Title: Leveraging Ground Sensor Networks to Evaluate Satellite-Based Land Surface Phenology in Smallholder Farming Systems

Authors: (1) Michael Cecil [a] (**corresponding author**) mcecil@clarku.edu;

(2) Natasha Krell [b] nkrell@ucsb.edu;

(3) Allan Chilenga [c] allanchilenga@gmail.com;

(4) John Gitonga [d] rajonny@ymail.com;

(5) Protensia Hadunka [e] hadunka2@illinois.edu;

(6) Andrew Zimmer [f] andrew.zimmer1@montana.edu;

(7) Adam Wolf [g] adam@eion.team;

(8) Tom Evans [h] tomevans@arizona.edu;

(9) Kelly Caylor [b] caylor@ucsb.edu;

(10) Lyndon Estes [a] LEstes@clarku.edu

[a] Clark University, Graduate School of Geography; 950 Main St., Worcester MA 01610, USA

[b] University of California Santa Barbara, Department of Geography; 1832 Ellison Hall, UC Santa Barbara CA 93106-4060, USA

[c] Zambia Agriculture Research Institute (ZARI); Mount Makulu Central Research Station, Private Bag 7, Chilanga, Zambia

[d] Princeton Ecohydrology Lab at Mpala Research Center; P.O. Box 555 - 10400 Nanyuki, Kenya

[e] University of Illinois Urbana-Champaign, Agricultural & Consumer Economics, 326 Mumford Hall, 1301 W. Gregory Drive, Urbana IL 61801, USA

[f] Department of Earth Sciences, Montana State University, P.O. Box 173480, Bozeman, MT 59717-3480, USA

[g] Eion Carbon; 82 Spruce St., Princeton NJ, 08542, USA

[h] University of Arizona, School of Geography, Development & Environment, 1064 E. Lowell Street, Tucson AZ 85721, USA

Title: **Leveraging Ground Sensor Networks to Evaluate Satellite-Based Land Surface Phenology in Smallholder Farming Systems**

Abstract: Satellite-derived estimates of land surface phenology (LSP) are important for quantifying the spatiotemporal dynamics of agricultural landscapes where LSP metrics can track crop growth and productivity. Previous LSP applications for agriculture have mostly been confined to developed countries, where fields are larger and more homogenous, and ground data networks are available. LSP studies from developing regions are limited and have focused on coarse scales without field-level comparison. This study used field-level multispectral sensors (Arable Marks) to quantify several key LSP metrics and compare satellite and ground-derived LSP timing in smallholder maize fields. The ground sensors were installed in crop fields in Kenya and Zambia that were managed according to typical smallholder practices, and used to estimate greenup, maturity, senescence, and dormancy dates. We then calculated differences between the ground LSP metrics and corresponding metrics from Sentinel-1, Sentinel-2, and VIIRS time-series, and the VIIRS Land Cover Dynamics product. We also compared the satellite-based LSP metrics to farmer planting and harvest dates for a subset of fields. Results showed that Sentinel-2 had the smallest differences among single sensor models relative to ground-based LSP metrics across all dates (bias-adjusted MAD 11-13 days), with multi-sensor models showing comparable correlation (Kendall's τ) with ground LSP metrics. Sentinel-1 and the VIIRS Land Cover Dynamics product were less comparable to ground-based measures due to high variability and missing observations, respectively. Across most sensors and models, correlations with ground LSP metrics were higher for late season dates (senescence, dormancy) compared to early season dates (greenup, maturity). For management events, single-sensor models had higher correlation for harvest date (highest $\tau = 0.75$), than planting date (highest $\tau =$

0.30). Overall the results show that satellite-derived LSP can estimate ground-based measures of LSP within a two week timeframe for the majority of sites. We identify multi-sensor models, extended ground sensor networks, and high-resolution satellite imagery as priorities for continued research of smallholder LSP monitoring. These results are the first evaluation of satellite-derived LSP metrics for smallholder agriculture comparing multiple satellite sensors and vegetation indices and provide a baseline performance for global, freely available satellite sensors.

1. Introduction

Satellite imaging has enabled repeated sensing of Earth's vegetation over broad spatial and temporal domains, enabling the assessment of crop temporal dynamics, including timing of key stages during annual growth cycles (Reed et al., 1994; Xin et al., 2002; Zhang et al., 2006, 2003, 2001). Land surface phenology (LSP), the study of vegetation growth and timing via remote sensing, relies on vegetation index time-series and is distinguished from biological phenology (e.g., emergence, anthesis, and other biological growth stages). Satellite sensors with daily revisit times, particularly AVHRR, MODIS and VIIRS, are especially well-suited for LSP studies and have been used in a wide range of applications, including climate impacts (Brown et al., 2010), biodiversity (Viña et al., 2016), crop type mapping (Zhong et al., 2011), and agricultural timing and productivity (Diao, 2020; Viña et al., 2004).

Given the importance of crop life cycle timing for understanding farmer management (Diao et al., 2021), crop productivity (Bolton and Friedl, 2013) , and within-field variability (Johansen et al., 2022), it is no surprise that agricultural applications of LSP are a primary focus of recent advances in the field. The field owes its foundations to studies based on the AVHRR and MODIS sensors (e.g., Zhang et al., 2003, 2001), which allowed for daily monitoring of crops, at a resolution of 250 m - 1 km. These studies included examination of curve-fitting methods (Beck et al., 2006)(Zhang et al., 2003); (Beck et al., 2006)(Zhang et al., 2003), the use of different vegetation indices (Klosterman et al., 2014; Rocha and Shaver, 2009; Zhang et al., 2018a), correlation with crop growth stages (Diao and Li, 2022; Shen et al., 2022; Zeng et al., 2020), and, more recently, approaches focused on image fusion (Frantz et al., 2016; Gao et al., 2017; Zhang et al., 2017) and near real-time monitoring (Gao and Zhang, 2021; Shen et al., 2023).

However, current research on agricultural LSP is primarily focused on developed countries, where fields are larger and more homogeneously managed, and extensive ground datasets exist for validation. LSP studies often use well-known ground validation datasets including crop progress reports (e.g., Diao, 2020; Gao et al., 2017; Shen et al., 2023; Wardlow et al., 2006), PhenoCam networks, (e.g., Diao and Li, 2022; Moon et al., 2019; Tran et al., 2022), and extensive ground-measured phenology datasets such as PEP725 in Europe (Ye et al., 2022).

While rich validation datasets in developed countries have enabled rapid development of new methods, LSP data and methods may be of even greater value in developing countries where smallholder agriculture predominates and ground-measured data are limited (Nakalembe and Kerner, 2023). Smallholder agriculture represents between 20 and 50% of global food production (Ricciardi et al., 2018), and the need for sustainable agricultural intensification may increase in the future. In sub-Saharan Africa (SSA) for example, population is expected to increase while available arable land is diminishing (Ittersum et al., 2016). Well-validated LSP products could provide insights into crop growth timing and the effectiveness of farmer management interventions (such as response to fertilizer and seed choice) when complemented with ground-surveys. Such information is especially valuable in smallholder contexts where farmer management practices are widely varied and have substantial and region-specific impacts on yield variance (Cecil et al., 2023).

Satellite-based LSP has been used at varying scales to study smallholder agriculture, including climate impacts (Adole et al., 2018a), national crop cover mapping (Htitiou et al., 2021), and the relation between planting date and satellite imagery using micro-satellites (Jain et al., 2016). However, there have been few comprehensive studies of LSP methods using ground sensor networks for field-level smallholder agriculture, analogous to the comprehensive examination of sensors, vegetation indices, and methods using the PhenoCam network in the US.

Several challenges exist for remote sensing-based smallholder agriculture monitoring using LSP. Beyond the aforementioned lack of ground sensor networks for validation, satellite-based LSP faces additional limitations in tropical or sub-tropical smallholder-dominated farming regions such as cloud cover, as fields in regions with a well-defined rainy season may have zero or few clear-sky observations for sensors with intermediate revisits times (e.g., Sentinel-2, Landsat) (Roy et al., 2021). In addition, smallholder regions may not be fully covered even by global satellite sensors. For example, gaps exist in MODIS coverage at the equator (e.g., Li et al., 2018). Also due to the expiration of one of the Sentinel-1 satellites, certain study areas (including sites in this study) lost coverage in 2021[1]. Smallholder field sizes are also significantly smaller, and more heterogeneously managed, with farmers using multiple management strategies for different fields, including inter-cropping (Sheahan and Barrett, 2017), differential fertilizer application (Tittonell et al., 2007), and a variety of planting dates and cultivars (Waldman et al., 2017). Smallholder fields are significantly smaller than coarse resolution sensors such as MODIS and VIIRS, and may result in an increase in mixed pixels for medium resolution sensors such as

[1] See https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/observation-scenario

Landsat (Jain et al., 2017). As such, data fusion algorithms, such as STARFM (Gao et al., 2006) may be difficult to implement given the underlying heterogeneity of fields and persistent cloud cover during the growing season.

However recent advances in smallholder agricultural monitoring suggest there is promise for measuring LSP in smallholder agriculture. Studies on yield mapping for smallholder maize farms in Kenya and Tanzania (Jin et al., 2019), sowing date and yield for smallholder wheat farms in India (Jain et al., 2016), yield gap assessments in India (Duncan et al., 2015), detecting crop rotations in Ethiopia (Kibret et al., 2020), and fodder biomass monitoring in Senegal (Diouf et al., 2015) suggest that satellite sensors can provide valuable insights on smallholder agriculture, even with the limitations outlined above. Moreover, land surface phenology metrics have specifically been used to map early sowing of wheat in India (Chakraborty et al., 2018), which may arise from farmers adapting to changing climate. In Africa, LSP signals may be difficult to interpret due to concurrence of climate and management effects in LSP metrics, and challenges in separating signals of natural and crop vegetation in rainfed systems (Bégué et al., 2020). Evaluating differences between satellite and local, ground-derived LSP metrics is thus crucial to assess whether satellite sensors can effectively monitor field-level LSP, and the degree of variation present in the satellite metrics.

In this study, we seek to advance the capabilities for smallholder agricultural monitoring over large areas by answering the following key question: how well can satellite-derived LSP measure field-level crop dynamics in smallholder-dominated systems? To do so, we employ a novel network of ground-based sensors that collect continuous in-field measures of crop growth and phenology, thereby overcoming the lack of available ground data, and use these data to evaluate the difference between ground- and satellite-derived LSP metrics for key events (e.g., greenup).

Specifically, we deployed multispectral Arable Mark sensors at maize fields in Kenya and Zambia that were managed under smallholder conditions. These sensors produce vegetation indices similar to those calculated for satellite-based multispectral sensors, at daily time-steps regardless of cloud conditions. The resulting VI time-series derived from Mark data provide precise timing of land surface phenology dates, including greenup, maturity, senescence and dormancy (which have precise definitions in an LSP context, see (Zhang et al., 2001). At each site, we calculated the difference of ground-derived LSP timing with LSP estimated from four freely available satellite based sources: Sentinel-1 synthetic aperture radar (SAR), Sentinel-2 and VIIRS surface reflectance, and the VIIRS Land Cover Dynamics product (Zhang et al., 2018c). We then calculated how the different satellite sensors, both separately and in combination, differ from ground-based Marks in estimating LSP throughout the growing season, including the timing of different management events. The results provide an evaluation of freely available sensors in their agreement with ground-measured LSP for smallholder agriculture, as well as estimates of sensor bias and uncertainty.

2. Methods

To develop field-level LSP metrics for smallholder agriculture, we drew on data collected from Arable Mark ground sensors deployed between 2017 and 2022 in maize fields in Zambia and Kenya (Table 1). Further detail about the Arable Mark sensors is explained in section 2.2. Several Zambia field sites are managed by the Zambia Agriculture Research Institute (ZARI) as part of field trials, but all sites used management practices (e.g., planting date, fertilizer, cultivar choice) representative of smallholder farmers.

Comparisons of satellite and ground-measured LSP were built around four LSP dates: greenup, maturity, senescence, and dormancy, that we extracted from smoothed VI curves using a threshold method. We calculated differences of LSP dates derived from Mark sensors and management events (planting and harvest dates) with LSP dates extracted from three satellite sensors: Sentinel-1, Sentinel-2, and VIIRS, as well as LSP dates from the VIIRS Land Cover Dynamics product. We also calculated differences of Mark LSP with LSP derived from three multi-sensor methods: mean, median, and a Random Forest model used to build a consensus VI. Finally, we outline the specific metrics used (median bias, median absolute deviation from median bias, and Kendall's τ) for comparison, and the steps used to assess difference when comparing satellite to Mark LSP dates.

