Developing Innovation in Sustainability, Resilience, and Adaptation (ISRA) Metrics for Agricultural Innovations in Western Kenya

Alliance of Bioversity International and CIAT
International Institute of Tropical Agriculture
University of Leeds
iShamba

March 31, 2026

1. Table of Contents

About Authors

Amal Chakhar is a postdoctoral fellow at the University of Leeds.

Chris Mwungu is a Postdoctoral Fellow (Agricultural Economist) at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Agnes Wanjau is a Data Analyst at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Rebecca Sarku is a research fellow at the University of Leeds.

Stephen Whitfield is a professor at the University of Leeds.

Benson Kenduiywo is a research specialist at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Tibebe Degefie is a research specialist at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Eduardo Garcia Bendito is a data scientist at the International Institute for Tropical Agriculture (IITA).

Lilian Kirwa is the iShamba product development manager at the Mediae Company.

Brian Mayanja is a research specialist at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Sabrina Rose is a Science and Implementation Officer at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Aniruddha Ghosh is a Senior Scientist at the Alliance of Bioversity International and the International Centre for Tropical Agriculture.

Andy Challinor is a professor at the University of Leeds.

To cite this report

Chakhar, A., Mwungu, C., Wanjau, A., Sarku, R., Whitfield, S., Kenduiywo, B., Degefie, T., Garcia Bendito, E., Kirwa, L., Mayanja, B., Rose, S., Ghosh, A., & Challinor, A. (2026). Developing Innovation in Sustainability, Resilience, and Adaptation (ISRA) Metrics for Agricultural Innovations in Western Kenya. Alliance of Bioversity International and the International Centre for Tropical Agriculture (CIAT).

Acknowledgements

This work was undertaken as part of the Innovation in Sustainability, Policy, Adaptation, and Resilience (iSPARK) Project in Kenya. iSPARK focuses on developing metrics for sustainable and resilient agriculture that span scales from farmers' fields to the national level.

Organisations participating in iSPARK include the Alliance of Bioversity International and CIAT, the University of Leeds, iShamba Limited and IITA. The UK International Development has funded this material through the UK-CGIAR Centre; however, the views expressed do not necessarily indicate the UK government's official policies.

This report has not undergone peer review. The opinions expressed herein are solely those of the author(s) and do not necessarily reflect the policies or views of The Alliance, Donors, or Partners.

Abstract

Climate variability and increasing environmental pressures boost the need for reliable metrics to assess agricultural innovations that enhance sustainability, resilience, and adaptation in smallholder farming systems. This study develops and applies the Innovation in Sustainability, Resilience, and Adaptation (ISRA) metrics to support evidence-based evaluation of climate-smart agricultural interventions in Western Kenya. Using a systematic review and scoring framework, we identified and assessed potential metrics derived from remote sensing, modelling, and socio-economic indicators to determine their relevance, scalability, and alignment with climate-smart agricultural goals. The analysis focused on three innovation areas: targeted intercropping for resilience, climate-smart irrigation management, and soil health improvement, selected through stakeholder consultations and technical feasibility assessments. For intercropping, satellite-based crop classification methods utilising Sentinel imagery and machine learning models were employed to distinguish between monocrop and intercropped maize systems, while drought monitoring involved Vegetation Indices such as the Vegetation Condition Index (VCI). Results show moderate classification accuracy and indicate that intercropped systems consistently maintain higher vegetation health during drought periods compared to monocrop systems. Concerning irrigation, a composite drought-monitoring framework integrating vegetation, temperature, precipitation, and soil moisture indicators was developed to support climate-smart irrigation advisories. For soil health, spatial nutrient gap analysis and soil fertility monitoring were proposed to guide targeted nutrient management interventions. Across these innovations, yield estimation emerged as a cross-cutting metric linking agronomic performance to resilience and adaptation outcomes. The findings highlight the potential of integrating Earth observation (EO) data, machine learning techniques, and socio-economic indicators to produce scalable metrics that support climate-resilient agricultural decision-making. The ISRA framework provides a practical approach to monitoring and evaluating agricultural innovations across levels, from farm management to regional policy planning, thereby strengthening evidence-based strategies for climate-smart agriculture in smallholder systems.

List of abbreviations and acronyms

Abbreviation/Acronym	Full Meaning
AfSIS	Africa Soil Information Service
CADCI	Comprehensive Agriculture Drought Condition Indicator
CDSE	Copernicus Data Space Ecosystem
CGIAR	Consultative Group on International Agricultural Research
CIAT	International Centre for Tropical Agriculture
CSA	Climate-Smart Agriculture
CWP	Crop Water Productivity
DEM	Digital Elevation Model
ECV	Essential Climate Variable
EO	Earth observation
ESA	European Space Agency
ET	Evapotranspiration
FAMEWS	Fall Armyworm Monitoring and Early Warning System
FAW	Fall Armyworm
FGDs	Focus Group Discussions
FVC	Fractional Vegetation Cover
GCVI	Green Chlorophyll Vegetation Index
GEE	Google Earth Engine
GEOBIA	Geographic Object-Based Image Analysis
GHG	Greenhouse Gas
GOSAT	Greenhouse gases Observing Satellite
GPP	Gross Primary Productivity
IITA	International Institute of Tropical Agriculture
ISRA	Innovation in Sustainability, Resilience, and Adaptation
iSPARK	Innovation in Sustainability, Policy, Adaptation, and Resilience in Kenya
LAI	Leaf Area Index
LST	Land Surface Temperature
LULC	Land Use/Land Cover
MAE	Mean Absolute Error
MCM	Participatory multi-criteria Mapping
MEL	Monitoring, Evaluation, and Learning
NDVI	Normalised Difference Vegetation Index
NPK	Nitrogen, Phosphorus, and Potassium
NPP	Net Primary Productivity
NUE	Nutrient Use Efficiency
OA	Overall Accuracy
PCI	Precipitation Condition Index
PET	Potential Evapotranspiration
QUEFTS	Quantitative Evaluation of the Fertility of Tropical Soils
RDM	Reference Data Module
RF	Random Forest
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SCYM	Scalable Crop Yield Mapper
SFI	Soil Fertility Index
SOC	Soil Organic Carbon
SOM	Soil Organic Models
SPI	Standardised Precipitation Index
SSA	Sub-Saharan Africa

STMs	Spectral Temporal Metrics
SVR	Support Vector Regression
TANSO	Thermal and Near-infrared Sensor for carbon Observation
TCI	Temperature Condition Index
TCWV	Total Column Water Vapour
VCI	Vegetation Condition Index
VegDRI	Vegetation Drought Response Index
VHI	Vegetation Health Index
VIs	Vegetation Indices
WP	Water Productivity

List of tables

List of figures

1. Introduction

Agricultural systems in Sub-Saharan Africa (SSA) are becoming increasingly vulnerable to climate variability, environmental degradation, and socio-economic pressures that threaten food security and the resilience of smallholder farming systems. At the farm level, poverty, limited access to resources, and high reliance on rain-fed agriculture reduce farmers' ability to manage climate risks, while access to climate-resilient technologies, financial services, and early warning systems remains limited (Alawode, 2025). On a broader scale, policy, investment, and advisory systems often struggle to translate scientific evidence into effective development actions due to weak data infrastructure, fragmented advisory services, and limited communication between researchers and policymakers (Damba et al., 2022). In Kenya, these challenges are reflected in increasing climate variability, land degradation, declining soil fertility, and growing pressure on agricultural systems that constrain productivity and rural livelihoods (Kogo et al., 2021). Western Kenya exemplifies these national dynamics, where rising temperatures and frequent droughts are increasingly affecting agricultural production and smallholder livelihoods (Wetende et al., 2018).

Addressing these challenges requires more than simply promoting individual agricultural technologies or practices (Barrett et al., 2020). Decision-makers need reliable and context-specific evidence to determine which combinations of innovations are suitable, where they are effective, and for whom. This evidence must also account for potential trade-offs among agricultural, environmental, and socio-economic outcomes and involve multiple stakeholders across various scales (Winowiecki et al., 2021). However, current evidence on climate-smart and sustainable agricultural interventions remains fragmented and often focuses on a limited range of outcomes. Some assessment tools emphasise biophysical indicators such as productivity, while paying less attention to social and economic dimensions, including food security, poverty, and gender (van Wijk et al., 2020). In some cases, adoption rates or levels of adoption are used as indicators of success without fully capturing behavioural, structural, and temporal factors that influence outcomes (Vallury et al., 2025). Consequently, policymakers, development practitioners, and investors often lack systematic and comparable metrics to guide the prioritisation, scaling, and financing of agricultural innovations.

This challenge highlights the need for scalable, spatially informed evaluation frameworks to assess how agricultural innovations support sustainability, resilience, and climate adaptation across diverse contexts and scales. Such frameworks must measure not only agronomic performance but also broader socio-economic and environmental outcomes, including trade-offs among productivity, resource-use efficiency, risk reduction, and equity. Several studies underline the importance of integrated approaches for evaluating agricultural systems. For example, Kanter et al. (2018) propose a framework for analysing trade-offs among agronomic, environmental, and socio-economic outcomes, while Inwood et al. (2018) emphasise agro-ecosystem approaches that utilise flexible, context-specific indicators. Other studies have developed multidimensional assessment frameworks that combine indicators of agricultural productivity, farmer income, climate resilience, and environmental sustainability (Zong et al., 2022). Reviews of resilience assessment tools further emphasise the challenge of developing operational metrics that capture the complexity of agricultural systems while remaining practical for decision-making (Douxchamps et al., 2017).

In response to this need, the Innovation in Sustainability, Policy, Adaptation, and Resilience in Kenya (iSPARK) project was established.[1]. Funded by the UK Foreign, Commonwealth and Development Office (FCDO) through UK International Development (UKAID), the project is co-implemented by the University of Leeds, the Alliance of Bioversity International and CIAT, the International Institute of Tropical Agriculture (IITA), and iShamba Limited. The project aims to generate evidence supporting sustainable and climate-resilient food and nutrition systems in Kenya. In particular, iSPARK seeks to connect farm-level innovations with evidence-based policy pathways and investment decisions that can support the large-scale transformation of agricultural systems. A key element of the project is the development of a metric-based framework called Innovation in Sustainability, Resilience, and Adaptation (ISRA). This framework enables systematic assessment of agricultural innovations using integrated indicators that capture agronomic, environmental, and socio-economic outcomes. The project also harnesses advances in geospatial data, Earth observation, modelling, and machine learning to develop scalable, spatially explicit metrics for monitoring agricultural systems across landscapes and over time. Combining spatial analytics with socio-economic evidence enables ISRA to provide decision-makers with robust, comparable indicators that inform advisory services, policy development, and investment strategies for climate-resilient agriculture.

[1] https://www.cabi.org/uk-cgiar-centre/confronting-climate-change-and-environmental-degradation-through-sustainable-crop-management-and-climate-smart-agronomic-practices/

2. Review of Existing Approaches and Indicators for Developing ISRA Metrics

Building on the earlier-highlighted need for structured and scalable evaluation frameworks, this chapter examines existing research and analytical tools relevant to the development of ISRA metrics. An expanding body of literature has proposed methods and indicators for evaluating sustainability, resilience, and climate adaptation in agricultural systems. These approaches offer vital insights that can guide the development of ISRA metrics within the iSPARK project.

Early work on ISRA metrics has focused on examining existing analytical methods and assessing their potential to develop indices of sustainability, resilience, and adaptation across spatial and temporal scales. The following sections provide an overview of key approaches and tools that can aid in developing ISRA metrics, including techniques based on remote sensing, drought monitoring, irrigation analysis, soil health evaluation, and socio-economic indicators.

2.1. Crop Classification

Mixed cropping systems have been widely shown to enhance soil health and agricultural productivity. For example, Hashakimana et al. (2023) found that mixed cropping systems significantly increase maize biomass and grain yields compared with monocropping systems. Their study concluded that mixed cropping practices can improve the sustainability and stability of farming systems and provide important benefits for smallholder farmers, particularly during drought conditions in Rwanda's Eastern Province, specifically in Kayonza District.

Monitoring the spatial distribution of such cropping systems requires reliable methods to distinguish between monoculture and intercropping practices. Crop classification, a remote sensing–based approach for identifying crop types and cropping patterns, provides an effective means of monitoring the spatial and temporal dynamics of agricultural systems. In particular, crop classification techniques can help differentiate monocropped fields from intercropped systems and track changes in cropping practices across agricultural landscapes.

Despite these advantages, mapping intercropping systems at field and regional scales remains challenging. Conventional agricultural surveys are often costly and typically provide limited spatial coverage, which constrains their effectiveness for large-scale monitoring. Advances in satellite-based Earth Observation (EO) technologies are helping to address these limitations. Satellite missions such as Sentinel-1, which provides synthetic aperture radar data, and Sentinel-2, which delivers optical imagery, provide consistent, high-resolution observations that enable scalable, cost-effective monitoring of cropping systems across large geographic areas.

Several studies demonstrate the potential of EO-based approaches for detecting intercropping systems. Mahlayeye et al. (2024) identify optimal crop growth stages, spectral bands, and Vegetation Indices that can effectively distinguish monocropped maize from intercropped maize systems. In their study, cropping patterns are classified using a Random Forest (RF) algorithm. Similarly, Ibrahim et al. (2021) developed a land cover map for the Jos Plateau in Nigeria to identify cropland areas and classify major crops such as maize and potato, as well as intercropping systems involving these crops. Their methodology integrates spectral-temporal metrics, Vegetation Indices, cloud-filtering techniques, and temporal feature engineering to detect maize and mixed-cropping systems. In another study, Liepa et al. (2024) generated a land-use and land-cover map for Kamuli District in eastern Uganda using harmonised Landsat and Sentinel-2 imagery. The authors integrate satellite observations with detailed in situ crop data to extract phenological metrics throughout the growing season and improve classification accuracy.

2.2 Yield estimation

Remote sensing–based classification of cropping systems plays a vital role in improving the accuracy of crop yield estimation, especially in SSA, where fragmented fields, diverse management practices, and widespread intercropping characterise smallholder farming systems. In such settings, satellite observations offer a valuable means of capturing spatial variability in crop distribution and productivity across large geographic areas. Jin et al. (2019) present an integrated framework that combines crop classification and yield estimation to create scalable maps of smallholder maize cultivation across western Kenya and Tanzania using Google Earth Engine. The study utilises 10-metre-resolution imagery from the Sentinel-1 and Sentinel-2 satellites to delineate cropland extents, identify maize-specific cultivated areas, and estimate yields. Crop classification employs Random Forest algorithms trained on seasonal composites derived from both radar and optical satellite data. The classification models achieve accuracies exceeding 85% in detecting cropland and between 63% and 79% in distinguishing maize fields. Notably, integrating Sentinel-1 synthetic aperture radar data significantly enhances classification performance in regions with persistent cloud cover, where optical imagery alone may be insufficient.

Yield estimation is conducted using the Scalable Crop Yield Mapper (SCYM) framework, which integrates satellite-derived Vegetation Indices with crop growth simulations to generate spatially explicit yield forecasts (Lobell et al., 2015). In this framework, regression models are trained to predict yields based on the Green Chlorophyll Vegetation Index (GCVI) derived from satellite imagery. Additionally, Jin et al. (2017) investigate the potential of various satellite sensors, including the commercial SkySat and RapidEye platforms and the publicly available Sentinel-2, to capture spatial variability in smallholder maize yields across approximately 40,000 km^2 in western Kenya. The study compares two distinct regression-based modelling strategies for yield prediction. The first, known as a calibrated model, utilises ground-measured yield data and weather observations to calibrate the regression model. The second, an uncalibrated approach, employs a process-based crop model to simulate daily Vegetation Indices and end-of-season biomass or yield, which then serve as pseudo-training data for the regression models.

2.3 Drought monitoring

Drought is a recurring hazard in SSA, and several indices have been reviewed to assess its spatial and temporal impacts. Musyimi et al. (2023) used the Sentinel-3 SLSTR sensor to derive Land Surface Temperature (LST), Fractional Vegetation Cover (FVC), and FVC anomalies for early drought warning in Lower Eastern Kenya. Barrett et al. (2020) showed that regional applications of the 3-month Vegetation Condition Index (VCI3M) in Kenya demonstrate that machine learning–based forecasting of drought 4–6 weeks in advance can achieve accuracies of 80–89%. Alkaraki and Hazaymeh (2023) developed the Comprehensive Agriculture Drought Condition Indicator (CADCI) to monitor agricultural drought in semi-arid environments. It was developed using spectral indices that represent drought conditions (for example, Vegetation Condition Index (VCI) and Temperature Condition Index (TCI)). The study employed Random Forest to identify the three spectral indices with the highest monthly short-term average relative importance for identifying agricultural drought in semi-arid environments.

2.4 Irrigation mapping

Detecting irrigated land in rainfed systems in SSA, where rainfall primarily dictates agriculture, requires high-temporal-resolution data and reliable methods to handle cloud cover. Vogels et al. (2019) employed object-based image analysis (OBIA) in the Horn of Africa and produced monthly Sentinel-2 Normalised Difference Vegetation Index (NDVI) composites to identify irrigated fields by comparing abnormal green-up with rainfall patterns. In addition to optical indices, microwave radar systems, particularly C-band Synthetic Aperture Radar (SAR) from Sentinel-1, are frequently used for soil moisture detection (Baghdadi et al., 2016; Gao et al., 2017), which can serve as a proxy for recent irrigation events, especially when analysed as anomalies or combined with phenology data. Attarzadeh et al. (2018) observed that SAR's ability to penetrate clouds and measure surface roughness and moisture renders it especially useful during peak cropping seasons in tropical regions. Hajj et al. (2017) verified SAR's ability to complement NDVI-based methods for irrigation mapping.

2.5 Water productivity and irrigation performance

Several indicators are commonly used to evaluate irrigation performance, including equity, adequacy, reliability, and Crop Water Productivity (CWP). Equity measures the uniformity of water distribution using the spatial coefficient of variation (CV) of evapotranspiration (ET) among different growers. Adequacy is assessed by comparing actual evapotranspiration with Potential Evapotranspiration (PET) through the ET/PET ratio, where lower values indicate water stress and potential yield reduction. Reliability evaluates the consistency of water supply across growers based on evapotranspiration patterns. At the same time, Crop Water Productivity (CWP) measures how efficiently water is utilised for crop production by relating crop yield to the amount of water consumed. Karimi et al. (2019) evaluated irrigation performance using equity, adequacy, and reliability indicators. They found that equity was relatively high among large-scale growers, whereas smallholders exhibited greater variability in water distribution. The adequacy indicator revealed critical water shortages in several regions, and reliability, assessed using relative ET, showed seasonal variations, with large-scale growers performing above target levels. Their analysis also indicated that Crop Water Productivity was highest in centre-pivot irrigation systems and estate farms, reaching 1.63 kg/m^3. In contrast, smallholder farmers and furrow irrigation systems exhibited significantly lower productivity levels. Similarly, Mekonnen et al. (2024) assessed the performance of a small-scale irrigation scheme using remote sensing and ground-truth data over two consecutive irrigation seasons in the Shimburit irrigation scheme in northwestern Ethiopia, generating comparable irrigation performance indicators.

2.6 Soil Organic Carbon (SOC)

Soil Organic Carbon (SOC) is a vital indicator of soil health, agricultural productivity, and climate mitigation potential. Recent research emphasises the increasing use of simulation models and remote sensing techniques to map and monitor SOC across various spatial and temporal scales. Mandal et al. (2022) reviewed recent advances in SOC assessment and noted that simulation models and optical remote sensing are becoming more effective tools for mapping soil carbon. Simulation models enable researchers to assess changes in SOC across different land-use and management scenarios. At the same time, optical remote sensing techniques provide spatial information to support SOC mapping at multiple scales and time periods. Soil organic matter models are also employed to understand SOC dynamics and to develop management strategies that optimise carbon inputs and outputs to promote SOC sequestration.

Remote sensing further improves the ability to capture spatial variability in SOC by providing reliable, high-resolution estimates of soil carbon stocks at scales ranging from local to regional. Radočaj et al. (2024) confirmed that SOC can be predicted using digital approaches such as digital soil mapping. Digital soil mapping integrates multiple data sources with statistical and machine-learning techniques to generate spatial predictions of soil properties. This approach can support environmental conservation by identifying and monitoring areas with high carbon sequestration potential, thereby contributing to climate change mitigation. In most applications, digital SOC mapping relies on three main groups of environmental covariates: topographic variables, climatic conditions, and spectral information derived from remote sensing, often complemented by field soil samples.

Applications of remote sensing and terrain covariates for SOC mapping are increasing in Sub-Saharan Africa. For example, Dahana et al. (2024) evaluated the potential of combining multi-temporal radar and optical remote sensing data to predict SOC in the Kaffrine region of Senegal. The study tested several scenarios using Sentinel-1 data, Sentinel-2 data, their combination, topographic features, and the integration of satellite data with terrain variables. Using 671 soil samples for model training and 281 for evaluation, the study compared three machine learning algorithms — Random Forest, XGBoost, and Support Vector Regression (SVR)—and found that the RF model consistently outperformed the other approaches across all scenarios. Similarly, Mponela et al. (2020) used MODIS data and topographic variables in Malawi to predict SOC and macronutrient levels with strong cross-validation performance. In their study, a Random Forest model was trained using soil sample measurements as response variables, including SOC and nutrient levels (Nitrogen, Phosphorus, and Potassium (NPK)), while environmental covariates served as predictors. The trained model was then applied to predict soil nutrient levels at unsampled locations across the study area.

Additional studies in Ethiopia, Kenya, and Nigeria have also applied advanced modelling techniques such as regression kriging, Random Forest, and Cubist models to estimate SOC distribution using environmental covariates, including Vegetation Indices (e.g., NDVI), topographic attributes, and climate variables (Leenaars et al., 2018; Hengl et al., 2015; Aynekulu et al., 2020). The Africa Soil Information Service (AfSIS) has played a key role in supporting these efforts by collecting harmonised soil datasets across multiple African countries, thereby enabling continental-scale SOC mapping. These initiatives have produced high-resolution SOC maps with spatial resolutions ranging from 250 m to 1 km, which are now used to support national soil health strategies, greenhouse gas inventories, and targeted soil fertility interventions. However, despite these advances, several challenges remain, including limited data availability in remote areas, insufficient ground-truth validation, and inconsistencies in sampling protocols. These limitations continue to affect the accuracy and reliability of SOC predictions across diverse agroecological zones in SSA.

2.7 Soil Fertility and Nutrients

Digital soil mapping for soil nutrients such as nitrogen (N), phosphorus (P), and potassium (K) is becoming increasingly feasible with the integration of machine learning methods and remote sensing data. Tunçay et al. (2021) investigated the spatial distribution of soil properties and examined changes in soil fertility under intensive cultivation using a combination of field data and remote sensing techniques. Their study introduced the Soil Fertility Index (SFI), designed to identify key soil parameters affecting crop yields in the arid ecology of central Anatolia. However, the efficiency classes defined by the index required validation. To address this limitation, the authors compared SFI classes with observed crop yield values and Vegetation Indices derived from satellite imagery. This comparison demonstrated the potential of combining remote sensing with ground-based measurements to monitor soil fertility. Similarly, Zhang et al. (2024) integrated Landsat 8 and Sentinel-2 satellite imagery using the Gram–Schmidt fusion algorithm and multiple linear regression to generate spatial maps of soil fertility classified into five SFI categories: very low, low, moderate, high, and very high.

In addition to remote sensing approaches, process-based models have also been widely used to assess soil fertility and nutrient dynamics. The Quantitative Evaluation of the Fertility of Tropical Soils (QUEFTS) model provides a robust framework for evaluating soil nutrient status and generating site-specific fertiliser recommendations. Developed by Janssen et al. (1990), the QUEFTS model extends beyond simple yield–fertiliser response relationships to capture the interactive effects of nitrogen, phosphorus, and potassium on crop productivity. Its effectiveness has been demonstrated in several empirical studies across Sub-Saharan Africa. For example, studies in Nigeria and Kenya show that the QUEFTS model can be successfully parameterised and validated to assess soil fertility and support balanced nutrient recommendations for maize intensification, with strong agreement between predicted and observed yields (Shehu et al., 2019).

Several additional empirical studies highlight the applicability of the QUEFTS model across diverse agroecological contexts. In northern Nigeria, nutrient omission trials conducted in the Sudan and Northern Guinea savanna zones demonstrate that QUEFTS can accurately predict maize yields, with coefficients of determination (R^2) of 0.69 and 0.75, respectively, and site-specific nutrient uptake ratios of approximately 6.1:1:7.9 (N:P: K) per tonne of grain. Similarly, on-farm trials conducted in Sidama, southern Ethiopia, during the 2019 cropping season show strong agreement between observed and QUEFTS-predicted maize yields (7.4 t/ha compared with 6.3 t/ha), with a Root Mean Square Error (RMSE) of 1.5 t/ha and a percent bias of 6.9% (Asfaw et al., 2024). In northern Ethiopia, the model was applied to barley production in the Alaje district, where QUEFTS-based fertilisation strategies yielded more than 4.7 t/ha and showed high agreement with observed field data (Tesfay et al., 2020). Furthermore, in northern Tanzania, the model was combined with participatory assessments involving smallholder farmers. The results reveal strong alignment between farmer-perceived soil fertility constraints and those identified by the model, particularly regarding potassium and phosphorus as limiting nutrients under different yield conditions (Kimaro et al., 2023).

