MONITORING CROP HEALTH FROM SPACE; IMPROVING DROUGHT RESILIENCE FOR FARMERS IN KENYA

GRAHAM DEANE[1], FABIO D. VESCOVI[1], ONYANGO OGEMBO[2]
[1]Airbus Defence and Space - Guildford, Surrey - UK
[2]Building Africa - University of Nairobi - Kenya

graham.deane@airbus.com

Paper prepared for presentation at the
"2020 WORLD BANK CONFERENCE ON LAND AND POVERTY"
The World Bank - Washington DC, March 16-20, 2020

Abstract

A high-resolution Vegetation Health Index (VHI) is proposed as an indicator of the water stress condition of vegetation cover. This index integrates satellite-derived vegetation and thermal information; it is a function of NDVI (Normalised Difference Vegetation Index, derived from Sentinel-2) and LST (Land Surface brightness Temperature, from MODIS). The VHI has already been tested experimentally for African environments (Rojas et al., 2011), exhibiting good performance, making it potentially suitable for insurance purposes.

This index informs a new satellite-based insurance scheme for Kenyan smallholder farmers. The proposed service is based on an online tool enabling rapid decision-making for compensations due to farmers. The tool shows VHI multi-temporal plots, tables and maps at 20m resolution, refreshed every 5 days. The service is undergoing an in-country trial for deployment to insurance providers, farmer cooperatives and Ministries in Kenya. This project is supported by grant funding within the UK Space Agency's International Partnership Programme.

Key Words: Agri-insurance, Drought, Kenya, Sentinel-2, Vegetation Health Index

1. Introduction

In Kenya the agriculture sector is a very important part of the economy and, according to FAO, the '*sector employs more than 40 per cent of the total population and more than 70 per cent of Kenya's rural people*' (FAO, 2014). However, many farmers lack robust information on their exposure to flood and drought risks. This partly reflects a lack of data on vulnerabilities, with meteorological data currently collected from expensive and sometimes unreliable meteorological stations, typically serving an area with a 10km radius, and little geographically granular understanding of the relationship between crop production and local rainfall variations.

The project described in this paper is addressing this information gap for agriculture in Kenya by analysing free-to-use Earth Observation (EO) data from the European Space Agency's (ESA) Sentinel-2 satellite for drought resilience monitoring. In addition, space-derived temperature data from NASA's MODIS (Moderate Resolution Imaging Spectroradiometer) mission are also used in a drought indexing mechanism that will be accessible to users online.

The satellite data inputs, essentially greenness and thermal information, are processed with an innovative algorithm, based upon a huge amount of multi-temporal input data, which generates a Vegetation Health Index (VHI). This is a simple, user-friendly product for end-users. The VHI is generated every five days throughout the growing season thus providing, with a resolution of 20m on the ground, a very detailed assessment of crop development and potential yields.

This project is supported by grant funding within the UK Space Agency's International Partnership Programme (IPP) (www.gov.uk/government/collections/international-partnership-programme). IPP is a five year, £152 million ($US195 million) programme run by the UK Space Agency. IPP is part of and is funded from the Department for Business, Energy and Industrial Strategy's (BEIS) Global Challenges Research Fund (GCRF): a £1.5 billion fund announced by the UK Government, which supports cutting-edge research and innovation on global issues affecting developing countries. The Programme, which encompasses forestry, land use, agriculture, marine environments, disaster resilience, health and education, and renewable energy, is being delivered in alignment with UK aid strategy and the United Nations' (UN) Sustainable Development Goals (SDGs).

The VHI product is designed to increase the information available to smallholder farmers on the health of their crops, so that they can make better decisions. The VHI can also be used to develop more accurate and cost-efficient agri-insurance products. The index is especially suitable to identify areas of water deficit at a field level. Particularly important is monitoring the flowering point: if the crop fails to generate flowers it is not likely to produce any yield. Therefore the index can also be used as early warning for a micro-insurance scheme to budget the expected payments to the identified farmers with failed crops in the following 6–8 weeks. Developments of the scheme will be designed to trigger insurance pay-outs automatically in the event of poor harvest, thus eliminating the traditional paperwork trail for insurance claims.

In this paper we present the process for calculation of the Vegetation Health Index and the validation activities undertaken in the selected study area. We then discuss how this can be used to stimulate both the agri-insurance market and longer term economic development. As part of the IPP process, we have also undertaken a Cost-Effectiveness Analysis in order to demonstrate how the use of satellite data does provide a cost-effective solution to the need for improved information for drought preparedness.

2. Description of the Project

Study area

The selected study area includes the Nyando Basin in the west of Kenya. According to FAO, *'the Nyando river basin covers an area of 3 500 km² of western Kenya and faces some of the most severe problems of agricultural stagnation, environmental degradation and deepening poverty found anywhere in Kenya. The river drains into the Winam gulf of Lake Victoria and is a major contributor of sediment, nitrogen and phosphorus to the lake.'* (Swallow et al., 2003) The principal crops of the area include maize, sorghum and beans with commercial production of sugar cane and irrigated rice. (Vescovi et al., 2009)

For the VHI calculation the selected area is based on the ESA Sentinel-2 grid tile 36MWE which is close to Lake Victoria, as illustrated in Figure 1. Each Sentinel-2 grid tile corresponds to an area of 100km by 100km which means that about 20 Sentinel-2 tiles are needed to cover the principal agricultural belt of Kenya.

(Figure 1)

Calculation of the Vegetation Health Index

After a review of the published scientific literature on drought indices (see Table 1), including several recent publications, a number of possible candidate indices were shortlisted. Evidence from their published experimental usage in different geographic and climatic contexts was reviewed in order to assess their robustness for potential operational use for agricultural insurance purposes.

(Table 1)

The Vegetation Health Index (VHI) was identified, tested and eventually selected because of its successful application in numerous case studies (Kogan, 1995a, 1995b). The VHI is a composite index which integrates the Vegetation Condition Index (VCI) and the Temperature Condition Index (TCI). The chosen method follows the principles established by Rojas et al. (2011) who produced a drought probability map for Africa using NOAA AVHRR data for the time window 1981 - 2010.

Their method is considered in this study as a framework for generation of a higher resolution map, at 20m resolution, using Sentinel-2 optical products and MODIS thermal information (from the MOD11 product), as basic inputs for the chlorophyll presence and thermal status of the crops, respectively. Following Rojas, et al. (2011), the VHI is defined as follows:

$$VHI_i = w_1 * VCI_i + w_2 * TCI_i \quad [1]$$

Where:

- $TCI_i = 100 * (T_{max} – T_i)/(T_{max} – T_{min})$
  Where: T_i is the temperature value of the considered 8-day period, while T_{max} and T_{min} are maximum and minimum brightness temperatures, respectively, calculated from daily time series for each MODIS product and each pixel. The MODIS product providing the brightness temperature is MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V005 (short name: MOD11A1).

- $VCI_i = 100 * (NDVI_i – NDVI_{min})/(NDVI_{max} – NDVI_{min})$
  Where: $NDVI_i$ is the Sentinel-2 Normalised Difference Vegetation Index (NDVI) of the considered day, while $NDVI_{max}$ and $NDVI_{min}$ are absolute maximum and minimum NDVIs, respectively, as calculated from the Sentinel-2 multi-temporal series.

- w_1 and w_2 are coefficients which can have various weights. Rojas et al. (2011), suggest 0.5 for both, however in this study, after validation, it was found that better results are given with w_1= 0.6 and w_2=0.4.

Essentially, the chlorophyll information is derived from the multi-temporal NDVI derived from Sentinel-2 data, atmospherically corrected, resampled at 20m and normalised across the previous two-year time-series within the maximum and the minimum NDVI. Whereas, the thermal information is derived from the MODIS Land Surface Temperature of the product MOD11A2, re-projected to UTM, then clipped and resampled to 20m to match the final product.

The two ESA Sentinel-2 platforms currently orbiting deliver product every five days, whereas MOD11A2 product is available every eight days. However, the processing is designed in such a way that a VHI map can actually be generated every five days.

The processing chain for calculating the VHI is illustrated in Figure 2. This processing chain is based on using the SeNtinel's Application Platform (SNAP) tool made available by ESA and is developed as a fully automated process from the basic data input to preparation of the final VHI product ready for delivery. Atmospheric corrections with the Sent2Cor module are applied before processing Sentinel-2 images, whereas MODIS products are already corrected. These atmospheric corrections can partially remove haze but not clouds.
(Figure 2)

It is not in the scope of this paper to describe all the details of each processing step because the majority of them are standard image processing operations. However the cloud removal process deserves a specific mention.

Cloud Removal Processing

The farming season in the study area starts at the onset of the rainy season. This implies that the imagery is usually cloudy during the farming period, when the vegetation index is most relevant; therefore a cloud removal algorithm must be applied. Airbus has developed a bespoke smoothing function to remove clouds almost completely from the VHI products and to clean the noise (residual haze) of the images. The algorithm assumes a continuous crop growth across consecutive cloud-free acquisitions and estimates the VHI value of the cloudy pixels for up to five consecutive cloudy acquisitions.

