Research article

A MODIS-based scalable remote sensing method to estimate sowing and harvest dates of soybean crops in Mato Grosso, Brazil

Minghui Zhang [a,*], Gabriel Abrahao [b], Avery Cohn [c], Jake Campolo [d], Sally Thompson [a,e]

[a] Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA
[b] Federal University of Vicosa, Vicosa, Minas Gerais, Brazil
[c] Tufts University, 419 Boston Ave, Medford, MA 02155, USA
[d] Department of Earth System Science, and Center on Food Security and the Environment, Stanford University, Stanford, CA, USA
[e] Department of Civil, Environmental and Mining Engineering, University of Western Australia, Western Australia, Australia

A B S T R A C T

Large-scale agriculture in the state of Mato Grosso, Brazil is a major contributor to global food supplies, but its continued productivity is vulnerable to contracting wet seasons and increased exposure to extreme temperatures. Sowing dates serve as an effective adaptation strategy to these climate perturbations. By controlling the weather experienced by crops and influencing the number of successive crops that can be grown in a year, sowing dates can impact both individual crop yields and cropping intensities. Unfortunately, the spatiotemporally resolved crop phenology data necessary to understand sowing dates and their relationship to crop yield are only available over limited years and regions. To fill this data gap, we produce a 500 m rainfed soy (*Glycine max*) sowing and harvest date dataset for Mato Grosso from 2004 to 2014 using a novel time series analysis method for Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery, adapted for implementation in Google Earth Engine (GEE). Our estimates reveal that soy sowing and harvest dates varied widely (about 2 months) from field to field, confirming the need for spatially resolved crop timing information. An interannual trend toward earlier sowing dates occurred independently of variations in wet season onset, and may be attributed to an improvement in logistic or economic constraints that previously hampered early sowing. As anticipated, double cropped fields in which two crops are grown in succession are planted earlier than single cropped fields. This difference shrank, however, as sowing of single cropped fields occurred closer to the wet season onset in more recent years. The analysis offers insights about sowing behavior in response to historical climate variations which could be extended to understand sowing response under climate change in Mato Grosso.

1. Introduction

The response of agricultural productivity to changing temperature and rainfall regimes will determine the health of economies, food systems, and societies as climate continues to warm. Climate change is projected to negatively impact food security by reducing arable land, decreasing mean crop yields and increasing vulnerability to extreme weather events (Schmidhuber and Tubiello, 2007); for example, in Africa and South Asia, temperature and rainfall shifts are expected to lower the mean yield of major crops by 8% by the 2050s (Knox et al., 2012).

This future agricultural productivity, however, will be influenced by how farmers adapt to changing weather regimes within their diverse physical, cultural, and economic contexts (Khanal et al., 2018; Reidsma et al., 2010). Management has such a strong influence on productivity

that farmers may be able to minimize or even reverse the harmful effects of climate change through a broad range of strategies, including cultivar selection, increased cropping intensity, improved water management, and nutrient and pest management (Alexander et al., 2018; Anwar et al., 2013; Borchers et al., 2014; Hassan and Nhemachena, 2008; Howden et al., 2007; Reidsma et al., 2010; Thornton et al., 2018). Frequently, a core element of strategies to mitigate climate impacts on agriculture involves changing the sowing and harvest dates of crops, in conjunction with changes in variety selection to adapt to the changes in photoperiod and growing season length. Examples spanning production systems from sorghum in Italy (Howden et al., 2007), to soybean in Austria (Alexandrov et al., 2002), the US south east (Baldwin and Cossar, 2009), and sub-Saharan Africa (Waha et al., 2013), rice in Sri Lanka (Dharmarathna et al., 2014) and corn in the US midwest (Kucharik and Serbin, 2008)

broadly demonstrate that adaptations in sowing dates could mitigate or reverse the negative effects of climate change on production. Typically, growers select a sowing date to optimize the exposure of crops to beneficial weather conditions and avoid exposure to harmful conditions during sensitive phenological stages. Such exposure can have large impacts on productivity - for example, a 1% decrease in irrigated wheat yield was observed for every day of delayed planting in northern India, a decline caused by heat stress during the grain filling period of crop growth (Ortiz-Monasterio et al., 1994).

