A Prototype online system to process and visualize phenology parameters for Kenya: Case study of Kitale and Bungoma

Conference Paper · April 2013
A Prototype online system to process and visualize phenology parameters for Kenya: Case study of Kitale and Bungoma

Erick Omwandho Opiyo, Thomas Ngigi, Eunice Nduati and and Moffat Mang’ondu

Abstract— Phenology has gained a lot of interest in the past decade as a method of studying biological events of vegetation and crop and how these events interact with both biotic and abiotic forces like climate change. Phenology parameters offer large potential for studying vegetation and crop cycles and its characteristics, but the method of deriving and disseminating this information has been hindered by the large dataset involved in processing and visualizing phenology parameters and the lack of a standardized tool for processing in a user friendly environment. An online phenology monitoring system will aid stakeholders in both governmental, non-governmental and private sector to access information that can be used to study anomalous agronomic conditions, feed into statistical models to carry out crop growth simulation and predict drought. In this study, a prototype phenology online monitoring system was developed to automate the process of downloading, processing and disseminate phenology time series and parameters, extracted from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery in an integrated web system. The focus was on developing software modules for automated downloading, processing, visualizing and disseminating phenology parameters in an online web system. The prototype online phenology system was able to show suitable methods for downloading and keeping track of new satellite images using a web crawler, processing, analyzing and disseminating phenology parameters using an integrated web system.

Keywords— Online, Phenology, Remote Sensing, Vegetation, MODIS

I. INTRODUCTION

A. Phenology

Phenology is a term that is mostly linked to agricultural events such as start of planting, emergence of crops, maturing of crops to harvesting. In many occasions phenological research is mostly geared towards ecosystem evaluation, as compared to whole ecosystem monitoring [36]. Over the last decade phenology has received a lot of focus as compared to whole ecosystem monitoring [36]. Phenology studies provide vital information that is important for various environmental applications. Example data from phenology can be used to diagnose environment response to climatic condition changes, it can be used to study seasonal variability in levels of water, carbon and energy between the atmosphere and land cover [16]. Remotely sensed images provide the only data source that can be used to monitor and collect large scale phenology at recurring interval. Satellite images from remote sensing offer an advantage because remote sensing cover large ecological areas and offer synoptic and repeated collection of data which facilitates multi-temporal monitoring of crop patterns [43], unlike

In the above definition phenomenology can be considered as the study of plant cycles, form the time a plant starts to germinate, through onset of the growing season, for crops when it is harvested or end of greening for forest vegetation. Phenology estimates the rate at which plant develops to maturity and the rate at which its greenness diminishes when it is harvested or senescence. In the definition seasonality was defined as "the occurrence of certain obvious biotic and abiotic events or groups of events within a definite limited period or periods of the astronomical year" [26].

In 1972, the committee of International Biological Program (IBP) suggested further refinement to the definition by adding spatial and temporal framework, therefore the definition was [22]:

“The unit of study may vary from a single species (or variety, clone etc.) to a complete ecosystem. The area involved may be small (for intensive studies on all phenostages of entire ecosystems) or very large (for interregional comparison of significant phenostages). The unit of time is usually the solar year with which the events to be studied are in phase. The events themselves may cover variable time spans, often shorter than the solar year”.

The second definition added two aspects to the study of phenology that is spatial and temporal framework, these two aspects are important in the context of studying phenology from satellite images.

Phenology studies provide vital information that is important for various environmental applications. Example data from phenology can be used to diagnose environment response to climatic condition changes, it can be used to study seasonal variability in levels of water, carbon and energy between the atmosphere and land cover [16]. Remotely sensed images provide the only data source that can be used to monitor and collect large scale phenology at recurring interval. Satellite images from remote sensing offer an advantage because remote sensing cover large ecological areas and offer synoptic and repeated collection of data which facilitates multi-temporal monitoring of crop patterns [43], unlike

Erick O. Opiyo, Department of Geomatic Engineering and Geospatial Information Systems, JKUAT ( phone: +2540724590982; e-mail: e.omwandho@gmail.com).

T. Ngigi, M. and E. Nduati and G. Magond, Department of Mechanical Engineering, JKUAT.
traditional grand phenology collection methods.