Cloud Removal Processing Results

The result of the Airbus cloud removal process is illustrated in Figure 3. The blue line represents the VHI values as output by the system without any cloud removal process: both clouds (0 values) and noisy haze (sudden VHI drops) make the resulting curve almost uninterpretable in any further process. However, the red line represents the algorithm result: cloudy and hazy values are identified and replaced with values which are a function of the crop growth and of cloud-free values taken from past and future satellite data acquisitions. The result shows clear representations of the rainy seasons with high VHI values, when crops are farmed, and the dry season with low VHI values.
(Figure 3)

Validation Process

The output VHI maps were validated in the field through ground truth campaigns carried out in June-July 2018 across the study area. The technical principles of this validation exercise observed CEOS (Committee on Earth Observation Satellites) Guidelines (CEOS, 2014). The spacing and location of the measurements in the field observed the sampling strategy outlined in the Copernicus Manual (Copernicus, 2015). Copernicus is the European Union's Earth Observation Programme supporting monitoring of the environment from satellite and in-situ ground data.

All collected field data were then re-mapped in a database sheet where each field observation could match the correspondent VHI pixel value on the map. A time and location correspondence was obviously maintained. Eventually, correlation statistics were calculated and scatterplots were created to show the relationships between field observations and the VHI based on satellite data.

Validation Results

Examples of scatterplots that have been prepared to show the correlation between field observations and VHI values are presented in Figure 4. Each point of the plot represents the field observation (measured percentage of green coverage, in this example) and the corresponding VHI map value. A similar scatterplot shows field values and corresponding NDVI values, the latter derived from the intermediate NDVI product generated in the VHI processing chain.
(Figure 4)

Figure 4 also reports basic correlation statistics for comparison of VHI and NDVI. The area of interest is generally in good water availability conditions so the expected vegetation indices are quite high. The VHI correlation exhibits a rather better fit than the NDVI suggesting that the VHI is, indeed, a good indicator of crop health. This improvement is essentially due to the following operations in the processing chain:

- The introduction of thermal information as the main indicator of water availability. When water is no longer available, the plants tend to reduce transpiration thus causing an increase in the surface brightness temperature of the canopy;
- Both NDVI and Land Surface Temperature are normalised against historical maximum and minimum values at pixel level. This operation minimises the effect of invariant objects in the pixel (e.g. an evergreen tree or a house in the middle of field, at the corner of several fields).

The project is still ongoing and another field campaign for further validation is planned for the dry season 2020. It is expected that these additional field / VHI data will populate the lower part of these scatterplots. This will enable validation of the VHI under a variety of conditions in Kenya and will also help to develop a validation mechanism for use in other areas where crop stress due to lack of water leads to lower yields and crop failure.

Final VHI maps and plots results

The final VHI map resulting from the overall process is illustrated in Figure 5, for a sample area observed on 30th August 2019. Examples of multi-temporal VHI plots for a rice farm close to Kisumu (Kenya) across the growing seasons of 2018-19 are represented and discussed in Figure 6. The VHI pixel values are averaged per field. Five of the six represented fields show similar regular VHI curves, just slightly shifted in planting time. Field number 6, however, shows an irregular crop growth; probably rice was not planted in this field. The insurance provider can use these plots on a field by field basis to derive information on the crop growth.

(Figure 5)

(Figure 6)

3. Product Development

Designing a Service for Insurance Providers

The majority of agri-insurance providers currently base their compensation payments to farmers on a crop performance index; this is in essence based on an assessment of the actual crop yield under critical conditions compared to the potential yield under normal conditions. Both measures encounter many challenges: the concept of an indicative yield measure in itself is hardly applicable over large areas, such as at a national level, because it involves a correct sampling strategy at field level, many measurements of product fresh and dry weight, specific estimation methods depending on the crop type, etc. In addition, generally among farmers, information on crop productivity is not easily disclosed or may be biased, especially when farmers previously obtained loans from a funding institution. Even more questionable is the determination of the potential crop yield which could have occurred under normal or ideal conditions: this operation cannot be considered a measure but rather an estimate, because that ‘ideal yield’ never actually materialised.

Currently the few insurance schemes commercially deployed in African environments either base the actual crop yield estimation on weather data, from the available meteorological stations (Kilimo Salama, in Kenya, TAKAFUL in East Africa, Nyala Insurance Company in Ethiopia, PepsiCo in India), or through field surveys (AMACO Insurance in Kenya) which are relatively expensive to implement. In the latter case, the sampling strategy also implies destruction of part of the yield, through cutting and weighing, in a small farm which may be already in critical condition. So the insurance companies tend to minimise the number of ‘destructive’ samples although this can be at the expense of data accuracy.

This current project gathered information on the above practices after running interviews in various workshops with both insurance providers and farmer communities in Kenya. However, the outcome is that their main goal is not really looking at the past, to try to measure what is lost, but looking at the future to secure whatever reasonable financial settlement can be granted after an under-performant growing season. The problem is not the accuracy measure of the yield loss, but the development of a method that is able to detect a general drop in crop performance, regardless of the crop type, and rapidly to trigger what is in effect a financial investment (rather than compensation) to re-start activities in the next season.

The method here proposed does not aim at trying to estimate yield losses. Instead, the decision of insurance providers is based on a comparison of the *actual* crop performance of a field in the current season against the *potential* crop performance of the same field assuming favorable farming conditions. In this context the performance is not necessarily the yield but the health status of the crop throughout the farming season as measured by a suitable vegetation index. This is ultimately considered the closest proxy to biomass and yield which we can obtain from satellite EO technologies. In the following sections, the crop actual and potential VHI curves are explained.

The Actual VHI Curve

As discussed, the VHI curve aggregates information on two key indicators of crop health: chlorophyll and surface temperature.

The presence of chlorophyll in the leaves is obviously one of the main driving factors for the production of organic matter in the plant. Novelli et al. (2019) used the Sentinel-2 derived Leaf Area Index (LAI), after calibration with the Erosion-Productivity Impact Calculator (EPIC) model, to estimate crop yield. They observed a generally good correlation between the yield estimated by the EPIC model (with assimilation of LAI data from Sentinel-2) and the observed yield. The authors also observed that the EPIC-LAI curve is influenced mainly by the temperature and not by the variable amount of Nitrogen applied in their field trial.

Also the thermal conditions and their duration are essential: a drop in temperature below a certain threshold may cause a growth delay or even a block; whereas an increment of the surface temperature of the canopy above a certain threshold may have destructive effects. This may be caused mainly by a lack of transpiration which, in turn, can be due to water deficit. Therefore the leaf temperature can provide information not only on the thermal but also the hydrological crop status.

It is common understanding among agronomists that the amount of heat required to complete a given plant stage, e.g. flowering (Cross et al., 1972), maturity (Chen, 1973) and even overall biomass (Edey et al., 1977), does not vary: the combination of temperature (between thresholds) and time will always be the same (Parthasarathi et al., 2013). This is measured in heat units called Degree Days (°D). For example, the cumulative thermal/time requirement for different development stages (Growing Degree Days) in corn and pearl millet is listed in Table 2. Some authors even calculate the first derivative of the thermal/time crop curves to determine parameters of agronomical interest before empirical evidence can be taken, such as maximum leaf elongation (Voorend et al., 2014).

Given the nature of the VHI, which aggregates the main biophysical factors (chlorophyll and temperature) driving the crop growth and eventually the biomass, the VHI curve across the growing season can be used to monitor the capacity of a crop to go through the development stages normally or to detect delays or failures in any of them. In principle, the wider and higher the area under this curve, then the longer and more efficient, respectively, is the time for the crop to benefit from chlorophyll at the optimal temperature range to produce organic matter. A correlation between that VHI area and crop grain yield (t/ha) is still to be examined; however, in recent studies some encouraging results were obtained which correlated one or more crop performance parameters (e.g. biomass, yield, etc.) with the area under the profile curve of the studied vegetation index (Liu et al., 2018, Novelli et al., 2019). In the following, some scientific advances from the vast literature on the subject are briefly reported in chronological order.

Labus et al. (2002) found strong relationships between wheat yields and integrated NDVI over the entire growing season. They used AVHRR-NDVI growth profiles and delivered the strongest yield estimates at regional level. However, at the farm level, the $1km^2$ spatial resolution had limited application. Kalubarme et al. (2003) observed a significant relationship between wheat yields and fractional area under the NDVI profile curve. Deosthali et al. (2006) found correlations between field measured LAI and crop yield in the crops grown in fragmented conditions using a crop growth model. Similarly, Lenz (2007) found correlations between biomass data, including yield, for various crops and modelled LAI throughout the growing season. Battude et al. (2016) estimate corn biomass and yield over large areas using high spatial and temporal resolution remote sensing data, similar to Sentinel-2. Some other authors find a good fit between various sigmoidal models and the crop growth to predict the amount of maximum biomass (Liu et al., 2018). Novelli et al. (2019) combined LAI data from the Sentinel-2 satellite with the EPIC model to estimate crop yield. LAI showed a good correlation between the Sentinel-2 estimates and the ground measurements. A correlating fit between satellite LAI curves and EPIC modelled LAI curve was also observed.

