An Investigation into Maize Lethal Necrosis Severity Mapping in Kenya using RapidEye and Landsat-8

Hosein Jafari Jozani (H.J.), Michael Thiel (M.T.), Michael Hahn (M.H.), Kyalо Richard (K.R.), and Tobias Landmann (T.L.)
Department of Remote Sensing, University of Würzburg, Oswald-Külpe-Weg 86, 97074 Würzburg, Germany; michael.thiel@uni-wuerzburg.de (M.T.)
Department of Geomatics, University of Applied Science Stuttgart (HFT), Schellingstraße 24, 70174 Stuttgart, Germany; michael.hahn@hft-stuttgart.de (M.H.); hoseinjafary@aol.com (H.J.)
International Center for Insect Physiology and Ecology (ICIPE), P.O. Box 30772, 00100 Nairobi, Kenya; rkyalo@icipe.org (K.R.);
Center for Development Research (ZEF), Department of Ecology and Natural Resources Management, University of Bonn, Walter-Flex-Str. 3, 53113 Bonn, Germany; cb@uni-bonn.de

Abstract
The spatial information of crops and cropping systems can provide useful information about disease outbreak mechanism in croplands. In September 2011, a severe outbreak of Maize Lethal Necrosis (MLN) reported at lower elevations (1,900 meter above sea level) in Southern Rift Valley, Longisa division of Bomet County, Kenya. The heterogeneous small scale farms and cloud cover in study area are some of challenges against application of remote sensing in this region. Aims of this study are; to classify maize fields, to discriminate mono/intercropping, continuous/rotation cropping systems, to identify severe MLN occurrence and finally to determine if any relevance between severe MLN occurrences with cropping system and ecological variables exist. Infield data collection was accomplished by African insect for food and health (icipe) (December 2014, January and August 2015), which provided GPS data, crop type, crop conditions (Type, Physical condition, growth stage, and MLN severity), cropping system (mono/inter cropping) and crop rotation (continuous/rotation). High-resolution RapidEye (RE) images, medium resolution Synthetic Aperture RADAR (SAR) Sentinel-1 (S1), Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) of 30 meter resolution and 12-months of Landsat-8 (L8) images were acquired. Several vegetation indices were calculated from RE and L8 adopted for classification. The Random Forest (RF) and One Class Classifier (OCC) were performed in the absence of complete and representative validation data of Land Cover Land Use (LCLU) samples, and the results compared. The RE and L8 images covered two cropping seasons. Also Normalized Difference Vegetation Index (NDVI) in regional-level time series calculated out of 12 L8 images to derive and analysed. As first step, RF classifier employed to produce a LULC map. Secondly, mono and intercropping maize fields identified via maize fields reclassification. The third step, severe MLN classified via adopting OCC classifier. As fourth step, cropping system determined (continuous/rotational) through applying raster analysis. Finally the relationship among cropping system, ecological variables and MLN severity was investigated if any.

Key words: Cropping systems, RF, OCC, MLN severity, NDVI, Ecological variables.

1. Introduction
Total population of the Kenya at 2014 was 45.5 Million individuals that the share of rural and urban area was 74.4% and 25.6% respectively. Within the crop income category, maize with horticultural crops, covering 14% of total household income in Kenya. At the same time, coffee and tea account for a combined 5.6% of total gross income. In the mid-1970s maize accounted for about 35% of the value of total crop production in Kenya while in the late 1990s, it reduced to 28% of the value of total crop production (Olwande, Sikei et al. 2009). Since 2008 the area under cereal production increased ~ 647,762 ha and the cereal production reached 1,446,328 tonnes.

Maize farming is the backbone of food security in Kenya. Maize fields increased from 1,700,000 ha at 2008 to 2,159,322 ha at 2012, which is the highest since 1961. Amid 2008 and 2012, maize production increased 1,382,643 tonnes. At 2014 around Kenya harvested 3,513,171 tonnes maize although it...
still imports tonnes of maize from neighbouring countries like Tanzania each year. In recent years, the price of maize has tripled rising from 700 Kenyan Shillings a bag to an average of 3000 Kenyan Shillings (2015) (kenya.info.ke). Maize consumption is estimated at 98 kilograms per person per year.

MLN disease is caused by the co-infection of Maize chlorotic mottle virus (MCMV) and the cereal viruses in the Potyviridae group, such as Sugarcane mosaic virus (SCMV), Maize dwarf mosaic virus (MDMV), or Wheat streak mosaic virus (WSMV). In combination, these two viruses rapidly produce a synergistic reaction that seriously damages or kills infected plants at any growing stage (Makumbi and Wangai 2013).