Traditional method of phenology data collection or ground based phenology data collection have several setbacks that the use of satellite images can address. Phenology data from ground based data collection methods often have incomplete data points and have limitation when it comes to temporal range. However advances in technology have enabled studying of the earth from space using remote sensing methods, which enable whole ecosystem observation in a global scale using proxy approaches [35], [46].

B. Phenology metrics
Lloyd [27] and Malingreaus [29] came up with a list of 12 phenological metrics that could be derived from Advance Very High Resolution Radiometer (AVHRR) Normalized Difference Vegetative Index (NDVI) data. The parameters can generally be categorized as shown in table 1 below [36].

<table>
<thead>
<tr>
<th>Metric</th>
<th>Phenological interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal NDVI metrics</td>
<td></td>
</tr>
<tr>
<td>Time of onset of greenness</td>
<td>Beginning of measurable photosynthesis</td>
</tr>
<tr>
<td>Time of end of greenness</td>
<td>Cessation of measurable photosynthesis</td>
</tr>
<tr>
<td>Duration of greenness</td>
<td>Duration of photosynthetic activity</td>
</tr>
<tr>
<td>Time of maximum NDVI</td>
<td>Time of maximum measurable photosynthesis</td>
</tr>
<tr>
<td>NDVI-value metrics</td>
<td></td>
</tr>
<tr>
<td>Value of onset of greenness</td>
<td>Level of photosynthetic activity at beginning of growing season</td>
</tr>
<tr>
<td>Value of end of greenness</td>
<td>Level of photosynthetic activity at end of growing season</td>
</tr>
<tr>
<td>Value of maximum NDVI</td>
<td>Maximum measurable level of photosynthetic activity</td>
</tr>
<tr>
<td>Range of NDVI</td>
<td>Range of measurable photosynthetic activity</td>
</tr>
<tr>
<td>Derived metrics</td>
<td></td>
</tr>
<tr>
<td>Time-integrated NDVI</td>
<td>Net primary production</td>
</tr>
<tr>
<td>Rate of greenup</td>
<td>Acceleration of photosynthesis</td>
</tr>
<tr>
<td>Rate of senescence</td>
<td>Deceleration of photosynthesis</td>
</tr>
<tr>
<td>Modality</td>
<td>Periodicity of photosynthetic activity</td>
</tr>
</tbody>
</table>

The first category is based on the onset of events; the second category is based on the NDVI value at which the events occur while the third category is based on time-series calculation.

In this study of crop phenology we will concentrate on four main parameters that can be derived from optical satellite for the study of “land surface phenology”, land surface phenology can be described as the process of development of spatial-temporal vegetation land surfaces as shown by satellite imagery form optical satellite [4], the main four parameters include:

- **Start of growing season** – Beginning of measurable photosynthesis and level of photosynthetic activity at beginning of growing season [36].
- **Onset of senescence** – Cessation of measurable photosynthesis activities and deceleration of photosynthesis [36].
- **Maximum of growing season** – This shows Time of maximum measurable photosynthesis [36].
- **Growing season length** – Duration of time between onset and offset of NDVI or Enhance Vegetation Index (EVI) it shows range of measurable photosynthetic activity

The above four metrics are summaries in the fig. 1 below.

Fig. 1 NDVI time series to depict phenology

The onset of growing season is the time when photosynthesis activities starts, onset of senescence is the time at which photosynthesis activities generally reduce, while maximum of growing season will be looked at as the maximum NDVI value that is recorded before senescence starts and finally growing season length will be calculated as the duration of photosynthesis activities

C. Objectives

Main Objectives

- To develop an online web based phenology monitoring system
Secondary Objectives

- Include development of automatic download procedures for Moderate Resolution Imaging Spectroradiometer (MODIS) imagery
- To develop scripts that would facilitate cloud cleaning from MODIS imagery
- To extract phenology parameters from EVI found on MODIS imagery
- To create a raster database to hold the imagery
- Create a web mapping application that would display phenology graph based on a point user clicked on a map

D. Importance of studying phenology

- Phenology can aid in showing anomalous agronomic conditions
- Data extracted from phenology can be fed into statistical models to help in predicting and simulating crop output
- Past crop performance data extracted from satellite imagery can be used by government and farmers to plan planting periods
- Phenology can be used to study regional to global trends that can help predict draught events [8].