2.8 Greenhouse Gas (GHG) Emissions Monitoring

Monitoring greenhouse gas (GHG) emissions using remote sensing is still in its early stages in SSA, although its importance for climate mitigation and sustainable agricultural management is expected to grow. Recent studies have begun examining GHG emissions from various cropping systems under smallholder farming conditions. For example, a study conducted in Upper Eastern Kenya measured soil GHG fluxes across several cropping systems, including maize, cereal–legume intercropping, coffee, tea, and vegetables, under on-farm conditions (Ezekiel Lemarpe Shaankua, 2024). Although the study did not utilise remote sensing data, it provides valuable empirical insights that could be further analysed using Earth observation (EO) datasets. The study had three main objectives: quantifying greenhouse gas fluxes (CH_4, CO_2, and N_2O) from selected cropping systems in Tharaka-Nithi County; assessing N_2O emissions from smallholder cropping systems, including the effects of organic and inorganic fertiliser use; and identifying environmental factors, climatic conditions, farm management practices, and soil properties that influence soil N_2O dynamics. The results indicated that intercropped fields produced lower N_2O emissions per unit of yield than monocropped systems, suggesting potential benefits for both agricultural productivity and environmental sustainability. The study also evaluated yield-scaled emissions (YSE) and N_2O emission factors (EF), which offer useful indicators for assessing the environmental efficiency of different cropping practices.

Beyond field-based measurements, several studies have investigated the use of remote sensing data to monitor greenhouse gas emissions. For example, Guo et al. (2012) developed a model based on satellite-derived environmental variables to estimate CO_2 concentrations worldwide. The model utilised temperature data from MOD11C3, vegetation cover from MOD13C2 and MOD15A2, and productivity data from MOD17A2, all derived from MODIS observations. This model, called the Temperature–Vegetation–Productivity (TVP) model, was validated using CO_2 concentration data obtained from the Thermal and Near-infrared Sensor for Carbon Observation (TANSO) aboard the Greenhouse Gases Observing Satellite (GOSAT). The satellite-based CO_2 measurements served as reference values to assess the accuracy of the TVP model.

Remote sensing techniques have also been utilised to estimate soil N_2O emissions at regional scales. Feudal et al. (2020) developed a methodology to monitor potential soil N_2O emissions using high-resolution optical satellite imagery. The approach enabled the creation of an inventory of potential N_2O emissions with annual updates without the need for extensive field measurements. The processing chain employed optical satellite imagery collected between 2006 and 2015 to estimate temporally consistent, spatially detailed emissions across agricultural landscapes. The study concentrated on a region in south-western France. It analysed seven major seasonal crops: wheat, barley, rapeseed, maize, sunflower, sorghum, and soybean. The first step involved land-use classification, achieving an Overall Accuracy (OA) greater than 0.8%. The second step estimated the annual potential N_2O emissions budget for the study area. The results indicated that annual potential N_2O emissions ranged from 97 to 113 tonnes, with an estimated relative error of less than 5.5%. Wheat, the dominant crop, contributed at least 48% of total annual emissions, while maize, cultivated on less than 8% of the area, was the second-largest source of emissions. A ten-year analysis of crop rotations revealed that the common wheat-sunflower rotation produced potential emissions of approximately 16 kg

N_2O per hectare. In contrast, monoculture maize systems generated the highest emissions, reaching around 28.9 kg N_2O per hectare.

2.9 Socioeconomic indicators

Socioeconomic indicators are often employed in agricultural research to evaluate the human and economic aspects of farming systems and to understand how innovations impact rural livelihoods. These indicators assess how the adoption of agricultural practices affects household well-being, economic sustainability, and equity in smallholder farming systems (Wohlenberg et al., 2022). Empirical evidence from impact assessments shows that household income and food security are crucial indicators of the success of agricultural innovations, as they reflect improvements in livelihoods and resilience. For instance, a review of 73 agricultural intervention evaluations found that 38 studies, accounting for 52% of the sample, used direct food security indicators to gauge impacts (Bizikova et al., 2020). Common measures include net household earnings from crop and livestock production, gross margins, and asset ownership, as well as food security indicators such as the Household Dietary Diversity Score (HDDS) (Antle et al., 2015).

In addition to income and food security indicators, other socioeconomic measures are increasingly integrated into agricultural assessments. These include labour productivity, gender equity, and access to agricultural services. Labour productivity is often measured as output per unit of labour or land, such as crop yield per hectare or the value of production per person day. It is used to assess the efficiency of farming systems and the economic returns from agricultural technologies. However, previous research highlights an ongoing lack of widely applicable indicators for monitoring the sustainability of agricultural productivity (Dorward et al., 2013). Gender equity indicators are also increasingly used to examine how agricultural innovations affect the distribution of resources and decision-making within households. For example, the Women Empowerment in Agriculture Index (WEAI) is commonly utilised to measure domains such as production decisions, access to resources, control over income, leadership, and time use (Quisumbing et al., 2023). Additionally, indicators of access to credit, extension services, insurance, and markets are frequently used to assess the enabling environment that supports the adoption of agricultural innovations. When combined with biophysical and spatial indicators, these socioeconomic measures offer a more comprehensive understanding of how agricultural innovations influence productivity, livelihoods, and resilience in smallholder farming systems.

3. Developing ISRA metrics

The first step in developing ISRA metrics was to compile a long list of agricultural interventions currently used by farmers in western Kenya. This was done through a review of the existing literature and several databases documenting agricultural technologies and practices. These sources included the Kenya Agricultural and Livestock Research Organisation Technology Inventory and Management Platform (KALRO TIMP), the Africa Adaptation Atlas, the Consultative Group on International Agricultural Research (CGIAR) database, and CAROB. The purpose of this review was to identify a broad range of innovations relevant to smallholder farmers in western Kenya. Following the literature and database review, a stakeholder consultation workshop was held in Nairobi in July 2024 to gather expert input on the development of ISRA metrics and to validate the long list of interventions. The workshop included researchers, policy experts, and development practitioners working in agriculture and climate adaptation. A detailed description of this process is provided by Wanjau et al. (2025).

The second step involved identifying and reviewing existing methods for measuring the outcomes of the listed interventions. These methods served as the foundation for developing ISRA metrics. Three criteria guided the review. First, the metrics needed to be derived from remote sensing observations (e.g., crop classification, yield estimation, or SOC mapping) or from modelling approaches (e.g., climate-smartness indices). Second, the metrics had to be relevant to the intervention being assessed. Third, they needed to be linked to Climate-Smart Agriculture (CSA) outcomes, meaning they should help evaluate sustainability, resilience, or adaptation. These criteria ensured that the proposed metrics were both scientifically measurable and pertinent to climate-related agricultural outcomes.

To operationalise this process, each intervention on the long list was systematically reviewed and associated with one or more potential ISRA metrics. For each intervention, relevant metrics were identified based on the three criteria described above. A combination of literature review, expert judgement, and insights from previous experience in related research projects guided this selection process. For each proposed metric, additional information was documented, including the justification for its selection, the anticipated confidence level, and the estimated time needed to obtain or generate the data. This structured approach offered a transparent basis for comparing metrics across interventions.

The compiled information also formed the basis for developing the shortlist of ISRA metrics. To do this, the analysis was approached from a different perspective. Instead of starting with interventions and assigning metrics, all identified metrics were first listed, and the number of interventions associated with each metric was counted. This frequency count served as the initial relevance score for each metric. The results indicated that some metrics, such as yield estimation and soil fertility monitoring, were relevant to more interventions. However, several metrics had similar frequency scores. Therefore, additional criteria were used to refine the selection process. Each metric was evaluated based on the number of interventions it supported, its data requirements and availability, its potential for scaling, its risk relative to expected benefits, and overall confidence in the metric. These aided in identifying the most feasible and relevant ISRA metrics for monitoring agricultural innovations in the iSPARK project.

Table 1. Summary of the ISRA metrics, data requirements and relevance for ISRA dimensions (S, R, or A).

Measurement=ISRA metrics	**How many interventions does this metric measure from?**	**Is this a metric of sustainability, resilience or adaptation**	**Required data.**	**Scaling potential Field County Country**	**Risk vs reward. Confidence, time Data**
Yield estimation/Yield prediction/ Biomass estimation	6	Resilience And can be considered as an adaptation	Ground truth data to perform crop classification and obtain a crop type map GPS or Coordinates of those fields Measured yield in the field official yield data measured Leaf Area Index (LAI), measured biomass Weather data (daily max temperature, daily min temperature, precipitation, solar radiation) The cultivar of the crop Soil data: soil profile (texture, rooting depth, organic carbon content) Rang of local sowing dates Soil moisture, fertiliser rates, and seeding rates. Remote sensing data	Field: The data are acquired at the field level County: medium Country: low	Medium
Crop classification	3	Adaptation and sustainability	Ground truth data about the cropping system or crops that we want to classify GPS coordinates of the fields where the collection of information was performed Delimitation or delineation of the fields of the obtained ground truth data. Remote sensing data.	Field: The data are acquired at the field level County: high Country: medium	Medium
Drought detection/ Vegetation Condition Index	3	Resilience/adaptation	The in-situ based drought monitoring indices are calculated using data from either agro-climatic stations that measure climatic variables (humidity, precipitation, temperature, atmospheric pressure) Estimated or measured soil moisture at the	Field: The data are acquired at the field level County: medium Country: medium	Medium

			regional or field scale obtained in situ or by remote sensing data Measured or estimated yield or biomass Microwave and optical Remote sensing data. Depends on the used methodology Possible requirements: Actual Evapotranspiration/ Potential Evapotranspiration		
Irrigation detection/ irrigated areas	2	Adaptation resilience sustainability	Ground truth data of fields that have an irrigation system or use an irrigation technique/ If possible, the total number of irrigations during a season GPS coordinates of the fields for which we have information about their irrigation system. Delimitation or delineation of the fields Microwave and optical Remote sensing data.	Field: The data are acquired at the field level County: low Country: low	Medium
Irrigation performance indicators and water productivity	1	Adaptation/resilience/sustainability	Delimitation or delineation of the fields In-situ data includes discharge, crop yields, and field sizes collected via GPS Estimated or measured actual Evapotranspiration/ Potential Evapotranspiration The amount of water used for irrigation Remote sensing data includes: • Surface reflectance and land surface temperature. • Shuttle Radar Topography Mission (SRTM) for digital elevation data. • Meteorological data. • Precipitation data.	Field: The data are acquired at the field level County: low Country: low	Medium
Soil Organic Carbon modelling & mapping/ Soil	3	Sustainability/Resilience	Coordinates of the fields Delimitation or delineation of the fields	Field: The data are acquired	Medium

Organic Carbon forecasting			Soil data: soil sample, Soil Organic Carbon SOC and topographic data Terrain attributes: Digital Elevation Models (DEMs) Climate data: Temperature and precipitation datasets Remote sensing data.	at the field level County: low Country: low	
Monitoring NPK availability/ Monitoring soil fertility	4	Adaptation/resilience/sustainability	Coordinates of the fields Delineation of the fields Field samples: soil particle size distribution by the hydrometer method, pH and electrical conductivity (EC), total nitrogen (Ntotal), available phosphorus (Pavb) d, exchangeable potassium (Kexc), calcium (Caexc), sodium (Naexc) and magnesium (Mgexc). Other soil properties can be measured. Data or information about Yield Remote sensing: optical remote sensing data in particular.	Field: The data are acquired at the field level County: low Country: low	Medium
Soil moisture analytics	1	Adaptation/resilience	At the field scale Coordinates of the fields from where we collected samples Delineation of the fields Meteorological observations Soil parameters: soil texture/ soil hydraulic properties For the retrieval models: roughness parameters, moisture data and canopy data Microwave and optical Remote sensing data At the regional/global scale Selection between: Soil Moisture and Ocean Salinity (SMOS) mission Soil Moisture Active Passive (SMAP) mission The Advanced Microwave Scanning Radiometer 2 (AMSR2) The European Space Agency (ESA Soil) Moisture Climate Change Initiative (CCI)	Field: The data are acquired at the field level County: medium Country: low	Medium

greenhouse gas (GHG) emissions	Soil Organic Carbon modelling can be complementary work, and the idea came from the Climate Smartness Index (CSI), based on Greenhouse Gas Intensity and Water Productivity.	Adaptation/resilience	*In-situ* Chamber Systems*:* Greenhouse gases can be sampled using static chambers; for example, gases can be analysed by gas chromatography (GC) fitted with a 63Ni electron-capture detector (ECD) for N2O and a flame ionisation detector (FID) for CH4 and CO2, with N2 as the carrier gas. Flux tower: eddy covariance method Remote sensing: MOD17A2 GPP/NPP, GOSAT TANSO, TROPOMI(Sentinel5) Where: GPP is the Gross Primary Productivity NPP is Net Primary Productivity	Field: medium County: medium Country: low	Medium

To identify the three main innovations for iSPARK, we analysed Table 1 in light of iSPARK baseline results and literature review, summarised in the Appendices. The reasoning is outlined in the following two paragraphs.

Crop classification was considered a useful ISRA metric because it can be linked to intercropping practice. According to the results from baseline farmer Focus Group Discussions (FGDs), we noticed that: (*i*) certified seed was the most mentioned practice (mentioned 38 times), but unfortunately, we could not easily link it to a specific ISRA metric based on remote sensing or modelling. For now, it is difficult to monitor seed quality or certification status using EO data. (ii) Intercropping was one of the top practices mentioned by farmers (mentioned 32 times). This helped us understand that crop classification could be used to assess intercropping from a remote sensing perspective. To support this idea, we are considering combining it with two other related metrics: drought detection, to create a full ISRA metrics group focused on "intercropping for resilience."

The same logic was used for the other innovations. Irrigation was mentioned 33 times in the focus groups, so we grouped the ISRA metrics that can help monitor irrigation practices into three categories: irrigation detection, irrigation performance, and soil moisture. For soil health assessment, we grouped metrics related to soil management, including Soil Organic Carbon, soil fertility, and GHG emissions. These also relate directly to the use of organic manure and fertiliser, which were among the top practices mentioned by farmers in the baseline FGD (33 and 32 times, respectively).

In addition to these three selected ISRA innovations, yield estimation can be considered as a cross-cutting metric. Yield is a very useful metric and can support many different types of interventions, which is why it had a high score in Table 1. It can help assess both resilience and adaptation outcomes. However, we decided not to focus on yield as the main innovation because it requires substantial ground-truth data. To estimate yield accurately, we need detailed field-level information such as measured yields, crop types, biomass, and other variables, which are not always available or easy to collect. Therefore, while yield remains important and can complement other ISRA metrics, we focused on intercropping, irrigation, and soil health assessment as our main areas of innovation, which can be monitored more directly using available remote sensing and modelling tools.

The above reasoning, along with Table 1, formed the basis for grouping related ISRA metrics towards a final choice of three focal innovations for iSPARK. Through this grouping, all potential ISRA metrics are presented in Table 1 were identified as relevant to at least one innovation area. The three innovation areas themselves emerged from a stakeholder-driven prioritisation process, reflecting the innovations of greatest interest and relevance to farmers. The results of this grouping are presented in Table 2.

Table 2. ISRA metrics grouped according to the three innovation areas identified by the analysis.

Innovation area	ISRA metrics	Total
Intercropping	Crop classification (3) + Drought detection (3)	6
Irrigation	Irrigation detection (2) + Irrigation performance (1) + Soil moisture (1)	4
Soil health assessment	SOC modelling (3) + Soil fertility monitoring (4)	7
Yield (cross-cutting metric, used in all three innovations)	Yield estimation/prediction/biomass estimation (6)	6

In conclusion, we selected these three ISRA innovations: intercropping for resilience, Targeted Climate-Smart Irrigation Advisory, and Climate-smart soil nutrient management for resilience, based on a combination of scoring from the intervention database and strong alignment with focus groups. This ensures that the selected metrics are not only technically possible to monitor but also meaningful for the local context.

The rest of this section is devoted to listing each iSPARK ISRA and presenting the specific methods employed for iSPARK. The resulting preliminary findings from applying the ISRAs to testing and/or target regions for iSPARK are presented in Section 4.

3.1 Operationalising Crop Classification for ISRA Metrics

This section details the methodological approaches employed to operationalise the crop classification as an ISRA metric within the iSPARK project. The methods aim to distinguish between maize monocropping and maize legume intercropping systems in Western Kenya using Earth observation data and machine learning techniques. Preliminary findings from applying these ISRA metrics in the study area are presented in Section 4.

Transfer learning has become a vital strategy in remote sensing and EO applications for tackling the challenge of limited labelled training data. It allows models trained on large datasets, such as extensive EO archives, to be adapted for new tasks where labelled examples are scarce (Risojevic & Stojnic, 2022). Recent research emphasises the growing importance of transfer learning in EO-based analysis, showing that it can support high-performing models even with limited ground-truth data (Ma et al., 2024). This method is becoming more common in applications such as land cover classification, crop type mapping, and environmental monitoring because it addresses ongoing data limitations. A major advantage of transfer learning is its ability to generalise across different geographic regions, making it especially suitable for satellite-based agricultural monitoring.

In crop mapping applications, transfer learning enables researchers to overcome data scarcity in regions with limited ground-truth observations. Pan & Yang (2010) emphasise that transfer learning facilitates the transfer of knowledge from a source domain with abundant labelled data to a target domain with limited labelled data. Several studies demonstrate the effectiveness of this approach in crop classification. For example, Koukos et al. (2024) applied transfer learning to crop mapping by pre-training a deep learning model on Sentinel-1 VH time-series data to detect paddy rice in South Korea. The model was subsequently fine-tuned for paddy rice detection in France and Spain and for barley detection in the Netherlands, showing strong performance when transferring models across regions. Similarly, Chang et al. (2024) evaluated the generalisability of foundation models pre-trained on multispectral satellite imagery such as Sentinel-2 for crop-type mapping across different geographic regions. Their results indicate that models pre-trained on EO-specific datasets, such as SSL4EO-S12, outperform those trained on general datasets, such as ImageNet, particularly when fine-tuned with limited local training samples.

The Pretrained Remote Sensing Transformer (Presto) model, developed under the WorldCereal project by the European Space Agency (ESA), represents a recent application of transfer learning for global crop-type mapping. Presto is a transformer-based model pretrained on satellite time-series data and designed to capture temporal patterns in Earth observation imagery (Tseng et al., 2023). By analysing sequences of satellite observations over time, the model learns relationships among land cover, weather conditions, and crop phenology, even in regions with limited ground-truth data. Presto uses a neural encoder–decoder architecture that integrates multi-temporal Sentinel-1 and Sentinel-2 data with additional environmental datasets. Its architecture supports fine-tuning with local ground-truth data, making it particularly useful in regions with limited training data while maintaining adaptability across agricultural systems.

In this study, the open-source Presto WorldCereal classification pipeline was evaluated for generating a customised crop map that distinguishes between maize monocropping and maize intercropping systems in Western Kenya during the long rain season. In addition to the Presto-based approach, a second classification strategy was implemented to refine crop system mapping. This alternative method focused on areas already identified as maize using the WorldCereal tc-maize-main mask generated for 2021. Within these areas, NDVI time-series data derived from Sentinel-2 imagery were used to train a Random Forest classifier to differentiate between monocropped maize and maize–bean intercropping systems.

3.1.1 Presto-based classification method

Presto is a transformer-based model designed for Earth observation time-series analysis (Tseng et al., 2023). Within the ESA WorldCereal crop mapping pipeline, Presto serves as a geospatial foundation model, generating robust representations of satellite time-series data. The WorldCereal system is developed as a Python-based framework that produces cropland and crop-type maps at various spatial scales using satellite imagery and auxiliary datasets. The system functions via the openEO platform and generally utilises the Copernicus Data Space Ecosystem (CDSE) for cloud-based processing.

WorldCereal combines globally pre-trained features generated by the Presto model with a lightweight CatBoost classifier to produce crop classification outputs. Because Presto is trained using self-supervised learning on large multimodal satellite time-series datasets, it can capture general spatiotemporal patterns in agricultural landscapes, enabling improved classification performance and transferability across regions.

The Presto architecture is specifically designed for pixel-level analysis of time-series data. It processes sequences of observations from multiple sensors over time, enabling the model to capture seasonal patterns and crop phenology. During pretraining, Presto uses a masked autoencoding strategy in which parts of the satellite data are intentionally masked, and the model learns to reconstruct the missing information. This approach helps the model understand relationships among temporal observations and sensor types without depending on labelled training data.

The pretraining dataset used to develop Presto included approximately 21.5 million pixel samples at 10 m spatial resolution covering a full 12-month period between 2020 and 2021. These samples were selected to represent diverse climates, land-cover types, and agricultural systems globally. Through this large-scale pretraining process, the model learns general spatiotemporal patterns such as crop phenological cycles and seasonal vegetation dynamics.

To apply the Presto-based classification workflow, ground-truth data describing maize monocrop and intercropped maize fields were first harmonised using the WorldCereal Reference Data Module (RDM)[2]. Field observations collected by GEOGLAM were cleaned to remove unrelated records and observations outside the study area before being uploaded to the RDM interface. The harmonisation process aligned spatial and temporal attributes and exported the dataset in GeoParquet format following the WorldCereal schema, which served as the input dataset for subsequent processing.

Earth observation time-series data were then obtained for each ground-truth location using the WorldCereal extraction framework. The extraction process retrieves spatially and temporally aligned datasets from the Copernicus Data Space Ecosystem via the openEO platform. The datasets included Sentinel-2 optical imagery, Sentinel-1 SAR backscatter data, AgERA5 climate variables, and terrain attributes derived from the Copernicus Digital Elevation Model. The extraction window was set to nine months before and after the observation date to ensure complete seasonal coverage.

[2] https://rdm.esa-worldcereal.org/create

After data extraction, the WorldCereal classification pipeline was customised to differentiate between monocropped maize and intercropped maize classes based on field observations from Western Kenya. Once the area of interest and the seasonal window for the long rain season (March–September) were defined, crop class labels were standardised into two categories: monocrop and intercrop maize. The Presto encoder then produced 128-dimensional feature embeddings representing the multi-source satellite time-series data for each ground-truth site. These embeddings were utilised to train a CatBoost classifier, which generated the final crop classification model and facilitated the creation of spatial crop maps for the study area.

3.1.2 Maize-mask NDVI classification method

A second classification method was used to improve crop system mapping, focusing specifically on areas already identified as maize using the WorldCereal tc-maize-main mask. The goal of this approach was to distinguish between monocropped maize fields and maize-bean intercropping systems during the long rainy season (April-August).

NDVI time-series data were initially derived from Sentinel-2 imagery for the 2021 season. Ground-truth observations obtained from GEOGLAM were used to train a Random Forest classifier to differentiate between the temporal NDVI signatures of monocrop and intercropped maize fields.

The analysis began with data acquisition and preprocessing. The study area was defined based on the spatial distribution of available ground-truth observations. A shapefile containing labelled monocrop and intercropped maize fields was used to extract NDVI time series data from Sentinel-2 imagery via Google Earth Engine. Missing values in the NDVI series were interpolated to ensure complete temporal profiles for each observation. The dataset was then divided into training (80%) and validation (20%) subsets.

Meanwhile, the WorldCereal tc-maize-main mask was utilised to identify pixels classified as maize within the study area. Sentinel-2 NDVI imagery for the same growing season was generated to enable pixel-level classification.

The Random Forest classifier was trained on the labelled NDVI time series and evaluated using standard classification metrics, including Overall Accuracy, confusion matrix results, precision, recall, and F1-scores. After validation, the trained model was applied to the NDVI image stack for the entire study area. Classification was limited to pixels identified as maize in the tc-maize-main mask. Each pixel was then classified as either monocrop maize or maize–bean intercropping based on its NDVI temporal profile, producing a spatial map of cropping systems across the study region.

3.2 Drought indices

Drought monitoring was conducted using satellite-derived Vegetation Indices to assess the effects of drought on maize monocrop and maize–bean intercropping systems. The methodology was implemented in two phases and at two spatial scales: field level and landscape (zonal) level.

In the first phase, field-level analysis was performed using data from 15 georeferenced plots collected during the baseline survey (Wanjau et al., 2025). For each plot, the Vegetation Condition Index (VCI) was calculated using two remote sensing products: MODIS and Sentinel-2. Comparing VCI derived from these datasets enabled an assessment of how spatial resolution influences the detection of drought stress in smallholder agricultural systems, particularly in highly fragmented landscapes typical of Western Kenya. VCI time series were analysed over four years to evaluate seasonal dynamics and interannual variability in vegetation conditions. In addition, the Vegetation Health Index (VHI) was calculated for the same locations to provide an additional measure of vegetation stress.

