data (30 m) is not enough to highlight inter-pixel heterogeneity.
- Open Tree Cover might be better identified using pixel aggregations (objects) as result of a segmentation process rather than pixel based approach.

Class	10	11	13	15	16	17	18	19	23	24	25	26	27	Total	Prod. Accuracy
10	138	51								1				190	72,6
11	10	39	1					2		1	1			54	72,2
13			81							1				82	98,8
15	1	1												2	0
16								17	2	9	19			47	0
17			1											1	0
18							2151							2151	100
19										3	1			4	0
23			6					764	96	100				966	79,1
24								153	215	30	54			452	47,6
25			1					89	48	68	96			302	22,5
26								13	6	5	132			156	84,6
27		1						1	1	2				5	0
Total	149	92	90	0	0	0	2151	0	1039	368	220	303	0	4412	
User Accuracy	92,6	42,4	90,0	0	0	0	100	0	73,5	58,4	30,9	43,6	0		81,3

Table 5 A Mahale Mountains National Park confusion matrix. Automatic classification land cover classes in columns. Ground truth observations in rows.

Class	10+11	13	15	16	17	18	19	23+24	25	26	27	Total	Prod. Accuracy
10+11	238	1						2	2	1		244	97,5
13		81							1			82	98,8
15	2											2	0
16							19	9	19			47	0
17		1										1	0
18						2151						2151	100
19									3	1		4	0
23+24	6							1228	130	54		1418	86,6
25		1						137	68	96		302	22,5
26								19	5	132		156	84,6
27	1							2	2			5	0
Total	241	90	0	0	0	2151	0	1407	220	303	0	4412	
User Accuracy	98,8	90,0	0	0	0	100	0	87,3	30,9	43,6	0		88,3

Table 5 B Mahale Mountains National Park confusion matrix. Automatic classification land cover classes with aggregation of classes 10+11 (Evergreen Tree Cover ,Close + Open) and 23+24 (Deciduous Tree Cover, Close + Open) in columns. Ground truth observations in rows.

Class	16	18	20	23	24	25	26	Total	Prod.Accuracy
16								0	0
18		223						223	100
20			2					2	100
23				122	22	1		145	84,1
24	1			57	336	20	6	420	80,0
25				1		229	8	238	96,2
26						6	53	59	89,8
Total	1	223	2	180	358	256	67	1087	
User Accuracy	0	100	100	67,8	93,9	89,5	79,1		88,8

Table 6 A Tsimanampetsotsa National Park confusion matrix. Automatic classification land cover classes in columns. Ground truth observations in rows.

Class	16	18	20	23+24	25	26	Total	Prod.Accuracy
16							0	0
18		223					223	100
20			2				2	100
23+24	1			537	21	6	565	95,0
25				1	229	8	238	96,2
26					6	53	59	89,8
Total	1	223	2	538	256	67	1087	
User Accuracy	0	100	100	99,8	89,5	79,1		96,0

Table 6 B Tsimanampetsotsa National Park confusion matrix. Automatic classification land cover classes with aggregation of classes 10+11 (Evergreen Tree Cover ,Close + Open) and 23+24 (Deciduous Tree Cover, Close + Open) in columns. Ground truth observations in rows.

Class	10	11	13	15	16	17	18	19	20	23	24	25	26	27	Total	Prod. Accuracy
10	138	51									1				190	72,6
11	10	39	1							2		1	1		54	72,2
13			81									1			82	98,8
15	1	1													2	0
16									17	2	9	19		47	0	
17			1												1	0
18							2374								2374	100
19											3	1		4	0	
20								2							2	100,0
23			6							886	118	101			1111	79,7
24										210	551	50	60		871	63,3
25			1							90	48	297	104		540	55,0
26										13	6	11	185		215	86,0
27		1								1	1	2			5	0
Total	149	92	90	0	0	0	2374	0	2	1219	726	476	370	0	5498	
User Accuracy	92,6	42,4	90,0	0	0	0	100	0	0	72,7	75,9	62,4	50,0	0		82,8

Table 7 A Mahale Mountains and Tsimanampetsotsa National Park confusion matrix. Automatic classification land cover classes in columns. Ground truth observations in rows.