2.1 Study Sites

2.1.1. Zambia

Zambia is a potential breadbasket for southern Africa (Phiri et al., 2020), as it has one of the larger stocks of potential arable land in Africa (Ittersum et al., 2016). The growing season for smallholder agriculture follows the regional rainy season, with planting in November or December and harvest between March and May. The study sites for Zambia come from three distinct Agro-ecological zones. During the 2017-18 season, Marks collected data from 15 smallholder maize fields in Choma district[2] in southern Zambia. During the 2020-21 season, Marks were installed in nine plots at four sites (Chilanga, Choma, Kabwe, Mufulira) managed by the Zambia Agriculture Research Institute in 2020-21 in order to examine different maturing cultivars in different regions. Management practices (e.g., cultivar, planting date, fertilizer) at these sites mirrored those of local smallholders. Finally, in 2021-22, Marks were installed at nine

[2] One site was in neighboring Kalomo district.

plots in a single site (Chilanga) to test fertilizer and cultivar interactions, using fertilizer levels and cultivars typical of smallholders. Planting date and harvest date were recorded for a subset of site-seasons.

2.1.2. Kenya

We installed 45 Marks in Kenya between March and May 2018 in Nyeri, Meru, and Laikipia counties in the Mount Kenya region. The Marks were located within maize fields of small-scale producers who also grew household-level crops such as ground nuts, leafy green vegetables, and fruit, as well as kept livestock (see Krell et al., 2022) for further detail on farmer demographics). Some farmers also had access to irrigation water from piped networks off Mount Kenya, depending on water rationing from local Water Resource User Associations.

The farmers selected to host the Marks on their farms were enrolled in a multi-year SMS survey program (Krell et al., 2022) and participated in annual household surveys (e.g., Guido et al., 2020; Krell et al., 2021). The Marks were roughly equally spatially distributed among the 75 communities and 500+ farmers who participated in the annual household surveys (see Krell et al., 2021) for information on farmers selected in the surveys). The farmers were provided a "farmer diary" in which they recorded information such as planting date, cultivar, amount planted, and intercropping. The farmers then recorded their farmer management, including planting and harvest date, and were compensated for their participation in the program at a rate of 600 Kenya Shillings per week (roughly \$4 USD per week).

Table 1: Field deployment sites

Site group	Number of site-seasons	Type of Farm	Start date/end date	Ancillary data
Zambia ZARI trials 20-21	13	Field trials at 4 sites using smallholder inputs. Also 3 smallholder farms.	2020-11-01 to 2021-05-31	Planting, harvest for some sites
Zambia ZARI trials 21-22	8	Field trials at 1 site using smallholder inputs.	2021-11-01 to 2022-05-31	Planting, harvest for some sites
Zambia Choma survey	14	Smallholders	2017-11-01 to 2018-05-31	
Kenya	109	Smallholders	2018-03-01 to present. Individual site-season start and end dates determined by curve-smoothing.	Irrigation in dry season. Planting, harvest for some sites

2.2 Arable Marks

Arable Marks are in-situ ground sensors mounted on 3-meter poles in crop fields, designed by Arable Labs Inc. Arable Marks are similar to other ground-based commercial sensors (e.g., Campbell). We opted for Arable Marks because of their ability to upload data via local cell phone networks and their cost-effectiveness for deployment. The Marks are designed to overcome limitations of cloud cover, revisit time, and between-date calibration of satellite sensors. The Marks include upward and downward facing radiometers, which can be used to estimate surface reflectance in seven spectral bands at a daily time-step, regardless of cloud conditions. The sensors include bands in the visible, red-edge, and near-infrared wavelengths, allowing for comparison with commonly used satellite sensors (Landsat, Sentinel-2, MODIS, VIIRS, see Figure 1). Marks use solar panels for recharging (when available), and sync data with local cellular networks, allowing for near real-time updates on crop growth. Mark devices also include sensors to measure meteorological variables, including temperature, precipitation and barometric pressure, but these data are not used in this study.

The spectral overlap for specific bands (e.g., red, near-infrared) is not perfect between the ground and satellite sensors used. We thus did not focus on direct comparison of surface reflectance or vegetation index (VI) values between sensors. Rather, we compared land surface phenology timing derived from these VI's, which should be more robust to small discrepancies in wavelengths measured for different sensors.

To process Mark data, we first calculated surface reflectance (ρ) in each band using equation (1), as the ratio of upwelling (uw) to downwelling (dw) radiation in band *i*, with values constrained between 0 and 1.

$$(1)\ \rho_i = uw_i/dw_i$$

To match typical satellite overpass times, we used hourly upwelling and downwelling data matched to the nearest hour of Sentinel-2 overpass times, typically 10 or 11 AM local time. After calculating surface reflectance, we derived three vegetation indices used for LSP time-series. These VI's were selected due to their long track-record, their use in agricultural applications, and their use in LSP products.

We used the two-band Enhanced Vegetation Index (EVI2, Equation 2a) (Jiang et al., 2008) as a baseline for most of the analyses, due to its prominence as the de facto choice for many LSP applications. EVI2 is preferred due to its dependence on only two bands (red and near-infrared), and its improved performance in areas of high vegetation relative to the

Normalized Difference Vegetation Index (NDVI). EVI2 is the index used in the land cover dynamics products for both MODIS (Gray et al., 2022) and VIIRS (Zhang et al., 2018c) . EVI2 has also been found superior to NDVI in LSP estimation when compared directly to NDVI (e.g., Klosterman et al., 2014; Zhang et al., 2018a).

NDVI (equation 2b) has been used extensively in LSP studies (e.g., Lloyd, 1990; Zhang et al., 2001, 2006, 2007), although it is known to saturate under high vegetation biomass (Huete et al., 2002), and have increased uncertainty when compared to LSP metrics derived from other vegetation indices (Klosterman et al., 2014; Zhang et al., 2018a).

We chose the Green Chlorophyll Index (GCVI, equation 2c) (Gitelson et al., 2003) as a third multispectral VI due to its different construction (using green instead of red band) and use in LAI estimation (Nguy-Robertson et al., 2012) and crop modeling (Lobell et al., 2015). GCVI has been used in smallholder agricultural contexts to study sowing date (Jain et al., 2017) and yield variance (Jin et al., 2017), as well as estimating LSP using high-frequency sensors in the US (Nieto et al., 2022).

These VI's use the green, red, and near-infrared bands from Mark sensors (bands 2, 4, 6 respectively). Their formulas are listed in equations (2a-c). The band names on the right-hand side of the equations represent surface reflectance (ρ) in that band.

$$\text{(2a) } EVI2 = 2.5 * (NIR - red)/(NIR + 2.4 * red + 1)$$

$$\text{(2b) } NDVI = (NIR - red)/(NIR + red)$$

$$\text{(2c) } GCVI = NIR/Green - 1$$

As a final filtering pass, we removed any observations with NA values for vegetation indices. Based on manual inspection of VI curves, we also removed any Mark observations with band 1 (blue) surface reflectance > 0.19, band 3 (yellow) surface reflectance > 0.2, or band 4 (red) surface reflectance > 0.4. These values may represent observations taken indoors, when Marks were charging, or otherwise invalid observations.

2.3. Satellite Sensors

We compared the Mark LSP dates with LSP dates calculated from the following satellite sensors: Sentinel-1, Sentinel-2, VIIRS surface reflectance (hereafter "VIIRS"), and the VIIRS Land Cover Dynamics (hereafter "VLCD") product (Table 2). These datasets were selected due to their history of use in LSP studies, their global coverage, their current and future data availability, and their being freely and publicly available. Other satellite sensors (e.g., Planet) have been effectively used in LSP studies, such as classifying crop stage growth (Nieto et al., 2022), but do not satisfy these conditions.

Sentinel-1, VIIRS surface reflectance, and VLCD data were downloaded from Google Earth Engine (Gorelick et al., 2017). The Sentinel-2 Level-2A data were downloaded from Digital Earth Africa (Yuan et al., 2021) due to a longer track record of Level-2 products being available on its site compared to Google Earth Engine.

Table 2: Ground and Satellite Sensors

Sensor	Spatial resolution	Revisit	Indices used	Data source
Arable Mark	10 m radius footprint	Hourly	EVI2, NDVI, GCVI	Arable API
Sentinel-2 Level-2A surface reflectance	10 m	5 days	EVI2, NDVI, GCVI	Digital Earth Africa
VIIRS surface reflectance (VNP09GA)	1 km	Daily	EVI2, NDVI, GCVI	Google Earth Engine
Sentinel-1	10 m	12 days for most sites	VH, VV	Google Earth Engine
VIIRS Land Cover Dynamics (VNP22Q2)	500 m	Extracted for 1-2 seasons per year	LSP dates extracted directly	Google Earth Engine

2.3.1. Sentinel-2

Sentinel-2 imagery provides a combination of spatial and temporal resolution and global coverage unmatched for freely available multispectral sensors. Since the launch of twin satellites in 2015 and 2017, Sentinel-2 has been widely applied for agricultural studies, including LSP (Meroni et al., 2021; Nieto et al., 2022), and has been used in smallholder agricultural studies such as yield estimation (Lobell et al., 2020). Sentinel-2 has several advantages for agricultural monitoring, including multiple visible, red-edge, and near-infrared bands, and a finer spatial resolution (10-20 meters) than Landsat or VIIRS. Cloud cover nevertheless remains a primary obstacle for Sentinel-2 imagery in regions with pronounced rainy seasons.

We used the Level-2A surface reflectance product, and the 10 meter green, red, and NIR bands (bands 3, 4, 8) to calculate the indices in equation (2). We used the Scene Classification

(SCL) band for cloud masking, filtering to SCL values of 4 or 5, representing clear skies and vegetation or bare soil. As a final filtering step, observations with surface reflectance in band 1 (blue aerosol) > 0.15 or surface reflectance in band 4 (red) greater than 0.40 were removed. These thresholds were determined by examining obvious outliers in VI curves at study sites.

2.3.2. VIIRS surface reflectance

VIIRS has succeeded AVHRR and MODIS as the predominant coarse resolution (250 m - 1 km) sensor used for global LSP studies due to its near-daily revisit time. While these sensors have provided an overall picture of vegetation phenology in Africa (Brown et al., 2010; Adole et al., 2018a, 2018b), their spatial resolutions may be too coarse for field-level LSP extraction in smallholder contexts where field sizes are often less than 2 ha (Ricciardi et al., 2018). We used the VNP09GA surface reflectance data product (v1) at 1 km resolution and bands M4, M5, M7 for red, green, and NIR in equation (2) to derive vegetation indices. Cloud masking was performed using the QF1 and QF2 bands, requiring imagery to be 'confident clear' with no cloud shadow or thin cirrus.

2.3.3 Sentinel-1

Sentinel-1 synthetic aperture radar (SAR) data has distinct advantages for crop monitoring in tropical and sub-tropical environments, due to its measurements not being affected by clouds. It has been previously been applied in smallholder contexts for crop/non-crop mapping (Jin et al., 2019), crop-type mapping (Kussul et al., 2017; Tricht et al., 2018), and for mapping crop growth stages in European crops (Khabbazan et al., 2019).

Sentinel-1 imagery does have limitations for crop monitoring. SAR imagery is sensitive to physical growth as opposed to crop health, and thus may not detect crop senescence as well as optical sensors. SAR imagery is prone to coherent speckle noise (Oliver and Quegan, 2004), requiring a larger footprint and pre-processing for analysis. Sentinel-1 backscatter is also affected by several factors beyond physical growth, such as soil moisture and non-crop vegetation.

We examined phenology metrics for the VH and VV backscatter (converted to float). We also considered additional indices (equation 3) designed to reduce effects of soil moisture and track the parameter of interest, vegetation growth. The cross-ratio index (equation 3a) has been shown to track crop seasonality (Veloso et al., 2017) and the Radar Vegetation Index (Kim and Zyl, 2009) has also been used for crop phenophase monitoring (Holtgrave et al., 2020; Kim and Zyl, 2009), using a dual-polarization formulation as seen in equation (3b).

$$\text{(3a) } CR = VH/VV$$

$$\text{(3b) } RVI = 4 * VH/(VH + VV)$$

We downloaded the Sentinel-1 Ground Range Detected (GRD) data product from Google Earth Engine using the Python API. We exclusively used imagery taken in "Ascending" mode as these data were available for all sites. We also filtered imagery to only consider the most common look angle at each site. For the VH and VV bands, data were then converted to float and passed through an iterative guided filter to despeckle images (He et al., 2012). The guided filter had a kernel of size three and three iterations. We then calculated the CR and RVI indices (equation 3). During the Zambia 2021-22 season, Sentinel-1 coverage was interrupted due to the failure of satellite Sentinel-1B. We thus do not include these sites in the Sentinel-1 and multi-sensor analysis.