In the second phase, the analysis was scaled up to the landscape level using the crop classification map developed in Section 3.1, which distinguishes between maize monocrop and maize bean intercropping systems. A pixel-based classification approach was applied to the VCI images derived from the maize mask classification method. Each pixel was assigned to one of four drought severity classes based on established VCI thresholds:

- Extreme drought (VCI < 10)
- Severe drought (10 ≤ VCI < 20)
- Moderate drought (20 ≤ VCI < 50)
- Healthy vegetation (VCI ≥ 50)

The VCI images were subsequently masked by crop type to isolate maize monocrop and intercropped maize areas. For each crop class, the number of pixels in each drought category was calculated. This process generated spatial drought severity maps and summary visualisations showing the distribution of drought classes during the long rain season. The resulting outputs enabled a comparative analysis of drought responses between monocropped maize systems and maize-bean intercropping systems, offering insights into the potential role of intercropping in enhancing resilience to drought stress.

3.3 Operationalising Socio-economic Indicators for ISRA Metrics

Socio-economic indicators were developed to evaluate the human, economic, and social outcomes associated with the selected innovations: intercropping, soil fertility management, and irrigation. These indicators were identified through a combination of Participatory Multi-Criteria Mapping (MCM) exercises with smallholder farmers and a review of relevant literature. In this context, indicators are measurable measures used to assess the performance and impacts of agricultural innovations. At the same time, metrics are the quantitative variables used to compute these indicators from household data. Metrics, their definitions, measurement methods, and data sources were chosen and adapted from the Sustainable Intensification Assessment Methods Manual (Musumba et al., 2017). A total of six interconnected metrics were identified to mirror the socio-economic priorities expressed during the MCM exercises with small-scale farmers regarding intercropping practices. These metrics reflect farmers' priorities related to income generation, food security, labour demands, and gender dynamics. The link between the priorities identified through the MCM exercises and the selected socio-economic metrics is summarised in Table 3.

The economic indicator focuses on the profitability and financial performance of agricultural innovations. This dimension was operationalised using two metrics derived from household survey data and complementary price information. The first metric, net income, was calculated by considering both farm revenues and production costs. Production costs included variable and fixed costs, such as labour, fertilisers, seeds, pesticides, livestock feed, veterinary services, transportation, and depreciation of agricultural equipment. Additional fixed costs included farm infrastructure costs such as housing, land rent, and capital investments. Data on input quantities, prices, and output quantities were collected through household surveys and complemented with market and e-commerce price data. Net income was estimated across multiple agricultural seasons and farms to capture variations in production outcomes. The second metric, variability of net income, was used to measure the stability of economic returns associated with agricultural innovations. This metric was calculated using the coefficient of variation (CV), defined as the standard deviation of profitability divided by the mean profitability. The CV indicates income volatility across seasons and allows assessment of how climate variability or other production risks influence the stability of farm income.

Table 3. Socio-economic indicators and associated metrics are used to evaluate agricultural innovations.

FGD priority	Metric	Definition	Data / Methods
Input Resource Costs and Income	Net income	The net income metric is calculated by considering both revenues and production costs, including variable and fixed (overhead) costs. Production costs include labour, fertilisers, seeds, feed, veterinary services, transportation, and depreciation of equipment.	Household survey and e-commerce price data[3]. Survey data include: (1) quantity of inputs used in production (fertiliser, pesticides, labour, feed); (2) price of each input; (3) other costs (veterinary services, transport, vaccination); (4) quantity and price of outputs (e.g., milk or crop production); (5) fixed costs such as housing, land rent, beginning stock, and depreciation. Net income is measured across multiple seasons and farms.
	Net income variability	Profitability variability is measured using the coefficient of variation (CV). Climate impacts on production may contribute to variation in profitability across years.	CV is calculated as the standard deviation of profitability divided by the mean profitability.
Food Security, Nutrition and Health	Food and nutrient availability	Measures the amount of food produced due to the intervention and the micronutrient content of the output. Production is converted into calories and micronutrients per unit area.	Household survey data and food/nutrient composition tables: (1) area used to produce food item; (2) total production from the area; (3) food composition tables showing calorie values (Musumba et al., 2017); (4) calculation of micronutrients produced per hectare using FAO nutritional conversion values (Musumba et al., 2017).
	Food access	Examines farmers' perceptions of the crop's ability to provide sufficient food year-round.	Participatory rating exercises with farmers comparing new technologies with conventional practices based on their ability to provide sufficient food for household consumption.
Labour and Gender	Labour requirements	Measures the quantity of time spent on agricultural activities related to the innovation.	(1) Survey data on the number of hours or days spent producing a crop; (2) direct observation of farmers performing agricultural activities; (3) participatory labour calendars to evaluate labour demand relative to labour supply.
	Time allocation by gender	Measures gender equity by quantifying differences in time spent on various tasks.	(1) Daily time-use exercise (CARE Daily Time Use method); (2) 24-hour time allocation recall surveys; (3) activity analysis.

[3] https://dataviz.vam.wfp.org/economic/prices?current_page=1&theme=10

The human indicator captures household welfare outcomes associated with agricultural innovations, particularly food security, nutrition, and household adaptive capacity. Two metrics were used to operationalise this indicator. The first metric, food and nutrient production (availability), measures the quantity of food produced by the household and the extent to which it contributes to household nutrition. Data on cultivated area, crop yields, and total food production were collected through household surveys. These data were then combined with food composition tables to convert crop production into per capita availability of calories and micronutrients. Nutrient estimates were calculated by converting production quantities to grams of micronutrients per hectare using standard nutrient conversion tables (Musumba et al., 2017). This approach enables assessment of whether innovations such as intercropping increase household-level dietary diversity and nutrient availability. The second metric, food access, measures households' ability to obtain sufficient food throughout the year. This metric was assessed using participatory evaluation approaches in which farmers rated the performance of new agricultural technologies relative to conventional practices in terms of their ability to ensure adequate food supply for household consumption. Household expenditure data complemented these participatory assessments to assess whether household incomes support the acquisition of a minimum-cost nutritious diet.

The social indicator captures the social implications of agricultural innovations, particularly labour requirements and gender equity. Two metrics were used to measure this dimension. The first metric, labour requirement, quantifies the time allocated to agricultural activities associated with different innovations. Data were collected through household surveys that recorded the total number of hours or days spent on crop production activities. These survey responses were complemented by direct observations of farmers engaged in agricultural activities to estimate the duration and labour intensity of specific tasks. In addition, participatory evaluation exercises using labour calendars were conducted with farmers to capture seasonal labour requirements and compare labour demand with labour availability within households. The second metric, time allocation by gender, was used to assess gender equity in labour distribution. This metric was measured using daily time-use surveys, including 24-hour recall exercises and activity analysis methods, to quantify the time men and women spent on agricultural and household tasks. These data allowed the identification of differences in labour burdens, including time spent on physically demanding activities and the availability of leisure time, thereby providing insights into how agricultural innovations influence gender dynamics within farming households.

The various ISRA metrics described in Sections 3.1–3.3 collectively support the evaluation of the three selected innovation areas within the iSPARK framework: intercropping for resilience, climate-smart irrigation management, and climate-smart soil nutrient management. These metrics combine spatial analysis from EO data with socio-economic indicators gathered from household and participatory data sources. Table 4 summarises the relationship between the innovations and the corresponding ISRA metrics used in the analysis.

Table 4. Overview of the relationship of ISRAs to innovations.

Innovation	ISRA metrics	Contribution to the innovation
Intercropping for resilience	Crop classification	Monitor the cropping system, Detect diversification: spatial distribution of cropping systems across the study area. Provide important information for yield estimation.
	Drought detection	Which intercropping systems demonstrate greater resilience to drought
Irrigation	Drought Detection	Which areas (pixels) are at higher risk of drought, defined by low precipitation, high temperature, low NDVI and low soil moisture?
	Targeted Climate-Smart Irrigation Advisory (T-CSIA)	Targeted advisories for farmers in high drought risk areas.
Climate-smart soil nutrient management for resilience	Nutrient gap estimation	Spatially resolved nutrient gap (NPK) in the soil provides information for targeting.
	Spatial and temporal resolved fertiliser rate and type determination	Which fertiliser types and rates are improving productivity and profitability, and reducing environmental footprints
	Nutrient use efficiency	Identify where nutrients are used efficiently and Support optimal rate targeting and management options that increase efficiency under drought/heat stress. Guide interventions that reduce nutrient losses and improve resilience.
	Profitability	Identify high-return, low-risk fertiliser options under climate variability. Support decision-making for scaling recommendations and prioritising farmer advisories.

4. Innovations Informed by ISRA Metrics

This section presents the results from applying the ISRA metrics framework to analyse cropping systems, drought patterns, and nutrient management conditions in the study area. The findings offer evidence for developing targeted climate-smart innovations to improve resilience and resource-use efficiency in smallholder farming systems. The following subsections present the analytical results for each innovation area.

4.1Targeted intercropping for resilience

This subsection presents the results of the analysis assessing the potential of intercropping systems to improve resilience in smallholder farming systems. The analysis combines crop classification outputs with satellite-based drought monitoring indicators to examine how maize monocrop and maize-bean intercropping systems respond to drought stress during the long rainy season. Crop classification results were used to map the spatial distribution of cropping systems across the study area. In contrast, vegetation-based drought indices were used to compare vegetation conditions under different drought scenarios.

4.1.1 Presto-based crop classification results

The performance of the Presto classification model was evaluated using ground-truth samples from 2021 and 2023. Classification accuracy was assessed under three training configurations: year-specific training, cross-year training, and combined multi-year training. These scenarios enabled evaluation of classification performance in individual years and assessment of the model’s ability to generalise across years.

For 2021, the year-specific training scenario achieved an Overall Accuracy (OA) of 0.68 (Table 5), with relatively balanced class performance (F1-score = 0.63 for intercrop and 0.71 for monocrop). When the model was trained using 2023 samples and applied to 2021 data (cross-year scenario), the OA increased slightly to 0.71 (Table 5). However, the resulting classification map was dominated by monocrop predictions (Figure 1) and recall for the intercrop class declined substantially to 0.38. This indicates that the apparent improvement in Overall Accuracy masked poorer detection of intercropped fields.

Table 5. Classification performance for 2021 under different training scenarios.

Scenario	Class	Precision	Recall	F1-score	Overall Accuracy
Year-specific training	Intercrop	0.65	0.62	0.63	
	Monocrop	0.70	0.73	0.71	0.68
Cross-year training	Intercrop	0.41	0.38	0.40	
	Monocrop	0.80	0.82	0.81	0.71
Combined multi-year training	Intercrop	0.56	0.46	0.51	
	Monocrop	0.73	0.80	0.77	0.68

The combined multi-year training scenario, which used ground-truth samples from both years, achieved an OA of 0.68 (Table 5), similar to that of the year-specific scenario. However, the spatial pattern of predicted intercropped fields appeared less realistic (Figure 1), and the F1-score for the intercrop class declined slightly. This suggests that increasing the training dataset size did not necessarily improve the model's ability to detect intercropping patterns.

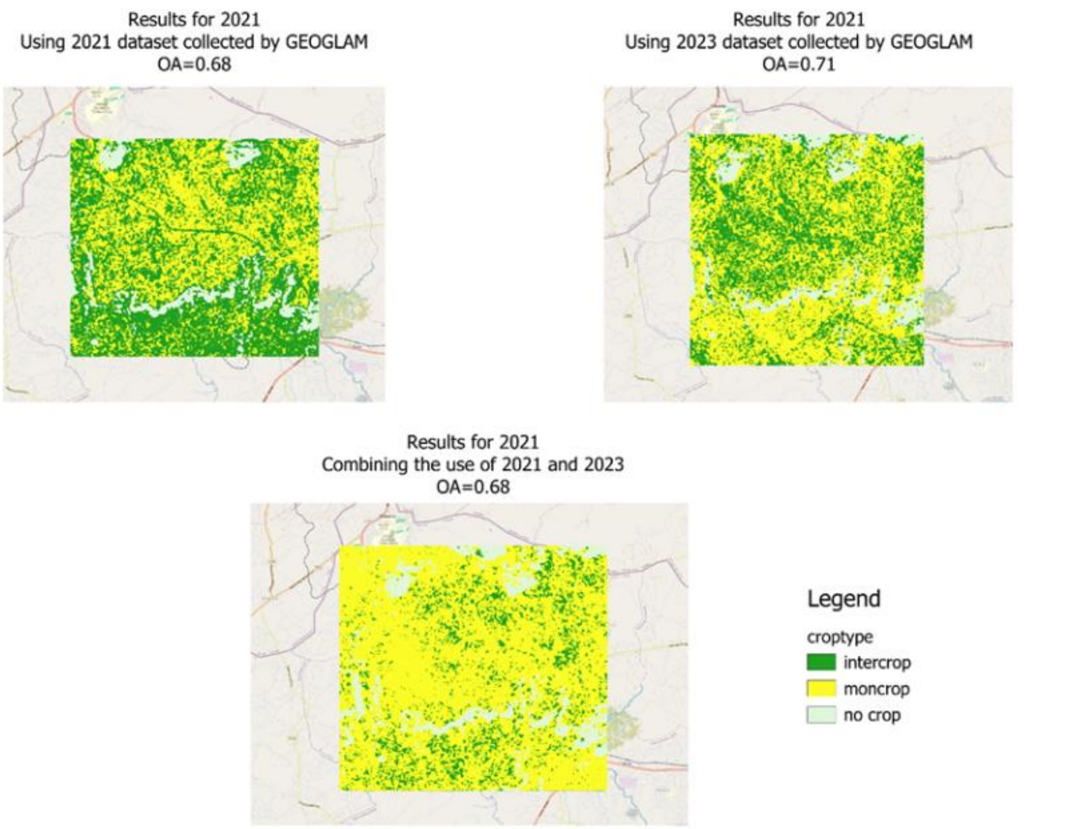

Figure 1. Presto-World Cereal classification results for 2021.

For 2023, both the year-specific training scenario (trained and applied to 2023 data) and the cross-year scenario (trained on 2021 and applied to 2023 data) achieved the same Overall Accuracy of 0.69 (Table 6). Despite this similarity in numerical accuracy, their spatial predictions differed noticeably (Figure 2). The combined multi-year training scenario produced a slightly lower OA of 0.67 (Table 6) and tended to underestimate intercropped areas.

Table 6. Classification performance for 2023 and the maize-mask scenario.

Scenario	Class	Precision	Recall	F1-score	Overall Accuracy
Year-specific training (2023)	Intercrop	0.65	0.66	0.66	
	Monocrop	0.72	0.71	0.71	0.69
Cross-year training (train 2021 → test 2023)	Intercrop	0.37	0.34	0.36	
	Monocrop	0.79	0.80	0.80	0.69
	Intercrop	0.54	0.47	0.50	

Combined multi-year training (2021+2023)	Monocrop	0.73	0.78	0.75	0.67

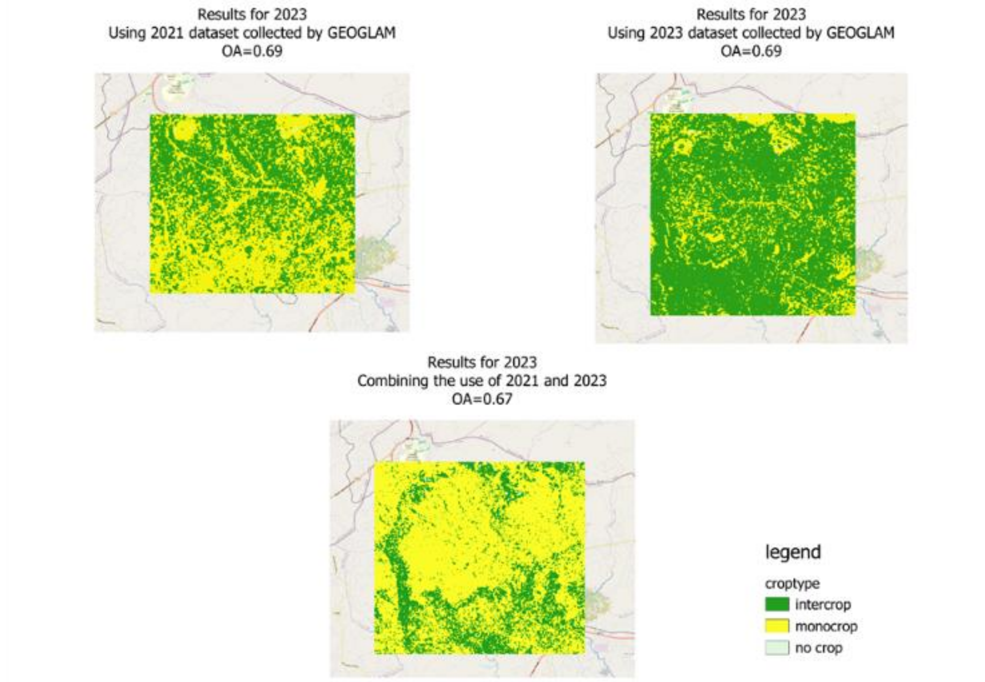

Figure 2. Presto-WorldCereal classification results for 2023.

Across all experiments, Overall Accuracy values ranged from 0.67 to 0.71. However, spatial patterns and class-specific metrics varied substantially, particularly for the intercropping class, which was sometimes underrepresented or overestimated. In addition, different training datasets produced different spatial predictions even when applied to the same year. These results highlight the challenges of developing generalisable classification models for heterogeneous cropping systems such as intercropping. They also emphasise the importance of evaluating spatial coherence alongside numerical accuracy metrics when assessing classification performance.

4.1.2 Maize-mask-based crop classification results

The maize-mask classification approach produced moderate classification performance, with an Overall Accuracy of 0.73 across the two classes: monocrop maize and maize–bean intercrop (Table 7). Precision and recall values were higher for monocrop fields (0.787) than for intercropped fields (0.630), indicating that the classifier identified monocrop maize areas more reliably.

Table 7. Performance Metrics for Maize-Mask-Based Classification of Monocrop and Intercrop Systems (2021).

Scenario	Class	Precision	Recall	F1-score	Overall Accuracy
Maize-mask classification (2021)	**Intercrop**	**0.630**	**0.630**	**0.630**	**0.73**
	Monocrop	**0.787**	**0.787**	**0.787**	

The spatial classification map further supports this observation. Most classified pixels are monocrop maize (yellow), whereas pixels identified as intercropped maize (green) are fewer, more scattered, and often spatially isolated (Figure 3). This spatial distribution is unlikely to represent the true cropping pattern in the study area and may instead reflect limitations in the maize mask or the available training data.

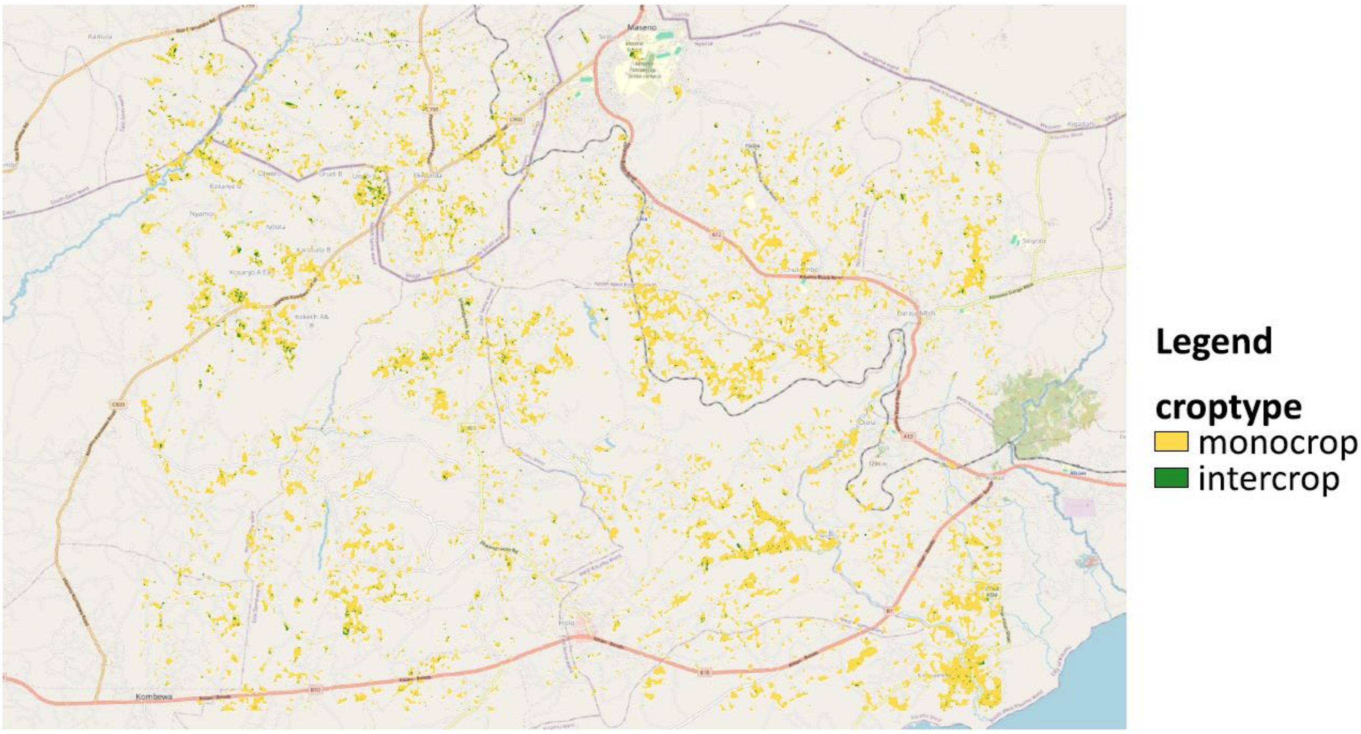

Figure 3. Maize-mask-based classification result for 2021.

Although the maize-mask approach achieved slightly higher Overall Accuracy than the Presto-based classification, its spatial outputs indicate that classification uncertainty remains substantial, particularly in detecting intercropped fields within fragmented smallholder landscapes.

4.1.3 Drought monitoring results

Drought monitoring was conducted using vegetation-based drought indicators derived from satellite imagery. The analysis examined temporal patterns of vegetation stress at two scales: field-level monitoring using 15 georeferenced plots and a broader zonal assessment of drought severity across monocrop and intercropped maize systems.

Vegetation Condition Index (VCI) time series derived from MODIS data show relatively smooth temporal patterns across the monitored fields (Figure 4). Most values fall within the “No Drought” and “Low Drought” categories, and differences between individual fields are limited. This smoothing effect likely results from the coarse spatial resolution of MODIS (~500 m), which averages signals across multiple fields and land-cover types.

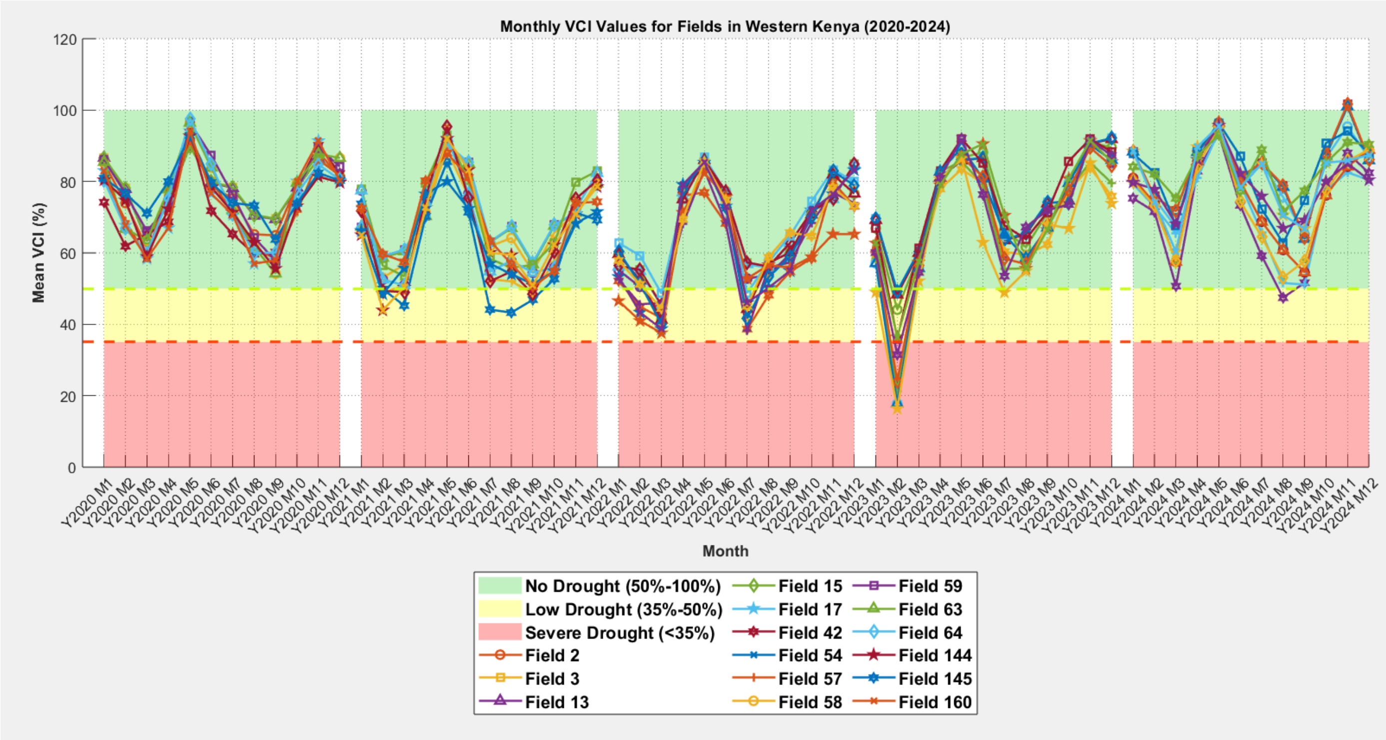

Figure 4. The MODIS-based VCI time series.