Class	10+11	13	15	16	17	18	19	20	23+24	25	26	27	Total	Prod. Accuracy
10+11	238	1							2	2	1		244	97,5
13		81								1			82	98,8
15	2												2	0
16									19	9	19		47	0
17		1											1	0
18						2374							2374	100
19										3	1		4	0
20								2					2	100
23+24		6							1765	151	60		1982	89,1
25		1							138	297	104		540	55,0
26									19	11	185		215	86,0
27	1								2	2			5	0
Total	241	90	0	0	0	2374	0	2	1945	476	370	0	5498	
User Accuracy	98,8	90,0	0	0	0	100	0	100	90,7	62,4	50,0	0		89,9

Table 7 B Union Mahale Mountains and Tsimanampetsotsa National Park confusion matrix. Automatic classification land cover classes with aggregation of classes 10+11 (Evergreen Tree Cover ,Close + Open) and 23+24 (Deciduous Tree Cover, Close + Open) in columns. Ground truth observations in rows.

A number of observations can be made after taking a cursory look at the confusion between the classes:

- confusion between evergreen and deciduous vegetation is limited to few observation and is concentrated on areas characterized by a semi-deciduous vegetation in Mahale Mountains as reported by the [44]
- the highest confusion is between Deciduous Forest and Shrub and between Deciduous Shrub and Herbaceous mostly due to difficulties in defining the right height from a 30 m resolution image; errors are localized at the edges of two classes;
- the overall accuracy can be considered higher if Land Use classes like Tree or Shrub Crop (15) or Arable/Humid Agriculture (16/27) and Urban Areas (19) are translated into correspondent Land Cover Classes such as Shrub or Herbaceous for Arable Agriculture or Evergreen Tree for Tree or Shrub Crop;

Hereafter, shown in Table 8, some images supporting the abovementioned statements.

Comments	Google Earth Images
<p>Discrimination between Open and Closed Deciduous forest is not always easy, especially in a dry environment as Tsimanampetsotsa.</p> <p>Automatic classification = Dec. Open Tree Cover Expert visual classification = Dec. Closed Tree Cover</p>	 <p>Image © 2014 DigitalGlobe</p>
<p>Confusion between Open and Closed Deciduous forest due to the interference of soil and bright shrub patches</p> <p>Automatic classification = Dec. Open Tree Cover Expert visual classification = Dec. Closed Tree Cover</p>	 <p>Image © 2014 DigitalGlobe</p>

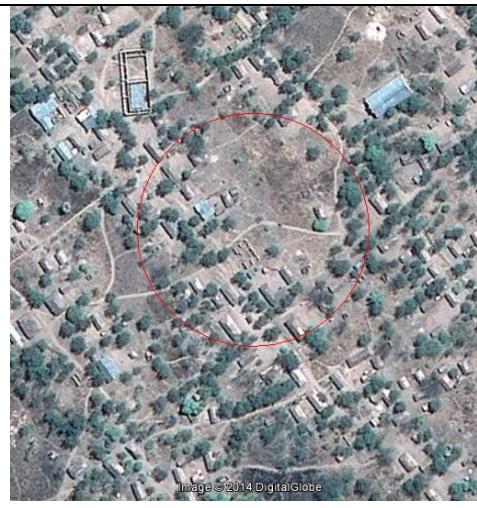
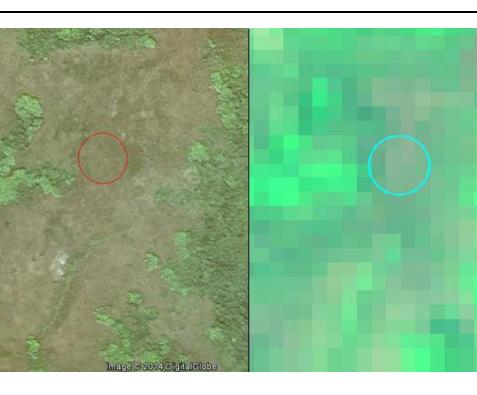
<p>Confusion between Close Deciduous Forest and Bare Area due to image resolution (30 m). Bright soil represents the majority class within the sample area of the Landsat 8 image (cyan circle).</p> <p>Automatic classification = Bare Area Expert visual classification = Dec. Closed Tree Cover</p>	 <p>Image © 2014 DigitalGlobe</p>
<p>Confusion between Deciduous Shrub and Herbaceous Bright soil represents the majority class within the sample area. That might be considered as an extreme example of Dry Shrub class.</p> <p>Automatic classification = Dec. Herbaceous Expert visual classification = Dec. Shrub Cover</p>	 <p>Image © 2014 DigitalGlobe</p>
<p>Confusion between Urban Areas and Deciduous Shrub Cover; at 30 m resolution the spectral response is characterized by a mixture of soil and vegetation typical of rural villages.</p> <p>Automatic classification = Dec. Shrub Cover Expert visual classification = Urban Areas</p>	 <p>Image © 2014 DigitalGlobe</p>
<p>Confusion between Herbaceous and Open Deciduous Tree Cover; the herbaceous corridor of 200/250 m is partially visible at 30 m resolution (only 2/3 pixel) due to the interference of the dense bright gallery forest that is creating a “blooming” effect on adjacent pixels.</p> <p>Automatic classification = Dec. Open Tree Cover Expert visual classification = Dec. Herbaceous</p>	 <p>Image © 2014 DigitalGlobe</p>