2.3.4. VIIRS Land Cover Dynamics Product

In addition to calculating LSP dates directly from VIIRS surface reflectance, we also compared results with the pre-calculated land cover dynamics product VNP22Q2 (v1) (Zhang et al., 2018c). The VLCD product uses EVI2 to track crop growth and a curvature based method for extracting phenological dates. The metrics used from the VLCD product are greenup onset, maturity onset, senescence onset, and dormancy onset, and are extracted based on site locations without any additional processing.

The VLCD site provides these metrics for up to 2 growing seasons per year. To account for Kenya sites with multiple seasons, we used the following steps to match VLCD dates to a given site-season. (1) Each Mark LSP date was matched with the closest VLCD date of the same type (e.g., greenup). (2) If the Mark LSP date was absent due to missing Mark data, a different LSP date was used (e.g., substituting maturity date for greenup). (3) If no VLCD date of a given type (e.g., greenup) was found within the defined season, then an NA value was returned. Notably, the VLCD product may have missing metrics for certain site-seasons, limiting its effectiveness for tracking smallholder seasonality.

2.4. Phenological date extraction

Curve-fitting and phenological date extraction were performed using the 'phenofit' package (Kong et al., 2022) in R. Our calculations followed the general steps outlined by Kong et al., including data pre-processing, growing season definition, rough and fine curve-fitting, and phenological metric extraction (Figure 2). Mark locations were defined by the median longitude and latitude of valid observations (to account for slight GPS drift). Satellite imagery were then extracted at Mark locations to create VI time-series after all pre-processing steps (e.g., cloud masking and guided filter).

For several steps (season definition, rough-curve fitting, removing weeding signal), we used the Whittaker smoother (Eilers, 2003) to remove the influence of outliers and identify local extrema. The Whittaker smoother provides a balance of data fidelity and smoothness based on a smoothness parameter λ (Atzberger and Eilers, 2011).

2.4.1. Season definition

Prior to extracting LSP dates, we first defined season start and end dates to determine which observations were included in curve-fitting. For Mark time-series, we generally used all relevant observations for a site-season (with the exception of Kenya sites, described below). For satellite imagery, we initially provide a padding of two months before and after the time-series of Mark observations, to account for dynamics both within the growing and shoulder seasons. These start and end-dates are listed in Table 1, and were based on Mark curves and common knowledge about the timing of local growing seasons.

For Kenya sites, Mark sensors collected data from 2018-03-20 to 2022-11-22 without separation for distinct seasons. To separate consecutive seasons, we first fit the Mark NDVI curve using a Whittaker smoother (described in section 2.4.3), with a smoothing parameter of $\lambda = 10000$. We then used a local minimum finder to separate seasons with a window size of 91 days. These parameters were selected as they effectively separated clear growing seasons in Mark NDVI curves (Figure S1). The resulting split seasons matched well with the two annual growing seasons in Kenya (Dimou et al., 2018).

2.4.2 Weeding events and background VI padding (Mark sensors)

To maximize the use of Mark sensor data (available for a limited number of fields), additional preprocessing was performed to account for early season weeding events and missing data observations. Early season weeding events are clearly present in some fields (see Figure 2), and were confirmed by our colleagues at ZARI as a common practice in smallholder fields, especially as smallholder farmers often do not use pre-emergence herbicides. To detect weeding events, we applied a Whittaker smoother to the Mark NDVI time-series with smoothing parameter $\lambda = 10000$, and detected early season local minima meeting the following conditions: (i) the local min is prior to and at least 0.20 units below the season's global max (ii) the local min has a local max preceding it which is at least 0.35 units and at least 0.10 units higher than the local min, and (iii) the local min is less than 0.60. In these cases, we truncated the Mark time-series by removing any observations prior to the local min. The smoothing parameter and conditions were selected based on visual inspection of VI curves before and after truncating the time-series. In total 9 of 144 sites had earlier observations removed for weeding events.

Mark time-series also had gaps in coverage due to timing of Mark installation and removal, as Marks were typically installed after planting is complete and removed before harvest. For this reason, raw Mark time-series may lack observations during transitions between the growing and off-seasons, a critical time for LSP-based analyses. To aid in curve-fitting and LSP date extraction, we padded VI time-series at the start and end of growing seasons with estimated background values for a VI. The background values represent the estimated VI value in the off-season, and are different for each VI. Background values were calculated for time-series judged 'complete' (by visual observation), meaning that the VI time-series covered the entire vegetation growing cycle. For each complete site, we calculated the mean of VI values within the bottom 10 percent of VI values (similar to methods in (Zhang et al., 2020), and then took the median of these background values across all complete sites. The final background values used are listed in Table 3.

Table 3: Background values used for padding in Mark curve-fitting.

EVI2	0.099
NDVI	0.214
GCVI	1.672

To pad VI time-series, we first added a 4 week gap of empty values before and after the Mark time-series, representing potentially missed observations. Then, 30 days of VI values at the dormancy level were added before and after the gaps. Padding occurred for all Mark time-series within the season start and end dates defined in Table 2, and occurred after weeding adjustments.

2.4.3. Curve-fitting and phenometric extraction

Curve-fitting VI time-series is a two step process. Rough-fitting occurred first to smooth the time-series and remove the effect of outliers using the Whittaker smoother. We used a more faithful (less smooth) smoothing parameter ($\lambda = 1000$) for the Mark data, which is daily and contains less variation than satellite data. For satellite imagery, we increase the smoothing parameter ($\lambda = 10000$) for all curves.

After rough-fitting, fine-fitting then occurred using a double-logistic function to ensure that LSP date extraction is well-defined (Kong et al., 2022). We used the Beck double-logistic curve (Beck et al., 2006), an asymmetric double-logistic model which assumes that before and after-season dormancy values are identical. This curve-fitting method has been used previously for LSP extraction for comparing ground and satellite sensors (Diao and Li, 2022). LSP date extraction is performed using the threshold extraction method, where LSP dates are defined based on the 15th and 90th percentiles before and after the seasonal maximum VI value. In order, the LSP dates are greenup onset (15th percentile before season max), maturity onset (90th percentile before season max), senescence onset (90th percentile after season max), and dormancy onset (15th percentile after season max). This phenological extraction method has been used extensively before, including in the MODIS land cover dynamics product (Ganguly et al., 2010; Zhang et al., 2006). The threshold method was selected instead of other methods (e.g., curvature-based extraction) because LSP dates are always defined using the threshold method.

We examined whether any sites extracted identical LSP dates for a satellite data product due to site proximity. For VIIRS surface reflectance, two sites (1.4%) had identical LSP dates. For VLCD, among sites that extracted at least one valid VLCD date, 16 sites (11.1%) had identical LSP dates with at least one other site. Four sites (2.8%) also had no valid VLCD dates extracted. All sites had distinct sets of LSP dates extracted for Sentinel-1 and Sentinel-2, except for 10 sites (6.9%) which did not have any Sentinel-1 LSP dates extracted due to the Sentinel-1B satellite failure.

2.5 Multi-sensor models

We also examined whether combining data from multiple sensors improved the ability to estimate LSP. We first assessed whether taking the mean or median of LSP estimates from separate sensors affected the differences in LSP compared to ground-based estimates. The mean and median were taken from Sentinel-2, VIIRS, and Sentinel-1 indices listed in Table 2 (8 indices total).

We also developed a Random Forests (Breiman, 2001) model to estimate LSP timing. The model used the rough-fit curves from the 8 satellite-based VI's as model predictors, and the rough-fit Mark EVI2 as the dependent variable. We included only the VH and VV derived LSP dates for Sentinel-1 due to their superior correlation with ground LSP dates compared to the CR and RVI LSP metrics. We grouped sites based on the four groupings in Table 1. To avoid over-fitting, we excluded sites in a specified group from model training when estimating the LSP for sites in the specified group. We then used the Random Forest model, built separately for each group, to estimate a 'consensus' VI using the rough-fit satellite VI's for that site. We extracted LSP dates from the 'consensus' VI using the same threshold method defined above.

2.6. Ground and Satellite LSP Comparison

We assessed LSP dates for the following scenarios. We compared LSP dates for different VI's of the same sensor, to assess consistency. We calculated differences of LSP dates of single-sensor satellite data with ground LSP derived from Arable Marks. Finally, we calculated differences of LSP dates from both satellite and ground LSP dates to management events, namely planting and harvest dates.

To assess differences between two sets of LSP dates (or between one set of LSP dates and management events), we designated one set of LSP dates as the reference set. For the within-sensor comparisons of different indices, the EVI2 derived LSP was the reference set for Mark, Sentinel-2, and VIIRS, and VH derived LSP was the reference set for Sentinel-1. For between-sensor comparisons, Mark EVI2 LSP was the reference set. Finally, for comparisons between LSP dates and management events, the management event dates were used as a reference. In this case, we compared the LSP greenup date with planting date, and the LSP dormancy date with harvest date, filtering to sites with valid planting (or harvest) dates.

We used multiple methods to compare LSP dates. First, we calculated the median bias as the median of (target VI date - reference VI date) across all sites, with negative bias representing a (median) earlier LSP date for the target VI LSP date (equation 4a). We then calculated median absolute deviation (MAD) to represent variation after adjusting for the median bias. This calculation was performed by subtracting the bias from target LSP observations, and calculating the median absolute deviation (equation 4b).

$$(4a)\ \text{bias} = \text{median}(LSP_{target}(\text{site}) - LSP_{reference}(\text{site}))\ \textit{across all sites}$$

$$(4b)\ \text{bias adjusted MAD} = \text{median}(\text{abs}(LSP_{target}(\text{site}) - \text{bias}))\ \textit{across all sites}$$

Previous studies have shown substantial bias in LSP dates between different sensors. For example, biases of 5-7 days were found between PlanetScope and PhenoCam estimates of crop emergence and maturity derived from LSP metrics (Diao and Li, 2022), up to 10 days for croplands between the MODIS and VIIRS Land Cover Dynamics products (Moon et al., 2019), and up to 18 days between VIIRS and PhenoCam LSP timing for different land cover types (Zhang et al., 2018a). Even larger biases may exist between LSP dates and management events like planting date (up to 19 days for planting date and 28 days for harvest based on US state progress reports for corn (Shen et al., 2022). Such lags are to be expected given the imperfect correlation between crop management and growth and LSP derived metrics (Gao and Zhang, 2021). To our knowledge, there has not been an extensive comparison of LSP timing based on multiple satellite sensors for smallholder agriculture.

To estimate how satellite LSP metrics correlate with corresponding Mark LSP dates, we calculated Kendall's rank correlation coefficient (τ) (Kendall, 1938) to compare LSP dates from different sensors, and to compare LSP dates with planting and harvest date timing. Kendall's τ is known to be more robust in distributions with many outliers than Pearson's correlation coefficient (Dehling et al., 2012), and has been used previously in phenological studies (Jong et al., 2011).

3. Results

3.1. Within-sensor comparison

Overall, LSP dates and season lengths were largely consistent when comparing indices from the same sensor. For multispectral satellite sensors (Sentinel-2 and VIIRS), the bias and MAD for LSP dates were 6 days or less between different indices (Table S1), and the bias and MAD for season and sub-season lengths were 9 days or less (Table S2). Small magnitude biases in LSP dates did exist for different VIs of the same satellite sensor, generally not exceeding 4 days (Table S1). LSP dates from NDVI had slightly longer season length and GCVI dates had slightly shorter length compared to EVI2 for Sentinel-2 and VIIRS (Table S2). For Sentinel-1, VH and VV backscatter had small biases (2-5 days) with moderate variability (MAD between 7-13 days).

The biases and MAD for different VI's of the Mark ground sensor were slightly higher, up to 8 days. For season length, the Mark NDVI had substantially longer season length; the full-season length for Mark NDVI was, at the median, 18 days longer than the Mark EVI2 full season length. Mark GCVI season and sub-season lengths were more comparable to Mark EVI2 (bias of 6 days or less). The larger shifts for Mark VI's may have been due to a larger discrepancy in time-series of different VI's for the ground Mark sensor compared to satellite sensor, or due to the effects of padding VI time-series.

