KSA/RC/2021/P1FORM

KENYA SPACE AGENCY

RESEARCH CHAIR

- **Small-scale Crop Mapping using Artificial Intelligence/Machine Learning (AI/ML)**

PROPOSAL APPLICATION by
JKUAT

2021/2022 FINANCIAL YEAR

(*Expand the spaces provided in this form to fit your content. The* ***CONCEPT NOTE*** *should not exceed* ***20 pages*** *(Candara 12, single line spacing) excluding appendices*

PROPOSED PROJECT DETAILS

TITLE OF RESEARCH: Small-scale Crop Mapping Using Artificial Intelligence/ Machine Learning

Project Summary

The project involves development of a web-based application for small-scale crop mapping using Artificial Intelligence (AI) and Machine Learning (ML). It is aimed at capitalizing on the technical expertise in JKUAT and other institutions of higher learning in exploitation of satellite image data acquisition and ingestion techniques for small-scale crop mapping towards estimation of the crop area, crop status, crop yields and assessment of the food security situation in the country.

The research chair activities will entail the design, testing and implementation of a web portal for automatic acquisition/ download, pre-processing, analysis and classification of satellite imagery using AI and ML algorithms. The proposed web portal will be designed and implemented with a view of providing users with varying levels of expertise and needs access to a variety of satellite image products, ranging from raw image data to scientific products.

The project is intended to enhance the capacity and capabilities of the country in taking a more proactive information-based approach towards food security. The project team consists of 6 students from a range of disciplines in which gender representation has been ensured. A faculty member from the JKUAT Department of Geomatic Engineering and Geospatial Information Systems (GEGIS) will guide the students in implementation of the project.

Proposed project budget: Ksh (500,000.)

1.0 Background to the project *(Provide information on the problem being solved).* The second sustainable development goal (SDG-2) on 'zero hunger' recognizes that, while the number of undernourished people has reduced over the years, there is an urgent need to promote and support sustainable agriculture, small-scale farmers and equal access to land, technology and markets (Assembly, 2015). The aspect of promotion and support of small-scale farmers and access to technology is particularly important in Africa, where approximately 80% of the available cropland is cultivated by smallholder or small-scale farmers (Aguilar et al., 2018). Doubling agricultural productivity of small-scale food producers, implementing resilient agricultural practices that strengthen the capacity for adaptation to climate change, and increasing investment in agriculture through agricultural research and technological development, are some of the targets of SDG-2 that need to be aligned with the four pillars of food security: availability, access, utilization, and stability (Gil et al., 2019).