Table 8 Confusion between different classes with Google Earth images support.

3 Conclusions

The proposed automated land-cover/land-use mapping method based on short time series analysis circumvents the limitations imposed by the single date classification such as the impossibility of discriminating evergreen vegetation from deciduous, dry forest from bare soil (at 30 m resolution) and removing/reducing atmospheric influence like clouds or haze.

The 30 m resolution land-cover map developed in this study provides opportunities for a large number of research projects. Our method has been tested only on Landsat 8 data but is fully compatible with Landsat TM and ETM family.

The high overall accuracy (89.9%) achieved on Mahale Mountains National Park and Tsimanampetsotsa National Park demonstrates the high potentiality of the proposed method in distinguishing between the major land cover types with a relatively small set of images. However, there is room for improvements: new decision rules could be implemented for identification of land use classes as agriculture and urban areas. On the bases of what already achieved, is possible to concentrate on strata (land cover classes) that might contain a specific land use class (e.g. Arable Agriculture could not be found on evergreen class but on deciduous either Shrub or Herbaceous) easing the analysis and reducing processing time.

The Land Resource Management Unit and the Global Forest Resource Monitoring Unit at JRC are willing to further test and produce an operational tool for their research mapping projects as well as conducting the accuracy assessment on the remaining sites.

According to Sentinel 2 satellite specification [61], the proposed methodology might be ready to operate with higher spatial and temporal resolution as soon as the satellite will become operational.

The used methodology has been presented at “Landsat Science Team Meeting”, (USGS EROS Center. Sioux Falls, South Dakota on October 29 – 31, 2013) and in “SENTINEL - 2 for Science Workshop”, Frascati in May, 2014.

The algorithm has been distributed at “Regional Centre for Mapping of Resources for Development” (RCMRD), Nairobi, at “Laboratório de Monitoramento Ambiental”, (EMBRAPA), Brazil and for internal use at JRC.

4 Acknowledgments

I must first express my gratitude towards my supervisor and brother, Dr. Dario Simonetti. His support, attention to detail, hard work, patience, motivation and enthusiasm have set an example I hope to match someday. His guidance helped me in all the time of research and writing of this article.

Besides my supervisor, I would like to thank Prof. Dr. Damiano Preatoni for reviewing this dissertation, Andrea Lupi (working at JRC) for providing data for the accuracy assessment and all JRC members who collaborated in this research project.

My sincere thanks also goes to Arcadia SIT srl, for offering me this stage opportunity, and JRC - Ispra, for sharing data and methods and leading me working on this exciting project.

I would also like to thank my entire family for the love, the patience and for all the support they provided me through my life. I must acknowledge my father Giovanni and my mother Liliana who encouraged me and expressed confidence in my abilities when I could only do the opposite, my brother Dario, all his family and Elena.