In contrast, the Sentinel-2–based VCI time series reveals much greater variability between fields (Figure 5). The higher spatial resolution of Sentinel-2 (10 m) enables clearer detection of field-level vegetation conditions and captures sharper temporal changes in vegetation stress. Several fields showed significant drops in VCI values during the 2023 growing season, entering the severe drought category.

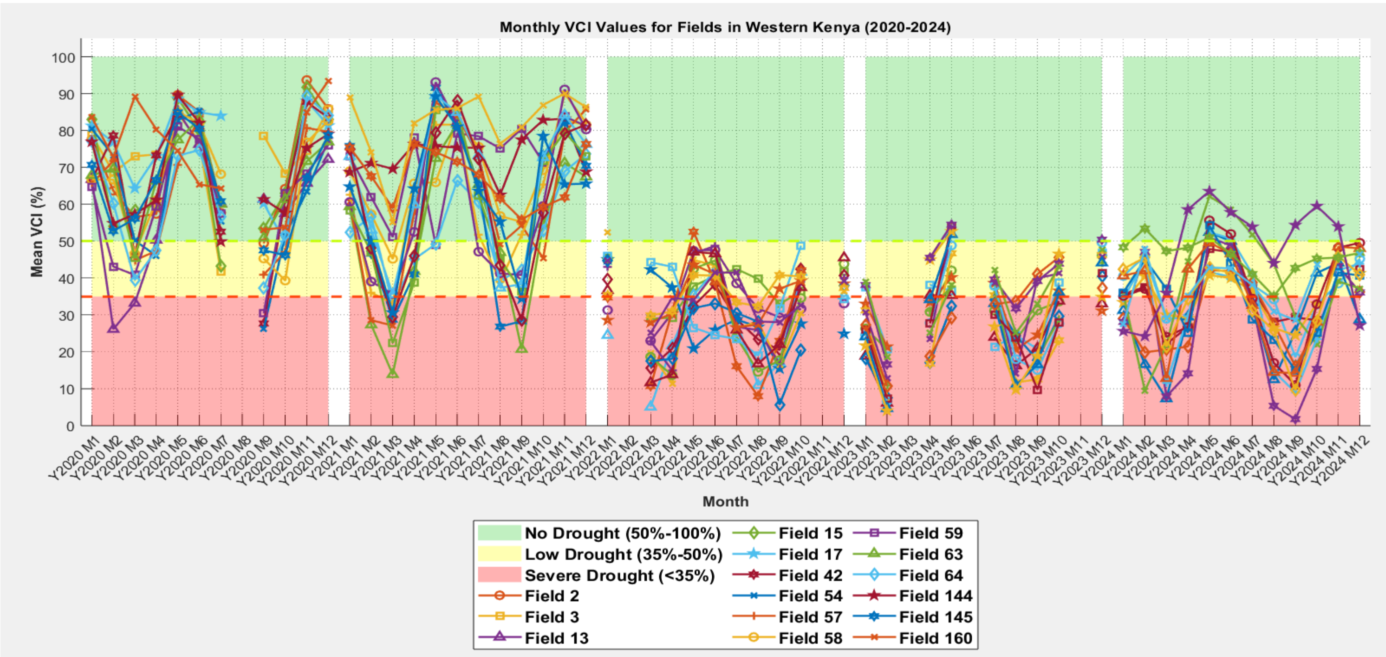

Figure 5. The Sentinel-2-based VCI time series.

A direct comparison of VCI values from MODIS and Sentinel-2 for the 2024 season further illustrates these differences (Figure 6). Sentinel-2 captures greater spatial detail and detects several instances of moderate-to-severe drought stress that are not visible in the MODIS signal. Conversely, MODIS tends to produce smoother and consistently higher VCI values, potentially masking localised drought conditions. These findings highlight the importance of spatial resolution when monitoring drought in fragmented smallholder agricultural systems.

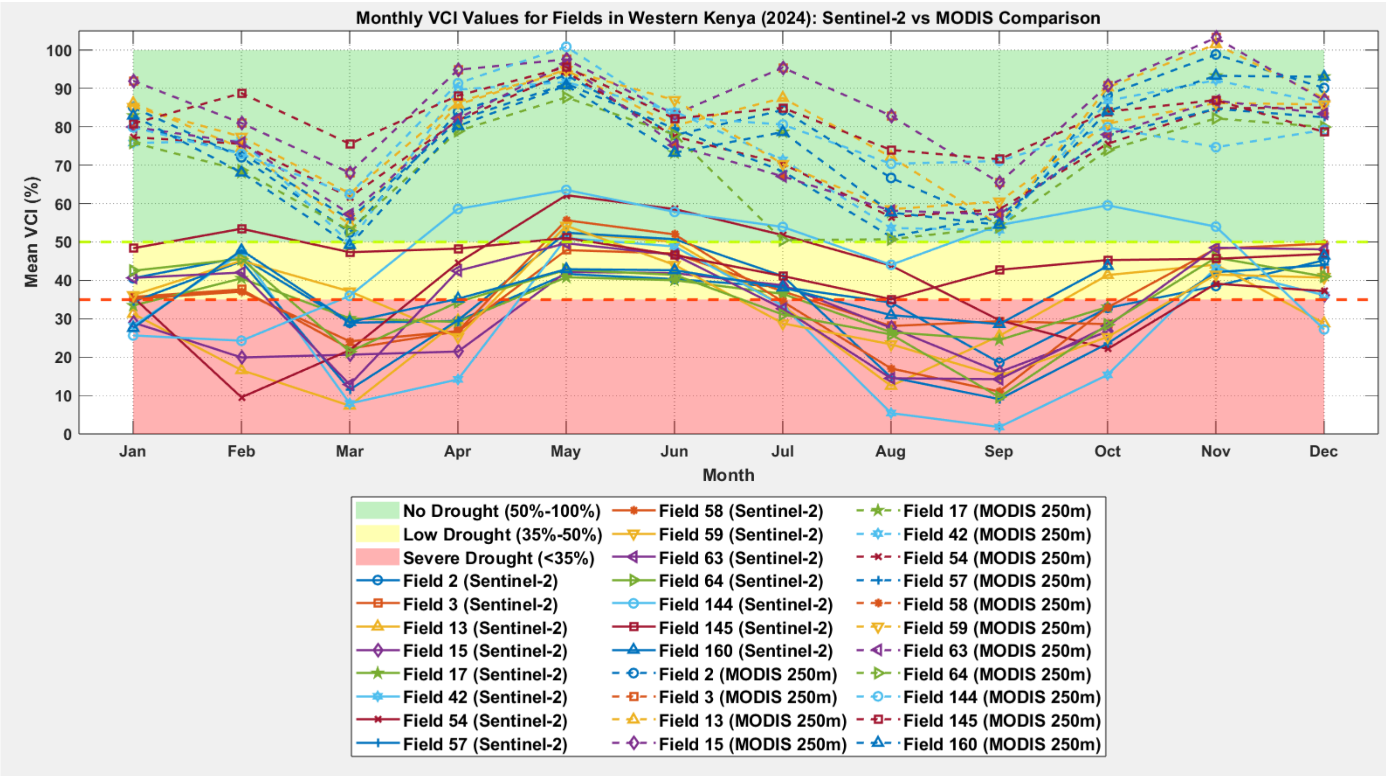

Figure 6. The MODIS and Sentinel-2-based VCI time series for the year 2024.

Vegetation Health Index (VHI) values derived from Landsat-9 imagery provide an additional perspective on drought conditions by incorporating both vegetation and thermal stress signals (Figure 7). Although Landsat offers moderate spatial resolution (30 m), its 16-day revisit interval and frequent cloud cover result in substantial data gaps during the peak growing season. These limitations reduce the reliability of Landsat-based indices for continuous seasonal drought monitoring.

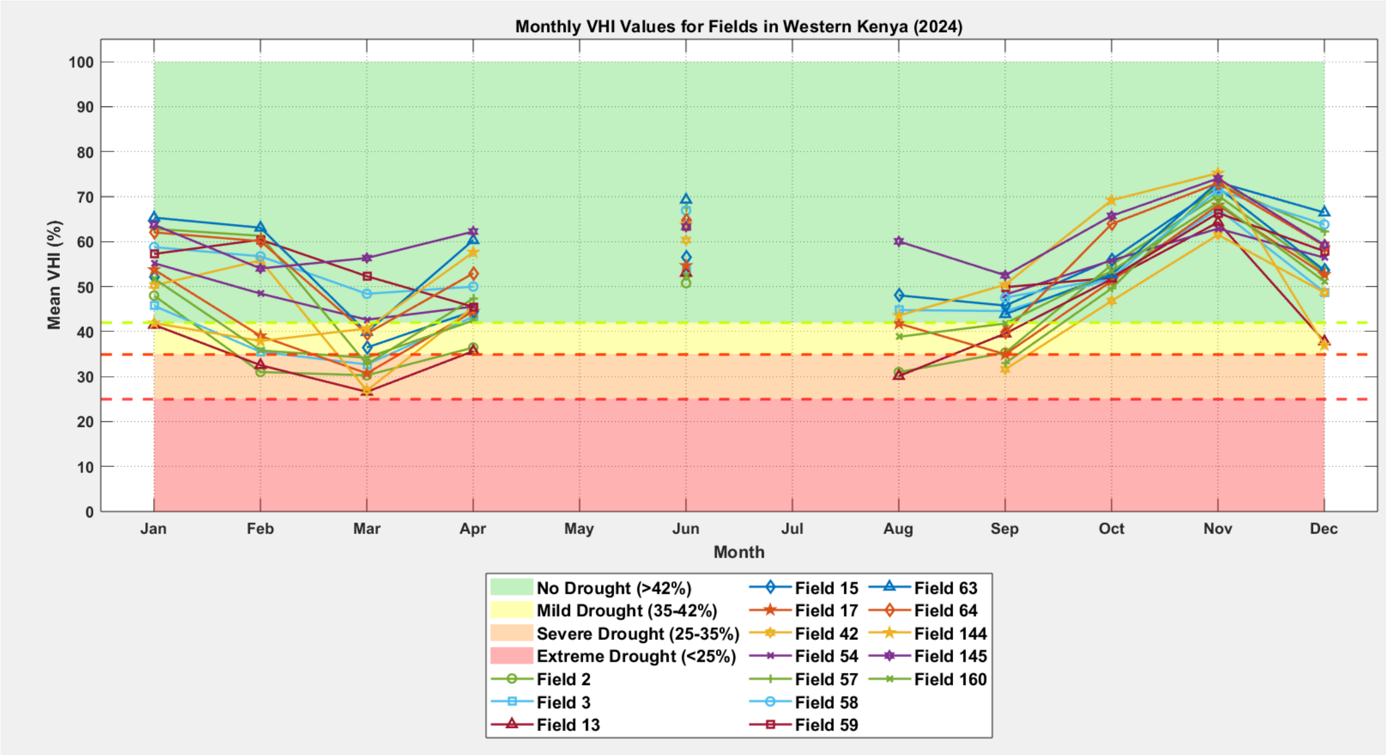

Figure 7. The Landsat-9-based VHI time series.

To assess drought patterns across cropping systems,

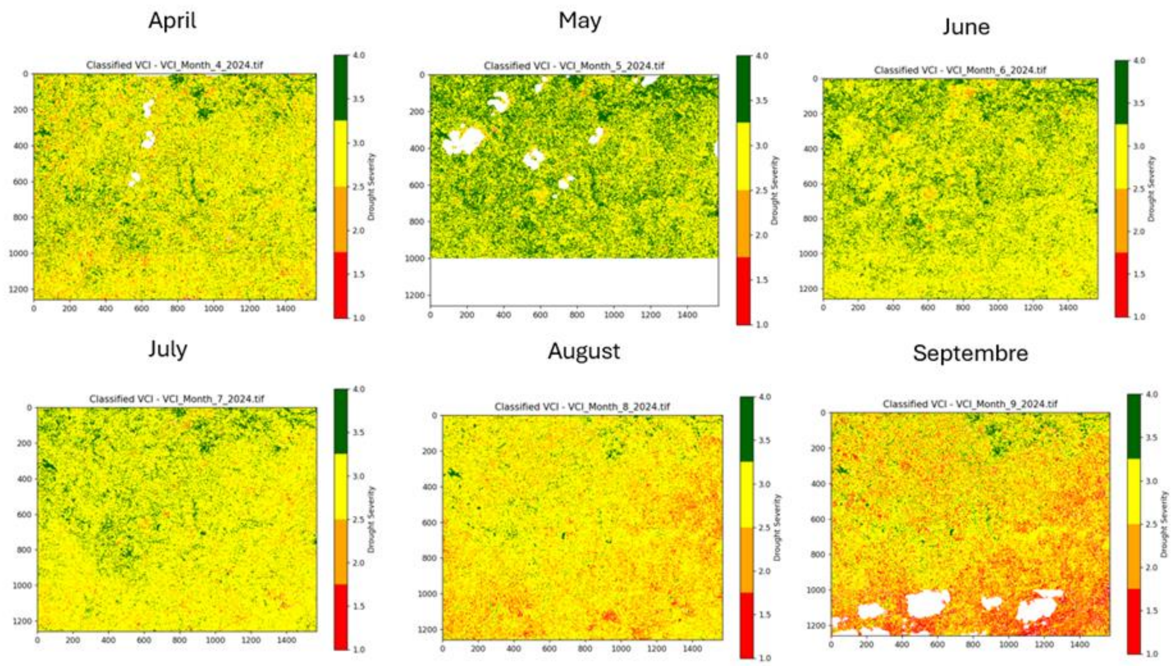

Figure 8. Drought severity maps using the VCI vegetation index for the long rain season of 2024 (April–September).

Monthly VCI images were classified into four drought severity categories: extreme drought (VCI < 10), severe drought (10 ≤ VCI < 20), moderate drought (20 ≤ VCI < 50), and healthy vegetation (VCI ≥ 50). The VCI maps were then masked using the crop classification layer to isolate drought conditions within monocrop and intercropped fields.

The resulting drought severity maps for the 2024 long rain season (April–September) show a progressive increase in vegetation stress during the season (Figure 8). Relatively healthy vegetation conditions characterised early-season months such as April and May, whereas July and August exhibited a clear expansion of severe and extreme drought zones.

Comparative analysis of monocrop and intercropped maize fields reveals consistent differences in vegetation condition during the season (Figure 9). Across May, June, and July 2024, intercropped fields maintained a higher proportion of healthy vegetation compared with monocrop maize fields. For example, in May 2024, 50.9% of intercropped maize pixels were classified as healthy vegetation compared with 38.7% of monocrop maize pixels. Similar differences were observed in June (34.0% versus 22.2%) and July (13.6% versus 9.0%), despite increasing mid-season drought stress.

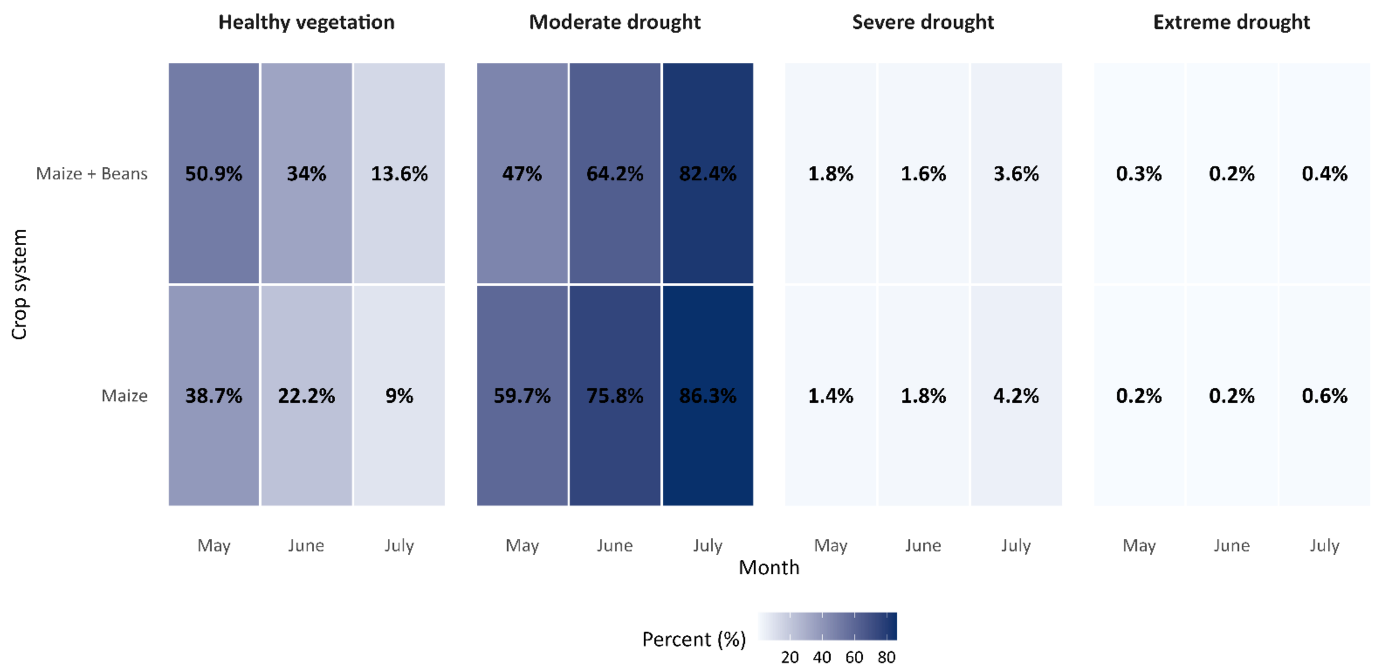

Figure 9. Heatmap showing the percentage distribution of drought severity categories for maize and maize–bean systems from May to July 2024. Values represent the proportion of area classified as healthy vegetation, moderate drought, severe drought, or extreme drought

Taken together, the crop classification and drought monitoring results indicate that intercropping systems may offer greater resilience to drought during critical crop growth stages. Although differences in vegetation condition are modest, the consistent pattern of healthier vegetation in intercropped fields supports the potential of intercropping as a climate-smart strategy to boost resilience in smallholder farming systems.

4.2 Targeted Climate-Smart Irrigation Advisory

The Targeted Climate-Smart Irrigation Advisory (T-CSIA) operationalises the ISRA metrics framework by pinpointing where irrigation support is most needed, especially under increasing climate variability. A detailed methodological overview of this innovation is provided by Kenduiywo et al. (2026). In western Kenya, agricultural production remains largely rainfed, but rising rainfall variability, more frequent droughts, and increasing temperatures are changing both water availability and crop water demand. These changes diminish the effectiveness of current irrigation planning methods that rely on a single indicator, such as rainfall alone.

To overcome this limitation, T-CSIA combines various climate indicators into a unified advisory system. It merges long-term rainfall and temperature data with seasonal probabilistic forecasts to create spatially detailed irrigation risk profiles. Rainfall indicators reflect variability, intra-seasonal distribution, and long-term trends, while temperature indicators indicate evaporative demand and heat stress. This integration offers a more complete assessment of irrigation requirements amid changing climatic conditions.

The composite irrigation advisory index categorises irrigation demand into operational levels from low to very high. These levels are mapped across the study area to pinpoint irrigation hotspots and zones with lower water stress. Including seasonal forecasts improves predictive capacity by providing forward-looking information, enabling anticipation of irrigation needs ahead of critical moisture-stress periods (Kenduiywo et al., 2026).

The findings reveal considerable spatial differences in irrigation needs. Areas characterised by high rainfall variability and elevated temperature-driven evaporative demand consistently have higher irrigation requirements and stand out as priority zones for targeted measures. Conversely, regions with more stable rainfall patterns display lower irrigation needs, although temperature still affects seasonal variations in water demand.

Temperature emerges as a crucial factor influencing irrigation demand. Higher temperatures increase evapotranspiration and can exacerbate water stress even in regions without notable rainfall deficits. This underscores the importance of considering both precipitation and temperature together in irrigation advisory systems. Furthermore, while seasonal forecasts do not significantly change the overall spatial pattern of irrigation risk, they tend to intensify existing hotspots, emphasising the need for early and targeted interventions.

Overall, T-CSIA demonstrates how ISRA metrics transform climate data into practical, location-specific advice. The approach supports proactive irrigation planning and improves the targeting of interventions, leading to more efficient water use, reduced exposure to climate risks, and increased resilience in smallholder farming systems. It is scalable and adaptable, and can be incorporated into extension services, digital advisory platforms, and policy processes to promote resilient agriculture (Kenduiywo et al., 2026).

4.3 Climate-smart soil nutrient management for resilience

The soil innovation focuses on strengthening smallholder resilience by improving site-specific nutrient management, moving from generic fertiliser advice to recommendations that reflect local soil constraints and seasonal climate risk. In Western Kenya, maize productivity is consistently limited by depleted soil fertility and widespread acidity, with substantial variability in soil texture, organic matter, nutrient reserves, and pH within and between farms. When soil pH is low, nutrient availability and root development can be reduced, increasing the likelihood of a weak fertiliser response even when nutrients are applied, which can make blanket recommendations inefficient and sometimes increase farmers' exposure to loss. This context underpins the Innovation 3 objective: to generate diagnostic, locally appropriate, and risk-aware fertiliser recommendations that can be operationalised through advisory channels.

Innovation 3 is tracked through four closely linked ISRA metrics that jointly strengthen the innovation pipeline and decision support. First, **nutrient gap estimation** produces spatially resolved maps of soil nutrient deficits (NPK), enabling the project to identify hotspots of deficiency and prioritise interventions. This is achieved by combining soil indicators, including acidity proxies and Soil Organic Carbon, with geo-referenced trial evidence to distinguish locations where nutrient limitations are most binding from locations where fertiliser responsiveness is likely constrained by secondary factors such as acidity. These diagnostics also provide a baseline for monitoring changes over time and better targeting advisory messaging and input planning. Second, **spatially and temporally resolved fertiliser rate and type determination** translates these diagnostics into practical, site- and season-specific fertiliser packages (type, rate, and timing), explicitly designed around the farmer's decision problem, selecting the option that is most likely to be both profitable and agronomically appropriate for a given location and season, rather than simply maximising yield. This approach helps reduce under- and over-application and supports productivity gains while minimising environmental footprints. Third, **Nutrient Use Efficiency (NUE)** quantifies and maps efficiency indicators (such as agronomic efficiency and partial factor productivity, where feasible) to identify where nutrients are used efficiently versus wasted, enabling more precise rate targeting and management adjustments that improve performance under climate stress and reduce nutrient losses. Finally, **profitability** links predicted yield gains with input costs to generate gross-margin or profitability insights, helping to identify high-return, lower-risk fertiliser options under climate variability and supporting scalable, evidence-based advisory recommendations.

Progress under Innovation 3 is implemented through an **ML- and optimisation-based pipeline (AgWise-aligned), as** documented in the progress report and depicted in Figure 10. The workflow integrates: (i) data assembly and harmonisation, (ii) ML-based prediction of yield and yield response, and (iii) recommendation optimisation under feasibility, efficiency, and risk constraints. Trial observations (yield and fertiliser rates) are augmented with spatial covariates describing soil properties (texture, Soil Organic Carbon, pH-related indicators), topography, and climate descriptors extracted at plot coordinates. A supervised ML model is trained to predict maize yield as a function of nutrient inputs and environmental context, and then used to simulate yield across a feasible grid of nutrient rates at each target location.