The first identification of MLN disease was in USA, 1976 (Naught, Styer et al. 1978). In Kenya, first reports of an unknown disease outbreak was in September 2011 at Bomet County. Further virological analyses identified the unknown disease as MLN (Adams, Miano et al. 2013). Kenya. MLN hotspots of Bomet reported at 0.8490 S 35.3874 E (Kusia, Subramanian et al. 2015). The MLN have threatened farmers income and food security (Center, Nairobi et al. 2014) as in 2012, farmers experienced up to 90% crop loss, equivalent to 126,000 metric tons valued at $52 million (Mahuku, Lockhart et al. 2015).

To overcome MLN outbreak, different approaches were conducted in farm-level (closed-season Observation, alternative crops, Chemical spraying, Uprooting and destruction, Early planting, intercropping), and in Seed-level (fungicide seed dressing etc.). Additionally, some international organizations such as International Centre for Maize and Wheat Improvement (CIMMYT) have set up a MLN screening facility at the Naivasha Research Station, Nakuru County (Gitonga 2014) to produce an MLN resistant variety.

Manual data collection in large geographical areas is expensive both in the matter of time and money. In this way, remote sensing is an option to evaluate the outbreak itself and besides effectiveness of the above mentioned approaches. Even thou remote sensing itself has many limitations related to sensors, weather, geographical, topological etc.

“Mono-cropping is the agricultural practice of growing a single crop year after year on the same land, in the absence of rotation through other crops or growing multiple crops on the same land. Inter-cropping system defined as sowing two or three crops together on the same land, one being the main crop and the others the subsidiaries. It ensures variety of food and brings more ecosystem stability” (agriwiki.org). “Continuous cropping defined as growing of a single crop species on a field year after year contrast crop rotation and inter cropping” (usda).

As reported by Bomet infield data, the cropping system at study area are consisting of mono-cropping (maize only)/inter-cropping (maize and legumes such as cow peas or beans) and continuous/rotation cropping systems. Crop rotation is an example which can control soil-borne diseases through limiting the MLN hosts for particular number of years (Divya Rani and Sudini 2013). According to infield data, maize crops were in different phenological stages during data collection and farm sizes varies widely. Previous works in Africa have specified small scale and highly fragmented farms as challenges for crop mapping. Moreover, discrimination of rain-fed crops from natural vegetation are difficult during the growing stage at rainy season (when both farms and surrounding vegetation are at the same phenological stages) (Conrad, Colditz et al. 2011).

According to (Löw, Schorcht et al. 2012, Ianninia, Molijn et al. 2013, Zillmann and Weichelt 2014), vegetation indices, multi-sensor integration and multi temporal images utilization can improve crop classification accuracy. In this work, RE images, in combination with S1 C-band (VV Single-polarization), L8, SRTM DEM data and several derived vegetation indices employed for classification.

Over other image classification methods, RF is a powerful option for classification purpose as non-parametric, capable of using continuous and categorical data sets, easy to parameterize, dealing with outliers in training data and not sensitive to over-fitting. RF calculates ancillary information such as classification error and variable importance. It performs better when dealing with several images (Cutler, Edwards et al. 2007).

On the other hand, dealing with insufficient LULC samples, the OCC is an option. It performs well dealing with few training data. It gives the target classes a label “Positive” and the remaining classes the label “Negative”. Therefore, OCC just classify the target class while RF will classify every single pixels (Heinl, Walde et al. 2009, Whiteside,
Vegetation indices are quantitative measurements which comparing to individual spectral bands are more sensitive for green vegetation discrimination. They could improve remote sensing image interpretation. In general, vegetation indices do not have a standard unique universal value. Several parameters such as atmosphere, sensor calibration, sensor viewing conditions, soil moisture, solar illumination, colour and brightness, affect vegetation indices. In addition, heterogeneous environment, with a mixture of vegetation and other ground elements in the pixels, makes the study of vegetation indices more challenging (Bannari, Morin et al. 1995).

In this work, 1) a general classification approach performed based on 5 classes (Maize, non-maize, trees, water, non-vegetation) 2) maize fields extracted 3) result of step 2 re-classified for the mono/inter cropping system 4) RE and L8 maize fields time series were analysed using ArcGIS raster analysis to discriminate continuous from non-continuous cropping system 5) an OCC classification performed to determine MLN severity, and finally, the relationship between

**Figure 1.** A brief description of the data processing and analysis.
cropping system, high MLN severity occurrence and ecological variables investigated if any.