The current state of the studies on the VHI precursor indices (i.e. VCI for Chlorophyll and TCI for thermal information) generally enable us to infer that the VHI area defined from planting time (first minimum) to harvesting time (second minimum, if no dramatic change is recorded before) can be assumed as a reasonable proxy of the crop production performance. However, it is our opinion that it is still too early to infer accurate yield or biomass data directly from satellite records and derived indices.

Figure 6 and Figure 7 show VHI sample curves for a number of fields visited during the field surveys. In order to ease comparison among plots in Figure 7 the time between planting and harvest was normalised. Please, note that these curves also show the success of the cloud removal algorithm during the rainy season, because no drops of VHI values due to clouds or haze are recorded in any sample plot at all times. (Figure 6, Figure 7)

The potential VHI curve

A potential crop growth curve is needed for reference to compare the measured crop performance (actual curve). This reference can be conceptualised as the best curve fitting an ideal crop growth cycle as if no crop adversity (e.g. pest, drought, disease, nutritional deficit, etc.) would have happened and under the best farming conditions.

Many authors proposed various crop growth models for this estimation (just to quote the most recent advances on this topic: Novelli et al., 2019; Berger et al., 2019; Kasampalis et al., 2018; Huang et al., 2013). The main advantage of a model approach is the capacity to describe the behaviour of the crop by predicting the physiological mechanisms of the crop growth (e.g., phenological development, photosynthesis, dry matter, portioning, organogenesis, etc.) using mechanistic equations (Gowda et al., 2013; Novelli et al., 2019). However, their applicability is complex at regional scales because they generally assume that the field conditions are uniform (Kasampalis et al., 2018). In addition, they need a considerable amount of input data on environmental factors (e.g. soil, weather data, etc.) and field management strategies (e.g. irrigation, nitrogen supply, etc.) which are unlikely to be available in African environments.

When undertaking research activities on this topic a common pitfall is to attempt to fit overly complicated models (Paine et al., 2012); inappropriate functional forms often fail to converge or yield unreasonable parameter estimates. African agricultural contexts, especially among smallholder farmers, do not have long and consistent data time series at high granularity to support robust statistics for crop growth dynamics. On the other hand, the time series of the most recent satellite missions (e.g. the Sentinel constellation), which are suitable for these kinds of applications, are also only 3 to 4 years long. So it is important to ensure that any model and statistical approach is as little data demanding as possible and to manage the unavoidable approximation of the results until more comprehensive time series are available.

The approach applied in this work follows essentially the experimental practice of many authors where various plant growth models are tried and compared with the observed field data with suitable statistical tests; eventually the best candidate curves explaining the field dynamics are selected (Jiang et al., 2017; Vorobiova et al., 2017; Paine et al., 2012; Liu et al., 2018). Generally the curve models are generated by linear, non-linear or exponential equations describing the change in biomass with a field measure or, more often, a vegetation remote sensed index. Then the fit of the model on the growth trajectory of field data is checked with an unbiased statistical approach, e.g. R^2 (Liu et al., 2018), RMSE (Novelli et al., 2019), or other indicators of crop dynamics, such as the Large Integral and Small Integral of the areas covered by the curves' trajectories (De Castro et al., 2018).

In our work, after simple observation of the multi-temporal data time series for different crops compared with the field observation for the same crops in Kenya in the growing season 2018-19 (Paine et al., (2012) recommend a preliminary data observation as essential for the choice of the best curve model) we identified the most interesting candidate curves possibly fitting the data, e.g. various sigmoidal growth trajectory (Liu et al., 2018), asymmetric Gaussian, cubic spline (Vorobiova et al., 2017) and cycloid curve. The equation of the cycloid (Ehrenborg, 2017; Hoheisel et al., 2009) seemed to fit particularly well the high performant fields on a number of tested crops in the best farming conditions, especially corn and rice. It also acceptably fits natural vegetation dynamics. The method only needs two inputs from the farmers or insurance providers about the planting / harvesting times and two inputs on the absolute historical maximum / minimum of the VHI time series, of the considered pixel or field (see Figure 8). The ideal VHI values on the Y axis are calculated with the following parametric equations (lenz, 2017):

$$\begin{cases} y = R - r\cos(t) \\ x = R\,t - r\sin(t) \end{cases} \qquad [2]$$

where:

- R = 0.5 (radius of external circle)
- r = radius of the internal circle generating the VHI arch; values, re-scaled from 0 to 1, are calculated as follows: (VHI max – VHI min)/200
- t = between $+\pi/2$ and $-\pi/2$

whereas the length of X axis is always π, so each column width depends on the number of points (satellite acquisitions) and is adjusted accordingly. In practical terms, the assumption is that the larger is r and the larger is the area under the potential curve, then the higher is the potential crop performance of that pixel (or field), with that crop, under the given soil fertility and bio-climatic conditions.
(Figure 8)

In order to test the goodness of fit of the cycloid trajectory, the Root Mean Square Error between the potential curve and its correspondent actual curve was calculated. In addition, the areas of the two curves were normalised and compared, integrating the function between planting and harvesting time points on both curves (Small Integral in the sense defined by Eklundh, et al., 2011 and de Castro et al., 2018).

Table 3 shows examples of cycloidal curves calculated for a number of crops compared with the relevant field actual curves. The fields for this comparison were not randomly chosen but were all surveyed in our field campaign in Kenya and selected for the best farming practices, as they had to check the assumption that the cycloid is to make them suitable for a comparison with an ideal VHI curve. All actual curves are re-scaled to π on the X axes, whereas the potential curves are normalised on the average between the two minima of the actual VHI trajectory. The table shows also the RMSE values and the area difference, calculated in percentage, between the two curves. All fractions of bio-physical index areas are above 90%.
(Table 3)

Similarly, Table 4 shows samples of cycloidal curves for visited fields under poor farming conditions and/or significant signs of water deficit. In this case, all fractions of bio-physical index areas are well below 60%. Generally, it is assumed that when the VHI index falls below about 30%-35% the vegetation is dead (Rojas et al. 2011).
(Table 4)

A practical example

The actual and the potential crop performances are measured and estimated, respectively, by the VHI curves explained in the previous paragraphs. The method as presented here calculates a ratio in percentage terms between the areas defined under the two curves (actual growth over potential) and ranks the fields in a list from the top high performant field to the bottom one. The list can be sorted by county, sub-county or any other relevant area, depending on the insurance needs. The insurance providers are used to working with thresholds which may or may not trigger the farmer payouts depending on various financial criteria of the insurance policy. For example, from the workshops conducted in Kenya it was confirmed that at least one insurance provider aims at guaranteeing 70% of the crop yield, so if the ratio between actual and potential performance falls below that, the insurance top up the financial equivalent up to 70%. Other insurances may apply different criteria on the VHI index rank.

Figure 9 shows a snapshot of the proposed tool which is described here.
(Figure 9)

The table at the top of the figure is composed of yellow columns, filled in with insurance and farmers' inputs; a green column, which lists the fraction of area between of the biophysical index from the most to the least performant field; and, the white columns show the individual VHI values per field, per date. The bottom left window shows the actual (orange) and the potential (blue) VHI curves in the period indicated by the farmer in Column C (planting) and D (harvesting time). In this example the rice field seemed to be slightly under-performant in the first half of the growing period, but it caught up quite well in the second half, resulting eventually in 75% of crop performance for the overall growing season, as shown in the green column. Decisions on farmer compensation are primarily based on the green column. In the bottom right window the relevant rice field is outlined in orange on the Google Earth map. The tool will allow dynamic videos of the VHI maps in false colour across the growing season.

These VHI maps in false colour, the VHI values, the areas under actual and potential curves and their shape and size are all transparent information to farmers, surveyors and insurers. The tool can be published online for all fields in the insurance scheme. This way the whole system is not exposed to manipulation or corruption. However, the field owners, the thresholds applied and the payouts are disclosed only to the relevant stakeholders after contractual agreement.

4. Service Development

Objectives

From the example shown in Figure 9, we have a method of providing information to the insurance providers, farmers and other stakeholders which will enable the performance of a crop, in any given growing season, to be monitored against the expected performance of the crop without 'adversity', such as drought stress. This satellite map based index provides the information to enable the insurance providers to compensate farmers where crop performance is significantly below the expectation. The method enables a reliable and consistent decision to be made more quickly as the traditional stages of making a claim can be streamlined, as illustrated in Figure 10.
(Figure 10)

Stimulating the agri-insurance market will also spur broader economic gains. More widespread uptake of insurance raises investment and long-term economic growth, thereby helping to reduce poverty, and the employment of intermediaries to advise farmers, based on the detailed information available, will support job creation. The project sees young women as ideal candidates to be trained as intermediaries between the technology (an app on their phone) and their village/community of farmers, to whom they bring the information in an understandable and unbiased way. This is, therefore, a very practical application of satellite Earth Observation (EO) technology that helps to develop resilience to irregular drought conditions.