Recommendation generation follows a transparent search-and-select logic: for each location and (where applicable) rainfall scenario, candidate nutrient packages are evaluated; yields are predicted; **efficiency (NUE-related)** and **profitability** are computed; plausibility checks are applied (e.g., agronomic realism and response consistency); and the best package is selected using a composite objective that balances **productivity, efficiency, and profit**. This design is intentionally replicable to support evaluation within iSPARK Work Group 3 and iterative improvement of the iSHAMBA/AgWise innovation bundle.

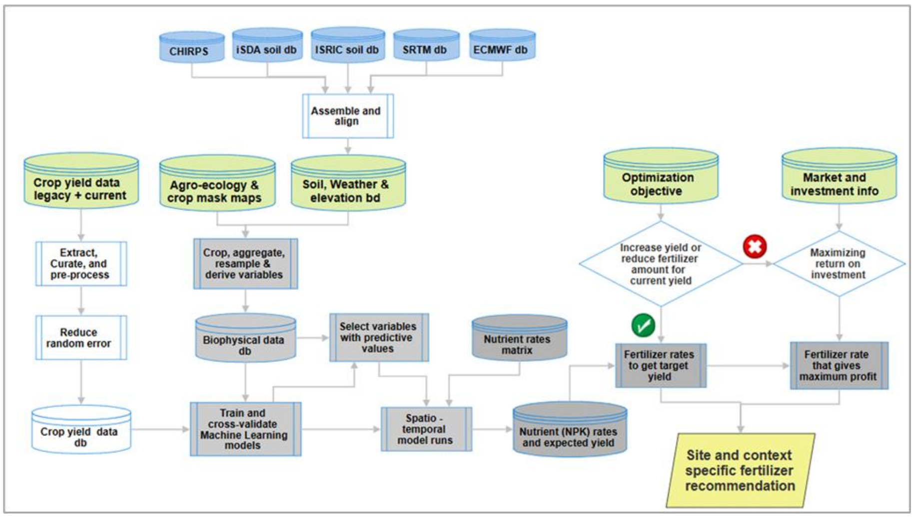

Figure 10. Overview of the Agwise ML-based framework linking data assembly, prediction, and optimisation.

Evidence-based and data readiness. The modelling dataset is built from a large compilation of geo-referenced maize experimental and trial records: approximately **19,000** plot-level observations were assembled, ranging from 2003 to 2024, and after quality control (unit harmonisation, completeness checks, coordinate validation, duplicate/outlier review, and covariate extraction checks), approximately **12,000** observations were retained for modelling. This cleaning step is critical because unstable or noisy agronomic records can lead ML models to learn spurious relationships and produce unreliable optimised recommendations.

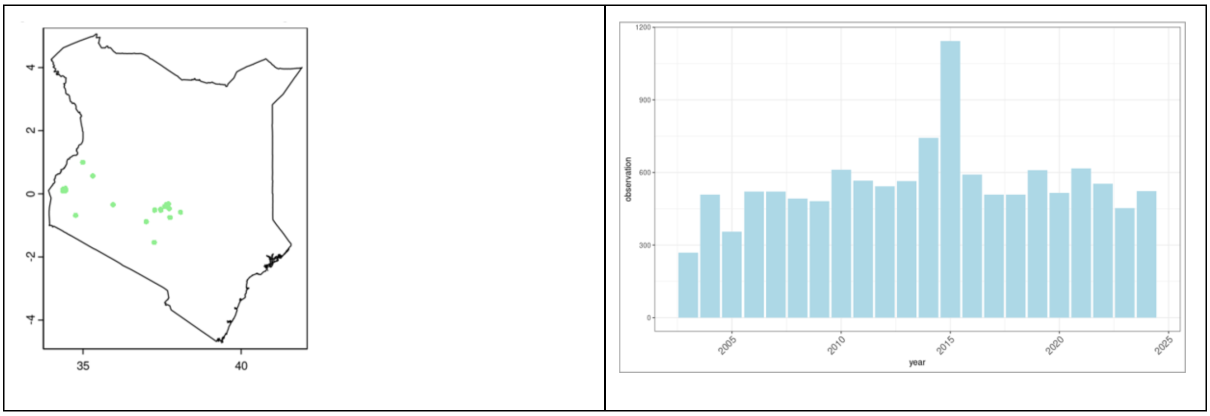

Figure 11. Spatial distribution of geo-referenced maize trial observations and annual number of trials used in model development.

Figure 11 shows geo-referenced maize trial observations and annual volume of maize trials used in model development. **Model performance and interpretability.** Held-out evaluation indicates good predictive skill, with test performance around $R^2 \approx 0.74$ and corresponding Root Mean Squared Error/Mean Absolute Error (RMSE/MAE) values reported for the train/test splits (Table 8). Figure 12 depicts the performance of the model through the scatter plot of yield prediction verses the observation. Accordingly, the model performed well as the points fall closely to the diagonal line.

Table 8. Model performance metrics on train/validation/test splits.

Dataset	RMSE	MAE	R2
Train	0.3915008	0.2826239	0.7558416
Test	0.4167853	0.2953593	0.7415956

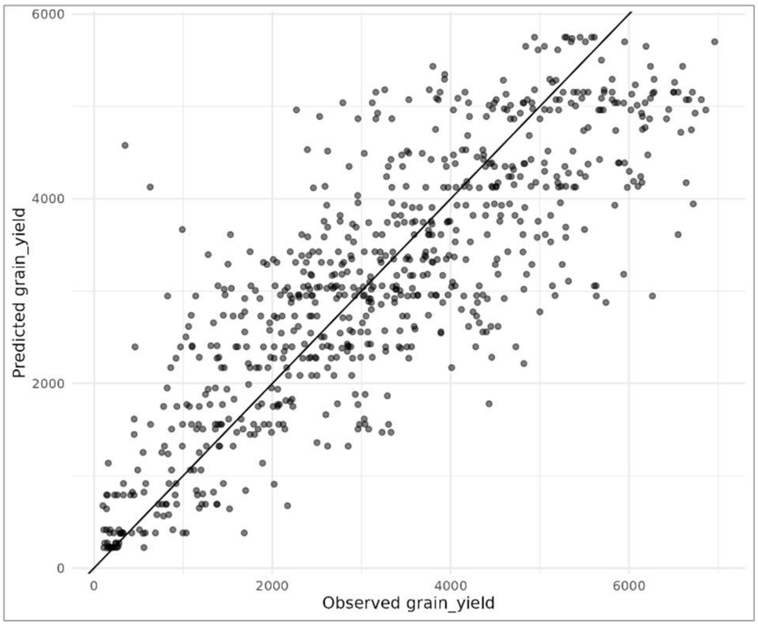

Figure 12. Model performance of predicted versus observed maize grain yield on the original scale (kg/ha) for the stacked ensemble model.

Interpretation tools confirm that fertiliser N and P rates, rainfall amount/variability, and key soil properties are among the most influential drivers of predicted yield response as depicted in Figure 13, supporting the logic of differentiated, context-aware advisory rather than uniform rates.

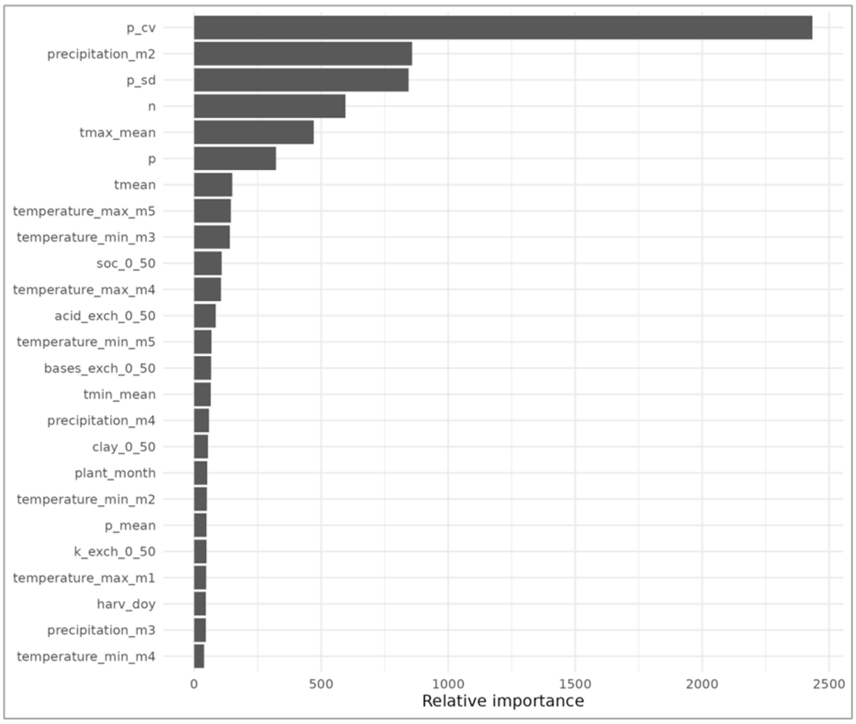

Figure 13. Variable importance from the gradient boosting model.

Scenario-based, risk-aware recommendations. A major output of this innovation is the production of gridded recommendation layers (best N, best P, expected yield) under rainfall tercile scenarios (below-normal / near-normal / above-normal). Relative to near-normal conditions, the below-normal scenario is expected to recommend more conservative rates, while above-normal conditions can support higher attainable yields and higher profitable rates. These scenario outputs provide the technical basis for “risk-graded” advisory packages (conservative vs opportunity) and directly support targeted guidance to farmers under different seasonal climate scenarios Accordingly, Figure 14 depicts that the optimal fertilizer rate of N and P for Maize crop considering the three climate scenarios

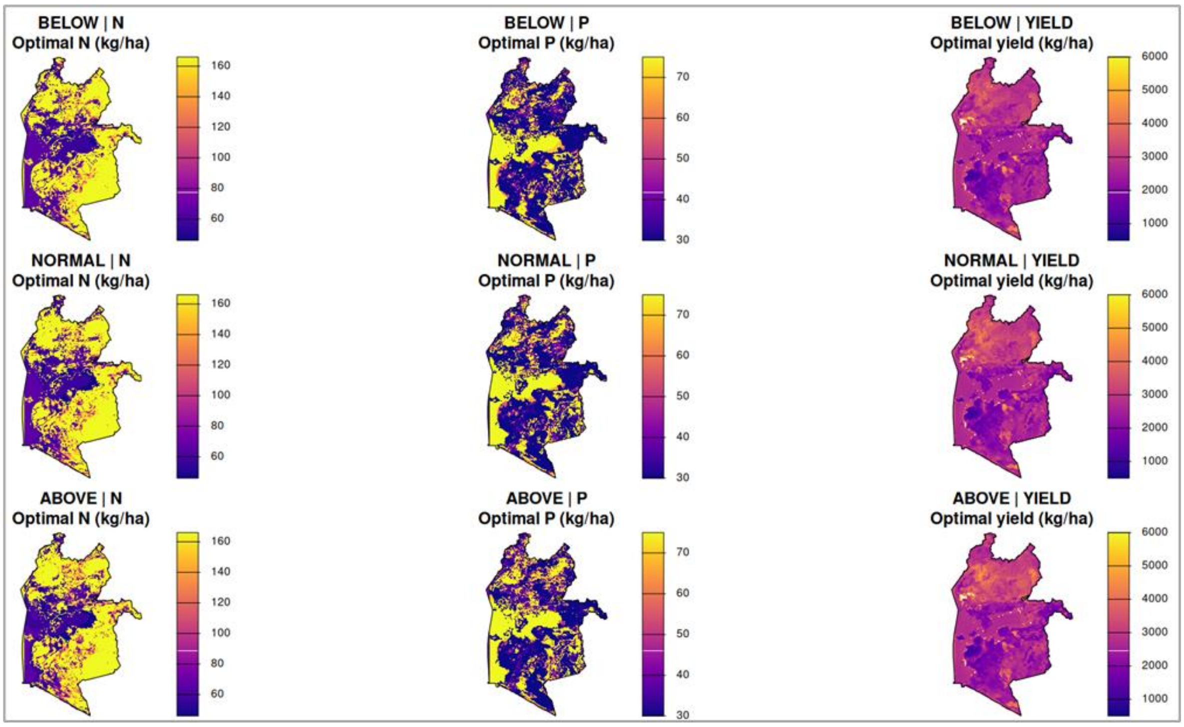

Figure 14. Best N (kg/ha), best P (kg/ha), and expected yield (kg/ha) under rainfall tercile scenarios (below/normal/above).

How this strengthens resilience in practice

Innovation 3 contributes to resilience in three practical ways:

1. **Targeting responsiveness and managing "non-response."** Because fertiliser response varies across space and seasons, driven by baseline fertility, acidity, moisture dynamics, and rainfall distribution, the system avoids the assumption that the average response applies everywhere. Where low responsiveness is likely (e.g., strong acidity or low SOC), the advisory is designed to pair fertiliser guidance with complementary interpretation notes (e.g., liming or organic amendments) to avoid inappropriate escalation of nutrient rates.
2. **Embedding climate risk into nutrient decisions.** Rainfall terciles serve as a simple yet operationally usable risk-management layer: conservative packages for below-normal seasons, default packages for near-normal seasons, and opportunity packages for above-normal seasons. This enables farmers and implementers to align fertiliser investment with risk tolerance and available climate information.
3. **Delivering actionable packages (products, timing, and tiers).** The advisory explicitly supports conversion from nutrient rates (kg/ha N and P) to locally available fertiliser products, with practical timing guidance. It also recommends packaging options aligned to affordability tiers (minimum investment, recommended, high opportunity) to improve usability and equity.

Operationalisation, monitoring, and next steps

For deployment, the report proposes an end-to-end seasonal workflow (refresh data → retrain/recalibrate → generate scenario layers → convert to product packages and messages → deploy → collect feedback and outcomes). A Monitoring, Evaluation, and Learning (MEL) framework is recommended to track adoption, yield, profitability, nutrient-use efficiency, model health (drift), and equity outcomes on a seasonal/annual basis.

Key planned enhancements include integrating operational seasonal forecast products to move from tercile scenarios to forecast-weighted recommendations, adding a QUEFTS module to strengthen agronomic plausibility constraints, implementing grouped environment-holdout validation and uncertainty quantification for risk communication, and conducting validation pilots in representative counties and soil-constraint classes (explicitly including acidic soils).

5. Monitoring Socio-Economic Outcomes of ISRA-Informed Innovations

The initial results from applying the ISRA metrics offer valuable insights into the potential of the identified innovations to improve resilience, productivity, and resource-use efficiency in smallholder farming systems. However, turning these insights into sustainable impact requires systematic monitoring of how these innovations perform once implemented in real farming settings. To facilitate this, a Monitoring, Evaluation, and Learning (MEL) framework is proposed to track the socio-economic and agronomic outcomes associated with ISRA-informed innovations over time.

The MEL framework is built on socio-economic indicators identified through the MCM exercises and the literature review conducted during the development of the ISRA metrics. These indicators reflect farmers' priorities and concerns, including profitability, food security, labour needs, gender equality, and climate resilience. By monitoring these aspects, the framework enables stakeholders to assess whether the innovations deliver tangible benefits to farmers while supporting broader sustainability and resilience objectives.

The monitoring system concentrates on several key outcome areas. Firstly, adoption indicators will measure the uptake of innovations among farmers, including the extent to which recommended practices such as targeted intercropping, climate-smart irrigation scheduling, and improved nutrient management strategies are adopted across various farming systems. Monitoring adoption offers insights into the practicality, relevance, and acceptability of innovations under real-world farming conditions.

Second, agronomic indicators will monitor changes in farm productivity and resource-use efficiency related to the innovations. These include crop yield, nutrient use efficiency, and water use efficiency. Tracking these indicators helps assess whether the innovations enhance agricultural performance and resilience under climate variability and environmental stress.

Third, economic indicators will evaluate the financial viability of the innovations. Key metrics include farm income, the variability of net income across seasons, and input costs for fertiliser, irrigation, and crop management practices. Monitoring these indicators helps determine whether the innovations generate economic benefits that can motivate sustained adoption by farmers.

Fourth, social indicators will reflect wider socio-economic outcomes related to the innovations. These include food and nutrient availability, household food access, labour requirements, and the distribution of labour across genders. Monitoring these indicators helps determine whether the innovations contribute to inclusive development and whether they generate unintended labour burdens or inequalities within farming households.

Alongside tracking farm-level outcomes, the MEL framework assesses the performance of the analytical tools used to generate innovation recommendations. Since several of the proposed innovations depend on EO data and machine learning models, regular evaluation of model performance is essential to ensure predictions remain accurate over time. Monitoring model performance helps identify potential model drift caused by environmental changes, evolving farming practices, or variations in data quality.

Monitoring activities will be carried out on a seasonal or annual basis, depending on the indicator's nature and data availability. Data sources may include household surveys, remote sensing products, field observations, and participatory feedback from farmers. Combining these data sources will support ongoing learning and the refinement of innovation strategies.

6. Conclusions

This study demonstrates the value of a structured, metric-based approach for evaluating agricultural innovations that enhance sustainability, resilience, and climate adaptation in smallholder systems. A systematic review of remote sensing and modelling approaches, together with clear selection criteria, yielded a set of robust, operational metrics. These metrics were further prioritised using a combination of frequency-based relevance and practical considerations, including data requirements, scalability, and expected confidence. This process identified three areas of innovation: intercropping for resilience, targeted climate-smart irrigation advisory, and climate-smart soil nutrient management.

The analysis of intercropping systems provides evidence of their potential to enhance resilience under drought conditions. Crop classification methods based on Earth observation and machine learning showed moderate performance, with overall accuracies ranging from 0.67 to 0.73. However, the results highlight an important methodological insight: classification accuracy alone is insufficient for evaluating complex cropping systems such as intercropping. Spatial consistency and realistic representation of cropping patterns are equally critical. Among the tested approaches, the maize-mask NDVI-based classification performed relatively well, though challenges in detecting intercropped fields in fragmented landscapes remain.

Drought monitoring results further strengthen the case for intercropping as a resilience-enhancing practice. High-resolution Vegetation Indices reveal substantial spatial variability in drought stress, with finer-resolution data capturing localised stress patterns that coarser datasets tend to smooth out. When combined with crop classification outputs, the analysis shows a consistent pattern: intercropped systems maintain better vegetation health than monocropped systems during periods of increasing drought stress. This provides empirical support for the role of diversification in buffering climate shocks in smallholder farming systems.

Beyond cropping systems, the study highlights the importance of integrating climate information into decision-making through targeted irrigation advisory services. By combining rainfall variability, temperature dynamics, and seasonal forecasts into a composite framework, irrigation needs can be spatially differentiated and proactively managed. The results reveal strong spatial heterogeneity in irrigation demand, driven not only by rainfall deficits but also by temperature-induced evapotranspiration. This underscores the limitations of single-indicator approaches and demonstrates the value of integrated, climate-informed advisory systems for improving water-use efficiency and reducing exposure to climate risk.

Similarly, the soil nutrient management analysis underscores the need to move from generic fertiliser recommendations to context-specific, data-driven strategies. By integrating soil properties, climatic conditions, and crop response data within a machine learning and optimisation framework, the approach generates site- and season-specific nutrient recommendations. These recommendations balance productivity, efficiency, and profitability while explicitly accounting for climate variability. The results highlight that addressing soil constraints is essential to unlocking the full benefits of other climate-smart interventions and to reducing the risk of non-responsive or inefficient use of inputs.

The findings demonstrate that resilience in smallholder agriculture is inherently multidimensional. Intercropping enhances ecological stability, irrigation advisory improves climate risk management, and soil nutrient management strengthens productivity and resource-use efficiency. Their combined application offers a more comprehensive pathway to improving agricultural performance under climate variability than any single intervention applied in isolation.

A key contribution of this work is demonstrating that integrated metrics derived from Earth observation, modelling, and socio-economic data can serve as a unifying framework for evaluating and targeting agricultural innovations. These metrics enable spatially explicit, scalable, and evidence-based decision-making, supporting the transition from generalised recommendations to context-specific advisory systems. The findings emphasise that effective climate-smart agriculture requires not only improved technologies but also robust systems for targeting, monitoring, and adapting interventions to local conditions. By linking data, analytics, and decision-making, the proposed approach provides a foundation for scaling resilient agricultural practices and informing policy and investment strategies in smallholder farming systems.

References

Alawode, A. (2025). Assessing climate change impacts on agricultural productivity and rural livelihoods in Sub-Saharan Africa- International Journal of Research Publication and Reviews (2025).

Alkaraki, K. F., & Hazaymeh, K. (2023). A comprehensive remote sensing-based Agriculture Drought Condition Indicator (CADCI) using machine learning. Environmental Challenges, 11(February), 100699.

Antle, J., Adhikari, R., & Price, S. (2015). An income-based food security indicator for agricultural technology impact assessment.

Asfaw, D. M., Asnakew, Y. W., Sendkie, F. B., Abdulkadr, A. A., Mekonnen, B. A., Tiruneh, H. D., & Ebad, A. M. (2024). Analysis of constraints and opportunities in maize production and marketing in Ethiopia. Heliyon, 10(20).

Attarzadeh, R., Amini, J., Notarnicola, C., & Greifeneder, F. (2018). Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at plot scale. Remote Sensing, 10(8), 1–18.

Aynekulu, E., Suber, M., Van Noordwijk, M., Arango, J., Roshetko, J. M., & Rosenstock, T. S. (2020). Carbon storage potential of silvopastoral systems of Colombia. Land, 9(9), 309.

Baghdadi, N., Hajj, M. El, & Zribi, M. (2016). Coupling SAR C-band and optical data for soil moisture and leaf area index retrieval over irrigated grasslands. International Geoscience and Remote Sensing Symposium (IGARSS), 2016-Novem, 3551–3554.

Barrett, A. B., Duivenvoorden, S., Salakpi, E. E., Muthoka, J. M., Mwangi, J., Oliver, S., & Rowhani, P. (2020). Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya. Remote Sensing of Environment, 248(April), 111886.

Barrett, C. B., Benton, T. G., Cooper, K. A., Fanzo, J., Gandhi, R., Herrero, M., ... & Wood, S. (2020). Bundling innovations to transform agri-food systems. Nature Sustainability, 3(12), 974-976.

Bizikova, L., Jungcurt, S., McDougal, K., & Tyler, S. (2020). How can agricultural interventions enhance contribution to food security and SDG 2.1?. Global Food Security, 26, 100450.

Chang, Y.-C., Stewart, A. J., Bastani, F., Wolters, P., Kannan, S., Huber, G. R., Wang, J., & Banerjee, A. (2024). On the Generalizability of Foundation Models for Crop Type Mapping.

Eichler Inwood, S. E., López-Ridaura, S., Kline, K. L., Gérard, B., Monsalue, A. G., Govaerts, B., & Dale, V. H. (2018). Assessing sustainability in agricultural landscapes: A review of approaches1, 2. Environmental Reviews, 26(3), 299-315.

Fieuzal, R., Sicre, C. M., & Tallec, T. (2020). Towards an improved inventory of n2 o emissions using land cover maps derived from optical remote sensing images. Atmosphere, 11(11).

Gao, Q., Zribi, M., Escorihuela, M. J., & Baghdadi, N. (2017). Synergetic use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution. Sensors (Switzerland), 17(9).

Guo, M., Wang, X., Li, J., Yi, K., Zhong, G., & Tani, H. (2012). Assessment of global carbon dioxide concentration using MODIS and GOSAT data. Sensors (Switzerland), 12(12), 16368–16389.

Hajj, M. El, Baghdadi, N., Zribi, M., & Bazzi, H. (2017). Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. Remote Sensing, 9(12), 1–28.

Hashakimana, L., Tessema, T., Niyitanga, F., Cyamweshi, A. R., & Mukuralinda, A. (2023). Comparative analysis of monocropping and mixed cropping systems on selected soil properties, soil organic carbon stocks, and simulated maize yields in drought-hotspot regions of Rwanda. Heliyon, 9(9).

Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the United States. Bulletin of the American Meteorological Society, 83(8), 1149-1166.

Hengl, T., Heuvelink, G. B., Kempen, B., Leenaars, J. G., Walsh, M. G., Shepherd, K. D., ... & Tondoh, J. E. (2015). Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. PloS one, 10(6), e0125814.

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), 1–24.

Janssen, B. H., Guiking, F. C. T., van der Eijk, D., Smaling, E. M., Wolf, J., & van Reuler, H. (1990). A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). Geoderma, 46(4), 299-318.

Jin, Z., Azzari, G., Burke, M., Aston, S., & Lobell, D. B. (2017). Mapping smallholder yield heterogeneity at multiple scales in eastern Africa. Remote Sensing, 9(9).

Jin, Z., Azzari, G., You, C., Di Tommaso, S., Aston, S., Burke, M., & Lobell, D. B. (2019). Smallholder maize area and yield mapping at national scales with Google Earth Engine. Remote Sensing of Environment, 228(September 2018), 115–128.

Kanter, D. R., Musumba, M., Wood, S. L., Palm, C., Antle, J., Balvanera, P., ... & Andelman, S. (2018). Evaluating agricultural trade-offs in the age of sustainable development. Agricultural Systems, 163, 73-88.