Methods

Basically, a hierarchical classification approach hired to map the cropping systems, using multi sensor, multi-temporal data classified by RF and OCC. In the first step, a general land use/land cover (LULC) classification map were created based on RE images and RF classifier. Later on, the second classification conducted based on L8 images and OCC classifier. Later on maize fields extracted from both RE and L8 classification to reduce data complexity. Afterward, the raster analyses on maize fields of two continuous season determined continuous/rotational cropping fields. In addition, maize fields were reclassified for mono/inter cropping system. Later on, fields which affected by the high severity MLN were classified using OCC classifier based on the last data collection on Aug. 2014. Ecological variables such as height, precipitation, aspect were extracted at study area. Finally, the relationship between cropping system, sever MLN occurrence and ecological variables were investigated. A brief description of the data processing and analysis employed visualized as Figure 1.

Study area

Bomet is amongst seven high potential maize production zone in Kenya and agriculture is the main economic activity (Olwande, Sikei et al. 2009). The county has a population of 730,129 (2009 census), an area of 1,997.9 km² and highest elevation of 1,962 m (6,437 ft.) (citypopulation.de). Average maize yield in eastern Africa is 2.03 ton/ha as compared to global average of 6.06 ton/ha. Due to biotic and abiotic constraints this amount in Kenya is ~ 1.7 ton/ha (FAO).

The most prevalent crops in the region are maize, beans, and cowpeas. In Bomet region, roughly 70% of households were selling maize (mean range of 3 tons), however, about 20% of small-scale households only purchase maize, or purchase more maize than they sell (Nyoro, Kirimi et al. 2004). At 2012 the Maize stalks provided 32% of livestock food in Bomet region (Turgut 2011).

With an average cumulative rainfall ranging from 500 to 2000mm (ranging from less than 20mm to more than 120mm average per month) (Figure 2) and mean annual maximum temperature of 28°C (Bryan, Ringler et al. 2013), it is located in a semi-arid climate. Rainfall peaks twice each year in Kenya: in March-May and September-November (Figure 2).

Vast majority of farmers consider the March rains to be in their major maize season (Hassan 1996). Respectively the cropping system, high MLN severity occurrence and ecological variables investigated if any.
The calendar divides to two cropping season, long rain and short rain, which are following different cropping system (FAO). The irrigation system in the target area is rain-fed.

The covered study area is ~ 61°61 Square Kilometre (01°18′16.8887″S, 034°52′04.4721″E, 00°44′43.1677″S, 035°25′23.3662″E) (Image 1) with elevation ranging from 1628 to 2194m above mean sea level.

Several decisions make an optimal planting regime such as date of planting, seeding rate and arrangement, selection of a suitable cultivar (variety), and cropping intensity (single vs. multiple cropping (per year)) and pattern (intercropping (maize and another crop) or (mono-cropping (only maize)) (Hassan 1996). Both mono-cropping and intercropping system are applying Bomet region. According to in-field data, 67% of maize fields were under mono-cropping system (just maize) and 33% were intercropped with beans or potatoes.

In those regions with uncertainties in rainfall patterns majority of farmers apply maize intercropping system with beans. In addition, irrigated farming is also practiced in locations neighbouring the Athi River with small scale farming of vegetables, tomatoes and chili peppers. (Macharia 2004).

Satellite data and pre-processing

**RapidEye**

Multi-temporal RE images provided from the RE Science Archive (RESA) of the German Aerospace Centre (DLR) on Dec. 2014 (09-12-2014, 12-12-2014), and Jan. 2015 (22-1-2015, and 23-1-2015) for Bomet region.

ENVI ATCOR3 was utilized for atmospheric correction of images (Richter and Schläpfer 2005, Forkuor, Conrad et al. 2014) SRTM DEM (30m, cropped for study area). The SRTM DEM resampled (bilinear interpolation) to 5m and transformed to UTM-wgs84-36S. The Band Sequential (BSQ) format DEM and RE image were used as input data for ATCOR-3. Two files of slope and aspect (BSQ format) were produced by the ATCOR3 automatically to be used for atmospheric correction process. This process have been repeated for each tile (12 tile) of each mosaic (totally 24 times). The detection of change over time is an important aspect of the analysis of digital satellite imagery. Iteratively re-weighted multivariate alteration detection (IR-MAD or iMAD) algorithm (Canty and Nielsen 2008) is such a method. Because of different acquisition date\time of RE mosaic tiles, IMAD (extension) in ENVI software (from Github) were adopted for normalization. To minimize the normalization
process, those central tiles were chosen as reference for normalization process. Chi-square product from overlapped area (target and reference image) applied on the target image. The final result is the radiometrically normalized image. Normalized tiles were mosaicked (~ 61*61 Square Kilometre).