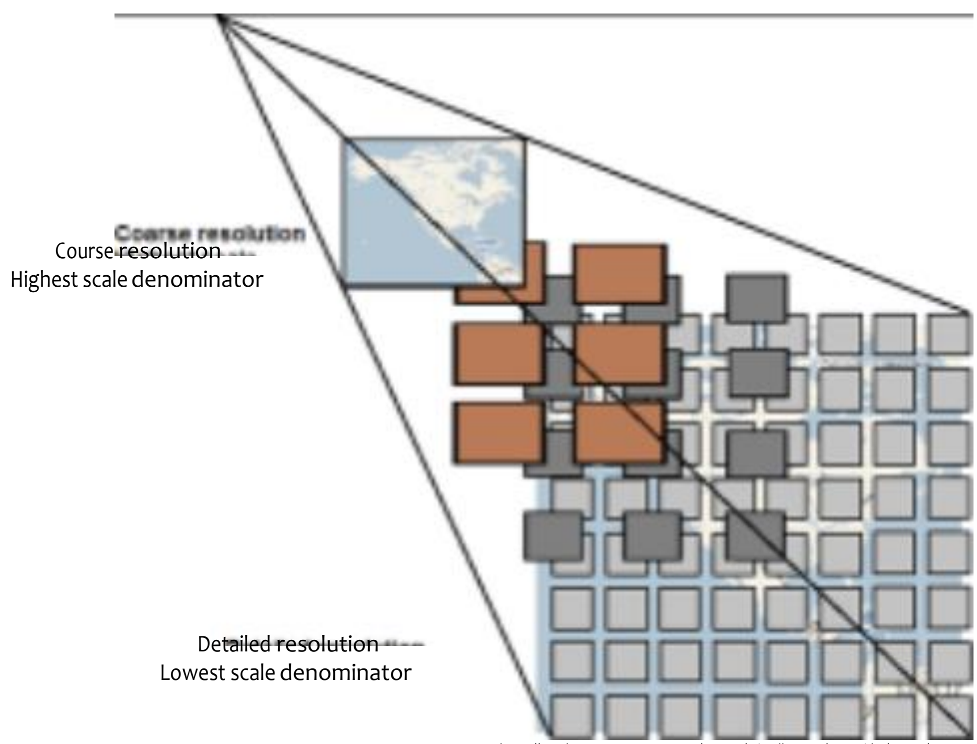
From an economic perspective crop cultivation is a significant and pertinent economic activity in Kenya, whose monitoring should be pursued. According to the Kenya Economic Survey (2021), the percentage contribution of crop cultivation to the Gross Domestic Product (GDP) in 2020 was 16.6%. In addition, the National Agricultural Investment Plan (NAIP, 2019 – 2024) highlights the limitations of the data that is currently available on agriculture, and the need for an increased capacity to use data from innovation, research and technology in agriculture. Data on the areas of crop cultivation, the types of crops that are under cultivation, and the status of crops during the growing season can aid in implementation of precision agriculture, thereby helping small-scale farmers to appropriately apply inputs and saving on production costs. Further, estimation of yield can help the country to plan for food shortages or surplus thus ensuring food security.

Crop mapping of smallholder agriculture is a non-trivial exercise, which requires procedures that can handle temporally, spatially, and spectrally complex data, since small- scale crop cultivation tends to be spatiotemporally heterogeneous and diverse (Ibrahim et al., 2021). Remote sensing imagery acquired by sensors aboard satellites and Unmanned Aerial Vehicles (UAVs) at varying spatial and temporal scales provide an excellent source of data for the mapping and continuous monitoring of small-scale crop cultivation. Satellite- derived estimations and predictions of crop features and outputs include crop monitoring models and applications centered around five main areas including: biomass and yields estimation, vegetation and water stress monitoring, crop acreage estimation, crop type proportion mapping and crop phenological development (Zhang et al., 2016). The proposed research will involve exploitation of AI and ML procedures including image signal processing for multi-spatiotemporal image correction and classification. A web-based application will allow access to derived estimates on cropland area and crop features using optical remote sensing satellite imagery. The application will be designed and implemented with a view of providing near-real-time data on crop cultivation areas and their status using a variety of satellite image products including Landsat-8/ 9, Sentinel-2A/2B and Sentinel- 1A/1B.

2.0. Justification *(A clear justification for the proposed application's relevance and uniqueness)*

Small-scale crop cultivation is characterized by varied crop types, crop varieties, tillage practices, and planting times, thus necessitating multi-temporal classification approaches which utilize time-series data (Nduati et al., 2019). In order to get up-to-date information of the land under crop cultivation and the crop growth and performance status of the crops at any instance in time, the most recently available satellite imagery is required. However, optical satellite imagery have varied temporal and spatial resolution, and the imagery may be severely affected by atmospheric artefacts, thus making information retrieval difficult or impossible due to missing data. In order to overcome this, active radar remote sensing missions, such as the Sentinel-1 C-band SAR active sensor, can be used since they can acquire data about the earth's surface at any time of the day or night, regardless of weather and environmental conditions. A major drawback of radar data is that the data require expert processing and interpretation thus limiting access to non-expert application users. In addition, optical sensors acquire data in a variety of spectral bands from which indices that are much better documented can be derived and directly applied by users with little to no experience with GIS and remote sensing. We propose the use of AI and ML algorithms that will be implemented and run on a server as per the data request made by a user. Since the datasets proposed, that is Landsat-8/ 9, Sentinel-2A/2B and Sentinel-1A/1B, have varied temporal resolution, the synthesis of these data sets using AI and ML will provide a continuum of images and products that can be used for monitoring crop production.