Karimi, P., Bongani, B., Blatchford, M., & de Fraiture, C. (2019). Global satellite-based ET products for the local level irrigation management: An application of irrigation performance assessment in the Sugarbelt of Swaziland. Remote Sensing, 11(6).

Kenduiywo, B., Mwungu, C., Wanjau, A., Tibebe, D., Kirwa, L., Mayanja, B., Rose, S., & Ghosh, A. (2026). Targeted climate-smart irrigation advisory using integrated climate indicators. Alliance of Bioversity International and CIAT.

Kimaro, O. D., Ojwang, D., Rotich, B., Desie, E., Verbist, B., Kimaro, D. N., ... & Vancampenhout, K. Land Use/Cover Change as a major driver of ecosystem services values variation in the Indigenous Agroforestry systems of North-Eastern Mountains, Tanzania.

Kogo, B. K., Kumar, L., & Koech, R. (2021). Climate change and variability in Kenya: a review of impacts on agriculture and food security: BK Kogo et al. Environment, development and sustainability, 23(1), 23-43.

Koukos, A., Jo, H.-W., Sitokonstantinou, V., Tsoumas, I., Kontoes, C., & Lee, W.-K. (2024). Towards Global Crop Maps with Transfer Learning. Dl, 1540–1545.

Leenaars, J. G., Claessens, L., Heuvelink, G. B., Hengl, T., González, M. R., van Bussel, L. G., ... & Cassman, K. G. (2018). Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. Geoderma, 324, 18-36.

Liepa, A., Thiel, M., Taubenböck, H., Steffan-Dewenter, I., Abu, I. O., Singh Dhillon, M., Otte, I., Otim, M. H., Lutaakome, M., Meinhof, D., Martin, E. A., & Ullmann, T. (2024). Harmonised NDVI time-series from Landsat and Sentinel-2 reveal phenological patterns of diverse, small-scale cropping systems in East Africa. Remote Sensing Applications: Society and Environment, 35.

Lindsay, W. L., & Norvell, W. (1978). Development of a DTPA soil test for zinc, iron, manganese, and copper. Soil science society of America journal, 42(3), 421-428.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. Remote Sensing of Environment, 164, 324-333.

Ma, Y., Chen, S., Ermon, S., & Lobell, D. B. (2024). Transfer learning in environmental remote sensing. Remote Sensing of Environment, 301(November 2023), 113924.

Mahlayeye, M., Darvishzadeh, R., & Nelson, A. (2024). Characterising maize and intercropped maize spectral signatures for cropping pattern classification. International Journal of Applied Earth Observation and Geoinformation, 128(January), 103699.

Mandal, A., Majumder, A., Dhaliwal, S. S., Toor, A. S., Mani, P. K., Naresh, R. K., Gupta, R. K., & Mitran, T. (2022). Impact of agricultural management practices on soil carbon sequestration and its monitoring through simulation models and remote sensing techniques: A review. Critical Reviews in Environmental Science and Technology, 52(1), 1–49.

Mekonnen, Y. G., Alamirew, T., Tadesse, K. B., & Chukalla, A. D. (2024). Monitoring small-scale irrigation performance using remote sensing in the Upper Blue Nile Basin, Ethiopia. Agricultural Water Management, 300(June), 108928.

Mponela, P., Snapp, S., Villamor, G. B., Tamene, L., Le, Q. B., & Borgemeister, C. (2020). Digital soil mapping of nitrogen, phosphorus, potassium, organic carbon and their crop response thresholds in smallholder managed escarpments of Malawi. Applied Geography, 124(July), 102299.

Musumba, M., Grabowski, P., Palm, C. A., & Snapp, S. S. (2017). Guide for the sustainable intensification assessment framework.

Musyimi, P. K., Sahbeni, G., Timár, G., Weidinger, T., & Székely, B. (2023). Analysis of Short-Term Drought Episodes Using Sentinel-3 SLSTR Data under a Semi-Arid Climate in Lower Eastern Kenya. Remote Sensing, 15(12), 1–26.

Nelson, D. W., & Sommers, L. E. (1982). Total carbon, organic carbon, and organic matter. Methods of soil analysis: Part 2 chemical and microbiological properties, 9, 539-579.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359.

Quisumbing, A., Cole, S., Elias, M., Faas, S., Galiè, A., Malapit, H., ... & Twyman, J. (2023). Measuring women's empowerment in agriculture: Innovations and evidence. Global Food Security, 38, 100707.

Radočaj, D., Gašparović, M., & Jurišić, M. (2024). Open Remote Sensing Data in Digital Soil Organic Carbon Mapping: A Review. Agriculture (Switzerland), 14(7).

Rhee, J., Im, J., & Park, S. (2024). Regional drought monitoring based on multi-sensor remote sensing. In Remote Sensing Handbook, Volume VI (pp. 87-108). CRC Press.

Risojevic, V., & Stojnic, V. (2022). Do We Still Need Imagenet Pre-Training in Remote Sensing Scene Classification? International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43(B3-2022), 1399–1406.

Rouse Jr, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Paper a 20. In Third Earth Resources Technology Satellite-1 Symposium: The Proceedings of a Symposium Held by Goddard Space Flight Center at Washington, DC on December 10-14, 1973: Prepared at Goddard Space Flight Center (Vol. 351, p. 309).

Shehu, B. M., Lawan, B. A., Jibrin, J. M., Kamara, A. Y., Mohammed, I. B., Rurinda, J., ... & Merckx, R. (2019). Balanced nutrient requirements for maize in the Northern Nigerian Savanna: Parameterization and validation of QUEFTS model. Field Crops Research, 241, 107585.

Tesfay, M. G. (2020). Does fertilizer adoption enhance smallholders' commercialization? An endogenous switching regression model from northern Ethiopia. Agriculture & Food Security, 9(1), 3.

Tseng, G., Cartuyvels, R., Zvonkov, I., Purohit, M., Rolnick, D., & Kerner, H. (2023). Lightweight, Pre-trained Transformers for Remote Sensing Timeseries. i, 1–21.

Tunçay, T., Kılıç, Ş., Dedeoğlu, M., Dengiz, O., Başkan, O., & Bayramin, İ. (2021). Assessing soil fertility index based on remote sensing and gis techniques with field validation in a semiarid agricultural ecosystem. Journal of Arid Environments, 190(April).

Vallury, S., Nelson, D. R., & Cook, N. J. (2025). Rethinking adaptation interventions in agricultural systems for sustainability. Current Opinion in Environmental Sustainability, 77, 101571.

van Wijk, M. T., Merbold, L., Hammond, J., & Butterbach-Bahl, K. (2020). Improving assessments of the three pillars of climate smart agriculture: current achievements and ideas for the future. Frontiers in Sustainable Food Systems, 4, 558483.

Vogels, M. F. A., de Jong, S. M., Sterk, G., Douma, H., & Addink, E. A. (2019). Spatio-temporal patterns of smallholder irrigated agriculture in the horn of Africa using GEOBIA and Sentinel-2 imagery. Remote Sensing, 11(2).

Wanjau, A.; Mwungu, C.; Kenduiywo, B.; Mayanja, B., Rose, S.; Ghosh, A. (2025) Understanding agricultural production and farmer livelihoods in Kisumu County: A baseline study. 68 p.

Wetende, E., Olago, D., & Ogara, W. (2018). Perceptions of climate change variability and adaptation strategies on smallholder dairy farming systems: Insights from Siaya Sub-County of Western Kenya. Environmental development, 27, 14-25.

Winowiecki, L. A., Bourne, M., Magaju, C., Neely, C., Massawe, B., Masikati, P., ... & Sinclair, F. (2021). Bringing evidence to bear for negotiating tradeoffs in sustainable agricultural intensification using a structured stakeholder engagement process. International Journal of Agricultural Sustainability, 19(5-6), 474-496.

Wohlenberg, J., Schneider, R. C., & Hoeltz, M. (2022). Sustainability indicators in the context of family farming: A systematic and bibliometric approach. Environmental Engineering Research, 27(1).

Zhang, M., Khosravi Aqdam, M., Abbas Fadel, H., Wang, L., Waheeb, K., Kadhim, A., & Hekmati, J. (2024). Evaluation of soil fertility using combination of Landsat 8 and Sentinel-2 data in agricultural lands. Environmental Monitoring and Assessment, 196(2).

Zong, X., Liu, X., Chen, G., & Yin, Y. (2022). A deep-understanding framework and assessment indicator system for climate-resilient agriculture. Ecological Indicators, 136, 108597.

Appendices

Appendix 1: Measurement: Yield estimation/Yield prediction

Reference/ paper Title	Smallholder maize yield estimation using satellite data and machine learning in Ethiopia
Remote sensing method OR Used method	Machine learning models used to estimate smallholder maize yield in Ethiopia, leveraging well-measured, broadly distributed ground-truth data and remote sensing.
Used input	remotely sensed data, Sentinel 2, rainfall data, temperature data (The Climate Hazards Group Infrared Precipitation with Station data (CHIRPS)), soil properties, soil moisture (NASA-USDA SMAP Global soil moisture dataset), and yield data from 2017 to 2018.
Expected accuracy	The scatter plot of observed and predicted yields (kg/ha) using the neural networks for both training R2=0.79 and for Testing R2=0.62
Can be used as an index or part of an index	Yes can be considered as an index or part of an index
Adaptation or resilience/ sustainability	Resilience

Reference/ paper Title	Maize Yield Estimation in Intercropped Smallholder Fields Using Satellite Data in Southern Malawi
Remote sensing method OR Used method	Estimation of crop yield in smallholder farming systems, using the currently available optical EO data. In particular, this study explored the reliability of PlanetScope and Sentinel-2 EO data for estimating maize yield in intercropped smallholder fields in Southern Malawi. Collection of in situ data on crop field characteristics (e.g., field size, intercropping practices, yield, biomass, harvest index, and LAI) from 150 maize-dominated fields in the 2019/2020 seasons. Empirical linear regression models linking field-measured yields to VIs/LAI derived from Sentinel-2 and PlanetScope data were developed to estimate maize yield.
Used input	150 maize fields were investigated: *Global positioning system (GPS) coordinates of the corners of all selected field plots were recorded and used to calculate the field size. *The LAI was measured in all 150 fields across the growing season *Maize biomass and yield were measured during the harvest season in late March. The intercrop yield was not measured due to varying planting densities and long harvest periods spanning 1–2 months

	*In situ LAI was used to evaluate the satellite-derived LAI *The ratio between yield and maize biomass was calculated to determine the harvest index, showing the proportion of the total aboveground biomass contributing to yield *Official Yield Data and Maize Mask Dataset *PlanetScope and Sentinel-2 satellite data were used: VIs, which proved capable of estimating crop yield in smallholder farming systems, were investigated *In addition to the LAI measurements from the DHP images, the LAI was also retrieved by using the radiative transfer model-based Simplified Level 2 Product Prototype Processor (SL2P) algorithm
Expected accuracy	Despite mixed pixels resulting from intercrops, the model based on the Sentinel-2 red-edge vegetation index (VI) could estimate maize yield with moderate accuracy ($R2$ = 0.51, nRMSE = 19.95%). At the same time, higher-spatial-resolution satellite data (e.g., PlanetScope 3 m) showed only a marginal performance improvement ($R2$ = 0.52, nRMSE = 19.95%). Seasonal peak VI values provided better accuracy than seasonal mean/median VI, suggesting that peak VI values may capture the signal from the dominant upper maize foliage layer and be less affected by understory intercrop effects. Still, intercropping reduces model accuracy, as model residuals are lower in pure maize fields (1 t/ha) than in intercropped fields (1.3 t/ha).
Can be used as an index or part of an index	Yes, prediction and estimation of maize and intercropped maize yields.
Adaptation or resilience/ sustainability	resilience

Reference/ paper Title	Smallholder maize area and yield mapping at national scales with Google Earth Engine
Remote sensing method OR Used method	• **For classification** The study aimed to classify maize areas by first isolating cropland (CRL) before performing maize-specific classification (MZL) across Kenya and Tanzania. The process used phenological and spectral data from Sentinel-1 and Sentinel-2, divided into three seasonal composites within each water year (October–September). For the crop classification: • **Feature Selection**: Phenological and spectral features were initially scored for relevance using mutual information (MI) and further refined by discarding highly correlated features (above a correlation threshold of 0.8). • **Classification methodology**: A Random Forest classifier was applied, using cross-validated grid searches to select optimal hyperparameters and features. The CRL map was used to mask out non-cropland areas, thereby focusing the MZL map on cropland. • **For yield estimation**

	The SCYM framework predicts maize yield by combining satellite data and simulated training data generated by the APSIM crop model. Unlike traditional methods that require actual field measurements, SCYM utilises simulated training data (seasonal phenology and yields) across a wide range of environmental conditions. This approach provides scalability and adaptability for different regions. **Improvements in SCYM:** 1. **Generalised Simulation Protocol**: The new SCYM version offers scalability for larger regions by streamlining data requirements and applying uniform protocols across areas. 2. **Peak Vegetation Index (VI) Modelling**: Instead of using the maximum VI across two seasonal windows, a parametric fitting approach calculates the peak VI. This method leverages high-temporal Sentinel-2 data, enhancing prediction accuracy by capturing the growth pattern more comprehensively. **Harmonic-Derived Peak VI** For each growing season, cloud-free Sentinel-2 GCVI data were analysed using a harmonic regression (Discrete Fourier Transform) to model crop growth over time. By decomposing the seasonal GCVI into the first two harmonics, the method captures essential growth dynamics without overfitting. This harmonic regression optimises yield predictions based on the summarised temporal pattern of canopy growth. **Yield Estimation Process** The yield estimation follows a two-step model to account for multicollinearity (strong correlations) between temperature, precipitation, and GCVI: 1. **Yield and GCVI Prediction Models**: Separate regression models are trained to relate yield and peak GCVI to key weather variables. 2. **Residual Modelling**: Residuals from both models are then combined in a secondary regression to estimate yield based on GCVI, enabling accurate yield prediction while mitigating collinear effects.
Used input	• **For classification** ✓ **Sentinel-1 (Radar) and Sentinel-2 (Optical) Imagery**: Sentinel-1 provided VV and VH radar bands for all-weather imaging, and Sentinel-2 offered multispectral bands for Vegetation Indices (NDVI, SWIR). ✓ **Feature Stacking**: Seasonal composites of Sentinel-1 and Sentinel-2 images were stacked, resulting in 24 Sentinel-1 features and 69 Sentinel-2 features. ✓ **Ground Data for Training and Validation**: Crop cuts were conducted for ~4,000 fields in each year, with a GPS point collected in the middle of each field by One Acre Fund (1AF). Training datasets consisted of labelled polygons from both countries. An 80% split for validation and 20% for testing was maintained, with geographic diversity achieved through district-based splitting. • **For yield estimation** **Training Dataset from APSIM Simulations**

	For yield prediction, the study area was divided into agroecological zones using the GAES map, which classifies regions by climate, soil, and vegetation. APSIM simulated crop growth for each zone across a variety of environmental conditions:
	
	• **Simulation Setup**: To capture environmental variability, APSIM simulated maize growth using local data on weather, soil, and crop phenology.
	• **Environmental Inputs**:
	    o **Weather Data**: From 2009-2017, daily temperature, precipitation, and solar radiation were sourced from aWhere, a weather platform with high-resolution data.
	    o **Soil Profiles**: Harvest Choice's HC27 soil profiles provided critical soil characteristics like texture and fertility.
	    o **Crop Cultivars and Sowing Dates**: Local maize varieties were assigned based on APSIM's cultivar map, with planting dates adjusted to mimic farming practices.
	    o **Simulation Variations**: Random adjustments were made to soil moisture, seeding, and fertilisation to reflect realistic farming scenarios.

Expected accuracy	• **For classification** ✓ **Cropland Classification (CRL)**: Achieved an accuracy of 85% for both countries, with the most important features for Kenya involving greenness indices and for Tanzania involving SWIR bands (reflecting soil wetness). ✓ **Maize Classification (MZL)**: Accuracy was slightly lower, with 67% for Kenya and 79% for Tanzania. The SWIR and NDVI features were most crucial for distinguishing maize, particularly in Kenya, where cloud cover posed challenges • **For yield estimation** SCYM estimates captured about 50% of the variation in district-level yields in Western Kenya, as measured by objective ground-based crop cuts.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience

Reference/ paper Title	Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa
Remote sensing method OR Used method	In this study, we investigated the utility of different sensors, including the commercial Skysat and RapidEye satellites and the publicly available Sentinel-2, for tracking smallholder maize yield variation across a ~40,000 km2 region in western Kenya. We tested the potential of two types of multiple regression models for predicting yield: (i) a "calibrated model", which required ground-measured yield and weather data for calibration, and (ii) an "uncalibrated model", which used a process-based crop model

	to generate daily vegetation index and end-of-season biomass and/or yield as pseudo training samples.
Used input	• **Field Data**: Household surveys, GPS-based plot outlines, crop cut data. • **Weather Data**: Historical and forecasted data from aWhere for the 2016 growing season. • **Satellite Imagery**: SkySat (2-m), RapidEye (5-m), and Sentinel-2 (10-m) images. • **Vegetation Indices**: NDVI, EVI, WDRVI, GCVI, ReCI, MTCI. • **Land Surface Temperature (LST)**: MODIS Aqua MYD11A2 1-km data. • **Soil Profiles**: Harvest Choice's HC27 soil profiles provided critical soil characteristics like texture and fertility
Expected accuracy	Results show that the "calibrated" approach captured a significant fraction (R2 between 0.3 and 0.6) of the variation in yield at aggregated administrative units (e.g., districts and divisions). In contrast, the "uncalibrated" approach performed only slightly worse. For both approaches, we found that predictions using the MERIS Terrestrial Chlorophyll Index (MTCI), which includes the red-edge band available in RapidEye and Sentinel-2, were superior to those from other commonly used Vegetation Indices.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Resilience

Appendix 2: Impact of mixed/intercropping system on yield

Reference/ paper Title	Comparative analysis of monocropping and mixed cropping systems on selected soil properties, Soil Organic Carbon stocks, and simulated maize yields in drought-hotspot regions of Rwanda, changing climate in Kayonza District, Eastern Province of Rwanda.
Remote sensing method OR Used method	This study compared monocropping and mixed cropping systems based on selected soil properties, SOC stocks, simulated maize biomass, and grain yields.
Used input	Data originated from the soil laboratory and soil profile study. Secondary data on climate trends for the 20 years (2001–2021) were collected from the Rwanda Meteorology Agency (RMA), and maize yield data were from the Rwanda Agriculture and Animal Resources Board (RAB), National Institute of Statistics of Rwanda (NISR) and Sector Agronomists, respectively. Herewith, the measured maize crop yield data of the past 20 years (2001–2021) were collected from the RAB station in Kayonza District, while observed maize yield data were from the NISR and sector agronomists.

Expected accuracy	The mixed cropping system was also criticised for producing poor sole crop yield, but many studies confirmed that it sustainably increases diverse yields and income. As shown by the results of this study, the selected soil chemical properties showed that the mixed cropping systems significantly increased soil pH, SOC, TN, C: N ratio, and SOC stocks compared to the monocropping systems. Mixed cropping had greater C sequestration and CO2 mitigation potential than monocropping because it stored a much larger amount of C in the soil. This improved soil health and quality under mixed cropping systems. Again, mixed cropping systems showed a highly significant increase in maize biomass and grain yields compared to monocropping systems. It is therefore that mixed cropping systems have proven to be more sustainable, stable, and beneficial to smallholder farmers than monocropping systems in the face of drought hazards in the Kayonza district.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Sustainability

Appendix 3: Yield estimation using GPP

Reference/ paper Title	MAIZE YIELD ESTIMATION IN KENYA USING MODIS
Remote sensing method OR Used method	Estimation of county-level maize yield for the 37 maize-producing counties in Kenya from 2010 to 2017 using Moderate Resolution Imaging Spectroradiometer (MODIS) data. Support Vector Regression (SVR) and Random Forest (RF) were used to model observed county-level maize yield as a function of Vegetation Indices.
Used input	Five MODIS Vegetation Indices were used: green normalised difference vegetation index, Normalised Difference Vegetation Index, normalised difference moisture index, gross primary production, and fraction of photosynthetically active radiation.
Expected accuracy	GPP was the most important metric in maize yield prediction, followed by NDVI. All the metrics show an asymptotic relationship with maize yields. Maize yields increase linearly with NDVI from around 0.1 to 0.5, corresponding to maize yields between 0 and 2 tons/ha. From 0.5, NDVI rises sharply to 0.7, corresponding to yields of 2–5 ton/ha.
Can be used as an index or	Yes: important relation between yield and NDVI, especially if we want to use Sentinel 2. The equation was not defined

part of an index	
Adaptation or resilience/ sustainability	Resilience

Reference/ paper Title	Estimating evapotranspiration and yield of wheat and maize croplands through a remote sensing-based model
Remote sensing method OR Used method	• The Remote Sensing-based cropland Ecohydrological Model (ReSEM) was created by integrating the PML (Penman-Monteith-Leuning) model with a crop growth module. • The PML model estimated ET and GPP using remote sensing canopy characteristics and meteorological inputs, while the crop growth module simulated crop biomass accumulation and yield based on the GPP series. • The PML framework utilised stomatal conductance and canopy assimilation models to simulate ET and GPP, taking into account environmental variables and soil moisture constraints. • GPP was used to calculate net primary productivity (NPP), which informed the crop growth module to simulate biomass and yield. • The model incorporated specific partitioning coefficients for different crop parts (e.g., leaves, stems, grains, roots), calibrated with in-situ data from the Yucheng site.
Used input	**Experimental Data**: • **Flux Tower Observations**: ET, GPP, net ecosystem exchange (NEE), sensible and latent heat, CO_2 and H_2O fluxes from Hetao, Weishan, and Yucheng sites. • **Crop Biomass and Growth Measurements**: Dry biomass, leaf area index (LAI), crop yield, and biomass partitioning at Hetao and Yucheng sites. **Meteorological Data**: • Air temperature, relative humidity, air pressure, precipitation, wind speed, and radiation from local meteorological stations. **Remote Sensing Data from MODIS.** • Canopy characteristics such as LAI, albedo, and emissivity. • Net radiation calculated using albedo, emissivity, and radiation data. **Model-Specific Inputs**: • Stomatal conductance, canopy gross assimilation coefficients, and vapour pressure deficit terms for simulating crop-water dynamics in the PML model. • Crop-specific parameters calibrated for maize and wheat to accommodate differences between C3 (wheat) and C4 (maize) photosynthesis.
Expected accuracy	Comparing with 20 years of measurements, the Root Mean Square Error (RMSE) of ET and GPP for winter wheat was 0.57 mm d-1 and 1.65 g C m- 2 d- 1, and for maize was 0.80 mm d- 1 and 2.92 g C m- 2 d-1, respectively. Besides, crop growth simulation,

	including biomass accumulation and allocation, agreed well with the measurements. Coefficient of determination was mostly larger than 0.66, and RMSE of yield was 554.7 and 1346.6 kg hm-2 for wheat and maize, respectively. Finally, the spatiotemporal Crop Water Productivity (CWP) of winter wheat and summer maize in the North China Plain during 2001–2018 was quantified. An increased Water Productivity (WP) in winter wheat, compared with the larger WP of maize, was observed. The developed model helps to examine regional crop ET and yield and provides critical information on agricultural water management.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Adaptation/resilience

Reference/ paper Title	Regional Crop Gross Primary Productivity and Yield Estimation Using Fused Landsat-MODIS Data
Remote sensing method OR Used method	We used a satellite data-driven LUE modelling approach based on the MODIS MOD17 algorithm to derive relatively high spatial (30-m) resolution GPP and crop yield predictions across the region Moreover, spanning the recent (2008-2015) satellite record. A data fusion approach was used for blending 30-m Landsat 5 and Landsat 7 NDVI with 250-m eight-day MODIS NDVI observations to produce harmonised 30-m eight-day fused NDVI data record for Montana extending from 2008 to 2015. The 30-m NDVI record was used as a primary LUE model input to estimate fPAR and GPP at a similar 30-m resolution and eight-day time step for the major Montana crop types. Both the LUE modelling framework and satellite NDVI data fusion were implemented within the Google Earth Engine (GEE) framework for this investigation. We derived 30-m annual crop yield maps encompassing Montana cropland areas using the resulting 30-m GPP record and crop-specific HI coefficients calibrated and validated using NASS county-level crop production data and field-level yield data.