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints or sensors. Registration is required in remote sensing (image mosaicking multispectral classification, environmental monitoring, change detection, etc.) (Srikrishnan, Araiza et al. 2001) and those produced mosaics were co-registered accordingly (image-to-image).

**Landsat**

Multi temporal images of L8 (30m) were acquired from USGS explorer for the Bomet region at Nov. 2014 to Aug. 2015 with less cloud-cover percentage in the study area (covered two cropping seasons). Preferably all of the time steps were chosen from L8 CDR while they were already corrected to the surface reflectance by USGS. These images were cropped to the size of RE mosaic products. Resized images were co-registered via ENVI software. The SRTM DEM arc 1-sec image were resampled with L8 images (bilinear interpolation). Cloud mask were created manually using ENVI classic software.

Both of RE and L8 images had cloud spots. Between, automatic cloud removal and manual approach, in this work manual method were chosen to detect and mask clouds and cloud shadows (Kross, McNairn et al. 2015). Recommended map projection is UTM WGS-84 ellipsoid Zone 36 for Bomet (Forkuor, Conrad et al. 2014). Projection was applied on SAR data (SRTM DEM arc 1-sec, Sentinel-1) via resampling in ENVI software.

**SAR Data**

Both SRTM DEM (30m) and Sentinel-1 (S1) data in Interferometric Wide Swath (IW) mode, with VV and VH polarizations utilized in this work. Available at European Space Agency (ESA) sentinel-1 data hub. S1A instrument, Interferometric Wide Swath (IW), Ground Range, Multi-Look Detected (GRD), with high resolution (H). Processing level is L and polarization mode is single VV and VH. Incidence Angle mid Swath of 39.115. Data are available with intervals of 6 days. For both time of data collection and optical data acquisition, it is used in this work, especially for assessment of its importance in classification accuracy improvement.

Radiometrically corrected image for this study is needed (El-Darymli, et al. 2014). Through absolute radiometric calibration, all those elements in the radiometric values that are not the result of the target characteristics were brought into account. In this way, the differences in the image radiometry were minimized. This enabled SAR images with different incidence angles, opposite look directions and ascending/descending geometries to be compared. Sigma Naught (σ0 ground range radar reflectivity/unit area), “Beta Naught (also called radar brightness (B0), representing the radar reflectivity/unit area in slant range) were calculated. Sigma Naught (σ0) deviation from the image pixel values (or Digital Number (DN)) or from Beta Naught, needs the local incidence angle (Eineder, et al. 2008).

Lee filter were used to preserve structure and texture (Ozdarici and Akyurek 2010) for agricultural fields. The filter was implemented by SNAP software from ESA to post filter the S1 data. A similarity window of 7 by 7 pixels will be perform, with the search window of 21 by 21 pixels. SRTM DEM 1 arc-second was used for terrain correction. It was automatically downloaded by SNAP sentinel-1 software, which is available for download at “ESA” S1 data hub. It is the local terrain incidence angle and additional flags demonstrating whether a pixel is changed by shadow / layover or not (Eineder, et al. 2008).

**In situ data**

Field campaigns were organized and performed at 22-Jan-2014 and 23-Jan-2014, 26-Feb-2014, 22-April-2014 and the last one at Aug. 2014 by icipe to collect reference data on non-cropland (water bodies, grasslands(13 field samples) soil, houses and tarmac roads) and croplands (62 maize mono-cropping, 18 maize inter-cropping, 2 Tea garden and two Napir grass). A stratified random sampling method were adopted to collect sample fields (Slob, Sato et al. 2010). A handheld Global Positioning System (GPS) device with an error of ±3m was used to locate the reference control points. Once a point was located, boundaries (polygon) within a minimum area of 30 by 30 m were delineated. To avoid the edge effect, polygon data were collected five meters away from the edge of each field crop. Geo-tagged photographs of each cropping system in the sample fields were recorded for further inspections of the cropping systems.
and crop age. Those maize fields that were about three weeks old or less at the first image acquisition date eliminated In order to mitigate the effect of soil background on the crops spectral features. The reference data were randomly divided into 70% training set and 30% validation set. The training set was used to train the RF classifier while the validation data set was used to evaluate the accuracy.