E. Study Area

The study area of the project will mainly focus on two MODIS scenes i.e, scene path 21 and row 08 and path 21 and row 09. The two scenes where chosen because of the relative large dataset needed to be able to show time-series of phenology of vegetation over 10 year time period that MODIS has been operational. The two sites were choose to display phenology as they are major maize growing areas in Kenya Western Province that is Bungoma and Kitale [20] show in figure 2 below.

Fig. 2 Map showing case study area

II. LITERATURE REVIEW

A. Extraction of phenology parameters from satellite imagery

Phenology is extracted from satellite imagery by studying time series generated from remotely sensed satellite imagery. Satellite imagery acquired from remotely sensing devices provide a viable source of both spatial and temporal observation of phenological patterns as they offer wide coverage in spatial location of vegetation to be monitored and temporal resolution which can be used to plot different occurrence of vegetation activities over time. In the past use of satellite imagery like NOAA AVHRR have been used to monitor growing season greenness in the northern hemisphere due to surface warming [30], [54].

Detecting phenological activities requires high temporal resolution which cannot be provided by Landsat based imagery because of its temporal resolution with revisit period ranging from 16-18 days in some areas; therefore more spatially coarse satellite imagery like NOAA AVHRR, MODIS and SPOT-Vegetation which has temporal resolution of 1 day are more preferred because they offer high temporal resolution.

Phenology is extracted from form satellite imagery by studying time series generated from remotely sensed satellite imagery. Satellite imagery acquired from remotely sensing devices provide a viable source of both spatial and temporal observation of phenological patterns as they offer wide coverage in spatial location of vegetation to be monitored and temporal resolution which can be used to plot different occurrence of vegetation activities over time. In the past use of satellite imagery like NOAA AVHRR have been used to monitor growing season greenness in the northern hemisphere due to surface warming [30], [54].

Detecting phenological activities requires high temporal resolution which cannot be provided by Landsat based imagery because of its temporal resolution with revisit period ranging from 16-18 days in some areas; therefore more spatially coarse satellite imagery like NOAA AVHRR, MODIS and SPOT-Vegetation which has temporal resolution of 1 day are more preferred.

To monitor phenology from satellite imagery vegetation indices (VI) are used, this is because they correlate well to green biomass and leaf area index [1], [3]. VI is an algorithm that combines one or two surface reflectance to highlight a given characteristic of vegetation. There are more than 150 vegetation indices but only a few have been tested to have significant biophysical basis [13]. Vegetation indices can be broadly categorized into various groups depending on similar characteristic of vegetation that they show. They include Broadband Greenness, Light Use Efficiency, Canopy Nitrogen, Dry or Senescent Carbon, Leaf Pigments, Canopy Water Content [13].

Study of phenology only concentrates on two indices that are found within the Broadband Greenness category above, they include NDVI and EVI. NDVI is one of the most established and frequently used vegetation indices; it applies a
normalized difference formula and uses high chlorophyll reflectance and absorption regions of the electromagnetic spectrum. Its formula is given by formula show in equation (i) below. The value of the index range is between -1 and 1 and green vegetation range is 0.2 to 0.8 [38], [48], [23], and [40].

\[
NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad \text{Equ (i)}
\]

The use of NDVI is affected in regions of saturated dense vegetation condition where leaf area index is high. EVI was developed to overcome some of the shortcomings of NDVI in dense saturated vegetation regions. EVI enhanced vegetation signal in high LAI regions by using the blue band of the electromagnetic spectrum this reduces backscatter from soil background and it also eliminates some of the atmospheric effects such as aerosol scattering [56]. Its formula is shown in equation (ii) below.

\[
EVI = 2.5 \left( \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6\rho_{RED} - 7.5\rho_{BLUE} + 1} \right) \quad \text{Equ (ii)}
\]

For most part of the last decades AVHRR indices was the only source of data for studying phenology, however since it was not meant for land application the shift now is moving to other data source like MODIS, this is because AVHRR was not adequate for examining vegetation application since it lacked calibration data, modest geometric cataloguing and effort needed in cloud screening [17].

B. Time series analysis of satellite imagery to study phenology

Time series can be defined as a collection of observations made consistently over time [5]. Time series data generated from satellite imagery have proved to be very valuable for studying global environment variables. Its applications range from climate studies, biochemical processes, land-cover changes to phenology studies [40], [44], [53], [28], [11]. The advantage that time series offers over single images spectral analysis is that time series enables continuous observation over a time period. Time series analysis is important because it allows detailed analysis of land cover changes over a period of time which can complement land cover studies [6].