In order to serve the data needs of users who have varied levels of remote sensing expertise, while at the same time providing high quality data and information for various locations across the country, it is prudent to implement a web-application which allows access to various types of data on-demand. We propose the design and implementation of a Web Map Tile Service (WMTS), which would allow any client to make requests and interact with the WMTS server. The WMTS will provide raster satellite imagery data at multiple resolutions in predefined imagery Tiles (PNG or JPEG formats), as depicted in Figure 2.1 below, thus making it more scalable and high-performing than a Web Map Service (WMS).



Source: https://gcs-docs.s3.amazonaws.com/EVWHS/Miscellaneous/DevGuides/WMTS/WMTS_MapTile.htm

Figure 2.1: WMTS service divided into tiles

3.0 Project objective(s) *(The objectives stated must be specific, measurable, realistic and attainable within the given project time frame)*

Main Objective

To design, develop and implement a web-based application for the access/ download, processing, classification and retrieval of small-scale cropland specific features and outputs.

Specific Objectives

a) To develop a web-based application for the automatic download of Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Kenya from their respective hubs/ portals.

b) To implement real-time image processing and archiving for retrieval of the Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Kenya, as they become available.

c) To develop and implement AI/ ML image classification algorithms for mapping of land cover and final estimation and retrieval of crop area, crop status indices and crop yield.

4.0 Expected outputs of the project *(Clearly state the project expected outputs which should be realistic and quantifiable. The stated expected outputs should be clearly linked to your project objectives).*

The following are the expected outputs from the research chair project:

a) A web-service designed to handle user-requests for search and discovery of Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images over selected area(s) within Kenya and selected time/ date range(s).

b) Backend-processing of the Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images for the selected area(s) and date(s) using AI/ML algorithms.

c) A web service archiving processed and classified imagery for future requests.

d) Delivery of cropland area estimation and crop status indices from the back-end processing as per user requests via the web application for download.

5.0 Research design and methodology *(Provide clear descriptions of software development model adopted, system design and modules and methodology)*

5.1 <u>Objective 1: To develop a web-based application for the automatic download of Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Kenya from their respective hubs/ portals</u>

A web-based application employing an integration of WMTS client and server applications will be designed, developed and implemented. The WMTS client will manage the interactions with WMTS interfaces through HTTP requests and dynamically generate HTML that can run in a web browser. The WMTS server will accept requests from the WMTS client and the viewer client in the form of HTTP URL strings, and returns results encoded as XML, PNG, GML, and so on. The WMTS exchange mechanism between the clients and servers will be defined by mandatory GetCapabilities and GetTile and optional GetFeatureInfo operations. The conceptual general architecture of the WMTS is as shown in Figure 3.1 below.