Vegetation indices

The importance of VI’s in crop classification have been mentioned in previous works (Bannari, Morin et al. 1995, Zillmann and Weichelt 2014). VI’s show better sensitivity to green vegetation detection than each spectral band separately. Indices selected for evaluation used a combination of visible, near-infrared and red-edge bands, including: NDVI, SR, red-edge normalized difference vegetation index (NDVire), red edge simple ratio (SRe), green NDVI (gNDVI), MTVI2. The NDVI, SR, gNDVI (Some of them for L8 and some others for RE). The original bands were used in combination with the above mentioned indices for RF classification and further vegetation analysis (Forkuor, Conrad et al. 2014). Important indices for RF classifier were selected based on RF classification accuracy results.

Classification algorithms

Random Forest

RF as a pixel base classification algorithm was adopted in this work due to high resolution (RE) and medium resolution(L8, SAR) (Whiteside, Boggs et al. 2011). Multi-spectral and SAR images cover one year (two farming season) in this study. For heterogeneous small scale farms with large dynamics RF as a pixel-based approach is preferable. It is considered as an efficient classification approach for crop mapping using high spatial resolution satellite data like RapidEye (Forkuor, Conrad et al. 2014). It is able to handle noisy and highly correlated predictor variables, which are commonly presented in remotely sensed data (Curran and Hay 1986). In addition, (Braun and Hochschild 2015) mentioned that RF perform better for multi-temporal dataset.

The RF method takes random subsets from a training dataset and constructs classification trees using each of these subsets. Trees consist of branches and leaves. Branches are often thresholds defined for the measured (known) variables in the dataset. Leaves are the class labels assigned at the termini of the trees. Classes are then assigned based on classes assigned by all of these trees. Based on a majority rule, as if each class assigned by a decision tree were considered to be a vote.

An Out-of-Bag (OOB) error estimate and an estimate of variable performance are performed in RF method. For each classification tree assembled, a fraction of the training data (one third of the data which is not included in the bootstrapped training sample) are left out and used to compute the error for each tree by predicting the class associated with that value and comparing with the already known class. This process results in a confusion matrix, which we will explore in our analysis.

RF classifier uses two user-defined parameters as number of trees and m-try (the square root of the total number of spectral variables (indices and/or bands)). To improve the classification accuracy, the two RF parameters were optimized using the internal OOB error rate with a grid search and a ten-fold cross validation method (Waske, Benediktsson et al. 2009). The optimal settings for RF parameters adopted to compose RF classifier.

For each tree in the forest, there is a misclassification rate for the OOB observations. To assess the importance of a specific predictor variable, the values of the variable are randomly permuted for the OOB observations, and then the modified OOB data are passed down the tree to get new predictions. The difference between the misclassification rate for the modified and original OOB data, divided by the standard error, is a measure of the importance of the variable. RF testify the variable importance of all the variables, this is against the elimination of good predictors. The mean decrease in accuracy for a variable is the normalized difference of the classification accuracy for the OOB data in the presence of data for that variable as observed, and the classification accuracy for the OOB data when the values of the variable in the OOB data have been permuted randomly. Higher mean decrease in accuracy indicate that the variables are more important for the classification (Cutler, Edwards et al. 2007). A clear interpretation of the absolute values of variable importance is hard to do well. Even some works for selection of optimal spectral variables, the variable importance from RF classification were utilized (Kuhn).
RF backward feature elimination method (Cutler, Edwards et al. 2007) were used here. “varSelRF” package (Diaz-Uriarte, Robin, Jean-Michel et al. 2010) in R statistical software (Kiriwa, Feyissa et al. 2016).

A .632+ bootstrap method using a leave-one-out cross-validation procedure with replacement with samples that are not used when fitting the RF was used to evaluate the selection process of the most significant spectral variables without any over fitting (Efron and Tibshirani 1997). The optimum spectral variables selected were therefore adopted for final classification.

A confusion matrix was constructed to assess the accuracy of the RF classification, using the overall accuracy (OA), producers’ accuracy (PA) and users’ accuracy (UA). Quantity and allocation disagreements (Pontius and Miliones 2011) were also calculated from the classification confusion matrix to evaluate the reliability of the classification map and to measure the agreement between the predicted classification features and the reference field data.