Overall, the magnitude of variations for different multispectral VI's was generally less than one week for both bias and MAD. For this reason, in subsequent sections we provide results for a single VI per sensor, except when noted otherwise. For Sentinel-2, VIIRS, and Mark sensors, we selected EVI2 for its consistency with other LSP studies and data products, and because LSP timing for EVI2 was intermediate between NDVI and GCVI LSP timing. For Sentinel-1, we selected VH backscatter converted to float as the reference index.

3.2. Satellite to Mark Comparison

3.2.1. Shifts in LSP Dates and Season Length

The shifts between satellite and Mark LSP were far larger than the within-sensor comparison in the previous section. Among satellite sensors, Sentinel-2 had the lowest MAD for all LSP dates (11-13 days, Table 4) and sub-season durations (9-20 days, Table 5), indicating that Sentinel-2 LSP was most consistent when compared to Mark LSP. Sentinel-2 LSP showed an increased season length compared to Mark LSP, with earlier greenup and maturity dates, and later senescence and dormancy dates. At the median level, Sentinel-2 length of season was 40 days longer than Mark length of season.

VIIRS LSP exhibited an even more extreme lengthening of season. At the median level, VIIRS full season length was 65 days longer compared to the Mark. VIIRS LSP dates had a higher MAD (18-25 days) than Sentinel-2 when compared with Mark LSP dates. Similar to Sentinel-2, early season VIIRS LSP dates (greenup, maturity) were shifted earlier, and late season LSP dates (senescence, dormancy) were shifted later compared to the Mark.

Sentinel-1 LSP dates had higher bias-adjusted MAD ranging from 22-34 days. This high variability is likely not usable for agricultural monitoring, implying that Sentinel-1 cannot provide consistent LSP timing on its own.

VLCD dates had the lowest bias (0-6 days) of any satellite-based LSP, and may provide the most accurate representation of ground-measured LSP without bias adjustment. VLCD also had low bias for season duration, with a full season length 1.5 days shorter than the Mark. The MAD values for VLCD LSP dates (16-21 days) (Table 4) were intermediate between Sentinel-2 (11-13-days) and VIIRS surface reflectance (18-25 days), and higher than either sensor for full season length (Table 5).

A critical limitation of VLCD data is that certain sites did not have any LSP dates extracted. This absence occurred because certain years had no data in the VLCD product, or because the VLCD only detected dates for a primary season that did not overlap with a site's growing season. VLCD coverage rate for greenup was 67%, implying that VLCD had no recorded LSP date for greenup at any time during the growing season for 33% of site-seasons. For the maturity, senescence and dormancy dates, the coverage rate was 80%, 85%, and 40%, respectively. In total 119 of 144 sites (82.6%) were missing a valid VLCD date for at least one of the four LSP dates extracted.

Table 4 Sensor timing (SSA) compared to Mark EVI2 timing. Median bias (MAD from bias) compared to Mark EVI2 LSP.

Sensor	**Sensor VI**	**Greenup**	**Maturity**	**Senescence**	**Dormancy**
S2	EVI2	-17 (11)	-6.5 (11.5)	12.5 (13)	17 (13)
VIIRS	EVI2	-34 (18)	-21 (24)	18 (20)	27 (19)
S1	VH	-32 (22)	-17.5 (22.5)	12 (34)	22 (24.5)
VLCD		-4 (19)	-5.5 (15.5)	-1 (18)	-1 (21)

Table 5 Sensor sub-season length (SSA) compared to Mark EVI2 timing. Median bias (MAD from bias) compared to Mark EVI2 LSP.

Sensor	**Sensor VI**	**Early (G to M)**	**Peak (M to S)**	**Late (S to D)**	**Full (G to D)**
S2	EVI2	10 (9)	22 (17)	23 (18)	39 (18)
VIIRS	EVI2	10 (14)	45.5 (26)	48 (24)	64 (23)
S1	VH	9 (14)	36.5 (28.5)	42 (23.5)	62.5 (26.5)
VLCD		-3 (19.5)	15 (19)	10 (25)	4 (42.5)

We also include the median absolute deviation *without* bias adjustment in Table S3 for LSP dates (analogous to the bias-adjusted metrics in Table 4). The non bias-adjusted MAD represents the typical difference between ground and satellite LSP estimates when we do not have ground sensors available for bias adjustment. The non bias-adjusted MAD is typically higher than bias-adjusted MAD, but this increase varies based on sensor and LSP date. Without bias adjustment, Sentinel-2 MAD increases from 11-13 days to 13-20 days, VIIRS MAD increases from 18-24 days to 27-35 days), and Sentinel-1 MAD increases from 22-34 days to 27-38 days. These increases are not surprising due to the typical biases for Sentinel-2 and VIIRS surface reflectance. Interestingly, the VLCD MAD does not shift considerably without bias-adjustment due to the lower biases for VLCD LSP dates (Table 4). VLCD MAD shifts do not exceed 1.5 days for any LSP date (Table S3). Thus without ground sensors needed for bias correction, the VLCD product (non bias-adjusted MAD of 17-20 days) performs similarly to Sentinel-2 when compared with reference LSP dates. Notably, the VLCD product is missing data for many sites as mentioned above.

3.2.2. LSP Proximity by Time Period

To provide a more concrete comparison of different sensors with Mark LSP, we calculated the percent of LSP dates for each sensor that were within a given time period of Mark dates. We adjusted for bias first given the large and consistent biases of satellite sensors for most LSP dates.

Sentinel-2 had the highest bias-adjusted proximity with Mark LSP for all LSP dates (Figure 3). VIIRS surface reflectance dates had fewer sites within a given proximity, especially for the mid-season LSP dates (maturity and senescence). The lower proximity for VIIRS surface reflectance for mid-season dates may be due to higher cloud rates during the peak of the rainy season, and higher within-pixel variability in mid-season crop timing compared to greenup and dormancy dates. Sentinel-1 had the lowest proximity, except for maturity, where VIIRS and Sentinel-1 performed comparably. The performance of VLCD is difficult to assess due to missing observations for certain site-seasons, meaning that the set of sites used for VLCD is substantially smaller (e.g., 102 sites for maturity) compared to other sensors (128 sites for Sentinel-2 and VIIRS maturity). Among these sites, VLCD did show comparable or slightly higher proximity with ground LSP when compared to VIIRS surface reflectance for maturity and senescence.

3.2.3. Patterns in Full Season Duration

A major pattern seen when comparing LSP-based season lengths is that full season length is 39 and 64 days longer for Sentinel-2 and VIIRS surface reflectance, respectively (at the median), when compared to Mark derived season length (Table 5). This pattern is fairly consistent for SSA sites with MAD values of 18 and 23 days respectively for Sentinel-2 and VIIRS surface reflectance. Accounting for these seasonal length shifts is thus essential to link field-measured and satellite LSP. A comparison of full-season lengths for all sites is shown in Figure 4 (top), which shows the fairly consistent offsets in season length between sensors. Dashed lines represent the median increase in season length above the 1:1 line: 39 days for Sentinel-2 and 64 days for VIIRS.

Figure 4 (bottom) displays the raw observations, rough-fit curve, and extracted LSP dates for Mark (orange), Sentinel-2 (green), and VIIRS surface reflectance (blue) EVI2 (rescaled) at a single Zambia site. The VIIRS time-series lacks observations near and before season peak, resulting in earlier dates for greenup and maturity compared to the Mark. These early LSP dates may in fact be capturing a weeding signal, and not the crop signal, as this site had Mark observations truncated due to weeding. The Sentinel-2 time-series also had a large gap during early and peak season, resulting in an earlier greenup date after curve-fitting. VIIRS and Sentinel-2 had similar late-season time-series and extracted senescence and dormancy dates. For both VIIRS and Sentinel-2, the late season decline in EVI2 occurred later and more gradually when compared to the Mark time-series. Mark sensors, potentially due to their proximity and clear crop signal, detect a much sharper decline in late season EVI2.

3.2.4. Explanatory Power of Satellite LSP

We calculated Kendall's τ to assess the strength of correlations between satellite and Mark LSP metrics. We included results for a single VI per sensor, as well as multi-sensor mean, median, and Random Forest models (Table 6). Extended results for all VI's, including red-edge and SWIR based VI's (and their formulas) are in Table S4. To assure we compared the same set of sites for all sensors, we excluded Zambia 2021-22 sites that did not have Sentinel-1 imagery and thus could not be included in multi-sensor models. We also did not include results for the VLCD product due to many sites lacking VLCD LSP data.

Sentinel-2 had slightly higher τ values than VIIRS EVI2 when compared with Mark LSP dates (Table 6). Sentinel-1 had the lowest correlation with ground-measured LSP averaged across all LSP dates (average $\tau = 0.38$), although it performed better in predicting late-season dates.

Multi-sensor LSP dates had comparable or slightly higher τ values compared to single-sensor LSP dates (Table 6). Satellite median and a Random Forest model including splined inputs had highest correlation with Mark LSP dates (average $\tau = 0.63$). These two models had slightly higher correlation compared to the single-sensor model with highest correlation, Sentinel-2 EVI2 (τ = 0.54). For all sensors/models, early-season LSP dates (greenup and maturity) had lower τ values than late-season dates (senescence and dormancy).

Table 6: Single-sensor τ for each LSP date and average across all LSP dates. In all cases, the sensor LSP was compared with Mark EVI2 LSP. The highest single-sensor (top) and multi-sensor (bottom) τ for each column is bolded.

Sensor/model	Greenup	Maturity	Senescence	Dormancy	Average
Single Sensor					
Sentinel-2 EVI2	**0.51**	**0.54**	**0.7**	**0.72**	**0.62**
VIIRS EVI2	0.43	0.46	0.66	0.71	0.57
Sentinel-1 VH	0.38	0.5	0.58	0.68	0.54
Multi-sensor					
Random Forest	**0.54**	**0.61**	0.68	0.69	**0.63**
RF Spline	0.53	**0.61**	0.67	0.69	0.62
Satellite Median	0.5	0.55	**0.72**	**0.76**	**0.63**
Satellite Mean	0.47	0.54	0.7	0.74	0.62

3.3. LSP Timing Compared to Management Events

A subset of sites had planting (n = 40) and harvest (n = 29) date recorded. This number of observations is small, but did allow for an initial assessment of differences between satellite LSP and both planting and harvest dates. Previous literature has linked planting dates to sensor greenup, and harvest dates to sensor dormancy (see Gao and Zhang, 2021). We thus used similar methods to the previous section, substituting planting and harvest dates as the "reference" dates for comparison. Table 7 lists the median bias and MAD comparing sensor greenup to planting date, and sensor dormancy to harvest date. A negative bias implies that the median LSP date (e.g., greenup) occurs before the management date (e.g., planting). We excluded VLCD from this analysis due to the number of VLCD seasons with missing observations.

Table 7: Sensor timing (SSA) compared to management events, displaying median bias (MAD from bias) for sensor LSP dates (greenup for Planting, dormancy for Harvest) compared to management events. Positive bias means that the LSP date (e.g., greenup) occurred after the management event (Planting) at the median level.

Sensor	**VI**	**Bias (MAD) Greenup - Planting Date**	**Bias (MAD) Dormancy - Harvest Date**
Mark	EVI2	26 (13)	0 (21)
S2	EVI2	0.5 (15)	13 (25)
VIIRS	EVI2	-8 (13.5)	27 (21)
S1	VH/VV	-13 (29.5)	21.5 (21.5)

For planting date, the Mark sensor had a large positive bias, indicating that greenup occurred, at the median level, 26 days after planting. This lag is to be expected if greenup is due predominantly to crop signal and not due to weeds or surrounding vegetation. Mark, Sentinel-2, and VIIRS had comparable MAD (13-15 days), while the high MAD for Sentinel-1 (29.5 days) indicated that it is not effective for tracking planting activities using LSP. Sentinel-1, Sentinel-2, and VIIRS all had lower bias (≤ 13 days) compared to Mark (26 days). However, this low bias should not be interpreted as higher accuracy, as a delay of several weeks between planting and LSP greenup is expected. For example, a delay of 19 days was seen between crop emergence and LSP greenup for US corn and soybean (Diao and Li, 2022), although this delay can be dependent on the LSP extraction method (Gao et al., 2021).