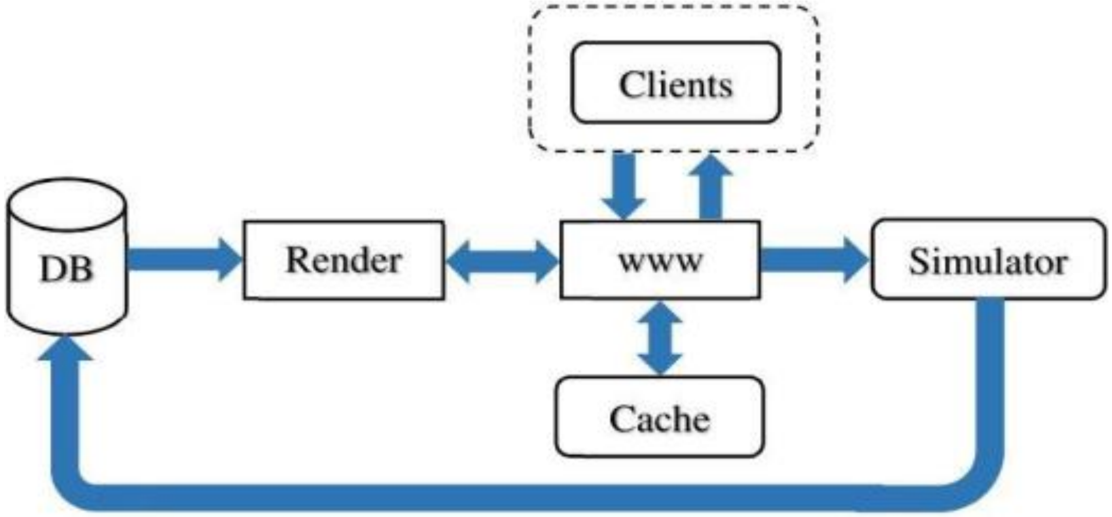


Figure 3.1: Tile map server architecture

Two approaches of satellite image data acquisition will be tested:

a) Access satellite data via URLs at sources that are product specific and include:

   i. Copernicus data hub: Sentinel-2A/2B and Sentinel-1A/1B data

   ii. Global Imagery Browse Services (GIBS) API for developers: Landsat 8

b) Access satellite imagery via Google Earth Engine WMTS Proxy

For imagery acquired via the GIBS API, imagery products in EPSG:4326 (Lat-lon/Geographic / WGS 84) and EPSG:3857 (Web Mercator / Spherical Mercator / "Google Projection") projections are supported. GIBS projections and their available resolutions are as shown in Tables 1 and 2 below.

Table 1: GIBS projections and their available resolutions for EPSG:4326

WGS 84 / Lat-lon / Geographic (EPSG:4326)

Resolution (per pixel)	Tile Matrix Set (WMTS)	# Zoom Levels	Max Resolution (deg/pixel)	Min Resolution (deg/pixel)
15.125m	15.125m	13	0.5625	0.0001373291015625
31.25m	31.25m	12	0.5625	0.000274658203125
250m	250m	9	0.5625	0.002197265625
500m	500m	8	0.5625	0.00439453125
1km	1km	7	0.5625	0.0087890625
2km	2km	6	0.5625	0.017578125

Table 2: GIBS projections and their available resolutions for EPSG: 3857

Web Mercator (EPSG:3857)

Resolution (per pixel)	Tile Matrix Set (WMTS)	# Zoom Levels	Max Resolution (m/pixel)	Min Resolution (m/pixel)
19.10925707129405 m	GoogleMapsCompatible_Level13	13	156543.03390625	19.10925707129405
38.21851414258810 m	GoogleMapsCompatible_Level12	12	156543.03390625	38.21851414258810
305.7481131407048 m	GoogleMapsCompatible_Level9	9	156543.03390625	305.7481131407048
611.4962262814100 m	GoogleMapsCompatible_Level8	8	156543.03390625	611.4962262814100
1222.992452562820 m	GoogleMapsCompatible_Level7	7	156543.03390625	1222.992452562820
2445.984905125640 m	GoogleMapsCompatible_Level6	6	156543.03390625	2445.984905125640

For the Earth Engine WMTS proxy data acquisition approach, we will first create an App Engine, and then deploy by running Google App Engine.