**One Class Classifier**

The purpose of an OCC for remote sensing applications is to map only one specific class of interest. Training these classifiers only requires ground truth for the class of interest, while training data for other classes is not required. Thus, the acquisition of reference data significantly reduces. However, one-class classification accompany with uncertainty. In addition, full automation is difficult, due to the limited training reference information that is available. Thus, a user-oriented one-class classification strategy is proposed, which is based among others on the visualization and interpretation of the one-class classifier outcomes during the data processing. A careful interpretation of the diagnostic plots such as the class separability and suitability of a specific threshold guarantee the understanding of the classification results. In the absence of complete and representative validation data, as the fact of a real one-class classification application, such information is important to evaluate and improve the classification (Mack, Roscher et al. 2014).

**Analysis**

Maize pixels were extracted out of RE classification (short-rain season) and utilized to mask L8 maize pixels (long-rain season). Maize pixels of each season were classified for mono-cropping and intercropping system. Raster analyses utilized to determine continuous cropping system among two main cropping seasons. In addition OCC classification were utilized to classify sever MLN among maize pixels. Ecological variables (precipitation and altitude), cropping system (mono/inter cropping) and MLN severity were analysed to investigate find the relationship among them if any.

**Results and discussion**

Classification of the RE images were performed with several different RF setting and Kappa values were used as a criteria to evaluate the optimum LULC classification result. The first results that were testified consisted of 6 classes (Maize, Grassland, Tree, Non-vegetation, Water and Non-maize crops) and different number of trees. The Kappa of 58.73 (total accuracy of 65.1%) and OOB error rate of 1.64% were achieved by using 500 trees (Oshiro, Perez et al. 2012). This setting comparing to the results of 100 and 1000 trees achieved better performances.

The small class of Non-maize-crops achieved total accuracy of 44.44% merged with Grass-lands class (total accuracy of 95.45%) to make a bigger class, named as Non-maize. A very high resolution base-map utilized to modify shapefiles’ borders and the RF composed to 500 trees. These new classes (Maize, Non-maize, Tree, Non-vegetation, and water) reached the total accuracy of 83.57%, 92.75%, 98.59%, 85.15%, and 92.65% respectively. Consequently the Kappa and overall accuracy of LULC classification reached to 85.74 and 89.63%.

**Regional-level NDVI time series extracted from L8 stack which cover 12 months. The results were visualized as**

Figure 3. According to the time series, the correlation is high at adjacent phenogenous stages. For example, the correlation is high between May-June at both 2014 and 2015. Towards the end of season (harvest time), the NDVI correlation drops too.

To classify mono/ inter cropping systems, RF were applied on maize class (extracted from first level RE classification (Jan. 2014)). The result reached the total accuracy of 89%, kappa of 0.75 and high OOB estimate of error rate 10.45%.
Figure 4  NDVI correlation at Regional-level and growing stages (days after plantation)

Growing stages:  day 0-30=fully emerged | day 30-60= 12-20 leaves & tasseling | day 66-100= maturity (day= days after plantation)

Figure 3  a) OCSVM puF (PU performance, Bootstrap), a1) the classification Histogram based on unrefined OCSVM puF.  b) OCSVM refined puF grid (0.1<\text{Sigma}<1 and 0.01<\text{nu}<0.1), b1) the classification Histogram based on refined puF grid.
In addition, OCSVM and BSVM classifiers were utilized for mono/inter cropping systems classification. The parameter space of the OCSVM model were determined 0.01<\nu<0.1 contains more powerful model which produced predictive values with better separability between two target classes comparing to first result (Figure 4 b). The low density area between two classes is wide but not thin which shows the confusion between two target classes. The one-class SVM as a P-classifier i.e., the classification model is trained with only positive samples, compared to the biased SVM as PU-classifiers i.e., they are trained on positive and unlabelled data.

The maximized Sigma of 0.1 and Multiplier cost of 4 on full training set. The prediction results of the rank 1 to 4 were visualized as Figure 5. Severity of the MLN were classified as high severity (4-5) and low severity (1-2) utilizing OCSVM classifier. The histogram shows separated classes at the Sigma between 0.2-0.7, and nu-values of 0.1-0.2 (Figure 6). (Osunga, Mutua et al. 2017) modelled the co-infection of viruses to determine best procedure for MLN control concluded that practicing crop rotation was successful in large farms but for small holders it will be less effective. Even MLN cannot be eradicated completely in both cases.
According to the analysis, at cropped study area, 52% of fields applied continuous cropping while 48% of fields applied rotation (out of two seasons). In addition, 32% of severe MLN infection detected in those fields which applied continuous cropping system. Maize fields that conducted crop rotation showed 68% of total severe MLN occurrences (Table 1, Image 5).