As well as helping to develop the agri-insurance market, the VHI data can also enable individual farmers and farmer cooperatives to understand the variation in their yields and potential crop value if they plant a different mix of crops on their land. The level of detail within the VHI data will help farmers and cooperatives to make decisions on future planting even when the average size of farms is only around 1ha. Such information, aggregated to the district or county level, is also of value to Kenya's central and local government agencies in helping to understand the wide range of climate-related impacts on agriculture and thereby assist them in developing strategies to mitigate these risks.

Cost Effectiveness Analysis

In accordance with the guidelines for the IPP projects supported by the UK Space Agency, Cost–Effectiveness Analysis (CEA) is the selected means of assessing the impacts of delivered space-related services. CEA is a type of 'Value-for-Money analysis that compares the costs of alternatives that achieve different amounts of the same impact.' (HM Treasury, 2013). The particular advantage of CEA compared with, for example, Cost-Benefit Analysis is that CEA provides a means of comparison without the need to put a monetary value to impacts that are difficult to quantify and is, therefore, significantly less time-consuming for comparing alternative strategies.

The CEA calculations for this project are designed to compare in detail the approaches based on the EO-derived VHI with existing monitoring systems which are typically based on the use of weather station data backed up with field visits by experienced surveyors and agronomists. This comparison is based on the total costs for preparation and delivery of the VHI products to end-users and the costs for making field visits to farms and for interpreting the information currently available.

However, the limitations of the existing monitoring systems are that the weather stations provide only one point of drought measurement over quite a wide geographic area, typically an area with a 10km radius, and farm visits can only be undertaken on an infrequent basis, due to cost. As a result, the detail available from existing systems is insufficient to provide individual famers, and their insurers, with information relevant to their specific circumstances; the information is only available as average data for groups of farmers in an area. In addition, the weather stations do not provide information on irrigated fields, where the farmers could prevent yield losses despite poor rainfalls; whereas the weather station data would deliver a false positive of crop failure.

The costs for providing the VHI data are based on the analysis of a large volume of data. Whilst processing these data is computationally quite heavy, the processing chain is quite stable and achieves consistent results, as indicated in the various illustrations presented in this paper. However the input data, both Sentinel-2 and MODIS are free to use and so all available data can be used to improve the quality of the VHI output.

The production costs, therefore, are based on the development of the automated image processing chain, running the process, quality checking of the output, to ensure consistency, and, delivery of data to end users. This delivery will be either as tabulated data or digital maps delivered via an easy to use online dashboard that will be available, as required, to insurance providers, farmer cooperatives and other interested users.

These CEA calculations will be finalised when the VHI testing period is completed within the project (currently planned for mid-2020) and will be undertaken by Vivid Economics Ltd (https://www.vivideconomics.com/) who are responsible for undertaking a Monitoring and Evaluation review of the current project and completing the CEA task. It is anticipated that this analysis will demonstrate that a VHI generated from satellite data does provide a cost-effective solution for generating detailed crop health information with a spatial resolution of 20m and a temporal resolution of five days when compared with existing methods. This resolution will generate around 1,000 VHI data points for a 1ha farm per growing season; significantly more than current methods.

Service design

The insurance companies can and must play a substantial role in agricultural risk management, especially in developing countries, which are mostly in need of an affordable insurance service. In turn, the insurers need new technologies and methods to monitor extensive farmed areas throughout the growing season. The designed VHI products, including drought maps of crops and flood damage mapping products, can be easily offered to both insurers and farmers in order to enable them to take decisions on the risk and manage the insurance schemes to make payments where crops have underperformed. The insurance contracts based on this new concept should feature annual premiums that are affordable enough to attract a mass participation amongst smallholder farmers. The aim, therefore, is to make the initial cost of the VHI component of such insurance less than 10 per cent of the annual insurance premium; these costs will be finalised once the service levels have been defined and agreed with the insurance providers. This target cost is based on the current size of the agri-insurance market. It is anticipated that expansion of this market will lead to significantly reduced costs per farmer in the future.

An additional advantage is the indemnity arrangement: after electronic delivery of the map an automatic payout process for the affected fields could potentially be triggered, subject to insurance approval, without significant paperwork for farmers and without an expensive local survey for insurers, as illustrated in Figure 10. The VHI map itself, in this case, becomes the field survey, enabling farmers to react quickly and, if appropriate, to re-plant a different crop if conditions are suitable.

The product, ready for distribution in a comprehensive end-to-end service, will also benefit from mobile technology, where the farmer communities, represented especially by women actors, will have access to technical information in an easy, user friendly way.

The main challenge encountered in the implementation of this project is not really the technical arrangement of the application. Indeed, after the demonstrated technical success of this product, the project team is now engaging in the dissemination of this new approach among local communities in Kenya. The project has already reached out to a number of local farmers in West and Central-East Kenya through bespoke workshops organised in their areas and with local translators. The VHI concepts were extremely well received. However, the insurance structure to implement the project and the engagement from governmental authorities, both at central and county level, are key to guarantee the social and market success of this project.

5. Conclusions

The satellite based vegetation health index, as presented here, has the potential of informing a new insurance scheme and the capability to impact directly on people. Even though the VHI calculation with Sentinel-2 and MODIS data involves significant computational efforts and an industrial infrastructure when covering vast areas (typically nationwide), the previous and the current studies have demonstrated an improved performance from VHI for crop condition monitoring than the traditional NDVI, VCI and other used indices.

The validation results showed encouraging improvements in the thematic accuracy due to the introduction of the thermal information in the index composition. This extra information, quite unusual in traditional satellite EO-based indices, seems to be key to describing the photosynthetic capacity of the leaves, the water availability and the overall crop status, even though the MODIS data has relatively low spatial resolution,.

The free availability of Sentinel-2 data regularly every 5 days enables unprecedented monitoring potential for the crops, which can be displayed in user friendly applications, and even on portable technologies, through the actual crop growth curve which can be compared to a potential growth under 'ideal' farming conditions in the considered field. In rural areas of countries with food security issues, like Kenya, the improved access to such crop health information will help farmers to improve their yields and support the micro-insurance industry to reduce the financial impact of drought events leading in turn to improved stability for farmers' livelihoods.

The two key pieces of information provided by this method, i.e. the multi-temporal actual and potential VHI curves from planting to harvesting time, only need a relatively limited number of inputs from farmers (specifically planting / harvesting dates), insurers (field identity) and the system itself (complete multi-temporal time series of satellite data from commencement of the archive). This makes the solution quite simple for the users (even though computationally complex) which can be applied, with acceptable approximations, to the majority of the cash crops in different agricultural and climatic environments.

The solution is designed especially for smallholder farmers who can typically farm more than one crop in a field (inter-crop). Even though not yet tested, there is evidence that the method could also be migrated to different environments and, possibly, beyond tropical-equatorial climates. Therefore this solution can inform a new, affordable insurance scheme, as currently under test in Kenya, which has the potential to attract significant participation from smallholder farmers in many parts of the world.

Acknowledgments

Just three weeks before submitting this paper one of the authors, Professor William Onyango Ogembo passed away after severe pneumonia at the age of 82. Professor Ogembo dedicated his entire professional life at the cause of the education of Kenyan students at Nairobi University and the technical and social development of his country, of which he was one of the leaders driving the process to independence. The authors are immensely grateful for his dedication and his social role for the here published work, which represents a fundamental achievement for the development of an affordable index-based insurance system for Kenyan farmers. Also, they wish to thank the surveyor Oliver Onyango and Prof. George Krhoda, members of Building Africa (BUA), founded by Professor Ogembo, and all involved staff for their invaluable support during the many meetings with the farmers, the validation field campaigns and all local activities undertaken in three years of our project.

List of references

Battude, M., Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., Le Dantec, V. &, Demarez, V. (2016) Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. Remote Sens. Environ., 184, 668–681.

Berger, A., Ettlin, G., Quincke, C. & Rodríguez-Bocca, P., (2019) Predicting the Normalized Difference Vegetation Index (NDVI) by training a crop growth model with historical data. Computers and Electronics in Agriculture, Volume 161, 305-311.

CEOS Global Leaf Area index Product – Validation Good Practices, V.2 Jan. 2014. (https://lpvs.gsfc.nasa.gov/DOC/CEOS_LAI_PROTOCOL_January 2014.docx - last seen May 2019).

Chen, C.S. (1973). Digital computer simulation of heat units and their use for predicting plant maturity. Int. J. Biometeorol., 17: 329– 335.

Copernicus manual 2015: IMPLEMENTING MULTI-SCALE AGRICULTURAL INDICATORS EXPLOITING SENTINELS - GUIDELINES FOR A FIELD CAMPAIGN, Annex I.

Cross, H. Z. & Zuber, M.S. (1972). Prediction of flowering dates in maize based on different methods of estimating thermal units. Agronomy Journal, 64:351– 355

De Castro, A.I., Six, J., Plant, R.E., & Peña, J.M. (2018) Mapping Crop Calendar Events and Phenology-Related Metrics at the Parcel Level by Object-Based Image Analysis (OBIA) of MODIS-NDVI Time-Series: A Case Study in Central California. Remote Sens. 2018, 10(11), 1745

Deosthali, V., Akmanchi, A.N. (2006). Yield estimation of summer crops grown in a mixed cropping area, a spectral growth curve profile approach. J Indian Soc Remote Sens 34, 279–288.