This job is dedicated to the memory of my beloved father Giovanni.

Your shining example is a constant source of inspiration. I will do my best to make you proud of me.

You are always with us.

5 References

- [1] Bonan, "A land surface model (LSM vansion) for ecological,hydrological, and atmospheric studies: technical description and user's guide.,," National center for atmospheric research, Boulder Colorado, 1996. [Online]. Available:
ftp://ftp.daac.ornl.gov/data/model_archive/LSM/lsm_1.0/comp/NCAR_LSM_Users_Guide.pdf.
- [2] Sellers et al., "Remote sensing of the land surface for studies of global change: models-algorithms-experiments.,," 1995.
- [3] Gower et al., "Direct and indirect estimation of leaf area index, f (APAR), and net primary production of terrestrial ecosystems.,," Vols. Volume 70, Issue 1, October 1999, Pages 29–51, 1999.
- [4] Sellers et al., "Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere.,," vol. Science, 1997.
- [5] Bonan, "Land-atmosphere interactions for climate system models:Coupling biophysical, biogeochemical, and ecosystem dynamical processes.,," Vols. 51(1): 57-73, 1995.
- [6] Kimes, "Optical Remote Sensing of Vegetation: Modeling, Caveats, and Algorithms. Rem. Sens. Environ," Vols. 51, Issue 1, January 1995, Pages 169–188, 1995.
- [7] Goetz et al., "Application of Multitemporal Landsat Data to Map and Monitor Land Cover and Land Use Change in the Chesapeake Bay Watershed, In Analysis of Multi-temporal Remote Sensing Images," vol. World Scientific Publishers, no. pp. 223-232, 2004.
- [8] Loveland et al., "Seasonal land-cover regions of the United States.,," Vols. 85, Issue 2, pages 339–355, June 1995, 1995.
- [9] Schowengerdt, "Remote Sensing: Models and Methods for Image Processing," in *E-Book*, Academic Press, 1997.
- [10] Vuolo et al., "Evaluation of time-series and phenological indicators for land cover classification based on MODIS data," *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII Vol. 8174*, 2011.
- [11] Knight et al., "Regional Scale Land-Cover Characterization using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology Based Approach," *GIScience and Remote Sensing*, pp. 1-23, 2006.
- [12] USGS, "Remote Sensing Phenology," 2013. [Online]. Available:
<http://phenology.cr.usgs.gov/overview.php>.
- [13] Campbell, *Introduction to Remote Sensing*, Taylor & Francis, 2002.
- [14] Lillesand et al., *Remote sensing and image interpretation.*, Chichester, UK: John Wiley & Sons Ltd, 2004.
- [15] Colwell, "History and place of photographic interpretation," in *Manual of photographic interpretation (2nd ed.) E-Book*, Bethesda, Maryland, 1997, pp. 3-23.
- [16] Jones et al., "Derived Heating Rates as Proxy Surface Wetness Data into a Regional Atmospheric Mesoscale Model. Part II: A Case Study.,," vol. Mon. Weath. Rev., 1998.
- [17] Slaton and Raymond-Hunt, "Estimating near-infrared leaf reflectance from leaf structural characteristics," Vols. 88 no. 2 278-284, 2001.
- [18] Peterson and Running, "Applications in forest science and management," *Theory and Applications of Optical Remote sensing*. Wiley-Interscience, pp. 429-453, 1989.
- [19] NASA, "NDVI," 2013. [Online]. Available:
http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php.
- [20] Blaschke, "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS," Vols. No. 6/2001, pp. 12–17, 2001.