Time series obtained from EVI of MODIS can be used to study vegetation characteristics; some of the properties that can be obtained include seasonal variation of vegetation and progress [38], [48], [21]. Time series obtained for long periods of time from EVI can also be used to show large scale shift in biophysical environment and land-use cover changes [32].

C. Existing phenology processing and analysis software

Currently there are only two published software that are used to study phenology or time-series related to vegetation.

TiSeg

TiSeg (time-series generator, shown on image below) [8] is software written in Interactive data language (IDL) [37]. TiSeg is used to analyse pixel level data from MODIS and is used to generate time-series data from varies MODIS products. It includes tools that give the user functionalities to visualize quality of some MODIS products by producing auxiliary data that include two indices that include invalid pixels layer and gap of invalid pixel that exist between temporal images [8], it can also be used to generate time-series that can be used to study phenology.

TIMESAT

TIMESAT is mainly a FORTRAN90 program for extracting vegetation seasonality characteristics but it also contains routines written in MatLab. It mainly uses Savitzky–Golay filtering algorithm to display time-series. TIMESAT implements mainly three processing methods to obtain vegetation characteristics which are mainly based on least square method that is used to fit upper cover of NDVI data. TIMESAT has timesat as the main driver subroutine which is always initialized once the application loads, gensincos that computes nine sine and cosine function and gengauss that computes Gaussian functions [32]. The main graphics user interface to TIMESAT is shown below [25].
Research project online phenology monitoring system

In this study an online system was proposed as the main method to automate the process of retrieving, extracting, processing and visualizing phenology parameters. Unlike in the above mentioned application an online system has got some advantages as opposed to desktop applications. Since processing and visualizing of data to extract phenology parameters is a lengthy process that entails downloading large volumes of satellite imagery, to processing this data to remove invalid pixels which in some cases can take days if not weeks depending on the scene to be processed.

In this study an automated system was developed and tested that downloaded MOD13Q1 products from MODIS website using a web-crawler [7] that scheduled download if an new image was uploaded for the test scene, once the data was downloaded automated scripts extracted the required dataset from HDF data array and saved them as TIFF imagery in a folder library. Once the downloading was finished the image were arranged in a composite of five images and filter for invalid pixel using a methods proposed by Xiao et al [55] and used by Fensholt et al. [15]. An online web interface was developed that used Google Map API (https://developers.google.com/maps/) that let the user click on any point and the coordinates of the point where extracted using mouse listeners that were sent back to the server for processing and respective time series of point clicked was returned to the user as graph images on the website. This is shown in figure 7 below.

Some advantages of web applications over desktop applications include [33]:
1. Web applications avoid the burden in deploying in each client machine.
2. Don't have to enforce version check in client machine.
3. Updates are easier.
5. No administrator rights checking.
6. Can access from anywhere.
7. Platform independent.
8. Support and maintenance are easier.
9. Adaptability in mobile applications.

D. Data Sources

Due to time series nature of phenology studies only a few satellite meet this requirement they include sparsely spectral resolution images from NOAA-AVHRR and MODIS. Satellite imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite was chosen for this project. This is because satellite imagery from MODIS offer wide coverage and repeated temporal cycle within a given area over different time periods this is beneficial for monitoring vegetation at global scale [30].

MODIS on board both the Terra and Aqua spacecraft provides improved capability to monitor vegetation applications. MODIS provides good frequency global coverage of vegetation frequency and temporal resolution that can be well suited for the study of phonological events from season to yearly patterns to local to international coverage [9], [36], [49], [53].

MODIS instrument is operated on both terra and aqua spacecraft. It has a view swath of 2,330 Km it views the entire earth every one to two days. It measures 36 bands at spatial resolution of 250, 500 and 1000 meters [31].
MODIS products are mainly in the sinusoidal grid tiling system. The tiles are 10 degree by 10 degrees at the equator. MODIS data are distributed at processing level 2 or higher [31]. MODIS instrument contains spectral bands that fall within the range 620–670 nm to 2105–2155 nm that are specifically suited for studying land applications and have a spectral resolution of 250m to 1 km [18].