Edey, S.N. (1977). Growing degree-days and crop production in Canada. Publication 1635, Department of Agriculture, Ottawa.

Ehrenborg, L. (2017) University of Mainz. Published online (last seen Feb. 2020): https://download.uni-mainz.de/mathematik/Algebraische%20Geometrie/Lehre/Sem-Ausgewaehlte-hoehere-Kurven-WS2016-17/Ehrenborg%20Lars_Zykloide_151217.pdf

Eklundh, L. & Jönsson, P. (2011) TIMESAT 3.1: A software package to analyse time-series of satellite sensor data—Software Manual; Lund University: Lund, Sweden.

FAO, 2014: FAO Country Programming Framework for Kenya 2014-2017

Gowda, P.T., Satyareddi, S.A., & Manjunath, S. (2013) Crop Growth Modeling: A Review. Research and Reviews: Journal of Agriculture and Allied Sciences, 2013, 2, 1–11.

HM Treasury: The UK Treasury's The Green Book, Appraisal and Evaluation in Central Government https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685903/The_Green_Book.pdft

Hoheisel, S., Knott, L. & Kurven, S. (2009) Ausarbeitung zum Vortrag im Proseminar Analysis. Ruprecht-Karls Universitaet. Heidelberg. Published online (last seen Feb. 2020): https://www.mathi.uni-heidelberg.de/~thaeter/anasem08/02Spezielle%20Kurven_01_11_2008.pdf

Huang, Y., Zhu, Y., Li, W., Cao, W., & Tian, Y. (2013) Assimilating Remotely Sensed Information with the WheatGrow Model Based on the Ensemble Square Root Filter For improving Regional Wheat Yield Forecasts. Plant Prod. Sci. 2013, 16, 352–364.

Jiang R., Wang, T., Shao, J., Guo, S., Zhu, W., Yu, Y., Chen, S. & Hatano, R. (2017). Modelling the biomass of energy crops: Descriptions, strengths and prospective. Journal of Integrative Agriculture, 16(6): 1197–1210.

Kalubarme, M. H., Potdar, M. B., Manjunath, K. R., Mahey, R. K. & Siddh, S. S. (2003). Growth profile based crop yield models: A case study of large area wheat yield modelling and its extendibility using atmospheric corrected NOAA AVHRR data. International Journal of Remote Sensing, Volume 24, 10

Kasampalis, D., Alexandridis, T., Deva, C., Challinor, A., Moshou, D. & Zalidis, G. (2018) Contribution of Remote Sensing on Crop Models: A Review. J. Imaging, 4(4), 52

Kogan, F. (1995a). Droughts of the late 1980's in the United States as derived from NOAA polar-orbiting satellite data. Bulletin of the American Meteorological Society, 76(5), 655−668

Kogan, F. (1995b). Application of vegetation index and brightness temperature for drought detection. Advances in Space Research, 15, 91−100

Labus, M. P., Nielsen, G. A., Lawrence, R. L., Engel, R. & Long, D. S. (2002) Wheat yield estimates using multi-temporal NDVI satellite imagery. International Journal of Remote Sensing, 23:20, 4169-4180

Lenz, V.I.S., (2007) A process-based crop growth model for assessing Global Change effects on biomass production and water demand. https://d-nb.info/986419168/34 (Zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln, 2007)

Liu Jun-He, Yan Yan, Abid Ali, Ming-Fu Yu1, Qi-Jie Xu, Pei-Jian Shi & Lei Chen (2018). Simulation of crop growth, time to maturity and yield by an improved sigmoidal model. Sci Rep 8, 7030

Novelli, F., Spiegel, H., Sandén, T., & Vuolo, F. (2019). Assimilation of Sentinel-2 Leaf Area Index Data into a Physically-Based Crop Growth Model for Yield Estimation. Agronomy 2019, 9, 255

Nyemeyer, S. (2008). IES, JRC, 1st international Conference on Drought Management, Zaragoza, June 2008.

Paine, D., Alexandridis, T., Deva, C., Challinor, A., Moshou, D. & Zalidis, G. (2018). Contribution of Remote Sensing on Crop Models: A Review. J. Imaging 2018, 4, 52

Parthasarathi, T., Velu, G., & Jeyakumar, P., (2013). Impact of crop heat units on growth and developmental physiology of future crop production: a review. Journal of Crop Science and Technology 2 (1), 2319-3395

Rojas, O., Vrieling, A. & Rembold, F., (2011) Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. Remote Sensing of Environment,115 (2), 343–352

Swallow, B., Onyango, L. & Meinzen-Dick, R. (2003), Catchment Property Rights and the Case of Kenya's Nyando Basin. Proceedings of the African Regional Workshop on Watershed Management, Ch. 10, 123-136

Vescovi, F.D, & Ogembo, O. (2009) Study of erosion phenomena by means of an integrated remote sensing approach. A case study for Nyando River, Kenya. Proc. of 33rd ISRSE, 4 - 8 May 2009, Stresa, Italy

Voorend, W., Lootens, P., Nelissen, H. et al. (2014). LEAF-E: a tool to analyze grass leaf growth using function fitting. Plant Methods, 10, 37

Vorobiova, N., & Chernova, A., (2017). Curve fitting of MODIS NDVI time series in the task of early crops identification by satellite images. Procedia Engineering 201 184–195. 3rd International Conference "Information Technology and Nanotechnology," April 2017, Samara, Russia

Tables

NMDC – Normalized Multi-Band Drought Index
VCADI – Vegetation Condition Albedo Drought Index
PDI – Perpendicular Drought Index
MPDI – Modified Perpendicular Drought Index
RDRI – Remote Sensing Drought Risk Index
VegDRI – Vegetation Drought Response Index
ADI – Aggregate Drought Index
SMDI – Soil Moisture Deficit Index
ETDI – Evapotranspiration Deficit Index
RDI – Reconnaissance Drought Index
RSDI – Regional Streamflow Deficiency Index
SDI – Sperling Drought Index
NDVI – Normalized Difference Vegetation Index
VCI – Vegetation Condition Index
NDVIA – Anomaly of Normalized Difference Vegetation Index
SVI – Standardized Vegetation Index
NDWI – Normalized Difference Water Index
NDII – Normalized Difference Infrared Index
LWCI – Leaf Water Content Index
DTx – agricultural drought index
DFI – Drought Frequency Index
TCI – Temperature Condition Index
VHI – Vegetation Health Index
SRWI – Simple Ratio Water Index
GVWI – Global Vegetation Water moisture Index
WI – Water Index
%N – percentage of normal
DECILES – deciles
RAI – Rainfall Anomaly Index
BMDI – Bhalme and Mooly Drought Index
SAI – Standardized Anomaly Index
DSI – Drought Severity Index
PAI – Palfai Aridity Index
EDI – Effective Drought Index
Q90 – low flow index
BFI – Base Flow Index
SWSI – Surface Water Supply Index
PHDI – Palmer Hydrological Drought Index
RDI – Reclamation Drought Index
CMI – Crop Moisture Index
SMDI – Soil Moisture Drought Index
CSDI – Crop Specific Drought Index
CDI – Corn Drought Index
SCI – Soybean Drought Index
KBDI – Keetch-Byram Drought Index
NBR – Normalized Burn Ratio
PDSI – Palmer Drought Severity Index
PMDI – Palmer Modified Drought Index
Z-Index – Palmer Z-Index

Table 1: List of Common Drought Indices (Source: S. Nyemeyer, 2008)

Stages	Corn	Pearl millet
Emergence – Coleoptiles	0	0
2 leaves fully emerged	213	251
4 leaves fully emerged	345	457
6 leaves fully emerged	476	659
8 leaves fully emerged	608	853
10 leaves fully emerged	739	-
12 leaves fully emerged	871	-
14 leaves fully emerged	1003	-
16 leaves fully emerged	1134	1013
Silking/Anthesis/Boot leaf	1397	1043
Kernal in blister stage/half bloom	1661	-
Kernal in dough stage	1924	-
Kernal begins to dent	2187	1262
Kernal fully dented	2450	
Physiological maturity	2713	1661

Table 2: Cumulative Degree Days or thermal requirement for different development stages (Source: Parthasarathi et al., 2013)

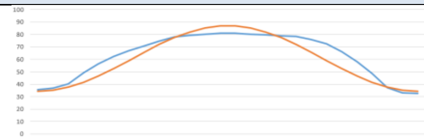
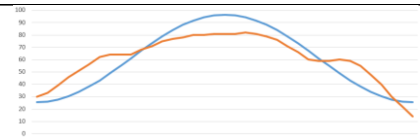
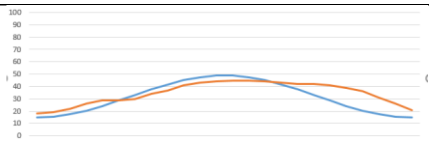
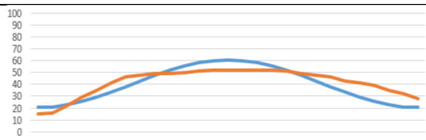
Actual and potential VHI curves	Crop type	RMSE	Actual VHI area	Potential VHI area	%
	Rice	6.7	737	655	112
	Cassava	11.6	1354	1278	105
	Corn	7.3	359	390	91
	Grassland	11.1	521	520	100