(Osunga, Mutua et al. 2017) also worked on a spatial modelling of MLN Disease in Bomet County, Kenya and concluded that Soil moisture, rainfall and slope are the most significant determinants of MLN severity index in Bomet. In this study ecological variable such as height, rainfall and aspect have been investigated. The results shows that 75% of the occurrences of sever MLN were at elevation 1800-1900 meter above sea level. In addition, the analysis of cropping system and

<table>
<thead>
<tr>
<th>Cropping system</th>
<th>Area</th>
<th>Sever MLN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation</td>
<td>48%</td>
<td>68%</td>
</tr>
<tr>
<td>Continuous</td>
<td>52%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Table 1 Percentage of the area covered by each cropping system and sever MLN affected each system

<table>
<thead>
<tr>
<th>Precipitation</th>
<th>&gt;135mm</th>
<th>≤135mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono cropping</td>
<td>51%</td>
<td>49%</td>
</tr>
<tr>
<td>intercropping</td>
<td>39%</td>
<td>61%</td>
</tr>
<tr>
<td>Severe MLN</td>
<td>51%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Table 2 Percentage of each cropping system and MLN infection according to average precipitation

<table>
<thead>
<tr>
<th>Elevation</th>
<th>Below 1800m</th>
<th>1800-1900 m</th>
<th>Above 1900 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono cropping</td>
<td>14%</td>
<td>70%</td>
<td>16%</td>
</tr>
<tr>
<td>intercropping</td>
<td>10%</td>
<td>72%</td>
<td>18%</td>
</tr>
<tr>
<td>Severe MLN</td>
<td>10%</td>
<td>75%</td>
<td>15%</td>
</tr>
</tbody>
</table>

Table 3 Percentage, Sever MLN and cropping system at different altitudes.

| Table 4 Percentage area covered by different aspects and MLN occurrences
<table>
<thead>
<tr>
<th>Region Aspect coverage %</th>
<th>10% North</th>
<th>9% South</th>
<th>9% East</th>
<th>11% West</th>
<th>61% Flat</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLN occurrences</td>
<td>11%</td>
<td>9%</td>
<td>8%</td>
<td>10%</td>
<td>62%</td>
</tr>
</tbody>
</table>

| Table 5 average precipitation at different altitudes
<table>
<thead>
<tr>
<th>Ave. precipitation period</th>
<th>Below 1800</th>
<th>1800-1900</th>
<th>Above 1900</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov.</td>
<td>121.5</td>
<td>113</td>
<td>108</td>
</tr>
<tr>
<td>Feb.</td>
<td>123</td>
<td>115</td>
<td>105</td>
</tr>
<tr>
<td>Apr.</td>
<td>238</td>
<td>234</td>
<td>231</td>
</tr>
<tr>
<td>Jul.</td>
<td>68</td>
<td>62</td>
<td>57</td>
</tr>
<tr>
<td>Nov. – Jul.</td>
<td>138</td>
<td>132</td>
<td>126</td>
</tr>
</tbody>
</table>

| Table 6 average precipitation and altitudes at eastern and western region of cropped area
<table>
<thead>
<tr>
<th>Western region</th>
<th>Eastern region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ave. Precipitation (mm)</td>
<td>140</td>
</tr>
<tr>
<td>Ave. altitude (M)</td>
<td>1789</td>
</tr>
</tbody>
</table>
elevation determined that more than 70% of both mono and intercropping located at this elevation. Altitude below 1800m received highest rainfall and the elevation above 1900m received the least precipitation and shows 8% more intercropping area (Table 3). The average precipitation of western region was 140mm and at eastern part 131mm (Table 6).

Table 2 shows the average precipitation of Nov.-Jul and the prevalence of cropping system and sever MLN above and below this average (135mm). The distribution of monocropping are just 2% more at region with >135mm. according to Table 2&6 and Error! Reference source not found. -5 towards western region of the study area rainfall is higher comparing to eastern side. At sampled region with lower precipitation intercropping increase 22% comparing to western side (Table 2, Table 6). Distribution of MLN in different aspects calculated as Table 4. 62% of MLN infection occurred at flat area. The region with aspect toward east showed the lowest infection by 8%.

Image 2 mono \ inter cropping overlaid by average precipitation Nov.-Jul.

Image 3 Overlay of the classified mono/intercropping system and DEM
Image 4 Closer look at a sample from western side.