Global MODIS indices provide consistent spatial and temporal comparison of vegetation around the globe. MODIS near-infra red, red and blue band are used to determine daily vegetation indices. MODIS offers two products the NDVI and the EVI. MODIS new EVI product enhances vegetation monitoring by reducing the effect of canopy cover variation in highly dense vegetation cover areas, it also uses blue band to remove cloud and aerosol contamination on the images. MODIS NDVI and EVI products have been generated from atmospherically corrected bi-directional images that have been masked for water, cloud and shadows [31].

III. METHODOLOGY

Data processing process chain as adopted by [8] was used. The processing chain involved downloading, processing and disseminating phenology parameters in an online system is shown in figure 8 below. The processing involves downloading and rearranging the satellite imagery in a library structure; a script would then run in the directory extracting only layers that are needed for extracting phenology parameters. After sub setting quality assurance is done on the images to filter for cloud pixels, then parameters are calculated and the images are arranges to be displayed to generate phenology time-series.

Since this is a real time system that will involve downloading new images whenever they are updated it would involve developing a web crawler program that would browse MODIS imagery websites, check and harvest new data while

<table>
<thead>
<tr>
<th>MODIS MOD13Q1 LAYERS</th>
<th>INVALID RANGE</th>
<th>SCALING FACTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>250m 16 days NDVI</td>
<td>-2000, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days EVI</td>
<td>-2000, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days VI Quality detailed QA</td>
<td>0, 65534</td>
<td>NA</td>
</tr>
<tr>
<td>250m 16 days red reflectance (Band 1)</td>
<td>0, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days NIR reflectance (Band 2)</td>
<td>0, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days blue reflectance (Band 3)</td>
<td>0, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days MIR reflectance (Band 7)</td>
<td>0, 10000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m 16 days view zenith angle</td>
<td>-9000, 9000</td>
<td>0.01</td>
</tr>
<tr>
<td>250m 16 days sun zenith angle</td>
<td>-9000, 9000</td>
<td>0.01</td>
</tr>
<tr>
<td>250m 16 days relative azimuth angle</td>
<td>-3600, 3600</td>
<td>0.1</td>
</tr>
<tr>
<td>250m 16 days composite day of the year</td>
<td>1, 366</td>
<td>NA</td>
</tr>
<tr>
<td>250m 16 days pixel reliability summary QA</td>
<td>0, 3</td>
<td>NA</td>
</tr>
</tbody>
</table>
updating local database of new images uploaded to the MODIS websites, while designing this program special attention would be made to use standard web policies avoid website overloading. An excerpt of policy used in the design of web crawlers as stated by [7], is a combination of policies that incorporate:

- Selection policy that states which pages to download,
- Re-visit policy that states when to check for changes to the pages,
- Politeness policy that states how to avoid overloading Web sites, and
- Parallelization policy that states how to coordinate distributed Web crawlers.

Web crawlers are types of software agents that are generally given a set of URLs that are placed in the crawler frontier and visited using a set of policies stated above. For the case of downloading satellite imagery a web crawler with the ability to login on websites would be designed and it will have a set of imagery hyperlinks would be used and constantly checked for new images. Sample architecture for the system would be as shown in figure 10 below:

![Data download flow diagram](image)

The crawler would have different set of arms one will be updating new information that would be used by the scheduler to spool information to download, another arm with multithreading capabilities would be used to download MODIS imagery in HDF format, convert it and reproject. Once the crawler has finished its task it will trigger another task to start cleaning MOIS imagery from cloud cover using Quality and Reliability bands.

### B. Data processing

Data processing procedure involved arranging the downloaded data in HDF data format into a well-defined map library, which was defined by the path and row of the scene where the image was located, further the library was subdivided into respective time of the year when the image was acquired this was arranged according to the image Julian date. After the data was arranged in the map library a script written using python programming language was used to read the raw HDF images and sub setting only the required layer defined by the table above. For the case of this study EVI, DOY, Reliability and Quality assurance layers were extracted from the raw MODIS MOD13Q1 data products and arranged in a root folder labeled with the respective data product, while the sub folders were arranged as discussed in the library above. The arranging of this data was to facilitate easier location, running of script that cleaned the images for cloud.