Used input	
Expected accuracy	The resulting GPP estimates capture characteristic cropland productivity patterns and seasonal variations. In contrast, the estimated annual crop production results correspond favourably with reported county-level crop production data (r = 0.96, relative RMSE = 37.0%, p < 0.05) from the U.S. Department of Agriculture (USDA). The performance of estimated crop yields at a finer (field) scale was generally lower, but still meaningful (r = 0.42, relative RMSE = 50.8%, p < 0.05). Our methods and results are suitable for operational crop yield monitoring at regional scales, suggesting the potential of using global satellite observations to improve agricultural management, policy decisions, and regional/global food security.
Can be used as an index or part of an index	
Adaptation or resilience/ sustainability	

Appendix 4: Measurement: Crop classification: classification between maize monocrop and intercropped

Reference/ paper Title	Characterising maize and intercropped maize spectral signatures for cropping pattern classification
Remote sensing method OR Used method	The goal is to identify the optimal crop growth phases, spectral regions and Vegetation Indices (VIs) that can accurately discriminate the two cropping patterns. p-values were computed for the spectral bands using the Mann-Whitney U test, and critical crop growth phases were identified. Classification of maize and imaize cropping patterns was performed using a Random Forest classifier.
Used input	This study integrates field surveys, farmer interviews, and temporal Sentinel-2 data across four crop growth phases and the post-harvest period for maize and intercropped maize (imaize).
Expected accuracy	The most suitable VIs contained red-edge and near-infrared spectral bands. Utilising spectral data and VIs from vegetative and flowering-yield phases, we achieved optimal discrimination during the vegetative phase (user's accuracy of 100 % and producer's accuracy of 100 %). However, accuracy decreased during the flowering yield phase (Overall Accuracy of 87% across all spectral bands). The highest classification results using all spectral bands at the flowering yield phase resulted in 80 % producer's accuracy for maize and 100 % for imaize.
Can be used as an index or part of an index	Yes

Adaptation or resilience/ sustainability	Resilience

Reference/ paper Title	Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery
Remote sensing method OR Used method	We first produced a **land-cover map to exclude other land-cover types and create a cropland mask for our crop-type mapping**. Model building and prediction were carried out in FORCE using the 2779 training points to train an RF model, with a selected set of 155 features obtained from recursive feature elimination (from the 770 Spectral Temporal Metrics (STMs)) and 500 trees in the RF. **We first generated five land cover classes: bare ground, cropland, water, vegetation, and built-up. After that, performing crop typing to map the main crops—maize and potato—and intercropping systems including these two crops** on the Jos Plateau, Nigeria.

Used input	**Image Data and Preprocessing**
	• **Data Sources**:
	  o Integrated Sentinel-2 (S2) imagery (A and B) and VHR SkySat imagery for validation.
	  o Downloaded S2 Level 1C data for 2019 with cloud cover < 70%, totaling 604 scenes across four footprints.
	  o SkySat imagery for four 25 km² areas of interest during five critical periods of the 2020 cropping season.
	  o Acquired Landsat Collection 1 Level 1 observations (2015-2020) to enhance multi-temporal co-registration of S2 data.
	• **Preprocessing Steps**:
	  o Used the FORCE (Framework for Operational Radiometric Correction for Environmental monitoring) tool for:
	    ▪ Atmospheric correction
	    ▪ Cloud masking
	    ▪ Topographic correction
	    ▪ Tiling
	    ▪ Image co-registration via the LSReg 2.0 algorithm, utilising Landsat NIR bands for co-registration over five years.
	Reference Data
	• **Field Data Collection**:
	  o Collected during the 2019 and 2020 cropping seasons.
	  o Conducted 119 semi-structured interviews focusing on farmers' cropping practices.
	  o Collected 530 samples of potato and maize fields in 2019 (180 potato samples, 352 maize samples).
	  o 2020 campaign added 690 samples (181 potato samples, 509 maize samples).
	• **Sample Distribution**:
	  o **Table 1: Reference Crop Type Samples** shows the distribution of samples for various crop types across years, indicating a total of 2779 samples:

	▪ Maize: 613 ▪ Potato: 338 ▪ Potato–maize: 673 ▪ Maize–legumes: 63 ▪ Others: 519 **Phenology and Spectral-Temporal Metrics** • **Crop Phenology**: o Defined four critical windows (CW-1 to CW-4) for mapping crop types in smallholder systems: ▪ **CW-1**: Sowing/emergence of potatoes (late March to early May). ▪ **CW-2**: Sowing/emergence of maize (3-5 weeks after potatoes). ▪ **CW-3**: Harvesting of potatoes. ▪ **CW-4**: Maize anthesis/maturity and legumes introduction. **Spectral-Temporal Metrics (STMs)** 1. **Aggregation**: STMs are created by aggregating all available observations within defined temporal windows into various statistical metrics. In this study, five critical time periods were defined for analysis: a. **Entire Calendar Year**: Days 001 to 365. b. **Entire Cropping Period**: Days 091 to 304. c. **Start of Season**: Days 121 to 181. d. **Peak of Season**: Days 213 to 273. e. **Off-Season**: Days 304 to 090. 2. **Calculated Metrics**: A total of 11 pixel-based statistics were calculated for each spectral band and selected Vegetation Indices: Minimum (Min), 25% Quantile (Q25),50% Quantile (Q50), 75% Quantile (Q75), Maximum (Max), Average (AVG), Standard Deviation (STD), Range (RNG), Inter-Quartile Range (IQR), Skewness (SQW), and Kurtosis (KRT). STMs were produced for each of the nine S2 spectral bands as well as the following indices: (i) Normalized Difference Vegetation Index (NDVI); (ii) Enhanced Vegetation Index (EVI); (iii) Normalized Difference Water Index (NDWI) and the Tasseled Cap transformations;(iv)Wetness (TCW); (v) Greenness (TCG); and (vi) Brightness (TCB). 3. **Total Features**: In total, 770 image features were produced from the different temporal windows and spectral bands.
Expected accuracy	Cropland identification achieved an Overall Accuracy of 84%, while the crop type map averaged 72% accuracy across the five relevant crop classes. Our crop-type map shows distinctive regional variations in crop distribution. Maize is the dominant crop, followed by mixed cropping systems, including maize–cereals and potato–maize cropping; potato was the least prevalent class.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience

Reference/ paper Title	Harmonised NDVI time-series from Landsat and Sentinel-2 reveal phenological patterns of diverse, small-scale cropping systems in East Africa.
Remote sensing method OR Used method	The methodological workflow consisted of four components (Fig. 3): (1**) preprocessing of satellite imagery (Sentinel 1 and 2) for Land Use Land Cover (LULC) classification; (2) LULC classification: derive a thematic LULC map for the study area based on comprehensive in situ crop data**; (3) sensor harmonization; and (4) phenological metrics extraction: over one season from the harmonized Landsat sensors and Sentinel-2 product using localized land use information. A machine learning approach was selected for land-use/land-cover classification. We used supervised Random Forest (RF). Different training-validation split ratios were tested to find the optimal split. The training data was split, **with 70% used for training and the remaining 30% for validation**. Stratified sampling was used to prevent disproportionate representation of classes during data splitting. Finally, a mask layer was derived containing the land use class, which included the **following classes of monocultures: sweet potato, rice, cassava, banana, cotton, coffee, maize and sugarcane; as well as the following intercrops: maize and groundnuts, banana and coffee, maize and cassava, maize and coffee, maize and beans.**
Used input	Remote sensing data for classification: Sentinel 1 and 2 Remote sensing data for harmonisation: the Landsat and Sentinel-2 images Ground Truth data: A field campaign in the study area took place in July of 2022. During this time, the growing season was at its peak vegetative state, allowing us to accurately identify the most prominent land cover and crop types in the area. Corner and centre coordinates of each agricultural field were captured using the GPS Coordinates App Version 4.71 (174). The data-gathering campaign concluded with farmer interviews to better understand agricultural practices in the area and farmers' perspectives on the local climate and its impact on farming.
Expected accuracy	The LULC classification yielded an Overall Accuracy of 87.9% and Cohen's kappa value of 0.83.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience

Appendix 5: Measurement: Drought detection

Reference/ paper Title	Analysis of Short-Term Drought Episodes Using Sentinel-3 SLSTR Data under a Semi-Arid Climate in Lower Eastern Kenya
Remote sensing method OR Used method	This study uses Sentinel-3 SLSTR data to analyse short-term drought events between 2019 and 2021. It investigates the crucial roles of vegetation cover, land surface temperature, and water vapour amount in drought over Kenya's lower eastern counties. Therefore, three Essential Climate Variable (ECVs) of interest were derived: Land Surface Temperature (LST), Fractional Vegetation Cover (FVC), and Total Column Water Vapour **(TCWV)**. These features were analysed across four counties during the wettest and driest episodes in 2019 and 2021. This study explores short-term drought episodes using Sentinel-3 SLSTR data in Kenya's lower eastern counties. The aims are to (i) analyse and interpret Sentinel-3 SLSTR-derived ECVs (LST, FVC, and TCWV) between wet and dry seasons; (ii) investigate Sentinel-3A SLSTR data usefulness in short-term moisture variation between 2019 and 2021; and (iii) investigate the association between soil moisture and short-term drought in eastern Kenya. **Part of the results on FVC: The study showed that Makueni and Taita Taveta counties had the highest FVC density** (60–80%) in April 2019 and 2021. Machakos and Kitui counties had the lowest FVC estimates, ranging from 0% to 20% in September for both periods and from 40% to 60% during wet seasons. As FVC is a crucial land parameter for carbon sequestration and for detecting losses in soil moisture and vegetation density, its variation is strongly related to drought magnitude. Taylor et al. [111] demonstrated FVC's fundamental role in determining changes in diverse ecosystems. Thus, vegetation responds rapidly to rainfall and moisture anomalies, influencing crop yield and food security [17]. The magnitude of vegetation density change contributes to environmental improvement in the arid and semiarid lower eastern regions of Kenya, as well as in other parts of the country.
Used input	Sentinel-3 SLSTR Data: The Sentinel-3 SLSTR product generates land surface parameters with a one km spatial resolution. It includes a measurement file with computed (i) Land Surface Temperature (LST) values for each pixel, associated with ancillary parameters, namely (ii) Normalised Difference Vegetation Index (NDVI), (iii) Glob Cover surface classification code (biome), (iv) Fractional Vegetation Cover (FVC), and (v) Total Column Water Vapour (TCWV). They used the 'Create Random Points' tool in ArcMap 10.3 to randomly generate a sample of 50,000 points from the entire region of interest to establish a correlation between ECVs. Once samples with null values were omitted, only 44,747 sampling points were left. LST, FVC, and TCWV were compared for 2019 and 2021.
Expected accuracy	Overall, Sentinel-3 SLSTR products provide an efficient and promising data source for short-term drought monitoring, especially when in situ measurement data are scarce. ECV-produced maps will assist decision-makers in better understanding short-term

	drought events. The study is also fundamental to future climate-variability preparedness in drought-prone or risk counties, resilience planning, and policy-making.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience/adaptation

Reference/ paper Title	A comprehensive assessment of remote sensing and traditional-based drought monitoring indices at global and regional scales **(Review)**
Remote sensing method OR Used method	Most drought indices calculated from surface reflectance or backscatter image bands are considered basic indices. The other category is the intermediate index, which is calculated using the numerical parameters of the minimum, maximum, average, or standard deviation of the long-term variables of these basic indices. The final category is the composite index, generated by combining the various basic or intermediate indices with the optical, thermal, and microwave remote-sensing drought indices. Vegetation Drought Response Index (VegDRI) is a hybrid drought index produced by combining satellite data-based calculations on climatic, biophysical features, and vegetation conditions. The concept of VegDRI was introduced by Rouse et al. (1974) and is based solely on NDVI, and it has proven to provide valuable information on plant health. This indicator represents vegetation stress associated with drought and can be used to perform calculations at regional scales with medium- to high-resolution. However, it is also possible that vegetation stress cannot be identified from NDV variability alone (Heim, 2002). The VTCI, TVDI, TWI, and TVI indices are designed using the scatter-plot approach (Triangular or trapezoidal) for temperature and Vegetation Indices, as shown in Figure 6. This approach mostly works at the regional scale, but not at the small scale. The VTCI, TVDI, TWI, and TVI indices combine NDVI (vegetation) and LST (thermal) characteristics and can be used to monitor drought based on temperature and vegetation conditions effectively. The Scaled Drought Severity Index (SDSI) was developed by Rhee et al. (2024) and is used as a multi-sensor index to monitor agricultural drought in arid, semi-arid, and humid areas. Here, the three derivatives of vegetation, precipitation, and temperature, such as VCI, Precipitation Condition Index (PCI), and TCI, are combined on a scale from 0 to 1. The Vegetation Soil-Water Deficit (VSWD) method differs from other drought index methods in that it combines precipitation, soil moisture, and Potential Evapotranspiration (PET) to generate a composite drought index. Therefore, using the VSWD index can identify the drought with greater accuracy.
Used input	Remote sensing (optical, thermal, microwave), AET – Actual Evapotranspiration, PET –

	Potential Evapotranspiration, P – Precipitation, T – Temperature, SM – Soil Moisture
Expected accuracy	A review of the most popular drought monitoring platforms used for the study also found that 90% of those platforms provide weekly or monthly drought information using Standardised Precipitation Index (SPI) indices, and about 60% use NDVI or its derivatives.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	adaptation

Reference/ paper Title	A comprehensive remote sensing-based Agriculture Drought Condition Indicator (CADCI) using machine learning
Remote sensing method OR Used method	Develop an indicator to monitor agricultural drought in semi-arid environments that incorporates all relevant factors influencing drought conditions, including LST, vegetation density, precipitation, SM, vegetation health, and ET. Additionally, using a more appropriate integration technique improves the reliability of the results and reduces the errors. In this study, the Comprehensive Agriculture Drought Condition Indicator (CADCI) was developed to monitor agricultural drought in semi-arid environments. It was conducted procedurally using spectral indices that represent drought conditions (i.e., VCI, TCI, VHI, PCI, SMCI, and ETCI) and employed the RF technique to identify the three spectral indices with the highest monthly short-term average relative importance for identifying agricultural drought in semi-arid environments. A methodology for developing a comprehensive remote sensing-based agriculture drought conditions indicator (CADCI) for semi-arid areas was developed. It had three primary components: (i) data collection and its preprocessing, (ii) development of the CADCI, and (iii) validation and evaluation of CADCI.
Used input	remote sensing data from several satellite systems (i.e., MODIS, SMAP/Sentinel, and GPM), and the second contains the climate data (i.e., precipitation) from agro-climate stations and satellite data
Expected accuracy	The SPI sets for 1 and 3 months (SPI-1 and SPI-3) were used to evaluate the CADCI's performance. The results showed that CADCI agrees well with SPI-1 classes in the study areas, with overall accuracies and Kappa values of 85% and 0.80 for study area A (Jordan) and 83% and 0.76 for study area B (Syria), respectively. Consequently, the CADCI shows its ability to monitor agricultural drought in semi-arid environments.

Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Resilience/adaptation

Appendix 6: Vegetation Condition Index

Reference/ paper Title	Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya
Remote sensing method OR Used method	The main goal of this paper is to explore machine learning techniques for forecasting Vegetation Indices commonly used in pastoral areas of Kenya to monitor drought. We focused in particular on the pastoral livelihood zones, as the VCI3M is more reliable at identifying drought conditions for grazing and browsing in the country's more arid regions. Here, we specifically focus on Gaussian Process (GP) modelling and linear autoregressive (AR) modelling to forecast NDVI and VCI3M, which are derived from both Landsat (every 16 days at 30 m resolution) and the Moderate Resolution Imaging Spectroradiometer. Forecasts of NDVI anomaly and VCI3M were made using two separate methods: Gaussian Process (GP) modelling and linear autoregressive (AR) modelling. GP forecasting was performed by fitting a GP to the forecast-mode aggregate time series for the index in question, and then using the GP to extrapolate.
Used input	Landsat and MODIS
Expected accuracy	As a benchmark, we predicted the NDMA drought alert marker (VCI3M < 35). Both of our models were able to predict this alert marker four weeks ahead, with a hit rate of around 89% and a false alarm rate of around 4% (81% and 6% six weeks ahead).
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	resilience

Appendix 7: Measurement: Detection of change in vegetation phenology dynamics

Reference/ paper Title	Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2
Remote sensing	We derived temporal NDVI profiles and fitted double hyperbolic tangent models (equivalent to commonly-used logistic functions), separately for the two rainy seasons locally referred to as the short and long rains. We estimated the start and end of the season for the series

method OR Used method	using a 50% threshold between the minimum and maximum levels of the modelled time series (SOS50/EOS50). We compared our estimates against those obtained from vegetation index series from two alternative sources, i.e. a) greenness chromatic coordinate (GCC) series obtained from digital repeat photography, and b) MODIS NDVI.
Used input	All available PlanetScope and Sentinel-2 imagery between March 2017 and February 2019
Expected accuracy	Both PlanetScope and Sentinel-2 series yielded acceptable phenology retrievals (RMSD of ~8 days for SOS50 and ~15 days for EOS50 when compared against the GCC series), suggesting that the sensors individually provide sufficient temporal detail. However, when applying the model to the entire study area, fewer spatial artefacts occurred in the PlanetScope results. This could be explained by PlanetScope's higher observation frequency, which becomes critical during periods of persistent cloud cover. We further illustrated that the PlanetScope series could distinguish the phenology of individual trees from grassland surroundings, with tree green-up occurring both earlier and later than that of grass, depending on location.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Resilience

Appendix 8: Measurement: Irrigation detection/ irrigated areas

Reference/ paper Title	Spatio-Temporal Patterns of Smallholder Irrigated Agriculture in the Horn of Africa Using Geographic Object-Based Image Analysis (GEOBIA) and Sentinel-2 Imagery
Remote sensing method OR Used method	The authors employed GEOBIA (Geographic Object-Based Image Analysis) for field-level segmentation using Sentinel-2 imagery. This approach allows for the identification of smallholder irrigated fields. They combined object-based segmentation with NDVI time series to differentiate between irrigated and rainfed agriculture, relying on the assumption that field boundaries do not restrict rainfall, while irrigation does. The study used process-based rules and three different spatial heterogeneity thresholds to assign vegetation growth patterns to either rainfall or irrigation. This enabled mapping monthly irrigated areas from September 2016 to August 2017. **eCognition Developer** is a specialised software tool designed for object-based image analysis (OBIA), commonly used in remote sensing and geographic information systems (GIS). Unlike traditional pixel-based image analysis, which classifies each pixel independently, OBIA groups neighbouring pixels into meaningful "objects" or segments based on their spectral, spatial, and contextual properties.

Used input	A mosaic of images was created. All available images between October 2016 and March 2017 (9065 images), which capture most of the dry season for the Horn of Africa, were sorted by the image-associated 'cloudy pixel percentage'. Monthly (September 2016 to August 2017) NDVI composites were computed by selecting the NDVI associated with the least-cloudy pixel from the Sentinel-2 collection in each month.
Expected accuracy	The study produced high-resolution maps (10 m) of irrigated agriculture. The results showed a significant increase in the estimated irrigated area compared to previous global estimates. For instance, the irrigated area estimates (27.96–37.13 Mha) were 2.8-3.7 times higher than those from global maps such as the Global Irrigated Area Map (GIAM). These findings suggest that traditional mapping efforts likely underestimate the extent of smallholder irrigation in the region. Cropland mapping can be validated through visual interpretation. However, rigorous validation of the spatio-temporal patterns of irrigated agriculture is not possible, as ground truth for the temporal dynamics of (smallholder) irrigation is unavailable and difficult to obtain. Accuracy statistics for the cropland classification are high: Overall Accuracy and kappa coefficient are 96% and 0.73, respectively, but decrease when monthly NDVI availability is reduced by cloud cover. As a result, some regions have better cropland classification than others.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience/Adaptation

Appendix 9: Measurement: Vegetation health

Reference/ paper Title	Use of time series Normalised Difference Vegetation Index (NDVI) to monitor fall armyworm - FAW (Spodoptera frugiperda)damage on maize production systems in Africa.
Remote sensing method OR Used method	The objective of this study is to use multi-date Landsat 8 images and NDVI datasets to assess the spatial and temporal status of FAW damage in maize-growing areas, covering pre- and post-FAW infestation periods. Additionally, the study seeks to measure the crop damage caused by FAW in both the West and South African study areas. **Step 1:** Data Acquisition and Preprocessing Landsat 8 imagery was used as the main source of remote sensing data. The study used Google Earth Engine (GEE) to process the images, specifically performing: –2020) was created using median pixel values. **Step 2:** NDVI Calculation

	The NDVI values were computed and then reclassified into four vegetation density classes (low, sparse, moderate, dense) using QGIS, with thresholds specific to each study region. **Step 3:** FAW Damage Identification The FAW infestation years were identified from the literature, and NDVI changes before and after FAW outbreaks were analysed for each region. The study used the Natural Breaks classification technique to reclassify NDVI values based on vegetation density. **Step 4:** Damage Quantification A comparison of annual NDVI values before and after FAW infestations was conducted to estimate the crop damage. This change was expressed in hectares and compared to a baseline (2013), a year with high NDVI values. **Step 5:** Validation Field data on FAW prevalence and crop types were obtained from the Food and Agriculture Organisation (FAO) for the year 2018/2019. The FAW-affected points were overlaid on NDVI raster datasets, and NDVI values were extracted for validation. Linear regression was used to assess the correlation between FAW prevalence and NDVI reductions based on scouting data.
Used input	Landsat 8 and Field data on FAW prevalence and crop types were obtained from the Food and Agriculture Organisation (FAO)
Expected accuracy	NDVI can be utilised as a proxy for monitoring FAW-induced crop damage. The validation results showed a correlation between FAW infestation and NDVI (R20.83).
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Sustainability

Reference/ paper Title	Regional Monitoring of Fall Armyworm (FAW) Using Early Warning Systems
Remote sensing method OR Used method	The methodology revolves around analysing Sentinel-2 NDVI time series to monitor maize fields before and after Fall Armyworm (FAW) infestation across Africa. It compares NDVI values from 2015 and 2018, using first-derivative analysis to detect changes in vegetation health, showing a negative correlation between FAW infestation and NDVI. The study uses cloud-free mosaics and Fall Armyworm Monitoring and Early Warning System (FAMEWS) app data to map infestation levels, revealing high accuracy in correlating NDVI declines with FAW severity, thereby aiding the development of an early warning system.

Used input	**FAW Sampling Protocols** • Data collected from farmers or extension officers included sowing dates, crop growth stage (vegetative or vegetative-reproductive), farming systems (seasonal, rotation, or intercropping), intercropped crops (if any), and pest control measures (pesticides, IPM, or no treatment). • **FAW Infestation Levels**: Fields were classified as **low**, **medium**, or **high** based on FAW infestation. Monocropping fields often exhibited higher FAW infestations, while intercropping with common beans showed lower infestations. **Satellite Data** • Sentinel-2 imagery. • **Cloud-Free Mosaics**: GEE cloud-masking algorithms produced cloud-free mosaics by combining multiple images. This enabled the generation of NDVI and time-lagged NDVI derivatives to identify FAW impacts. • **NDVI Derivative**: ΔNDVI (first derivative of NDVI) was used to capture FAW impacts on maize fields by detecting reduced leaf area during early vegetative stages. • Commercial microsatellite PlanetScope. • UAV data. **Data Validation**: Field data from 2019 was used to validate satellite-based results, with additional contributions from the FAO FAMEWS app.
Expected accuracy	1. **Sentinel-2 Time Series Analysis:** • A comparison of NDVI time series averages before and after FAW invasion showed a logical negative correlation, with an R^2 **= 0.401** between the NDVI values and FAW infestation level. This suggests a moderate correlation. • For the 2018 fall season in Kenya, the analysis using the first derivative of the NDVI time series resulted in an R^2 **= 0.81** with FAW infestation levels, indicating a strong negative relationship. 2. **Multiscale Comparison (PlanetScope, multispectral camera, fisheye lens):** • For the LAI (Leaf Area Index) measured with a fisheye lens and NDVI from PlanetScope (nanosatellite), the R^2 **= 0.737**, showing a good level of agreement. • The LAI vs NDVI from a multispectral camera had an R^2 **= 0.617**, showing moderate correlation. • For GA index vs. NDVI from PlanetScope, the R^2 **= 0.936**, indicating an excellent correlation. • GA index vs NDVI from the multispectral camera showed an R^2 **= 0.708**, suggesting a good relationship. Despite these relatively high R^2 values, the study noted challenges with cloud cover, which affected the utility of satellite data and limited the reliability of some results, particularly for large-scale FAW monitoring. Nonetheless, the first-derivative NDVI analysis proved promising for better correlating with FAW severity and timing.

Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Sustainability

Appendix 10: Measurement: Soil Organic Carbon modelling & mapping/ Soil Organic Carbon forecasting

Reference/ paper Title	Open Remote Sensing Data in Digital Soil Organic Carbon Mapping: **A Review**
Remote sensing method OR Used method	Digital soil mapping is a process that involves creating spatial predictions of soil properties using various data sources and statistical or machine-learning methods. Three primary environmental covariate groups commonly used for digital SOC mapping were topography, climate, and spectral, along with other data sources.
Used input	
Expected accuracy	
Can be used as an index or part of an index	
Adaptation or resilience/ sustainability	facilitate environmental conservation by identifying and monitoring areas with high carbon sequestration potential, thereby contributing to climate change mitigation => Sustainability

Reference/ paper Title	Synergistic Use of Multi-Temporal Radar and Optical Remote Sensing for Soil Organic Carbon Prediction
Remote sensing method OR Used method	This study evaluates the suitability of the synergistic use of multi-temporal and high-resolution radar and optical remote sensing data for SOC prediction in the Kaffrine region of Senegal. For this purpose, various scenarios were developed: Scenario 1 (Sentinel-1 data), Scenario 2 (Sentinel-2 data), Scenario 3 (Sentinel-1 and Sentinel-2 combination), Scenario 4 (topographic features), and Scenario 5 (Sentinel-1 and Sentinel-2 with topographic features). The findings from comparing three algorithms (Random Forest (RF), XGBoost, and Support Vector Regression (SVR)) on 671 soil samples for training and 281 for model evaluation highlight that RF outperformed the other models across scenarios.
Used input	Time series of Sentinel 1 and 2, soil sample SOC and topographic data.

	Soil data: a stratified random sampling design was applied, and the entire study area was partitioned into blocks (10 × 10 km), enabling systematic organisation of the sampling effort and ensuring extensive coverage. Of these blocks, 45 were selected at random. Between 2018 and 2019, soil samples were collected at each site from the top 20 cm. After collection, the soil samples were transported to the laboratory for preparation and analysis. The preparation involved drying the soil, removing all plant debris, and sieving through a 2 mm mesh to obtain a uniform soil fraction for analysis. The SOC content was measured using the Walkley–Black method.
Expected accuracy	Including topographic features (Scenario 5) achieved the highest accuracy, reaching an R2 of 0.7, an RMSE of 0.012%, and an RPIQ of 5.754 for the RF model. **Remote sensing data:** VH, VV, Brightness Index (BI), Colouration Index (CI), Modified Normalised Difference Water Index (MNDWI), MERIS Terrestrial Chlorophyll Index (MTCI), Normalised Difference Vegetation Index (NDVI), Normalised Difference Water Index (NDWI), Redness Index (RI), and Soil-Adjusted Vegetation Index (SAVI). The digital elevation model was obtained from ASTGTM and used to extract topographic features, including elevation, slope, aspect, Topographic Wetness Index (TWI), profile curvature, plan curvature, and Multi-Resolution Index of Valley Bottom Flatness (MRVBF).
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Sustainability/Resilience Orienting sustainable land management practices can also contribute to improving agricultural resilience.

Reference/ paper Title	Digital soil mapping of nitrogen, phosphorus, potassium, organic carbon and their crop response thresholds in smallholder managed escarpments of Malawi.
Remote sensing method OR Used method	Random Forest (RF) Algorithm: The Random Forest (RF) machine learning model was employed to predict the spatial distribution of soil nutrients based on the environmental covariates. The RF model was trained using soil sample data as the response variables (NPK and SOC levels) and environmental covariates as predictors. The model output was a prediction of the nutrient levels at unsampled locations across the study area. After training the Random Forest model, it was used to predict soil nutrient levels for the entire study area. The model interpolated nutrient values for locations without direct soil sample data using relationships learned from covariates and sample data.
Used input	**Soil Sample Data**: Ground-truth soil data were collected via field sampling on smallholder farms. The samples were tested for nitrogen (N), phosphorus (P), potassium (K), and

	organic carbon (SOC) levels. This dataset formed the basis for training and validating the models. Environmental Covariates: Various spatial and environmental covariates were used as predictors in the modelling process. These included: **Remote sensing data**: Satellite imagery (Sentinel-2) was used to derive spectral indices, such as the Normalised Difference Vegetation Index (NDVI), which help understand soil properties linked to vegetation and moisture levels. **Terrain attributes**: Digital Elevation Models (DEMs) were used to derive slope, elevation, and other topographic features that can influence soil nutrient distribution. **Climate data**: Temperature and precipitation datasets were also integrated as covariates to capture climatic influences on soil nutrient levels.
Expected accuracy	Validation: The predicted maps were validated using a subset of soil samples not used in model training. This step involved calculating accuracy metrics, such as Root Mean Square Error (RMSE), to assess how well the model predicted nutrient levels at validation points. Threshold Mapping for Crop Response: Once the nutrient maps were generated, they were further analysed to determine crop response thresholds. These thresholds indicate nutrient levels at which crops would either suffer from deficiency or be sufficiently nourished, guiding optimal fertiliser application strategies. It was observed that the combination of all the covariates was the best specification for predicting SOC. Validation using an independent sample (3-fold) shows that the Random Forest model predicted SOC, TN, and P with moderate accuracy but had low prediction accuracy for K.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	sustainability

Appendix 11: Measurement: Irrigation performance indicators and water productivity

Reference/ paper Title	Monitoring small-scale irrigation performance using remote sensing in the Upper Blue Nile Basin, Ethiopia
Remote sensing method OR Used method	This study aimed to assess the performance of a small-scale irrigation scheme using remote sensing and ground-truth data for the 2021/22 and 2022/2023 irrigation seasons, employing the Shimburit irrigation scheme in Northwestern Ethiopia, predominantly cultivated with wheat, as a case study. The performance indicators, including equity, adequacy, overall consumed ratio (OCR), and productivity, were assessed. **Overview of the Approach**

	The study uses the SEBALI evapotranspiration model and field data to assess irrigation performance. Three primary steps are employed to derive performance indicators: • Estimate monthly reference evapotranspiration (ETo), actual evapotranspiration (ETa), and water productivity using SEBALI. • Calculate seasonal ETa, crop evapotranspiration (ETc), and productivity between the start (SOS) and end of the season (EOS). • Analyse irrigation performance indicators by integrating CHIRPS satellite precipitation data, SEBALI outputs, and ground-based data. **Study Area** The study focuses on the Shimburit small-scale irrigation scheme in the Upper Blue Nile Basin, covering 270 ha, constructed in 2015, benefiting 780 households. Wheat is the dominant crop. **1. Estimation of Evapotranspiration and Water Productivity** The SEBALI model estimates reference evapotranspiration (ETo), actual evapotranspiration (ETa), and Water Productivity (WP) at 30-m resolution. It incorporates the surface energy balance equation to calculate latent heat flux, improving satellite-based estimates by reducing input data requirements and automating pixel identification. **2. Validation of Results** The results from the SEBALI model were validated against those from SWAT+ hydrological models and FAO's WaPOR portal. Validation also used RMSE, MAE, and R^2 to assess accuracy. **3. Irrigation Performance Indicators** **Adequacy:** Evaluated using relative evapotranspiration (RET) and relative water supply (RWS). RET is calculated as the ratio of ETa to ETc, whereas RWS assesses water-supply sufficiency. **Equity:** Measured using the coefficient of variation (CV) of ETa, with thresholds defining good, fair, or poor equity based on irrigation distribution. **Depleted fraction** **Overall consumed ratio (Efficiency)** **Water and land productivity:** Measured by production per hectare and Crop Water Productivity.

Used input	The study uses remote sensing data combined with in-situ monitoring to assess irrigation performance:
	• Remote sensing data includes:
	  o Landsat 8 OLI-TIRS for surface reflectance, NDVI, and land surface temperature.
	  o Shuttle Radar Topography Mission (SRTM) for digital elevation data.
	  o MODIS for surface reflectance.

	○ ERA5-Land for climate data. ○ CHIRPS for precipitation estimates. • In-situ data includes discharge, crop yields, and field sizes collected via GPS.
Expected accuracy	The irrigation performance indicators revealed moderate accuracy in the Shimburit irrigation scheme. Coefficients of variation (CV) ranged from 1.90% to 6.34% in the 2021/22 season and 1.63% to 3.89% in 2022/23, indicating relatively low variability in water use. The scheme exhibited stable water storage (DFs of 0.65 and 0.60) and adequate water supply (RETs of 1.00 and 1.04) in both seasons. However, the on-crop ratio (OCR) was low at 0.54 and 0.43, indicating inefficiencies and excess water release during irrigation, particularly in the early and middle stages of the season.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Resilience/sustainability

Reference/ paper Title	Global Satellite-Based ET Products for the Local Level Irrigation Management: An Application of Irrigation Performance Assessment in the Sugarbelt of Swaziland
Remote sensing method OR Used method	1 Potential Evapotranspiration (PET): • PET, calculated using the Penman-Monteith equation, was compared to actual ET values to assess the adequacy of water use. It represents ideal growing conditions for well-watered crops. 2 Irrigation Performance Assessment: Four performance indicators were used to assess irrigation performance: 1. Equity: Measured the uniformity of water distribution using the spatial coefficient of variation (CV) of ET among different growers. 2. Adequacy: The adequacy of water was measured by comparing actual ET to PET (ET/PET ratio). Lower ET/PET ratios indicated water stress and yield reduction. It was used to predict yield reductions due to water stress. 3. Reliability: Assessed by examining the consistency of water supply across different growers, based on ET patterns. 4. Crop Water Productivity (CWP): Measured how effectively water is used for crop production, relating yield to water consumed. • Sugarcane yield benchmarks were defined for performance analysis: critical yield (80 TCH(tons of cane per hectare)), target yield (100 TCH), and maximum yield (130 TCH), and the corresponding ET/PET ratios were determined as 0.68, 0.8,

	and 1.0, respectively. This helped categorise performance into poor, acceptable, and operational levels.
Used input	**Evapotranspiration (ET)** data were derived from two satellite products: **SSEBop** (Operational Simplified Surface Energy Balance) and **CMRSET** (MODIS Reflectance-based ET), downscaled to 250m resolution. These ET products were combined into an **Ensemble map**, which was validated against irrigation water data, revealing an 11% overestimation bias. **Precipitation** data from the **CHIRPS** dataset (high-resolution, 5km grid) were used, based on its reliability in nearby Mozambique. **Farm boundaries** were obtained from multiple sources to accurately delineate sugarcane areas, grower clusters, and irrigation methods. **Sugarcane yield** data in tonnes of cane and sucrose per hectare were sourced from key industry bodies.
Expected accuracy	The accuracy of irrigation performance was evaluated using equity, adequacy, and reliability indicators. Equity was high among large-scale growers, but smallholders exhibited greater variability. Adequacy revealed critical water shortages in some regions, while reliability was assessed using relative ET, which showed seasonal variations, with large-scale growers performing above target levels. Irrigation performance assessment indicated high equity, especially in large-scale operations, while adequacy was below target in many regions, with smallholder farmers facing severe water shortages. Crop Water Productivity (CWP) was highest in centre pivot systems and estates, reaching 1.63 kg/m^3, while smallholder growers and furrow systems exhibited significantly lower CWP.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Resilience/sustainability

Appendix 12: Measurement: Soil fertility

Reference/ paper Title	Assessing Soil Fertility Index based on remote sensing and gis techniques with field validation in a semiarid agricultural ecosystem
Remote sensing method OR Used method	Our objective was to quantify the spatial distribution of soil properties and, accordingly, determine changes in soil fertility under intensive cultivation using ground-based data and remote sensing techniques for validation. For this purpose, a modern approach was presented using the SFI, which was designed by selecting parameters known to affect product yield in the arid ecology of central Anatolia. However, the efficiency classes determined as a result of the index need to be validated in order to be included in the current literature. For this reason, the research aimed to compare SFI (Soil Fertility Index) classes with both yield

	and vegetation index values derived from satellite images. Thus, the possibilities of using remote sensing technologies, as well as ground truth, were discussed for monitoring soil fertility.
Used input	The study region was divided into a 1 km2 grid, and a total of 281 soil samples were collected from the surface soil to a depth of 0–30 cm. • All the **soil samples** were transferred to the laboratory, air-dried, and crumbled by hand to remove the root materials and passed through a 2- mm sieve. The following soil indicators were analyzed using the following methods: soil particle size distribution by the hydrometer method, pH and electrical conductivity (EC) in 1:2.5 (w/v) in soil: water suspension by pH-meter and EC-meter and CaCO3 content by the volumetric method, total nitrogen (Ntotal) by the Kjeldahl method, available phosphorus (Pavb) by 0.5 M NaHCO3 extraction method, exchangeable potassium (Kexc), calcium (Caexc), sodium (Naexc) and magnesium (Mgexc) by the 1 N ammonium acetate extraction method (Soil Survey Staff, 1992). All soil samples were sieved through a 150-µm mesh and then analysed for total organic matter content by the wet oxidation method (Walkley-Black) using K2Cr2O7 (Nelson & Sommers, 1982). Available soil micronutrients (Feavb, Cuavb, Znavb, Mnavb) were determined (Lindsay & Norvell, 1978). Micronutrients (Feavb, Cuavb, Znavb, Mnavb) were analysed using an Atomic Absorption Spectrophotometer. **Summary**: **chemical and physical analysis of soil properties**. Soil properties were calculated to determine the Soil Fertility Index (SFI). The two approaches were used to validate SFI. First, the mean values of winter wheat yields for each plot over the three years (2017–2019) were obtained from the General Directorate of Altınova State Farm. Thus, the relationship between the rainfed wheat yields and the soil fertility was revealed. Second **RE-OSAVI values** calculated from **Sentinel-2A satellite images for April 2017–2019 were compared with** the yield values and SFI classes.
Expected accuracy	The comparisons show that the SFI distributions of the plots indicate a strong relationship between RE-OSAVI and average yield (r^2 > 0.75). This suggests that the parcels were strongly represented by both the RE-OSAVI and yield averages, and that the SFI model successfully represented the land. Strong relationship (83%) between the RE-OSAVI values and the plot average values of the SFI classes
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Adaptation/resilience/sustainability

Reference/ paper Title	Evaluation of soil fertility using a combination of Landsat 8 and Sentinel-2 data in agricultural lands

Remote sensing method OR Used method	We combined Landsat 8 and Sentinel-2 satellite images using the Gram-Schmidt algorithm. Finally, SFI was calculated from satellite data using multiple linear regression, and its spatial distribution was classified into very low, low, moderate, high, and very high classes.
Used input	Soil sampling was performed irregularly at depths of 0–30 cm at 216 points; 11 soil properties were measured and used to calculate the Soil Fertility Index (SFI). Landsat 8 and Sentinel 2
Expected accuracy	The combination of Landsat 8 and Sentinel-2 data using the Gram-Schmidt algorithm yields a higher correlation with SFI than using these data individually. Therefore, Landsat 8 and Sentinel-2 data were combined to model SFI. Using model selection indices (including Cp, AIC, and ρc), the visible-range bands, notably blue (r=0.65), green (r=0.63), and red (r=0.61), yield the best model for estimating SFI (R2=0.43, Cp=3.34, AIC=-277.4, and ρc=0.44).
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Adaptation/resilience/sustainability

Appendix 13: Measurement: Soil moisture

Reference/ paper Title	A roadmap for high-resolution satellite soil moisture applications – confronting product characteristics with user requirements **Review**
Remote sensing method OR Used method	Summarises existing applications of satellite-derived soil moisture products and identifies gaps between the characteristics of currently available soil moisture products and the application requirements from various disciplines.
Used input	➢ **High resolution**: **Synthetic aperture radar (SAR) Sentinel-1** can provide much higher spatial resolution than other satellites, but with more challenges in the retrieval of soil moisture, due to the combined effects of vegetation structure, water content on the backscattering coefficients, and surface roughness. Sentinel-1 is currently the most advanced SAR mission to support the systematic generation of a surface soil moisture product at **high resolution and regional/continental scale**. However, the accuracy of these simulated soil moisture products depends on the quality and availability of **meteorological observations, soil texture, soil hydraulic properties, and the physics of the models involved (e.g., roughness parameters, moisture data, and canopy data).** ➢ **Coarse resolution:** • The ESA Soil Moisture Climate Change Initiative (CCI): CCI data is provided at a global scale with a spatial resolution of 0.25° (roughly 25 km) and daily temporal resolution from 1978 to 2022-12-31

	• Soil Moisture and Ocean Salinity (SMOS) mission: Provides global soil moisture data with a spatial resolution of 50 km and a temporal resolution of 1-3 days. The data can indicate soil moisture down to about 5 cm. • Soil Moisture Active Passive (SMAP) mission: Provides global soil moisture measurements with a spatial resolution of around 36 km and a temporal resolution of 2-3 days. • NASA merges the SMAP L-band radiometer with Sentinel-1 C-band backscatter data to • produce soil moisture maps at 3-km and 1-km resolutions • The Advanced Microwave Scanning Radiometer 2 (AMSR2): Near-global coverage every day. Its spatial resolution varies by frequency, from about 5 km at the highest frequencies to 50 km at the lowest. • NASA-ISRO Synthetic Aperture Radar (NISAR): Global coverage, revisiting the same location every 12 days; satellite launch has been delayed to February 2025
Expected accuracy	
Can be used as an index or part of an index	
Adaptation or resilience/ sustainability	

Appendix 14: Measurement: Carbon sequestration

Reference/ paper Title	Impact of agricultural management practices on soil carbon sequestration and its monitoring through simulation models and remote sensing techniques**: A review**
Remote sensing method OR Used method	Simulation models and optical remote sensing are recent, effective tools for mapping the Content. Simulation models can perform numerical evaluations of SOC in response to changes in land use or management practices. At the same time, optical remote sensing techniques can map SOC across different temporal and spatial scales. As a research tool, Soil Organic Models (SOM) models provide insight into SOC dynamics, and management strategies can be formulated to optimise C inputs and outputs for SOC sequestration. Again, remote sensing enhances the ability to capture the spatial distribution of SOC by providing reliable, high-precision estimates of local-to-regional soil C sequestration.
Used input	
Expected accuracy	
Can be used as an index or	

part of an index	
Adaptation or resilience/ sustainability	

Reference/ paper Title	Predicting the Potential Impact of Climate Change on Carbon Stock in Semi-Arid West African Savannas
Remote sensing method OR Used method	The current study was undertaken in semi-arid savannas in Dano, southwestern Burkina Faso, with the threefold objective of: (i) identifying the main land use and land cover categories (LULCc) in a watershed; (ii) assessing the carbon stocks (biomass and soil) in the selected LULCc; and (iii) predicting the effects of climate change on the spatial distribution of the carbon stock.
Used input	Dendrometric data (Diameter at Breast Height (DBH) and height) for woody species and soil samples were collected in 43 plots, each measuring 50 × 20 m. Tree biomass carbon stocks were calculated using allometric equations, while Soil Organic Carbon (SOC) stocks were measured at two depths (0–20 and 20–50 cm). To assess the impact of climate change on carbon stocks, data on geographical location, carbon stock records, remote-sensing spectral bands, topographic data, and bioclimatic variables were used.
Expected accuracy	Carbon stocks in woody biomass were higher in woodland (10.2 ± 6.4 $Mg{\cdot}ha^{-1}$) and gallery forests (7.75 ± 4.05 $Mg{\cdot}ha^{-1}$), while the lowest values were recorded in shrub savannas (0.9 ± 1.2 $Mg{\cdot}ha^{-1}$) and tree savannas (1.6 ± 0.6 $Mg{\cdot}ha^{-1}$). The highest SOC stock was recorded in gallery forests (30.2 ± 15.6 $Mg{\cdot}ha^{-1}$) and the lowest in the cropland (14.9 ± 5.7 $Mg{\cdot}ha^{-1}$).
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Adaptation/sustainability.

Appendix 15: Measurement: Greenhouse gas (GHG) emissions

Reference/ paper Title	Assessment of Global Carbon Dioxide Concentration Using MODIS and GOSAT Data

Remote sensing method OR Used method	A model based on temperature (MOD11C3), vegetation cover (MOD13C2 and MOD15A2), and productivity (MOD17A2) from MODIS (which we have named the TVP model) was developed in the current study to assess CO2 concentrations globally. We assumed that the CO2 concentration measured by the Thermal And Near infrared Sensor for carbon Observation (TANSO) aboard the Greenhouse gases Observing SATellite (GOSAT) is the true value, and we used these values to assess the TVP model's accuracy.
Used input	MOD11C3 LST, MOD13C2 NDVI/EVI, MOD15A2 LAI/FPAR, MOD17A2 GPP/NPP, GOSAT TANSO
Expected accuracy	The results indicate that the TVP model's accuracy varies across continents: the highest Pearson's correlation coefficient (R2) was 0.75 in Eurasia (RMSE = 1.16) and South America (RMSE = 1.17); the lowest R2 was 0.57 in Australia (RMSE = 0.73). Compared with the TANSO-observed CO2 concentration (XCO2), we found that the accuracy across the world ranges from -2.56 to 3.14 ppm. Potential sources of uncertainty in the TVP model were also analysed and identified.
Can be used as an index or part of an index	Yes
Adaptation or resilience/ sustainability	Resilience/adaptation

Reference/ paper Title	QUANTIFICATION OF GREENHOUSE GAS FLUXES FROM SELECTED CROPPING SYSTEMS UNDER ON-FARM CONDITIONS IN THARAKA-NITHI, KENYA. Master
Remote sensing method OR Used method	This study quantified soil GHG fluxes across different cropping systems on two smallholder farms in Upper Eastern Kenya. Did not use remote sensing methods, but valuable information was obtained. The main objectives were: 1.To quantify greenhouse gas fluxes (CH4, CO2 and N2O) from selected cropping systems in Tharaka-Nithi County. 2. To evaluate N2O emissions from smallholders' cropping systems. Details about the 2nd objective: The soil N2O emissions from smallholders' cropping systems in SSA. The cropping systems of interest were maize, cereal-legume intercropping, coffee, tea, and vegetables. It also touches on soil N2O emissions from the use of organic and inorganic fertilisers. Yields scale, Emissions (YSE), and N2O emissions factors (EF) are also covered. 3. To determine the environmental factors, climatic conditions, farm management practices, and soil properties that influence soil N2O dynamics
Used input	Five cropping systems on two smallholder farms: sole maize, maize-bean intercrop, coffee, banana, and agroforestry. Greenhouse gases were sampled using three static chambers per cropping system. Gases were analysed using gas chromatography (GC) fitted with a 63Ni electron-capture detector (ECD) for N2O and a flame ionisation detector (FID) for CH4 and

	CO2, with N2 as the carrier gas. Soil samples were analysed for soil texture, total nitrogen, Soil Organic Carbon, and pH. During harvesting, a 2 m x 2 m sub-plot near each chamber was selected, and all the crops within the area were harvested. Both above-ground and below-ground biomass were harvested for food crops. Fresh weight for the plant components (grains, leaves, stems, and roots) was determined using an electronic balance. A sub-sample of each plant component (leaves, stem, root, and grain) was weighed, air-dried for three weeks, re-weighed, and measured again. Soil moisture was determined using the gravimetrical method, and temperature was measured using a thermometer during gas sampling. Meteorological data (precipitation and Temperature)
Expected accuracy	The maize yields ranged from 0 to 3.38 Mg ha-1. The nitrous oxide yields scaled emissions ranged from 0.10 to 0.26 g kg-1 for maize and from 0.68 to 1.30 g kg-1 for beans.
Can be used as an index or part of an index	yes
Adaptation or resilience/ sustainability	Sustainability/adaptation

Reference/ paper Title	Towards an Improved Inventory of N2O Emissions Using Land Cover Maps Derived from Optical Remote Sensing Images
Remote sensing method OR Used method	The proposed method provides access to an inventory of potential N2O emissions (the term potential refers to possible but not yet actual) at a fine scale, with annual updates, without a heavy deployment tied to a collection of field measurements. The processing chain is applied to optical satellite images regularly acquired at high spatial resolution during the 2006–2015 period, enabling higher-resolution and temporally consistent estimates of potential N2O emissions from crops. The annual potential N2O emissions inventory is estimated for a study site in southwestern France, considering seven main seasonal crops (i.e., wheat, barley, rapeseed, corn, sunflower, sorghum, and soybean). The first step of the study was to classify lane use. The second step was to estimate the annual potential budget of N2O for the study area.
Used input	In Situ Data—Registre Parcellaire Graphique (RPG): displays detailed information on land use Satellite data: Landsat 8, Spot 2/4//5 for 2006 and 2014/2015.
Expected accuracy	The land-use classification has an Overall Accuracy greater than 0.8. The yearly potential budget of N2O emissions ranges from 97 to 113 tons, with an estimated relative error of less than 5.5%. Wheat, the main cultivated crop, is associated with the maximum cumulative emissions regardless of the considered year (with at least 48% of annual emissions), while maize, the third crop in terms of the allocated area (grown on less than 8% of the study site), has the second highest cumulative emissions. Finally, the analysis of a 10-year map of the potential N2O budget shows that the predominant crop rotation (i.e., alternating wheat and sunflower) yields potential emissions of approximately 16 kg N2O per hectare. In comparison, the monoculture maize is associated with the maximum value (close to 28.9 kg per hectare).
Can be used as an index or part of an index	Yes

Adaptation or resilience/ sustainability	Sustainability/adaptation