For harvest date, Mark dormancy had a low bias (0 days) and a higher MAD when compared to planting (21 days). The higher MAD for Mark in harvest date indicates that it is more difficult to track than planting date. Among satellite sensors, VIIRS EVI2, Mark EVI 2, and Sentinel-1 VH had the lowest MAD (21 days). Sentinel-2 had a higher MAD for harvest date (25 days), indicating the difficulty of detecting harvest at the field level. Interestingly, the MAD for Sentinel-1 is lower for harvest (21.5 days) than for planting (29.5 days).

We also calculated the percent of sites where the LSP dates are within a given time period to the management event, after bias adjustment (Figure 5). Sentinel-2 and VIIRS had similar proximity for planting date (behind Mark), while VIIRS had the highest proximity among satellite sensors for harvest date. The high proximity of VIIRS for tracking harvest date (even higher than Marks for time periods of 10, 20 days), is worth noting, although its cause is unclear. Sentinel-1 again had the fewest sites within a given proximity, although Sentinel-1 did perform better at tracking harvest compared to planting date.

Finally, we examined the correlation of Mark LSP dates with planting and harvest dates. The τ values for harvest were far higher than for planting dates for all sensors/methods (Tables 8, S5). The highest τ for planting date was for the single-sensor VIIRS GCVI, while Sentinel-2 EVI2 had highest correlation with harvest date (EVI2, $\tau = 0.75$).

Table 8: τ for each planting and harvest date using satellite derived greenup and dormancy date, respectively. The highest single-sensor (top) and multi-sensor (bottom) τ for each column is bolded.

VI/Model type	Planting	Harvest
Single sensor		
S2 EVI2	**0.23**	**0.75**
VIIRS EVI2	0.2	0.66
S1 CR	0.05	0.66
Multi-sensor		
Random Forest	0.2	**0.66**
RF Spline	0.21	**0.66**
Satellite Median	0.21	0.64
Satellite Mean	**0.22**	0.65

4. Discussion

This study leads to several clear findings on LSP at smallholder maize sites in SSA, and provides new insight into the differences in how satellite-based LSP measurements vary compared to ground measurements. For regional policy-makers, the results detail consistent biases of satellite-derived LSP, allowing for a more robust interpretation of satellite-derived LSP dates and their uncertainty. In particular, satellite-derived LSP shifts can be used to identify where farmers are shifting management practices (Chakraborty et al., 2018), particularly when using higher-resolution imagery (e.g., Sentinel-2) less prone to mixed-pixel signals in areas with high management heterogeneity (Waldman et al., 2017; Bégué et al., 2020). We also assessed how satellite based vegetation indices using different spectral bands (e.g., near infrared in combination with red, green, red-edge, or short-wave infrared) compare with ground-based LSP, as previous studies have recommended examining VI's from different spectral regions (Misra et al., 2020).

4.1 VI and Sensor comparison

LSP based on different VI's from the same satellite sensor had differences in timing of less than three days (Table S1). Our study reinforced the use of EVI2 as the reference VI for agricultural LSP, such as its use data products like VLCD (Zhang et al., 2018c). For Sentinel-2 and VIIRS, EVI2 and NDVI had nearly equivalent correlation with ground-based LSP metrics (Table S4), similar to previous studies that found minimal differences in explanatory power for VIIRS EVI2 and NDVI (Zhang et al., 2018a). For both Sentinel-2 and VIIRS, EVI2 LSP dates had intermediate growing season length compared to NDVI (longer) and GCVI (shorter) (Table S2). Previous agricultural focused studies have found advantages of green-band indices such as GCVI (Peter et al., 2020). However, our results showed that Sentinel-2 and VIIRS LSP dates derived from red-based indices (EVI2, NDVI) had slightly higher correlation with ground-measured LSP dates as compared to LSP dates derived from GCVI for the same sensor (Table S5).

For Sentinel-2, we also examined red-edge and SWIR based indices, per recommendations based on a previous review of phenological studies (Misra et al., 2020), but did not find improved correspondence with Mark LSP metrics for these indices. The red-edge based Leaf Chlorophyll Index (LCI, average $\tau = 0.61$) had similar relationships with Mark LSP as those calculated from Sentinel-2 EVI2 and NDVI (average $\tau = 0.62$) (Table S4). SWIR-based indices had slightly lower correlations, for example the Normalized Difference Moisture Index

(NDMI, average $\tau = 0.59$). Overall, we found similar correlation with Mark LSP for Sentinel-2 indices using several spectral bands (green, red, red-edge, SWIR) in addition to NIR bands, with average τ ranging from 0.59 to 0.62 (Table S4). with highest τ for Sentinel-2 NDVI and EVI2.

For Sentinel-1, VH backscatter had the highest average correlation with ground-measured LSP ($\tau = 0.54$), followed by VV backscatter (average $\tau = 0.49$) while the calculated indices CR and RVI had lower correlation (average $\tau = 0.38$, 0.37 respectively). These results are similar to a previous study on maize growth stages in Germany (Holtgrave et al., 2020) where VH backscatter had higher correlation than VV backscatter or RVI with optical indices for all maize growth stages.

Compared to differences in LSP dates arising from the VI choice, far larger differences were found when comparing LSP metrics from different sensors and data products. A novel aspect of our study is the comparison of LSP metrics extracted from both individual sensors and from multi-sensor methods. A previous review of cropping practice mapping found that most studies only used a single sensor (Bégué et al., 2018). The satellite sensors used in this study have differing spatial resolution, revisit time, and cloud-penetrating ability, and our results demonstrated clear differences in the biases and variations among these sensors for LSP monitoring.

Sentinel-2 based LSP generally had the smallest differences among satellite sensors when compared to ground-measured LSP metrics or planting and harvest dates. The MAD for Sentinel-2 LSP was lowest after adjusting for bias (11-13 days) compared to VIIRS (18-24 days), and Sentinel-1 (22-34 days), indicating its LSP metrics have the lowest difference with ground-measured metrics, potentially due to the higher spatial resolution of Sentinel-2 compared to VIIRS and VLCD.

Length of season calculated from LSP planting and harvesting dates also progressively increased from ground-sensors to Sentinel-2 to VIIRS, indicating that the sensors with finer spatial resolutions may yield shorter calculated season lengths. A similar comparison for US corn fields found slightly longer season length for MODIS as compared to Sentinel-2 (4 day increase), although the magnitude was far larger for winter wheat fields (40 day increase) (Pipia et al., 2022). For comparison, our study showed a 25 day increase in season length from Sentinel-2 to VIIRS for smallholder maize fields (Table 5).

Sentinel-1 had the highest MAD of any sensor or data product (Table 4). Like other satellite sensors, Sentinel-1 had stronger correlation with ground LSP metrics for late season events (senescence, dormancy, harvest) compared to early season events (greenup, maturity, planting) (Table S4). Sentinel-1 has been used in LSP studies for European crops (Holtgrave et al., 2020; Khabbazan et al., 2019; Meroni et al., 2021). Meroni et al. compared LSP metrics from Sentinel-1 CR and Sentinel-2 NDVI for European maize and found Pearson's correlations of 0.57, 0.52, and 0.42 for start, end, and peak season timing between the two sensors, albeit with biases generally of 10-20 days between LSP dates for the two sensors (Meroni et al., 2021). Holtgrave found that VH backscatter correlated highest with Sentinel-2 indices (NDVI, NDWI, and a red-edge index, PSRI[3]), with Pearson's correlations of values of 0.69-0.83 for different maize growth stages, substantially higher than the τ correlation between Sentinel-1 VH and ground LSP metrics in our study (0.38-0.68). For reference, we also calculated Pearson's correlation for Sentinel-1 VH and found values of 0.39 to 0.85, with an average Pearson's correlation of 0.65, still below the range found in Meroni's study. Our study found that VH and VV outperformed the cross-ratio (Table S4), but did not have τ values as high as most Sentinel-2

[3] PSRI = Plant Senescence Reflectance Index. (B4-B2)/B6

and VIIRS vegetation indices. Sentinel-1 derived LSP may be less effective in smallholder environments as compared to the above European-based studies, as small field size can complicate data cleaning steps like speckle filtering or taking field-level medians (as in Holtgrave et al., 2020; Khabbazan et al., 2019; Meroni et al., 2021). We conclude that for our study's smallholder sites, Sentinel-1 and Sentinel-2 LSP are not directly comparable or interchangeable.

VLCD results notably had the lowest magnitude biases, albeit with high missing data rates (15%-60% of site-seasons, depending on the specific LSP metric). Low magnitude biases are desirable in use cases where ground-measured LSP metrics are not available for bias-correction. In such cases, VLCD may be an effective data product for LSP monitoring. The low biases may be due to effective processing methods for VLCD (Zhang et al., 2018c) that minimize the influence of cloud cover and missing data, despite the coarse resolution of VLCD that may result in multiple fields being included within each VLCD pixel. We also note that VLCD provided a relatively accurate representation of season lengths (Table 5), without the stretched seasonality seen for Sentinel-2 and VIIRS surface reflectance.

When calculating MAD values without bias-adjustment (Table S3), the VLCD MAD values (17-20 days) were similar to those for Sentinel-2 (13-20 days). Thus VLCD may be a viable alternative for LSP monitoring when consistent biases between satellite and ground LSP dates are unknown. End-users seeking to use VLCD for smallholder monitoring should first examine the historical VLCD observation rate for previous seasons before relying upon it for monitoring. In our study, only 25 of 144 sites (17.4%) had valid VLCD observations for all four LSP dates.

We place our study's results in the context of US based studies due to the lack of previous field-level LSP studies in SSA. Notably the results of these US studies do not perfectly align with our findings on bias direction and early-season vs. late-season timing. A US study using PhenoCam found a high R^2 for croplands (between 0.69 and 0.83), and essentially no difference in using NDVI or EVI2 based LSP from VIIRS (Zhang et al., 2018b). Notably, in that study the highest R^2 was for greenup phase (SOS), while in our study late-season LSP dates (senescence, dormancy) had higher τ values. Other US based studies have found that multi-sensor image fusion improved LSP estimates (e.g., Moon et al., 2019; Zhang et al., 2020; Shen et al., 2022; Diao and Li, 2022; Tran et al., 2022). A study across 12 Midwestern US states found high association (typically $R^2 \geq 0.9$) between HLS-VIIRS LSP and field crop progress reports provided by the National Agricultural Statistics Service (NASS) (Shen et al., 2022). The authors found a bias of 10 days between planting and greenup, and -18 days between harvest and dormancy (i.e., dormancy occurs before harvest). These results do not match our study's findings, where Sentinel-2 had a small positive bias (2 days, Table 7) between planting and greenup, and a larger positive bias (18 days, Table 7) between harvest and dormancy, indicating dormancy occurred after harvest. These discrepancies between US studies and our study's results could be due to different management practices in the US vs smallholder sites, including early weeding events and later harvest (after crop has yellowed) for smallholder sites.

Synthesizing the above studies, we see that the biases between ground and satellite LSP dates are vastly different for US and smallholder agriculture, indicating that approaches used in US studies may not be transferable to smallholder settings. Several of the US studies found substantial improvements due to sensor-level image fusion methods beyond the scope of this paper. We used simpler fusion methods that found small improvements in LSP estimation (average $\tau = 0.63$ for spline Random Forest and satellite median, compared to $\tau = 0.62$ for Sentinel-2 EVI2). Additional research and data could help identify for which use-cases and which model type (e.g., Random Forest) would best exploit the complementary advantages of multiple sensors for smallholder agriculture.

4.2 Limitations and Next Steps

Land surface phenology and downstream products for smallholder systems will require a combination of (i) high quality ground validation datasets, (ii) appropriate satellite imagery, and (iii) appropriate EO models. The combination of these three factors has led to rapid advancement of LSP modeling and applications, particularly in the US, while the development of these methods in smallholder contexts is more limited. There are several ways in which future studies can build upon this study's results and expand our understanding of LSP metrics for smallholder agriculture monitoring.

4.2.1 Ground Data Networks and Timely Management Observations

This study used a network of Arable Marks to provide continuous measurements of crop growth. While the Marks provided valuable field-level measurements, our network only included fields in two countries and a limited number of site-seasons (n = 144). Also, as with all field-based sensors, the Marks have a limited sensing footprint (10 meter diameter), and furthermore had to be installed after planting and removed before harvest, which, combined with the need to be recharged, resulted in data gaps. We used curve-fitting and a padding method to provide realistic estimates of VI curves for these gaps (Figure 2), but ideally the sensors would also record observations during shoulder seasons and throughout the growing season without interruption. We also lacked ground data on crop growth stages and comprehensive management decisions at a number of sites (e.g., weeding, fertilizer, maize cultivar, irrigation). Planting and harvest dates were also recorded only for a limited number of sites.