#### C. Cloud cleaning

Data generated from satellite imagery are often tainted with invalid pixel values, this largely due to acquisition systems, contact with the atmosphere and during image production [39]. Several techniques have been suggested to reduce effect of contaminated pixel in time series generation. For remotely sensed data like NOAA-AVHRR [21].

Some of the data pixels to be used to extract and display of phenology parameters contained pixels that were contaminated with cloud, a smoothing filtering algorithm was developed similar to a method developed by Xiao et al. [55] and used by Fensholt et al. [15] for smoothing SPOT-vegetation and AVHRR GIMMS and MODIS respectively. MODIS contains layers with quality assurance to help users make a decision on where best to use the data. This quality assurance (QA) information gives clues on the usability that is the application on particular purpose and usefulness to achieve given purpose [31]. QA within the MOD13Q1 is found within two scientific pixel-level metadata layers which are more reliable than file-based metadata quality information, they include: VI Quality and Pixel Reliability QA layers. The first layer contains multiple quality information sources encoded in binary encoding and the second one contains single information sources like pixel reliability in vegetation index products [31]. This is shown in the diagrams below.

<p>| TABLE 3 |
| MODIS MOD13Q1 QUALITY LAYER |
|---|---|---|</p>
<table>
<thead>
<tr>
<th>bit</th>
<th>Long Name</th>
<th>Value</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1</td>
<td>MODLAND_QA</td>
<td>00</td>
<td>VI produced, good quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>VI produced, but check other QA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>Pixel produced, but most probably cloudy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>Pixel not produced due to other reasons than clouds</td>
</tr>
<tr>
<td>2–5</td>
<td>VI usefulness</td>
<td>0000</td>
<td>Highest quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0001</td>
<td>Lower quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0010</td>
<td>Decreasing quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0100</td>
<td>Decreasing quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>Decreasing quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1001</td>
<td>Decreasing quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1010</td>
<td>Decreasing quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1100</td>
<td>Lowest quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1101</td>
<td>Quality so low that it is not useful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1110</td>
<td>L1B data faulty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1111</td>
<td>Not useful for any other reason/not processed</td>
</tr>
</tbody>
</table>
### D. Cloud filter algorithm

The cloud filtering algorithm used was similar to method used by Xiao et al. [55] and used by Fensholt et al. [15] it involved arranging the EVI and Reliability layer of the MOD13Q1 into five images composite of 16-days temporal resolution the images composites as shown below the resultant smoothed image was labelled EVI_n-2 Filter, where n is the nth image in the time series R is reliability layer, EVI is enhanced vegetation index layer and WM is water mask, the procedure involves subtracting image from number shown in the sequence of five images as it propagates through the time series. This is shown in figure 10 below.

![Fig. 10 Cloud cleaning algorithm](image)

### E. Phenology extraction

For this study a comparison would be made from different phenology calculation methods listed below [22] and suitable one was be chosen to be best used to predict start of growing season

Phenology parameters usually consist of:
- Onset of greening
- Onset of senescence
- Timing of the maximum of the growing season
- Growing season length base on the analysis of Vegetative index curve – [36].

Method used in calculating phenology can generally be categories into
- Threshold
- Derivatives
- Smooth Algorithms
- Model fit

Threshold – This is the most common used method to determine phenology parameters. Where the threshold value is arbitrarily set at a given value then the onset of the greening period is determined as that day of year (DOY) that the VI would crosses this threshold. The threshold can be based on different approaches applied by different authors.
- Threshold based on Long-term mean VI [24]
- Threshold based on baseline year [42]
- Threshold based on NDVI ratios [49].
- Threshold based on Normalized difference water index (NDWI) [11]

Derivatives – This approach generally accepts maximum increase and decease in NDVI as start of greening. Two methods are described by Hudson and Keatley [22] and they include:
- Greatest increase and decrease in vegetation index – this method was applied by [47]. They determined the derivative
based on three consecutive ten-day composites and deduced that increase in NDVI is designated as start of greening.

2. Another method used under derivative is camelback phenology algorithm [2]. Used a moving window consisting of five composite passed over a time series for every pixel then the final derivative is determined as the point where the second derivative reaches a local maximum within a 13-composite window.