Table 3: Comparison between actual and potential VHI curves in good performing fields (Source: Project Report, Airbus, 2019)

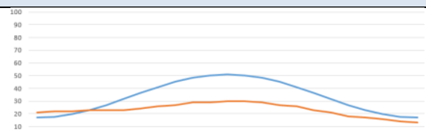
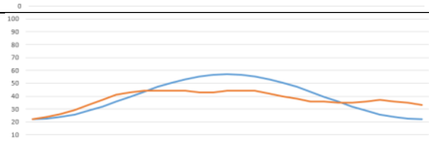
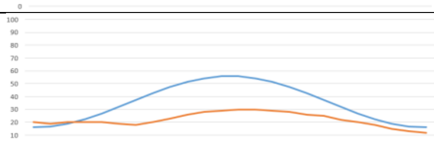
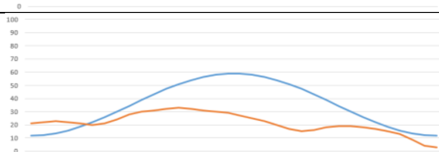
Actual and potential VHI curves	Crop type	RMSE	Actual VHI area	Potential VHI area	%
	Corn1	13.3	142	374	38
	Cow pea	8.5	290	490	59
	Sorghum	16.9	146	460	32
	Corn2	21.8	422	775	41

Table 4: Comparison between actual and potential VHI curves in poorly performing fields (Source; Project Report, Airbus, 2019)

Figures

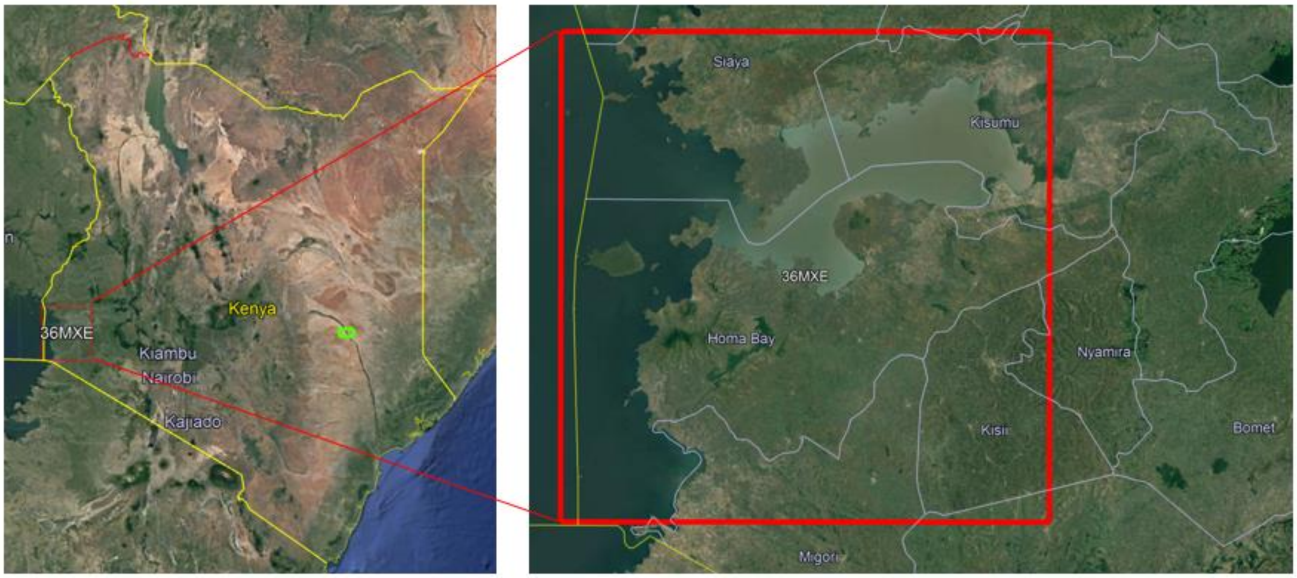


Figure 1: Location of the study area (Source: snapshot from GoogleEarth)

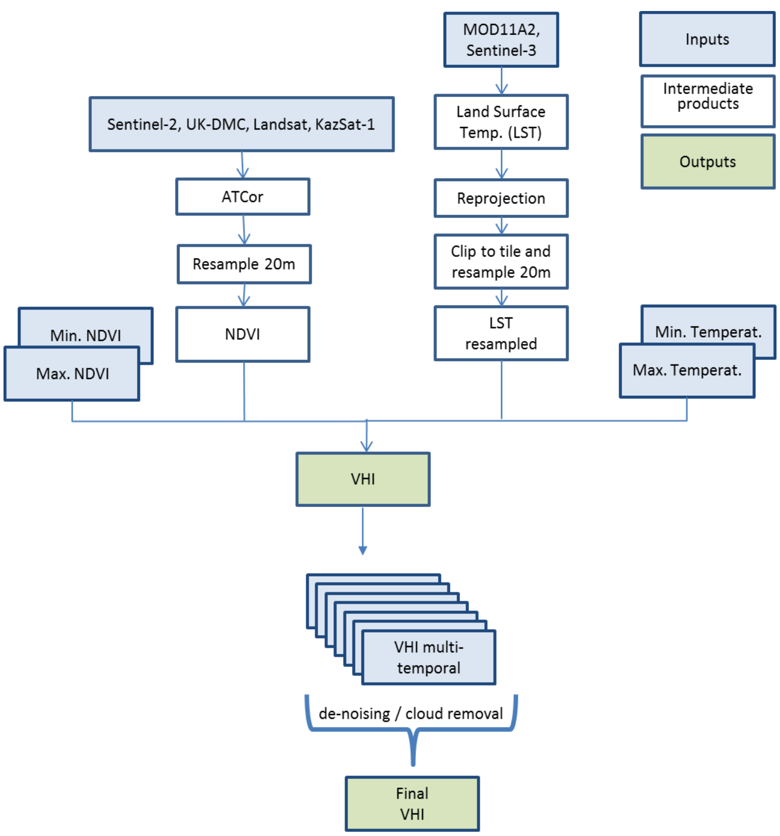


Figure 2: VHI image processing chain (Source: Project Report, Airbus, 2019)

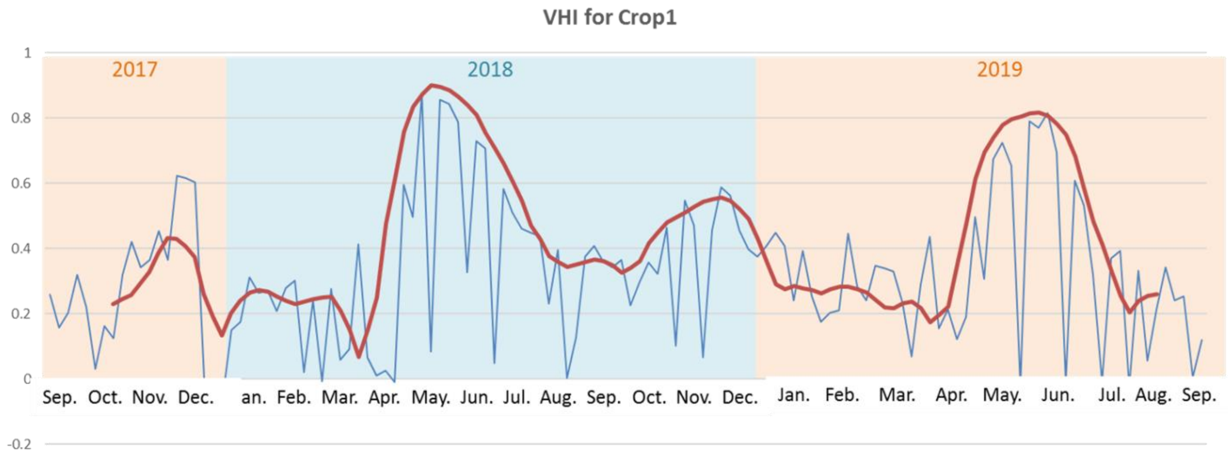


Figure 3: Multi-temporal VHI on a farmed field. Blue line: VHI without cloud removal. Red line: VHI with cloud removal (Source: Project Report, Airbus, 2019)