As far as is possible, we do not foresee the need for any commercial software for the implementation of the WMTS. However, the server operating system may be Windows (Windows Server 2019) or Linux (Ubuntu 18+/ Debian 10) based depending on the design functionalities that will be implemented at the development stage. The hardware requirements are as below:

- 1 CPU - NVIDIA GPU (32GB Memory); 128GB RAM; 1TB SSD Hard disk

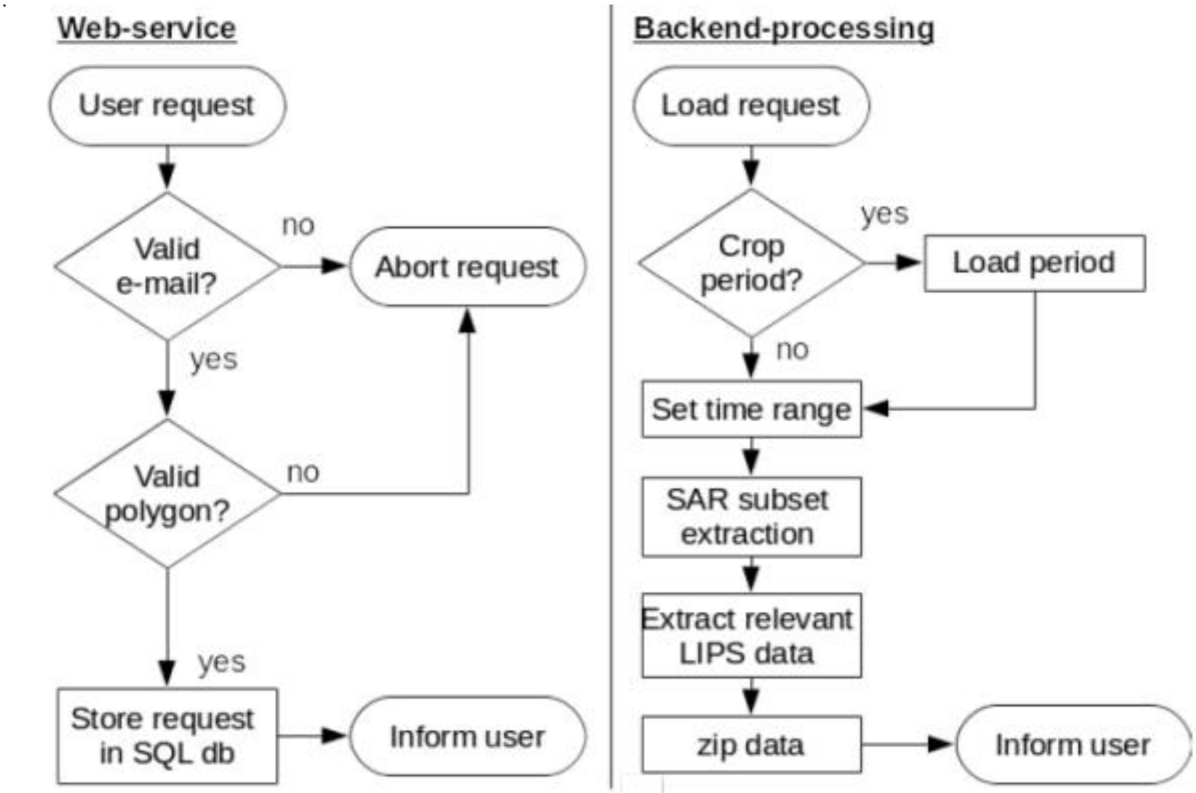
5.2 Objective 2: To implement real-time image processing and archiving for retrieval of the Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Kenya, as they become available.

In addition to the WMTS GetTile services, running on the server, background image processing algorithms will be running and processing the image data to various levels, as may be requested by the user. The main processing algorithms include:

a) Landsat 8 and Sentinel-2A/2B: At-sensor radiance and TOA reflectance, Surface reflectance, Compositing, Cloud Masking and Mosaicking.

b) Sentinel 1A/1B: Pixel-wise derivation of the backscatter coefficient through application of orbit file, GRD border noise removal, thermal noise removal, radiometric calibration, and terrain correction (orthorectification).

Figure 5.1 below shows an example of how the back-end image processing will be implemented in relation to the WMTS client.