- [21] Ballantine et al., "Mapping North African landforms using continental scale unmixing of MODIS imagery," *Remote Sensing Environment* 97(4), vol. Vol. 97, no. 4, pp. 470-483, 2005.
- [22] Giri et al., "A comparative analysis of the Global Land Cover2000 and MODIS land cover data sets.," *Remote Sensing Environment*. 94(1), pp. 123-132, 2005.
- [23] Kang et al., "A regional phenology model for detecting onset of greenness in temperate mixed forests, Korea: an application of MODIS leaf area index.," *Remote Sensing Environment* 86(2), pp. 232-242, 2003.
- [24] NASA, "EOS," 2014. [Online]. Available: <https://earthdata.nasa.gov/data/near-real-time-data/rapid-response>.
- [25] NASA, "About MODIS," 2014. [Online]. Available: <http://modis.gsfc.nasa.gov/>.
- [26] E. C. JRC, "TREES-3 project," 2010. [Online]. Available: <http://bioval.jrc.ec.europa.eu/index.php/pages/GlobalForestResourceMonitoring/4>.
- [27] USGS, "MODIS MRTWeb," 2010. [Online]. Available: URL: <http://mrtweb.cr.usgs.gov/>.
- [28] GDAL, "GDAL - Geospatial Data Abstraction," 2014. [Online]. Available: <http://www.gdal.org/>.
- [29] Xu and Bing, "A Novel Compound Smoother—RMMEH to Reconstruct MODIS NDVI Time Series," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, VOL. 10, NO. 4., 2013.
- [30] IDL, "IDL, scientific programming language," 2014. [Online]. Available: <http://www.exelisvis.com/ProductsServices/IDL.aspx>.
- [31] USGS, "NDVI," 2014. [Online]. Available: http://phenology.cr.usgs.gov/methods_metrics.php.
- [32] Roy et al., "Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data," *REMOTE SENSING OF ENVIRONMENT*, 112(6), pp. 3112-3130, 2008.
- [33] Nguyen, "EVALUATION OF MODIS PRODUCTS OVER FOUR EUROPEAN ECOLOGICAL STUDY SITES," in *FACULTY OF BIOLOGY, CHEMISTRY AND GEOSCIENCES*, UNIVERSITY OF BAYREUTH, GERMANY, 2008.
- [34] USGS, "Landsat mission," 2014. [Online]. Available: <http://landsat.usgs.gov>.
- [35] USGS, "Landsat 8 (LDCM)," 2014. [Online]. Available: <http://landsat.usgs.gov>.
- [36] NASA, "Landsat 8, Data," 2014. [Online]. Available: http://landsat.gsfc.nasa.gov/?page_id=2372.
- [37] NASA, "Landsat 8, application," 2013. [Online]. Available: http://landsat.gsfc.nasa.gov/?page_id=2298.
- [38] USGS, "LandsatLook Viewer," 2014. [Online]. Available: <http://landsatlook.usgs.gov>.
- [39] E. C. JRC, "GLC," 2010. [Online]. Available: <http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php>.
- [40] E. C. JRC, "APAAT," 2013, 2013. [Online]. Available: <http://bioval.jrc.ec.europa.eu/APAAT/>.
- [41] MMNP, "about Mahale," 2012. [Online]. Available: <http://www.mahalepark.org/aboutmahale.html>.
- [42] MMNP, "Vegetation and area map," Mahale Wildlife Conservation Society, Tanzania, 2012.
- [43] Moyer, "Biodiversity of Mahale Mountains National Park, Tanzania.," Wildlife Conservation Society Tanzania Program., 2006.
- [44] ZGF, "Mahale Mountains National Park General Management Plan (2006-2016) .]," 2006. [Online]. Available: www.zgf.de/download/1373/Mahale+GMP+FINAL+print+version.pdf.
- [45] "Mana Pools National Park," 2013. [Online]. Available: <http://www.zimparks.org/index.php/parks-overview/national/mana-pools>.
- [46] "Mana Pools National Park," 2014. [Online]. Available: http://www.zambezi.com/location/mana_pools_national_park.