Ideally, continental scale networks like PhenoCam and Fluxnet could be constructed for smallholder sites in Africa, as there are few, if any, active sites in Africa currently for these networks.[4] More flexibly, data sensors within the internet-of-things (IoT) umbrella could provide opportunities for decentralized data collection for smallholder agriculture (Antony et al., 2020; Bayih et al., 2022; Routray et al., 2019). Several examples of IoT data collection for smallholder agriculture exist, including the use of smartphone sensors for estimating crop phenology (Hufkens et al., 2019; Tonnang et al., 2020), and monitoring crop progress and management through mobile phone-based surveys (Giroux et al., 2019).

In an overview of Earth Observation agricultural monitoring systems in SSA, Nakalembe stresses the importance of investing in "consistent representative digital ground data through networks of observers to provide the data required to train EO-based systems" (Nakalembe, 2020). To the extent possible, ground datasets should be well-documented, spatially referenced, and, when feasible, shared. Standardizing and sharing observations from ground-based sensors and surveys (to the extent possible due to privacy considerations) can improve comparisons between different sensor types, and expand the pool of ground observations for use in LSP studies.

We also note that some of the planting and harvest dates recorded (for Kenya sites in particular) were self-reported, and previous studies have found that such self-reported observations may be unreliable if not collected soon after the event (see Jain et al., 2016). Self-reported observations may also result in unreliable downstream models if used for calibration

[4] See maps at https://fluxnet.org/sites/site-summary/ , https://phenocam.nau.edu/webcam/network/map/

(Paliwal and Jain, 2020). We thus interpret results on planting and harvest dates with caution, and recommend more expansive and timely collection of management practices in conjunction with extended ground sensor networks.

4.2.2. Improved sensor resolution and image fusion

Our study used freely available satellite sensors and primarily derived LSP dates from single sensor methods. However even the best performing sensor, Sentinel-2, still may have too coarse a resolution (10 meters) for small fields due to mixed pixel effects for field boundary pixels and the presence of buildings, trees and other vegetation cover. The revisit time for Sentinel-2 (5 days) also still allows for large gaps in time-series due to frequent cloud cover. Commercial satellite sensors like PlanetScope and image fusion methods improve spatial and temporal resolution, and have proven effective for LSP monitoring in the US and Europe, but need to be examined for smallholder agriculture.

The high-resolution (2m) SkySat sensor has been used to estimate sowing date and yield for smallholder wheat fields in India, using GCVI (Jain et al., 2016). The study found R^2 values of 0.41 and 0.62 for sowing date estimation, based on GCVI values at specific observation dates, improving on the low correlation of Sentinel-2 GCVI with planting date in this study ($\tau = 0.39$, Table S5), potentially due to SkySat's higher resolution. In smallholder contexts, PlanetScope imagery, with a near daily revisit and 3-4 meter resolution, has been used to provide high-resolution crop field boundaries in Ghana (Estes et al., 2022). There thus is promise to use the Planet constellation for LSP monitoring in smallholder contexts, particularly the Planet Fusion product which is designed to gap-fill cloudy or missing scenes (Planet, 2021). Planet Fusion imagery has been used to classify maize growth stages in the US, improving upon results based on Sentinel-2 imagery (Nieto et al., 2022). Future studies could examine if high-resolution sensors or data fusion products effectively decrease the variability (MAD) in satellite LSP compared to the methods used in this study. While such products are not freely available, they can provide target levels of performance for satellite-based LSP, or conversely confirm that satellite LSP has consistent variability vis-a-vis ground measure LSP regardless of the satellite source. Currently the Planet Fusion product is not freely available, and has not to our knowledge been used to examine LSP for smallholder agriculture.

Other image fusion methods have been applied extensively in the US to study LSP, typically merging coarse sensors (VIIRS, MODIS) with high revisit times with medium-resolution sensors (Sentinel-2, Landsat) that are more infrequent (Gao et al., 2006; Luo et al., 2018). In smallholder contexts, ESTARFM (Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model; Zhu et al., 2010) has been modified for use in high cloud regions (Knauer et al., 2016), and used to calculate LSP metrics in Burkina Faso (Knauer et al., 2017). While an investigation of image fusion based LSP for smallholder agriculture is beyond the scope of this paper, these results show there is potential for image fusion methods, and that additional testing and validation in smallholder contexts is needed.

4.2.3. Appropriate Models for Smallholder LSP

The LSP methods used in our study, including curve-fitting, VI choice, and metric extraction, were based on common LSP practices that have been used in the production of MODIS and VIIRS Land Cover Dynamics products (Ganguly et al., 2010; Zhang et al., 2006, 2018c). We primarily used single sensor methods and simple multi-sensor models to provide baselines of sensor performance. However, these standard practices may not be ideally suited for smallholder agriculture. High cloud cover and small field size require methods that (i) do not depend on a large spatial footprint, (ii) can handle large gaps in optical time-series, and (iii) ideally can use multiple sensors to compensate when a single sensor has missing data.

Machine learning models, properly designed, can meet the above requirements, and have shown promise in smallholder contexts, provided they are sufficiently well trained. For example, deep learning models have been used for crop type mapping in Ghana and Sudan (Rustowicz et al., 2019) and Pakistan (Khan et al., 2023), crop boundary delineation in Nigeria and Mali using very high resolution WorldView-3 imagery (Persello et al., 2021), and smallholder maize yield estimation in China (Zhang et al., 2021).

However, a review of deep learning studies focused on phenological stage mapping (Katal et al., 2022) found the majority of studies in the US, Europe, Australia, and New Zealand, and no studies based in Africa. This may be due to the high ground data requirements for phenological modeling, requiring data throughout the growing season. Geospatial foundation models (GFM) offer promise for applications with limited data on a variety of downstream tasks (see Jakubik et al., 2023), and may have potential for LSP modeling due to their ability to reconstruct missing observations. In particular, foundation models that do not require large spatial footprints, like the single-pixel based PRESTO (Pretrained Remote Sensing Transformer; Tseng et al., 2023) model, may be well-suited for small fields. Above all, these models must be evaluated in smallholder settings, and rely upon labeled data representing a range of realistic crop growth and management. This data dependence of new machine learning models again highlights the need for high quality ground observations in smallholder LSP monitoring.

5. Conclusion

This study provides a multi-sensor evaluation of the ability to estimate LSP over smallholder maize fields in Kenya and Zambia, comparing LSP timing from *in situ* sensors with measurements derived from Sentinel-1, 2, and VIIRS. Sentinel-2 LSP performed best among individual sensors, although multi-sensor methods performed comparably to single sensor models (similar τ values) in many cases and warrant further study. To support development of LSP models for smallholder systems, we recommend the development of ground-based sensor networks and datasets to create standardized reference data resources, similar to the PhenoCam network and Crop Progress Reports in the US. Existing datasets from crop trials and previous studies could provide field observations to such a database. New sensor networks based on ground-based phenological cameras or multispectral sensors (like the Marks in this study) can provide continuous reference data as well for crop growth. The variability of differences between Sentinel-2 and ground LSP (MAD of 11-13 days) indicates that LSP estimates may only be accurate within two weeks (at the median). However, this study's methods still provide a foundational comparison of LSP from multiple single-sensors and multi-sensor fusion for smallholder agriculture and identifies potential next steps in multi-sensor modeling, machine learning, and ground sensor networks. Similar to the development of LSP methods in the US, improved smallholder LSP methods can provide critical metrics to understand regional crop timing, management practices, and climate variability at broad scales, provided sufficient ground data sets are available.

References

Adole, T., Dash, J., Atkinson, P.M., 2018a. Characterising the land surface phenology of Africa using 500 m MODIS EVI. Appl. Geogr. 90, 187–199. https://doi.org/10.1016/j.apgeog.2017.12.006

Adole, T., Dash, J., Atkinson, P.M., 2018b. Major trends in the land surface phenology (LSP) of Africa, controlling for land-cover change. Int. J. Remote Sens. 39, 8060–8075. https://doi.org/10.1080/01431161.2018.1479797

Antony, A.P., Leith, K., Jolley, C., Lu, J., Sweeney, D.J., 2020. A Review of Practice and Implementation of the Internet of Things (IoT) for Smallholder Agriculture. Sustainability-basel 12, 3750. https://doi.org/10.3390/su12093750

Atzberger, C., Eilers, P.H.C., 2011. Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. Int. J. Remote Sens. 32, 3689–3709. https://doi.org/10.1080/01431161003762405

Bayih, A.Z., Morales, J., Assabie, Y., By, R.A. de, 2022. Utilization of Internet of Things and Wireless Sensor Networks for Sustainable Smallholder Agriculture. Sensors 22, 3273. https://doi.org/10.3390/s22093273

Beck, P.S.A., Atzberger, C., Høgda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. Remote Sens Environ 100, 321–334. https://doi.org/10.1016/j.rse.2005.10.021

Bégué, A., Arvor, D., Bellon, B., Betbeder, J., Abelleyra, D. de, Ferraz, R.P.D., Lebourgeois, V., Lelong, C., Simões, M., Verón, S.R., 2018. Remote Sensing and Cropping Practices: A Review. Remote Sens-basel 10, 99. https://doi.org/10.3390/rs10010099

Bégué, A., Leroux, L., Soumaré, M., Faure, J.-F., Diouf, A.A., Augusseau, X., Touré, L., Tonneau, J.-P., 2020. Remote Sensing Products and Services in Support of Agricultural Public Policies in Africa: Overview and Challenges. Front. Sustain. Food Syst. 4, 58. https://doi.org/10.3389/fsufs.2020.00058

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84. https://doi.org/10.1016/j.agrformet.2013.01.007

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

https://doi.org/10.1023/a:1010933404324

Brown, M.E., Beurs, K. de, Vrieling, A., 2010. The response of African land surface phenology to large scale climate oscillations. Remote Sens. Environ. 114, 2286–2296. https://doi.org/10.1016/j.rse.2010.05.005

Cecil, M., Chilenga, A., Chisanga, C., Gatti, N., Krell, N., Vergopolan, N., Baylis, K., Caylor, K., Evans, T., Konar, M., Sheffield, J., Estes, L., 2023. How much control do smallholder maize farmers have over yield? Field Crop. Res. 301, 109014. https://doi.org/10.1016/j.fcr.2023.109014

Chakraborty, D., Sehgal, V.K., Dhakar, R., Das, D.K., Sahoo, R.N., 2018. Trends and Change-Point in Satellite Derived Phenology Parameters in Major Wheat Growing Regions of North India During the Last Three Decades. J. Indian Soc. Remote Sens. 46, 59–68. https://doi.org/10.1007/s12524-017-0684-8

Dehling, H., Vogel, D., Wendler, M., Wied, D., 2012. Testing for Changes in Kendall's Tau. arXiv. https://doi.org/10.48550/arxiv.1203.4871

Diao, C., 2020. Remote sensing phenological monitoring framework to characterize corn and soybean physiological growing stages. Remote Sens. Environ. 248, 111960. https://doi.org/10.1016/j.rse.2020.111960

Diao, C., Li, G., 2022. Near-Surface and High-Resolution Satellite Time Series for Detecting Crop Phenology. Remote Sens. 14, 1957. https://doi.org/10.3390/rs14091957

Diao, C., Yang, Zijun, Gao, F., Zhang, X., Yang, Zhengwei, 2021. Hybrid phenology matching model for robust crop phenological retrieval. ISPRS J. Photogramm. Remote Sens. 181, 308–326. https://doi.org/10.1016/j.isprsjprs.2021.09.011

Dimou, M., Meroni, M., Rembold, F., 2018. Development of a national and sub-national crop calendars data set compatible with remote sensing derived land surface phenology. European Commission Joint Research Centre (JRC).