Source: https://www.researchgate.net/publication/327496231_Current_potentials_and_challenges_using_Sentinel-1_for_broadacre_field_remote_sensing

Figure 5.1 Flowchart of the server processing of requests and image processing

5.3 Objective 2: To develop and implement AI/ ML image classification algorithms for mapping of land cover and final estimation and retrieval of crop area, crop status indices and crop yield.

During the client data access request process, users will be required to stipulate which kind of data output they wish to view or download. This research will implement automatic AI/ML-driven classification algorithms to generate:

a) Land cover classification (general), depicting all classes as outlined by the UN Land Cover Classification System's initial 'Dichotomous Phase', in which eight major land cover types are defined: (1) Cultivated and Managed Terrestrial Areas, (2) Natural and Semi-Natural Terrestrial Vegetation, (3) Cultivated Aquatic or Regularly Flooded Areas, (4) Natural and Semi-Natural Aquatic or Regularly Flooded Vegetation, (5) Artificial Surfaces and Associated Areas, (6) Bare Areas, (7) Artificial Waterbodies, Snow and Ice, and (8) Natural Waterbodies, Snow and Ice (Di Gregorio, 2005).

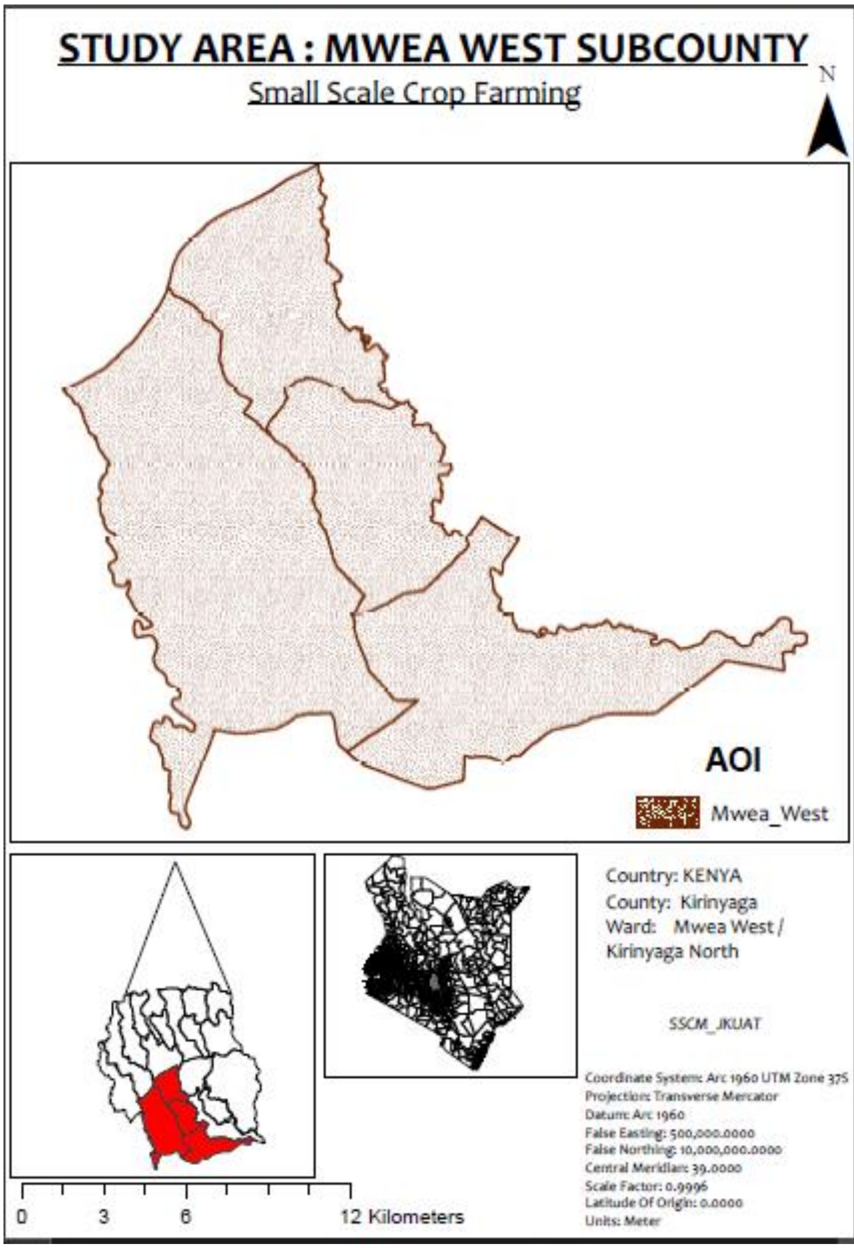
b) Cropland area including Cultivated and Managed Terrestrial Areas and Cultivated Aquatic or Regularly Flooded Areas