- [47] "ISRIC," 2012. [Online]. Available: <http://www.isric.org>.
- [48] SOTER, "SOTER Programme," 2012. [Online]. Available: <http://www.isric.org>.
- [49] "West Lunga National Park," 2013. [Online]. Available:
<http://www.zambiatourism.com/destinations/national-parks/west-lunga-national-park>.
- [50] "Tinley," 1977. [Online]. Available: http://www.gorongosa.org/sites/default/files/research/056-tinley_-_gorongosa_1977_0.pdf.
- [51] "Gorongosa National Park," 2013. [Online]. Available: <http://www.gorongosa.org/>.
- [52] Stalmans and Beilfuss, "Long-term plan for Gorongosa National Park vegetation monitoring at multiple scales.,," 2007. [Online]. Available:
http://www.gorongosa.org/sites/default/files/research/050-gorongosa_longterm_vegetation_monitoring_plan_final.pdf.
- [53] Stalmans and Beilfuss, "Tinley's plant species list for the Greater Gorongosa Ecosystem, Moçambique.,," Report by International Conservation Services to the Carr Foundation and the Ministry of Tourism, 2006.
- [54] Stalmans and Beilfuss, "Landscapes of the Gorongosa national park," Report by International Conservation Services to the Carr Foundation and the Ministry of Tourism, 2008.
- [55] Cunliffe and Lynam, "Preliminary vegetation classification and mapping of Gorongosa National Park, Mozambique," Ministry of Tourism of the Government of Mozambique and Carr Foundation, 2005.
- [56] Müller, "VEGETATION SURVEY OF MOUNT GORONGOSA,," Biodiversity Foundation for Africa Zimbabwe, 2012.
- [57] "Tsimanampetsotsa National Park," 2013. [Online]. Available:
<http://www.travelmadagascar.org/PARKS/Tsimanampetsotsa-National-Park.html>.
- [58] Simonetti and Szantoi, "Fast and Robust Topographic Correction Method," *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING* vol. 6 no. 4 p. 1921-1933, 2012.
- [59] Bodart et al., "Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics," *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (2011) 555–563, vol. 66 (2011) 555–563, 2011.
- [60] LCCS, Roma: Food and Agriculture Organization of the United Nations, 2005.
- [61] S.-2. ESA, 2014. [Online].

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*): 00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.

It can be accessed through the Europa server <http://europa.eu/>.

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu>), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.

You can obtain their contact details by sending a fax to (352) 29 29-42758.

European Commission

EUR 26841 EN – Joint Research Centre – Institute for Environment and Sustainability

Title: Phenology-based land cover classification using Landsat 8 time series

Author(s): Edoardo Simonetti, Dario Simonetti, Damiano Preatoni

Luxembourg: Publications Office of the European Union

2014 – 57 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424

ISBN 978-92-79-40844-1

doi: 10.2788/15561

Abstract

This article describes the methodology and results of a new JRC phenology -based classification algorithm able to generate accurate land cover maps in a fully automatic manner from Landsat 8 (L8) remote sensed data available since 12th April 2013 at no charge throughout the USGS website. A preliminary study aiming to bypass the single date classification inaccuracy (mainly due to seasonality) using long term MODIS time series as a “driver” to fill gaps between high resolution data, has been carried out. The high global acquisition frequency (~16 days) and distribution policy are making Landsat 8 product extremely suitable for near real time land cover mapping and monitoring.

Five national parks in east Africa have been selected as study areas (Mahale Mountains, Mana Pools, West Lunga, Gorongosa, Tsimanampetsotsa); they are covering diverse eco-regions and vegetation types, from evergreen to deciduous. A buffer of 20 km around each park has been considered as well.

Selected single date images were first preprocessed in order to convert raw DN values to top of atmosphere (TOA) reflectance and minimizes spectral differences caused by different acquisition time, sun elevation, sun-earth distance, and after processed by the algorithm to generate a thematic raster map with land cover classes. Is worth noting that the single date classification accuracy is closely related to the acquisition date of the image, the status of the vegetation and weather conditions such as cloud and shadows often present in tropical regions; here the need of developing a phenology based algorithm that considers the vegetation evolution and generates a more accurate land cover map including evergreen and deciduous discrimination on the basis of “frequency” rules.

Land cover maps have been created for all parks and an exhaustive accuracy assessment has been carried out on Mahale Mountains and Tsimanampetsotsa. The combined overall accuracy of 82.8% demonstrates the high potentiality of this method and makes it usable at either local or regional scale.

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

*Serving society
Stimulating innovation
Supporting legislation*