Diouf, A.A., Brandt, M., Verger, A., Jarroudi, M.E., Djaby, B., Fensholt, R., Ndione, J.A., Tychon, B., 2015. Fodder Biomass Monitoring in Sahelian Rangelands Using Phenological Metrics from FAPAR Time Series. Remote Sens. 7, 9122–9148. https://doi.org/10.3390/rs70709122

Duncan, J.M.A., Dash, J., Atkinson, P.M., 2015. The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes. Front. Environ. Sci. 3, 56. https://doi.org/10.3389/fenvs.2015.00056

Eilers, P.H.C., 2003. A Perfect Smoother. Anal. Chem. 75, 3631–3636. https://doi.org/10.1021/ac034173t

Estes, L.D., Ye, S., Song, L., Luo, B., Eastman, J.R., Meng, Z., Zhang, Q., McRitchie, D., Debats, S.R., Muhando, J., Amukoa, A.H., Kaloo, B.W., Makuru, J., Mbatia, B.K., Muasa, I.M., Mucha, J., Mugami, A.M., Mugami, J.M., Muinde, F.W., Mwawaza, F.M., Ochieng, J., Oduol, C.J., Oduor, P., Wanjiku, T., Wanyoike, J.G., Avery, R.B., Caylor, K.K., 2022. High Resolution, Annual Maps of Field Boundaries for Smallholder-Dominated Croplands at National Scales. Front. Artif. Intell. 4, 744863. https://doi.org/10.3389/frai.2021.744863

Frantz, D., Stellmes, M., Röder, A., Udelhoven, T., Mader, S., Hill, J., 2016. Improving the Spatial Resolution of Land Surface Phenology by Fusing Medium- and Coarse-Resolution Inputs. IEEE Trans. Geosci. Remote Sens. 54, 4153–4164. https://doi.org/10.1109/tgrs.2016.2537929

Ganguly, S., Friedl, M.A., Tan, B., Zhang, X., Verma, M., 2010. Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. Remote Sens. Environ. 114, 1805–1816. https://doi.org/10.1016/j.rse.2010.04.005

Gao, F., Anderson, M.C., Johnson, D.M., Seffrin, R., Wardlow, B., Suyker, A., Diao, C., Browning, D.M., 2021. Towards Routine Mapping of Crop Emergence within the Season Using the Harmonized Landsat and Sentinel-2 Dataset. Remote Sens. 13, 5074. https://doi.org/10.3390/rs13245074

Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., Mueller, R., Johnson, D.M., Prueger, J.H., 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sens. Environ. 188, 9–25. https://doi.org/10.1016/j.rse.2016.11.004

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance. Ieee T Geosci Remote 44, 2207–2218. https://doi.org/10.1109/tgrs.2006.872081

Gao, F., Zhang, X., 2021a. Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities. J Remote Sens 2021, 1–14. https://doi.org/10.34133/2021/8379391

Giroux, S.A., Kouper, I., Estes, L.D., Schumacher, J., Waldman, K., Greenshields, J.T., Dickinson, S.L., Caylor, K.K., Evans, T.P., 2019. A High-Frequency Mobile Phone Data Collection Approach for Research in Social-Environmental Systems: Applications in Climate Variability and Food Security in Sub-Saharan Africa. Environ Modell Softw 119, 57–69. https://doi.org/10.1016/j.envsoft.2019.05.011

Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys. Res. Lett. 30, n/a-n/a. https://doi.org/10.1029/2002gl016450

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gray, J., Sulla-Menashe, D., Friedl, M., 2022. User Guide to Collection 6.1 MODIS Land Cover Dynamics (MCD12Q2) Product. USGS.

Guido, Z., Zimmer, A., Lopus, S., Hannah, C., Gower, D., Waldman, K., Krell, N., Sheffield, J., Caylor, K., Evans, T., 2020. Farmer forecasts: Impacts of seasonal rainfall expectations on agricultural decision-making in Sub-Saharan Africa. Clim. Risk Manag. 30, 100247. https://doi.org/10.1016/j.crm.2020.100247

He, K., Sun, J., Tang, X., 2012. Guided Image Filtering. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1397–1409. https://doi.org/10.1109/tpami.2012.213

Holtgrave, A.-K., Röder, N., Ackermann, A., Erasmi, S., Kleinschmit, B., 2020. Comparing Sentinel-1 and -2 Data and Indices for Agricultural Land Use Monitoring. Remote Sens. 12, 2919. https://doi.org/10.3390/rs12182919

Htitiou, A., Boudhar, A., Chehbouni, A., Benabdelouahab, T., 2021. National-Scale Cropland Mapping Based on Phenological Metrics, Environmental Covariates, and Machine Learning on Google Earth Engine. Remote Sens. 13, 4378. https://doi.org/10.3390/rs13214378

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213. https://doi.org/10.1016/s0034-4257(02)00096-2

Hufkens, K., Melaas, E.K., Mann, M.L., Foster, T., Ceballos, F., Robles, M., Kramer, B., 2019. Monitoring crop phenology using a smartphone based near-surface remote sensing approach. Agric. For. Meteorol. 265, 327–337. https://doi.org/10.1016/j.agrformet.2018.11.002

Ittersum, M.K. van, Bussel, L.G.J. van, Wolf, J., Grassini, P., Wart, J. van, Guilpart, N., Claessens, L., Groot, H. de, Wiebe, K., Mason-D'Croz, D., Yang, H., Boogaard, H., Oort, P.A.J. van, Loon, M.P. van, Saito, K., Adimo, O., Adjei-Nsiah, S., Agali, A., Bala, A., Chikowo, R., Kaizzi, K., Kouressy, M., Makoi, J.H.J.R., Ouattara, K., Tesfaye, K., Cassman, K.G., 2016. Can sub-Saharan Africa feed itself? Proc National Acad Sci 113, 14964–14969. https://doi.org/10.1073/pnas.1610359113

Jain, M., Singh, B., Srivastava, A.A.K., Malik, R.K., McDonald, A.J., Lobell, D.B., 2017. Using satellite data to identify the causes of and potential solutions for yield gaps in India's Wheat Belt. Environ Res Lett 12, 094011. https://doi.org/10.1088/1748-9326/aa8228

Jain, M., Srivastava, A.K., Balwinder-Singh, Joon, R.K., McDonald, A., Royal, K., Lisaius, M.C., Lobell, D.B., 2016. Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data. Remote Sens-basel 8, 860. https://doi.org/10.3390/rs8100860

Jakubik, J., Roy, S., Phillips, C.E., Fraccaro, P., Godwin, D., Zadrozny, B., Szwarcman, D., Gomes, C., Nyirjesy, G., Edwards, B., Kimura, D., Simumba, N., Chu, L., Mukkavilli, S.K., Lambhate, D., Das, K., Bangalore, R., Oliveira, D., Muszynski, M., Ankur, K., Ramasubramanian, M., Gurung, I., Khallaghi, S., Hanxi, Li, Cecil, M., Ahmadi, M., Kordi, F., Alemohammad, H., Maskey, M., Ganti, R., Weldemariam, K., Ramachandran, R., 2023. Foundation Models for Generalist Geospatial Artificial Intelligence. arXiv. https://doi.org/10.48550/arxiv.2310.18660

Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112, 3833–3845. https://doi.org/10.1016/j.rse.2008.06.006

Jin, Z., Azzari, G., Burke, M., Aston, S., Lobell, D.B., 2017. Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa. Remote Sens-basel 9, 931. https://doi.org/10.3390/rs9090931

Jin, Z., Azzari, G., You, C., Tommaso, S.D., Aston, S., Burke, M., Lobell, D.B., 2019. Smallholder maize area and yield mapping at national scales with Google Earth Engine. Remote Sens Environ 228, 115–128. https://doi.org/10.1016/j.rse.2019.04.016

Johansen, K., Ziliani, M.G., Houborg, R., Franz, T.E., McCabe, M.F., 2022. CubeSat constellations provide enhanced crop phenology and digital agricultural insights using daily leaf area index retrievals. Sci. Rep. 12, 5244. https://doi.org/10.1038/s41598-022-09376-6

Jong, R. de, Bruin, S. de, Wit, A. de, Schaepman, M.E., Dent, D.L., 2011. Analysis of monotonic greening and browning trends from global NDVI time-series. Remote Sens. Environ. 115, 692–702. https://doi.org/10.1016/j.rse.2010.10.011

Katal, N., Rzanny, M., Mäder, P., Wäldchen, J., 2022. Deep Learning in Plant Phenological Research: A Systematic Literature Review. Front. Plant Sci. 13, 805738. https://doi.org/10.3389/fpls.2022.805738

Kendall, M.G., 1938. A NEW MEASURE OF RANK CORRELATION. Biometrika 30, 81–93. https://doi.org/10.1093/biomet/30.1-2.81

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Arntz, L.R., Marinetti, C., Valk, D. van der, Iannini, L., Molijn, R., Westerdijk, K., Sande, C. van der, 2019. Crop Monitoring Using Sentinel-1 Data: A Case Study from The Netherlands. Remote Sens-basel 11, 1887. https://doi.org/10.3390/rs11161887

Khan, H.R., Gillani, Z., Jamal, M.H., Athar, A., Chaudhry, M.T., Chao, H., He, Y., Chen, M., 2023. Early Identification of Crop Type for Smallholder Farming Systems Using Deep Learning on Time-Series Sentinel-2 Imagery. Sensors 23, 1779. https://doi.org/10.3390/s23041779

Kibret, K.S., Marohn, C., Cadisch, G., 2020. Use of MODIS EVI to map crop phenology, identify cropping systems, detect land use change and drought risk in Ethiopia – an application of Google Earth Engine. Eur. J. Remote Sens. 53, 176–191. https://doi.org/10.1080/22797254.2020.1786466

Kim, Y., Zyl, J.J. van, 2009. A Time-Series Approach to Estimate Soil Moisture using Polarimetric Radar Data. IEEE Trans. Geosci. Remote Sens. 47, 2519–2527. https://doi.org/10.1109/tgrs.2009.2014944

Klosterman, S.T., Hufkens, K., Gray, J.M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M.A., Richardson, A.D., 2014. Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11, 4305–4320. https://doi.org/10.5194/bg-11-4305-2014

Knauer, K., Gessner, U., Fensholt, R., Forkuor, G., Kuenzer, C., 2017. Monitoring Agricultural Expansion in Burkina Faso over 14 Years with 30 m Resolution Time Series: The Role of Population Growth and Implications for the Environment. Remote Sens. 9, 132. https://doi.org/10.3390/rs9020132

Knauer, K., Gessner, U., Fensholt, R., Kuenzer, C., 2016. An ESTARFM Fusion Framework for the Generation of Large-Scale Time Series in Cloud-Prone and Heterogeneous Landscapes. Remote Sens. 8, 425. https://doi.org/10.3390/rs8050425

Kong, D., McVicar, T.R., Xiao, M., Zhang, Y., Peña-Arancibia, J.L., Filippa, G., Xie, Y., Gu, X., 2022. phenofit: An R package for extracting vegetation phenology from time series remote sensing. Methods Ecol. Evol. 13, 1508–1527. https://doi.org/10.1111/2041-210x.13870

Krell, N., Davenport, F., Harrison, L., Turner, W., Peterson, S., Shukla, S., Marter-Kenyon, J., Husak, G., Evans, T., Caylor, K., 2022. Using real-time mobile phone data to characterize the relationships between small-scale farmers' planting dates and socio-environmental factors. Clim. Risk Manag. 35, 100396. https://doi.org/10.1016/j.crm.2022.100396

Krell, N.T., Giroux, S.A., Guido, Z., Hannah, C., Lopus, S.E., Caylor, K.K., Evans, T.P., 2021. Smallholder farmers' use of mobile phone services in central Kenya. Clim. Dev. 13, 215–227. https://doi.org/10.1080/17565529.2020.1748847

Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geosci. Remote Sens. Lett. 14, 778–782. https://doi.org/10.1109/lgrs.2017.2681128

Li, F., Zhang, X., Kondragunta, S., Csiszar, I., 2018. Comparison of Fire Radiative Power Estimates From VIIRS and MODIS Observations. J. Geophys. Res.: Atmos. 123, 4545–4563. https://doi.org/10.1029/2017jd027823

Lloyd, D., 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int. J. Remote Sens. 11, 2269–2279. https://doi.org/10.1080/01431169008955174

Lobell, D.B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., Murray, S., 2020. Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis. Am J Agr Econ 102, 202–219. https://doi.org/10.1093/ajae/aaz051

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens Environ 164, 324–333. https://doi.org/10.1016/j.rse.2015.04.021

Luo, Y., Guan, K., Peng, J., 2018. STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product. Remote Sens Environ 214, 87–99. https://doi.org/10.1016/j.rse.2018.04.042

Meroni, M., d'Andrimont, R., Vrieling, A., Fasbender, D., Lemoine, G., Rembold, F., Seguini, L., Verhegghen, A., 2021. Comparing land surface phenology of major European crops as derived from SAR and multispectral data of Sentinel-1 and -2. Remote Sens. Environ. 253, 112232. https://doi.org/10.1016/j.rse.2020.112232

Misra, G., Cawkwell, F., Wingler, A., 2020. Status of Phenological Research Using Sentinel-2 Data: A Review. Remote Sens. 12, 2760. https://doi.org/10.3390/rs12172760

Moon, M., Zhang, X., Henebry, G.M., Liu, L., Gray, J.M., Melaas, E.K., Friedl, M.A., 2019. Long-term continuity in land surface phenology measurements: A comparative assessment of the MODIS land cover dynamics and VIIRS land surface phenology products. Remote Sens. Environ. 226, 74–92. https://doi.org/10.1016/j.rse.2019.03.034

Nakalembe, C., 2020. Urgent and critical need for sub-Saharan African countries to invest in Earth observation-based agricultural early warning and monitoring systems. Environ. Res. Lett.