c) Crop-related indices including Normalized Difference Vegetation Index (NDVI), Normalized Difference Red Edge (NDRE) index, Modified Soil Adjusted Vegetation Index (MSAVI), Normalized Difference Moisture Index (NDMI) and Red edge Chlorophyll (ReCl) index.

d) Field-scale time series analytics derived from the indices above.

The classification of the satellite images will require ground reference data collected at regular intervals over the course of the research project. JKUAT possesses two state of the art drones (multi-spectral and LiDAR) and RTK-GNSS survey equipment which will be used for the ground truth data acquisition exercises.

6.0 Study Area



The Region of Interest (ROI) was chosen as Kirinyaga County due to its closeness to Kenya's weather conditions as well as its small-scale farming productivity. Kirinyaga County has an average annual temperature of 25° with that of Kenya ranging between 20°C and 28°C. A favorable study area temperature would be considered to be at 24°C, thus Kirinyaga would prove sufficient in its temperature conditions.

Additionally, Kirinyaga's rainfall is estimated to be 574 mm annually, with Kenya's average rainfall standing at 680mm.

Mwea West Subcounty was chosen as the AOI (Area of Interest), under which research would be undertaken.

This was under the supported fact that apart from the Mwea Irrigation scheme, a variety of small-scale crop growing including Tomatoes, onions, French Beans, Maize and other horticultural crop do score well in the said region.

7.0 Workplan & estimated budget *(Provide a clear project work plan of the activities, timelines and detailed budget cost estimate for the development of the mobile and web applications in the template below. This should include personnel, equipment and other related cost).*

Note: Administration and personnel cost should not exceed Ksh 100,000 of the total award.

Expected outcome 1: A web-service for search and discovery of Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images.

Module Development	Expected Output	Timeline	Estimated Budget
WMTS design, development and implementation	*Web-based application*	*2months*	*Purchase of Hardware and cloud subscription - 400,000*

Expected outcome 2: Backend-processing of the Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images for the selected area(s) and date(s) using AI/ML algorithms.

Compilation and testing of AI/ML Google Earth Engine processing algorithms	*Processing algorithms*	*3months*	*Covered within Expected outcome 1*

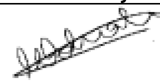
Expected outcome 3: A web service archiving processed and classified imagery for future requests.

Expected outcome 4: Delivery of cropland area estimation and crop status indices from the back-end processing as per user requests via the web application for download.

Processed imagery archiving	*Standards and definition of archive database structure*	*3months*	*Covered within Expected outcome 1*
Image classification using AI/ML algorithms	*Processing algorithms; Land cover maps; Cropland area maps; Crop-related indices*	*3months*	*Covered within Expected outcome 1*
Ground reference data collection survey	*Ground reference data*	*1month*	*Research expenses for personnel (airtime, internet bundles, transport)-100,000*