Image 5 Closer look at a region at eastern side of cropped area

Image 6 Severe MLN affected area and Mono/inter cropping overlap
Conclusion
This study demonstrated the utility of medium and high spatial resolutions optical (L8 and RE), SAR (S1) and VIs for obtaining maize fields information in a heterogeneous landscape at Bomet, Kenya. RF and OCC algorithm were utilized for classification. Based on the results, RF performed better in the case of balanced classes comparing to imbalanced classes. In addition, RF algorithm showed acceptable performance for two balanced classes of mono and intercropping. Both OCSVM and BSVM have been performed to classify mono and intercropping too. The BSVM was not able to separate these two classes. On the other hand, OCSVM results shows that the two classes are separable even there is confusion between two classes. OCSVM was able to separate high and low severe MLN affected maize fields even with few and imbalanced samples.

In the case of NDVI time series, correlation between adjacent phenological stages was the highest between final growing stages, tasselling and maturation stages and the correlation drop during harvest season. Even producing space born NDVI time series in this region especially during the rainy season is very challenging issue due to high number of cloudy days. In addition any anomaly detection needs higher resolution.

The result of the cropping system analyses showed that covered area by continuous and rotational cropping doesn’t show big differences, even the agricultural management system advised farmers to apply crop rotation to control MLN occurrences but still near half of farmers prefer continuous cropping. On the other hand, the occurrences of MLN at rotational cropping system was 36% higher than continuous cropping system. This result attained from two season investigation. Even this result is against the effectiveness of crop rotation for MLN control, but it proves previous works that mentioned long term crop rotation as an effective practices to control diseases outbreak. In addition previous literatures doubted the effectiveness of crop rotation in those regions with small size fields. In this manner, to control MLN, crop rotation should be applied as long term practice and more than two seasons. Still because of large number of small size fields MLN vectors can survive. Yet more investigation is needed to cover longer than two cropping season’s period to determine the effectiveness of this practice in Africa.

Based on previous works, the ecological variable can affect the MLN outbreak. Some of ecological variables such as height, precipitation and aspect investigated in this study. The results showed that the MLN prevalence is high at elevation 1800-1900 but at the same time the maize cropping in this elevation is more than other elevations. Also, investigating the region aspect shows that most of maize fields are located at so called flat regions and MLN occurrences are more prevalent in this aspect too. To have better understanding of how the ecological variables influence MLN occurrences, further studies should be performed.

Mono/intercropping system and MLN occurrences were analysed in a selected region divided by precipitation higher and lower than the average rainfall. The prevalence of MLN didn’t show more than 2% difference in mono and inter cropping fields. On the other hand the intercropping was 22% more prevalent at region with lower than average rainfall. This shows that uncertainties in rainfall patterns in the region have encouraged inter cropping system.

MLN ends up with necrosis which could be confused with maturation stages. So further works on MLN detection can focus on very precise VI’s products and growing stages of the maize plant itself. Due to high cloud covers at the study area, very high resolution SAR images for structural-change detection can be an advantage for further works. In this way, interferometry analysis can be adopted to determine growing stages. Combining very high resolution air born optical images can rectify information which produced by SAR data analysis and lead to anomaly detection in maize fields.

All to all, further studies is needed to consider some of very important factors such as distribution of maize growing regions, alternative hosts, vectors, and the seasonality of maize distribution in the region and farm practices e.g. use of pesticides as a MLN severity mitigation measure. Also it is needed to determine alternative hosts for the viruses causing MLN- MCMV and SCMV for Africa. As (Rogers et al., 2002; Peterson and Shaw, 2003; Lane and Jarvis, 2007; Peterson, 2009) mentioned, it is possible that MCMV and MLN are spreading west and south and might eventually get there, with implications on the respective area suitability the MLN risk in Africa is high on East and Central Africa, which are and will remain hotspots in the future.
Thus further research based on MLN presence and absence and the spatiotemporal analysis is suggested to show MLN trends from the initial year of observation to date for a clearer understanding of MLN occurrences and distribution.

Acknowledgements

We appreciate the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) for funding of the project “Better implementation of crop season breaks for management of Maize Lethal Necrosis (MLN) Virus in East Africa – Can remote sensing be an option?” We are grateful to Julius-Maximilians-University Würzburg, Department of Remote Sensing for their cooperation and helping in data preprocessing. Our appreciation also extends to the department of plant health International Center for Insect Physiology and Ecology (ICIPE) for helping in field identification and data collection.

References


Kuhn, M. "Variable importance using the caret package."


retrieving satellite and radar images. Systems, Man, and Cybernetics, 2001 IEEE International Conference on, IEEE.