15, 121002. https://doi.org/10.1088/1748-9326/abc0bb

Nakalembe, C., Kerner, H., 2023. Considerations for AI-EO for agriculture in Sub-Saharan Africa. Environ. Res. Lett. 18, 041002. https://doi.org/10.1088/1748-9326/acc476

Nguy-Robertson, A., Gitelson, A., Peng, Y., Viña, A., Arkebauer, T., Rundquist, D., 2012. Green Leaf Area Index Estimation in Maize and Soybean: Combining Vegetation Indices to Achieve Maximal Sensitivity. Agron J 104, 1336–1347. https://doi.org/10.2134/agronj2012.0065

Nieto, L., Houborg, R., Zajdband, A., Jumpasut, A., Prasad, P.V.V., Olson, B.J.S.C., Ciampitti, I.A., 2022. Impact of High-Cadence Earth Observation in Maize Crop Phenology Classification. Remote Sens. 14, 469. https://doi.org/10.3390/rs14030469

Oliver, C., Quegan, S., 2004. Understanding synthetic aperture radar images. SciTech Publishing.

Paliwal, A., Jain, M., 2020. The Accuracy of Self-Reported Crop Yield Estimates and Their Ability to Train Remote Sensing Algorithms. Frontiers Sustain Food Syst 4, 25. https://doi.org/10.3389/fsufs.2020.00025

Persello, C., Wegner, J.D., Hänsch, R., Tuia, D., Ghamisi, P., Koeva, M., Camps-Valls, G., 2021. Deep Learning and Earth Observation to Support the Sustainable Development Goals. arXiv. https://doi.org/10.48550/arxiv.2112.11367

Peter, B.G., Messina, J.P., Carroll, J.W., Zhi, J., Chimonyo, V., Lin, S., Snapp, S.S., 2020. Multi-Spatial Resolution Satellite and sUAS Imagery for Precision Agriculture on Smallholder Farms in Malawi. Photogramm. Eng. Remote Sens. 86, 107–119. https://doi.org/10.14358/pers.86.2.107

Phiri, J., Malec, K., Majune, S.K., Appiah-Kubi, S.N.K., Gebeltová, Z., Maitah, M., Maitah, K., Abdullahi, K.T., 2020. Agriculture as a Determinant of Zambian Economic Sustainability. Sustainability-basel 12, 4559. https://doi.org/10.3390/su12114559

Pipia, L., Belda, S., Franch, B., Verrelst, J., 2022. Information and Communication Technologies for Agriculture—Theme I: Sensors. Trends in Satellite Sensors and Image Time Series Processing Methods for Crop Phenology Monitoring., in: Springer Optimization and Its Applications. pp. 199–231. https://doi.org/10.1007/978-3-030-84144-7_8

Planet, 2021. Planet Fusion Monitoring Technical Specification Version 1.0.0-beta.3, March 2021.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. J Veg Sci 5, 703–714. https://doi.org/10.2307/3235884

Ricciardi, V., Ramankutty, N., Mehrabi, Z., Jarvis, L., Chookolingo, B., 2018. How much of the world's food do smallholders produce? Global Food Secur 17, 64–72. https://doi.org/10.1016/j.gfs.2018.05.002

Rocha, A.V., Shaver, G.R., 2009. Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. Agric. For. Meteorol. 149, 1560–1563. https://doi.org/10.1016/j.agrformet.2009.03.016

Routray, S.K., Javali, A., Sharma, L., Ghosh, A.D., Sahoo, A., 2019. Internet of Things Based Precision Agriculture for Developing Countries. 2019 Int. Conf. Smart Syst. Inven. Technol. (ICSSIT) 00, 1064–1068. https://doi.org/10.1109/icssit46314.2019.8987794

Roy, D.P., Huang, H., Houborg, R., Martins, V.S., 2021. A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery. Remote Sens. Environ. 264, 112586. https://doi.org/10.1016/j.rse.2021.112586

Rustowicz, R., Cheong, R., Wang, L., Ermon, S., Burke, M., Lobell, D., 2019. Semantic Segmentation of Crop Type in Africa: A Novel Dataset and Analysis of Deep Learning Methods. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.

Sheahan, M., Barrett, C.B., 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010

Shen, Y., Zhang, X., Yang, Z., 2022. Mapping corn and soybean phenometrics at field scales over the United States Corn Belt by fusing time series of Landsat 8 and Sentinel-2 data with VIIRS data. ISPRS J. Photogramm. Remote Sens. 186, 55–69.

https://doi.org/10.1016/j.isprsjprs.2022.01.023

Shen, Y., Zhang, X., Yang, Z., Ye, Y., Wang, J., Gao, S., Liu, Y., Wang, W., Tran, K.H., Ju, J., 2023. Developing an operational algorithm for near-real-time monitoring of crop progress at field scales by fusing harmonized Landsat and Sentinel-2 time series with geostationary satellite observations. Remote Sens. Environ. 296, 113729. https://doi.org/10.1016/j.rse.2023.113729

Tittonell, P., Vanlauwe, B., Ridder, N. de, Giller, K.E., 2007. Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms: Soil fertility gradients or management intensity gradients? Agr Syst 94, 376–390. https://doi.org/10.1016/j.agsy.2006.10.012

Tonnang, H.E.Z., Balemi, T., Masuki, K.F., Mohammed, I., Adewopo, J., Adnan, A.A., Mudereri, B.T., Vanlauwe, B., Craufurd, P., 2020. Rapid Acquisition, Management, and Analysis of Spatial Maize (Zea mays L.) Phenological Data—Towards 'Big Data' for Agronomy Transformation in Africa. Agronomy 10, 1363. https://doi.org/10.3390/agronomy10091363

Tran, K.H., Zhang, X., Ketchpaw, A.R., Wang, J., Ye, Y., Shen, Y., 2022. A novel algorithm for the generation of gap-free time series by fusing harmonized Landsat 8 and Sentinel-2 observations with PhenoCam time series for detecting land surface phenology. Remote Sens. Environ. 282, 113275. https://doi.org/10.1016/j.rse.2022.113275

Tricht, K.V., Gobin, A., Gilliams, S., Piccard, I., 2018. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. Remote Sens-basel 10, 1642. https://doi.org/10.3390/rs10101642

Tseng, G., Cartuyvels, R., Zvonkov, I., Purohit, M., Rolnick, D., Kerner, H., 2023. Lightweight, Pre-trained Transformers for Remote Sensing Timeseries. arXiv. https://doi.org/10.48550/arxiv.2304.14065

Veloso, A., Mermoz, S., Bouvet, A., Toan, T.L., Planells, M., Dejoux, J.-F., Ceschia, E., 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens Environ 199, 415–426. https://doi.org/10.1016/j.rse.2017.07.015

Viña, A., Gitelson, A.A., Rundquist, D.C., Keydan, G., Leavitt, B., Schepers, J., 2004.

Monitoring Maize (Zea mays L.) Phenology with Remote Sensing. Agron J 96, 1139–1147. https://doi.org/10.2134/agronj2004.1139

Viña, A., Liu, W., Zhou, S., Huang, J., Liu, J., 2016. Land surface phenology as an indicator of biodiversity patterns. Ecol. Indic. 64, 281–288. https://doi.org/10.1016/j.ecolind.2016.01.007

Waldman, K.B., Blekking, J.P., Attari, S.Z., Evans, T.P., 2017. Maize seed choice and perceptions of climate variability among smallholder farmers. Global Environ Change 47, 51–63. https://doi.org/10.1016/j.gloenvcha.2017.09.007

Wardlow, B.D., Kastens, J.H., Egbert, S.L., 2006. Using USDA Crop Progress Data for the Evaluation of Greenup Onset Date Calculated from MODIS 250-Meter Data. Photogramm. Eng. Remote Sens. 72, 1225–1234. https://doi.org/10.14358/pers.72.11.1225

Xin, J., Yu, Z., Leeuwen, L. van, Driessen, P.M., 2002. Mapping crop key phenological stages in the North China Plain using NOAA time series images. Int J Appl Earth Obs 4, 109–117. https://doi.org/10.1016/s0303-2434(02)00007-7

Ye, Y., Zhang, X., Shen, Y., Wang, J., Crimmins, T., Scheifinger, H., 2022. An optimal method for validating satellite-derived land surface phenology using in-situ observations from national phenology networks. ISPRS J. Photogramm. Remote Sens. 194, 74–90. https://doi.org/10.1016/j.isprsjprs.2022.09.018

Yuan, F., Lewis, A., Leith, A., Dhar, T., Gavin, D., 2021. Analysis Ready Data for Africa. Presented at the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 1789–1791. https://doi.org/10.1109/igarss47720.2021.9554019

Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. Remote Sens. Environ. 237, 111511. https://doi.org/10.1016/j.rse.2019.111511

Zhang, L., Zhang, Z., Luo, Y., Cao, J., Xie, R., Li, S., 2021. Integrating satellite-derived climatic and vegetation indices to predict smallholder maize yield using deep learning. Agricultural and Forest Meteorology 311, 108666. https://doi.org/10.1016/j.agrformet.2021.108666

Zhang, X., Friedl, M.A., Schaaf, C.B., 2006. Global vegetation phenology from Moderate

Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. J. Geophys. Res.: Biogeosciences 111. https://doi.org/10.1029/2006jg000217

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84, 471–475. https://doi.org/10.1016/s0034-4257(02)00135-9

Zhang, X., Hodges, J.C.F., Schaaf, C.B., Friedl, M.A., Strahler, A.H., Gao, F., 2001. Global Vegetation Phenology from AVHRR and MODIS Data. Igarss 2001 Scanning Present Resolv Futur Proc Ieee 2001 Int Geoscience Remote Sens Symposium Cat 01ch37217 5, 2262–2264. https://doi.org/10.1109/igarss.2001.977969

Zhang, X., Jayavelu, S., Liu, L., Friedl, M.A., Henebry, G.M., Liu, Y., Schaaf, C.B., Richardson, A.D., Gray, J., 2018a. Evaluation of land surface phenology from VIIRS data using time series of PhenoCam imagery. Agric. For. Meteorol. 256, 137–149. https://doi.org/10.1016/j.agrformet.2018.03.003

Zhang, X., Jayavelu, S., Liu, L., Friedl, M.A., Henebry, G.M., Liu, Y., Schaaf, C.B., Richardson, A.D., Gray, J., 2018b. Evaluation of land surface phenology from VIIRS data using time series of PhenoCam imagery. Agr Forest Meteorol 256, 137–149. https://doi.org/10.1016/j.agrformet.2018.03.003

Zhang, X., Liu, L., Liu, Y., Jayavelu, S., Wang, J., Moon, M., Henebry, G.M., Friedl, M.A., Schaaf, C.B., 2018c. Generation and evaluation of the VIIRS land surface phenology product. Remote Sens Environ 216, 212–229. https://doi.org/10.1016/j.rse.2018.06.047

Zhang, X., Tarpley, D., Sullivan, J.T., 2007. Diverse responses of vegetation phenology to a warming climate. Geophys. Res. Lett. 34. https://doi.org/10.1029/2007gl031447

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., Henebry, G.M., 2017. Exploration of scaling effects on coarse resolution land surface phenology. Remote Sens Environ 190, 318–330. https://doi.org/10.1016/j.rse.2017.01.001

Zhang, X., Wang, J., Henebry, G.M., Gao, F., 2020. Development and evaluation of a new algorithm for detecting 30 m land surface phenology from VIIRS and HLS time series. ISPRS J. Photogramm. Remote Sens. 161, 37–51. https://doi.org/10.1016/j.isprsjprs.2020.01.012

Zhong, L., Hawkins, T., Biging, G., Gong, P., 2011. A phenology-based approach to map crop types in the San Joaquin Valley, California. Int. J. Remote Sens. 32, 7777–7804. https://doi.org/10.1080/01431161.2010.527397

Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sens. Environ. 114, 2610–2623. https://doi.org/10.1016/j.rse.2010.05.032