8.0. References (*Provide all references used in your proposal*)

1. Aguilar, R., Zurita-Milla, R., Izquierdo-Verdiguier, E., & A De By, R. (2018). A cloud-based multi-temporal ensemble classifier to map smallholder farming systems. Remote sensing, 10(5), 729.
2. Assembly, G. (2015). Sustainable development goals. SDGs Transform Our World, 2030.
3. Gil, J. D. B., Reidsma, P., Giller, K., Todman, L., Whitmore, A., & van Ittersum, M. (2019). Sustainable development goal 2: Improved targets and indicators for agriculture and food security. *Ambio*, *48*(7), 685-698.
4. Bosc, P. M., Berdegué, J., Goïta, M., van der Ploeg, J. D., Sekine, K., & Zhang, L. (2013). *Investing in smallholder agriculture for food security* (No. 6). HLPE.
5. https://www.agck.or.ke/Downloads/NATIONAL-AGRICULTURE-INVESTMENT-PLAN-NAIP.pdf
6. https://www.knbs.or.ke/wp-content/uploads/2021/09/Economic-Survey-2021.pdf
7. Ibrahim, E. S., Rufin, P., Nill, L., Kamali, B., Nendel, C., & Hostert, P. (2021). Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery. *Remote Sensing*, *13*(17), 3523.
8. Nduati, E., Sofue, Y., Matniyaz, A., Park, J. G., Yang, W., & Kondoh, A. (2019). Cropland Mapping Using Fusion of Multi-Sensor Data in a Complex Urban/Peri-Urban Area. *Remote Sensing*, *11*(2), 207.
9. Zhang, X., Zhang, M., Zheng, Y., & Wu, B. (2016). Crop mapping using PROBA-V time series data at the Yucheng and Hongxing farm in China. *Remote Sensing*, *8*(11), 915.
10. Di Gregorio, A. (2005). *Land cover classification system: classification concepts and user manual: LCCS* (Vol. 2). Food & Agriculture Org..

Faculty lead	Name	Dr. Eunice Nduati
	Email	enduati@jkuat.ac.ke
	Mobile	0759 81 44 99
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Signature	
	Date	2021/12/01
Student researcher 1	Name	Macheche Timothy Simiyu
	Email	macheche.timothy@students.jkuat.ac.ke
	Mobile	0705 85 20 91
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	B.Sc. Geomatic Engineering and Geospatial Information Systems
	School	School of Civil, Environmental and Geospatial Engineering
Student researcher 2	Name	Orera Anita Carolyne
	Email	orera.anita@students.jkuat.ac.ke
	Mobile	0780 873 246
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	B.Sc. Geomatic Engineering and Geospatial Information Systems
	School	School of Civil, Environmental and Geospatial Engineering
Student researcher 3	Name	Wainoi Evalyne Nyawira
	Email	wainoi.evalyne@students.jkuat.ac.ke
	Mobile	0701 01 03 94
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	B.Sc. Geospatial Information Science
	School	School of Civil, Environmental and Geospatial Engineering
Student researcher 4	Name	Otieno Ibrahim
	Email	odhiambo.ibrahim@students.jkuat.ac.ke
	Mobile	0703 81 83 11
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	B.Sc. Computer technology
	School	School of Computing and IT
Student researcher 5	Name	Michael Njoroge
	Email	wanjiku.njoroge2019@students.jkuat.ac.ke
	Mobile	0729 84 29 98
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	Bachelor of Business Information Technology (BBIT)
	School	School of Business
Student researcher 6	Name	Raphael Migwi Njuguna
	Email	njuguna.raphael@students.jkuat.ac.ke
	Mobile	0718 30 87 33
	Institution	Jomo Kenyatta University of Agriculture and Technology
	Course	B.Sc. Computer Technology
	School	School of Computing and Information Technology (SCIT)

Guidelines on Evaluation of Project *To be filled by the judging panel

The following will be the Judging Criteria for the project:

Item	Score	Remarks
1. Research Problem	10 points	
2. Design and Methodology	20 points	
3. Execution: Proposed Development and Testing	20 points	
4. Creativity	10 points	
5. Presentation	30 points	
6. Budget	10 points	