Other methods that were researched included:

Methods based on Smooth functions and model fits
- Autoregressive Moving Average Model fit
- Logistic models
- Gaussian local function
- Growing degree-days

The threshold method was used for the process of extracting phenology parameters.

IV. RESULTS
A. Vegetation time series

The vegetation time-series of the two selected sites were as shown in figure 11 and 12 below. It can be clearly shown that time series can be used to study the trend that can be depicted by seasonality of a crop; this is shown by the movement of up and down of the EVI values as time progresses which can be shown on the online phenology monitoring system web interface.

Fig. 11 Time series for Bugoma showing relative change of seasonality

Fig. 12 Time series for Kitale showing relative change of seasonality

B. Phenology parameters

The following maps where derived to show phenology parameters that was show on the website

Fig. 13 Map showing number of growing seasons
B. Data dissemination

Fig. 17 Map showing maize phenology in Bungoma

map interface with click and drop marker
Phenology time series graph
V. CONCLUSION

Phenology presents validated metrics that can be used to study vegetation growth, agronomic conditions, how vegetation performs within its environment, seasonality, plant life cycles and timing of biological events which are important in crop growth. Since information that is derived from phenology is important for studying vegetation its wide use has been hindered by inadequate research in temporal features of vegetation from satellite imagery and lack of functionality in major remote sensing processing software. This research has shown that an online application system would enable wide usage as opposed to desktop software that require expertise to process the data before use. Some of the indicators that have been show in this study that can be derived from phenology metrics include start of growing season, rate of photosynthesis activity, peak of greenness, end of photosynthesis activity and duration of greenness. This information is important to both stakeholders in the private and governmental institutions because it can be used in planning start of growing seasons, planning for food security when phenology metrics are available over a long time series period. Phenology can also find suitable application in studying how the land cover is affected by climatic changes. It has been show from this study that remote sensing data acquired from satellite imagery offer better area coverage in both space and time which can be advantages when studying different species of vegetation over long time periods, since data from satellite imagery like MODIS has been acquired and archived over long time periods. An online system would deliver phenology information more quickly to users who are not acquainted to processing satellite data and programming that is required of existing software.

APPENDIX

Cloud filtering algorithm:

if \((WM == 0)\) 
then \(EVI_{n-2} \_Filter = -3000\)
else if \((WM == 1 \text{ and } R_{n-2} == 0)\) 
then \(EVI_{n2} \_Filter = EVI_{n-2}\)
else if \((R_{n-2} == 1)\) { 
if \((R_{n-3} \text{ and } R_{n-1})\) 
then \(EVI_{n-2} \_Filter = \frac{EVI_{n-3} + EVI_{n-1}}{2}\)
else if \((R_{n-3} == 0 \text{ and } R_{n-1} != 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n-3}\)
else if \((EVI_{n-3} != 0 \text{ and } EVI_{n-1} == 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n-1}\)
else if \((R_{n-4} == 0 \text{ and } R_{n-3} != 0 \text{ and } R_{n-1} == 0 \text{ and } R_{n} != 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n-4} + EVI_{n} / 2\)
else if \((R_{n-3} != 0 \text{ and } R_{n-1} != 0 \text{ and } R_{n} == 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n}\)
else if \((R_{n-4} == 0 \text{ and } R_{n-3} != 0 \text{ and } R_{n-1} == 0 \text{ and } R_{n} != 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n-4}\)
else if \((R_{n-4} != 0 \text{ and } R_{n-3} == 0 \text{ and } R_{n-1} != 0 \text{ and } R_{n} == 0)\) 
then \(EVI_{n-2} \_Filter = EVI_{n-2}\)
else if \((R_{n-2} > 1)\) 
then \(EVI_{n-2} \_Filter = -3000\).

ACKNOWLEDGEMENT

I would like to thank Dr. Thomas Gumbricht and Dr. Tor Vagen, who supervised me while I was doing projects related to phenology at the World Agroforestry Centre.

REFERENCES

[2] Baltzer H, Gerard F, George C et al. (2007) Coupling of vegetation growing season anomalies and fire activity with hemispheric and...
and land surface temperature. Theoretical & Applied Climatology, 105(1/2), 37-50. doi:10.1007/s00704-010-0374-8


[37] Research Systems Inc., 2004. IDL versions 5.6 for Microsoft Windows and 6.0.1 for Linux x86 m32. URL http://www.rsinc.com