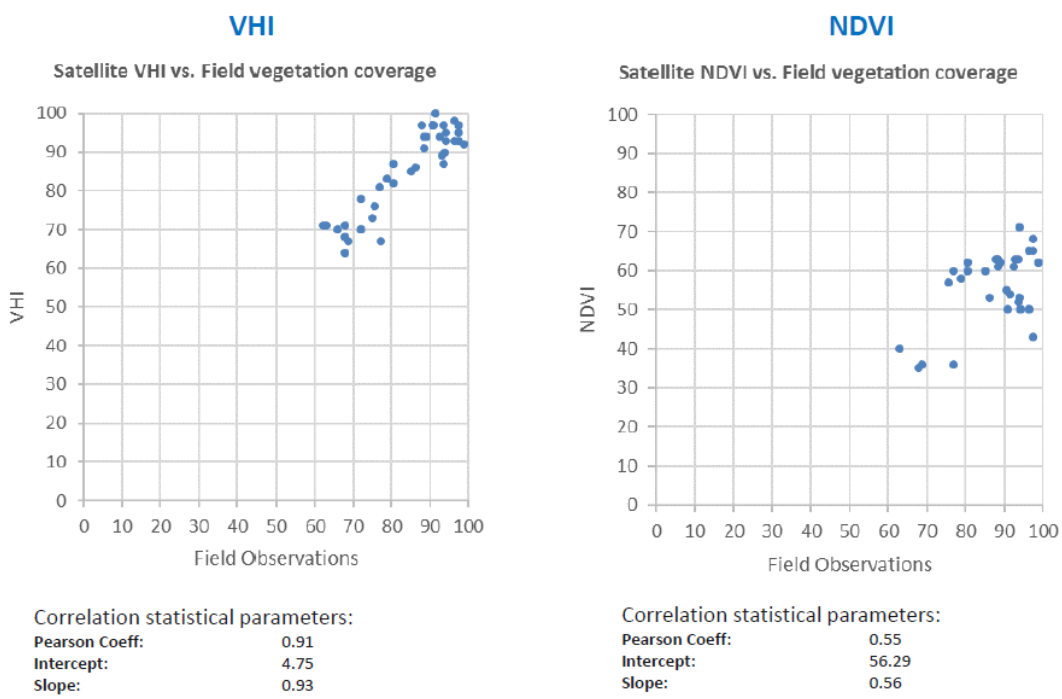


Figure 4: Validation results: VHI and NDVI comparison (Source: Project Report, Airbus, 2019)

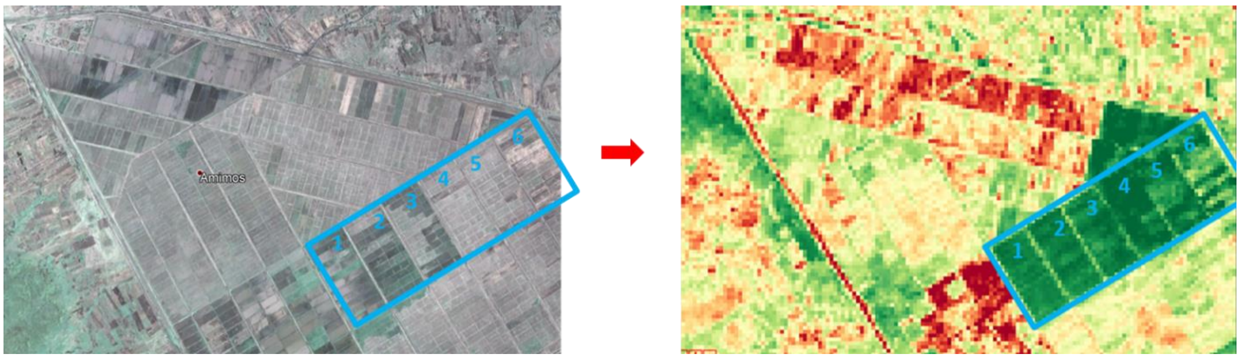


Figure 5: *Left*: satellite image (Source: GoogleEarth) before VHI processing on a large rice farm close to Kisumu (Amimos). *Right*: VHI of the same area mapped at the end of August 2019 (rice before harvesting time). The farmers grew rice in the six fields outlined in the blue box and different crops elsewhere, in a crop rotation scheme (Source: Project Report, Airbus, 2019)

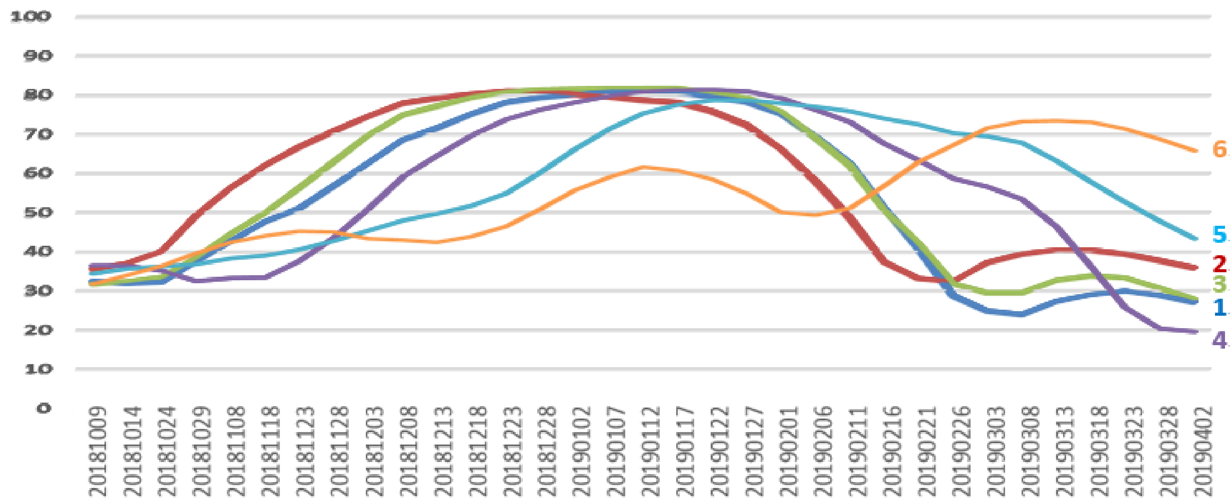


Figure 6: VHI plots of the six rice fields represented in Figure 5**Error! Reference source not found.** across the growing season in 2019. Fields numbered 1 to 5 show normal crop growth curves, with slight shifts in planting/harvesting times due to their farming practices. Field number 6, however, shows a sub-normal crop growth; possibly indicative of a different crop in that field (Source: Project Report, Airbus, 2019)

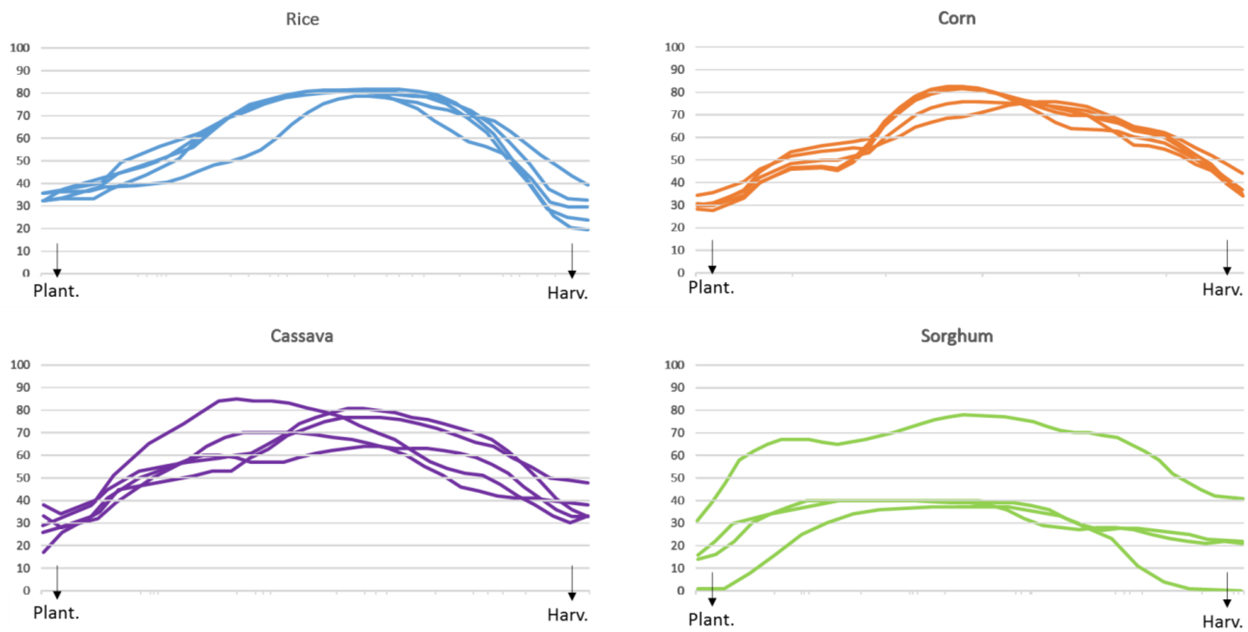


Figure 7: Example of VHI curves: each curve represents the crop cycle in one field. Planting and harvesting times are normalised for easy comparison (Source: Project Report, Airbus, 2019)

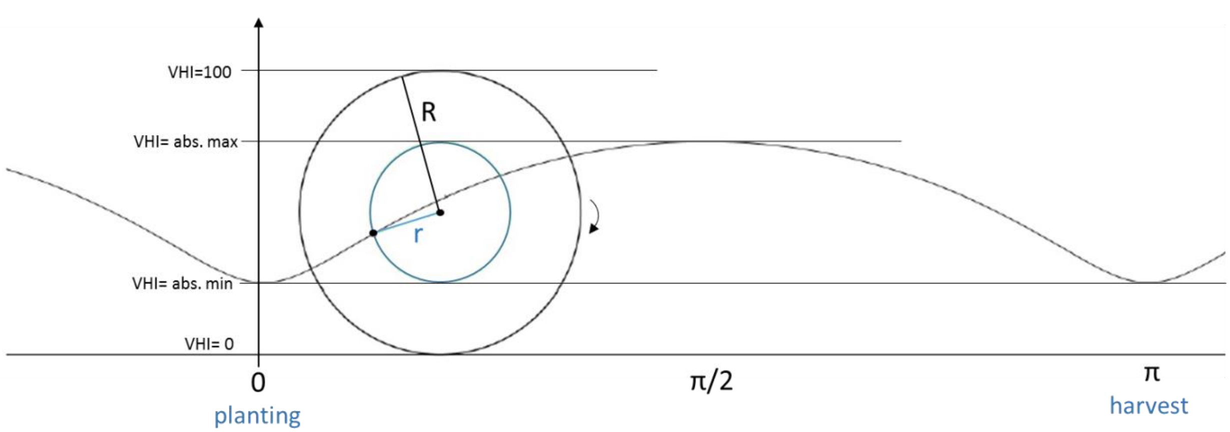
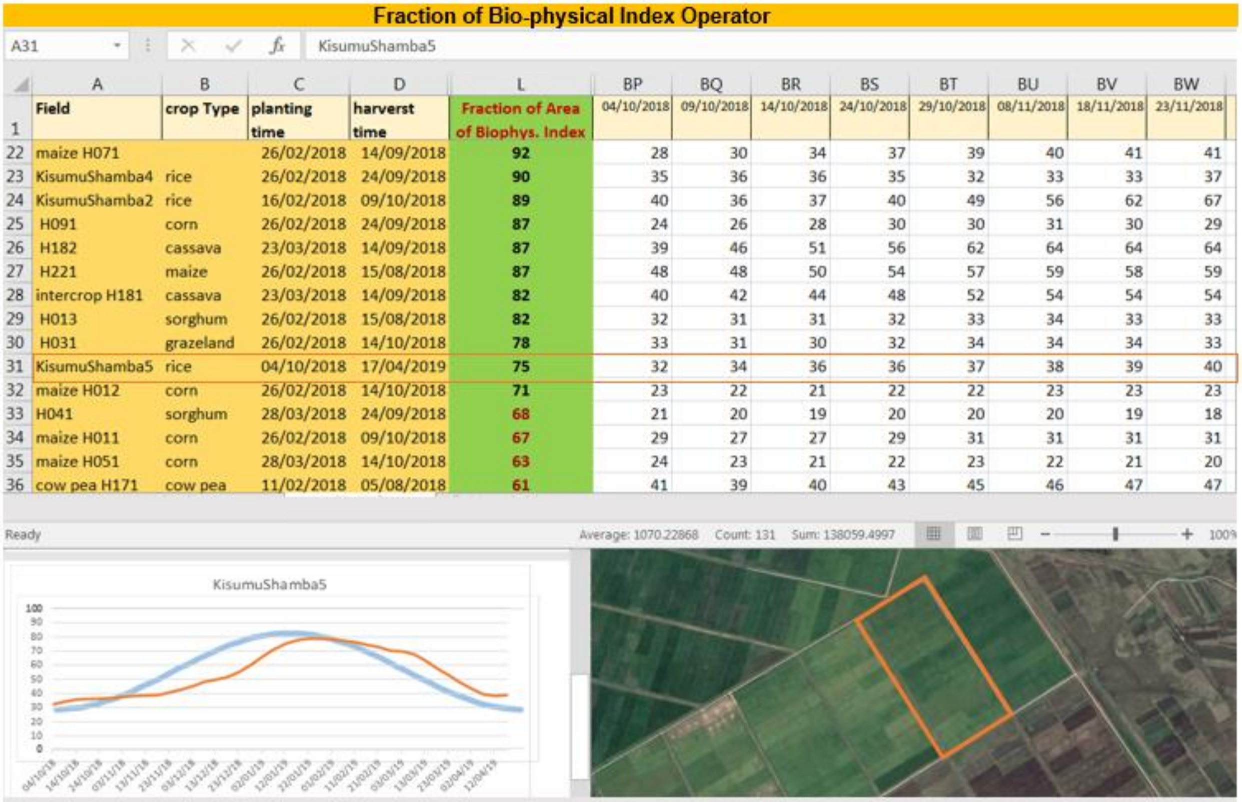


Figure 8: A cycloid is the curve traced by any point on the radius R of a wheel as it rolls along a straight line. The radius of the internal circle r is defined as the difference between the absolute historical maximum and minimum in the VHI time series. (Source: Ehrenborg, 2017, adapted by Airbus 2019)



Field	crop Type	planting time	harverst time	Fraction of Area of Biophys. Index	04/10/2018	09/10/2018	14/10/2018	24/10/2018	29/10/2018	08/11/2018	18/11/2018	23/11/2018
maize H071		26/02/2018	14/09/2018	92	28	30	34	37	39	40	41	41
KisumuShamba4	rice	26/02/2018	24/09/2018	90	35	36	36	35	32	33	33	37
KisumuShamba2	rice	16/02/2018	09/10/2018	89	40	36	37	40	49	56	62	67
H091	corn	26/02/2018	24/09/2018	87	24	26	28	30	30	31	30	29
H182	cassava	23/03/2018	14/09/2018	87	39	46	51	56	62	64	64	64
H221	maize	26/02/2018	15/08/2018	87	48	48	50	54	57	59	58	59
intercrop H181	cassava	23/03/2018	14/09/2018	82	40	42	44	48	52	54	54	54
H013	sorghum	26/02/2018	15/08/2018	82	32	31	31	32	33	34	33	33
H031	grazeland	26/02/2018	14/10/2018	78	33	31	30	32	34	34	34	33
KisumuShamba5	rice	04/10/2018	17/04/2019	75	32	34	36	36	37	38	39	40
maize H012	corn	26/02/2018	14/10/2018	71	23	22	21	22	22	23	23	23
H041	sorghum	28/03/2018	24/09/2018	68	21	20	19	20	20	20	19	18
maize H011	corn	26/02/2018	09/10/2018	67	29	27	27	29	31	31	31	31
maize H051	corn	28/03/2018	14/10/2018	63	24	23	21	22	23	22	21	20
cow pea H171	cow pea	11/02/2018	05/08/2018	61	41	39	40	43	45	46	47	47

Figure 9: Example of tool designed for use by insurance providers (Source: Project Report, Airbus, 2019)

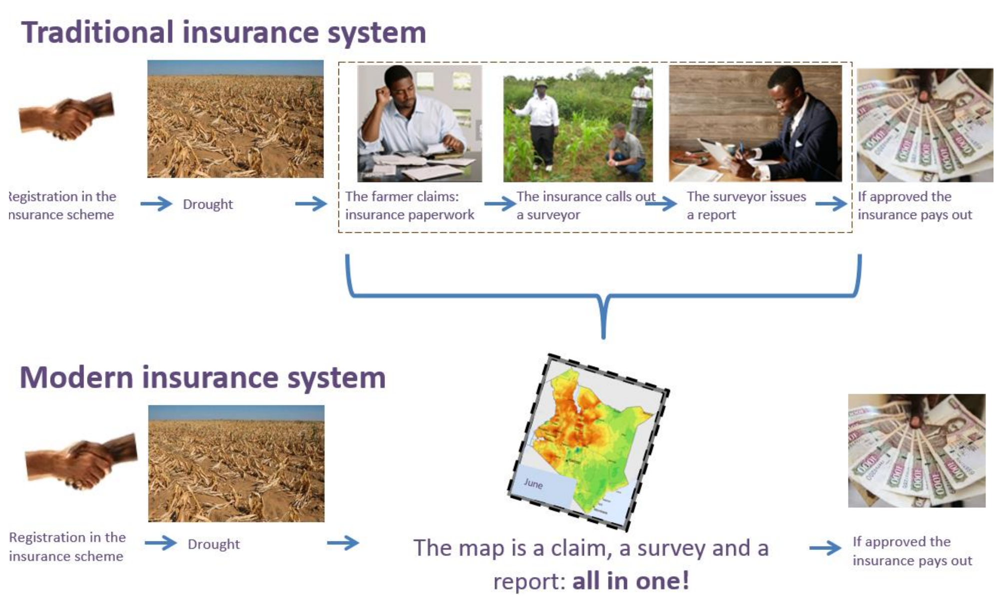


Figure 10: Comparison of the traditional insurance flow of operations with the new satellite index based method (Source: Project Report, Airbus, 2019)