In tropical regions like Brazil, where crops are heavily reliant on rainfall and already experience the upper limit of their temperature tolerance (Challinor et al., 2014; Rosenzweig et al., 2014), projections of a hotter, drier climate may limit farmers' ability to adapt by adjusting sowing dates. The state of Mato Grosso in central-west Brazil, is a center of soy production that accounts for 27% of Brazil's and 10% of global soy supplies (Victoria et al., 2012; Abrahao and Costa, 2018; Cohn et al., 2016). Soy's productivity in Mato Grosso depends in part on a long rainy season that facilitates double cropping systems, in which soy is the first of two crops grown in succession during the same year (Abrahao and Costa, 2018; Correa and Schmidt, 2014; Arvor et al., 2012b). However, sustaining this high productivity depends in part on whether its intensive cropping practices and high crop yields can be maintained under the shorter wet seasons and higher temperatures expected with climate change (Arvor et al., 2012b,a, Costa and Pires, 2010, Dubreuil et al., 2012, Gourdji et al., 2013). Sowing dates form a crucial link between climate and agricultural productivity in Mato Grosso. Since adjusting sowing dates offers an effective climate adaptation strategy, understanding the current timing of soy planting is essential to determine whether and how Mato Grosso's agricultural systems can be sustained in the future (Abrahao and Costa, 2018; Pires et al., 2016).

High-resolution sowing data are essential to quantify these climate risks to soy production, but until now were unavailable in Mato Grosso. In this paper, we produce a 500 m dataset of soy sowing and harvest dates over Mato Grosso from 2004 to 2014. Spatiotemporal patterns in the sowing dates reveal drivers of sowing behavior (such as cropping intensity and wet season onset) that can be incorporated into models to improve yield projections under climate change.

Sowing dates play two key roles in Mato Grosso's response to a hotter and drier climate (Pires et al., 2016; Fu et al., 2013). First, as a primary control on the weather experienced by crops, changing sowing dates is one of the most cost-effective adaptations to climate change (Alexandrov et al., 2002; Baldwin and Cossar, 2009; Howden et al., 2007; Kucharik and Serbin, 2008; Waha et al., 2013; Waongo et al., 2014). Farmers may select sowing dates to avoid water and heat stress during important phenological stages (Ortiz-Monasterio et al., 1994), increasing biomass accumulation and yields (Soltani and Sinclair, 2012). Second, sowing dates constrain choices in cropping intensity or crop variety. For example, farmers who practice rainfed double cropping may prefer to sow as early as possible in the wet season (Pires et al., 2016). Delays in the sowing could render double cropping impossible, effectively halving the region's agricultural productivity (Fu et al., 2013; Spera et al., 2020). In Mato Grosso, a drier wet season has been associated with increased soy (first crop) yields but reduced corn (second crop) yields (Cohn et al., 2016). Cohn et al. (2016) attribute this to a switch from double cropped systems in which a short-cycle, lower-yield soy variety is planted in combination with corn, to single cropped systems using long-cycle, higher-yield soy varieties. Higher temperatures can also impact cropping intensity in Mato Grosso: a temperature anomaly of $+1\ ^{\circ}C$ was associated with a decline in cropping intensity (3.2% overall production loss) (Cohn et al., 2016).

In addition to cropping choices, sowing dates in Mato Grosso are subject to a range of climatic, legal, and socio-economic constraints – all of which can interact to produce spatially complex variations in when and where sowing occurs. The onset of the wet season is important because soy in Mato Grosso is predominantly rainfed (Abrahao and Costa, 2018; Pires et al., 2016). Farmers are also legally constrained from

sowing in a period close to the wet season onset (June 15 - September 15/30) by a prescribed sanitary period imposed to prevent pathogen outbreaks (Pires et al., 2016). Farmers are also subject to a host of socio-economic constraints that lead to dynamic sowing behavior (Waha et al., 2012). Variations in sowing dates can occur over spatial scales as small as individual fields and reflect diverse factors such as risk aversion, access to equipment and labor, use of irrigation, desired cropping intensity, and soil type (Begue et al., 2010; Bussmann et al., 2016; Dounias et al., 2003; Feola et al., 2015; Kala, 2015; Pires et al., 2016), while interannual variations reflect volatility in climate, crop price, equipment availability, or technological development (Iizumi et al., 2019; Kucharik, 2006; Kucharik and Serbin, 2008; Ma et al., 2012). Interactions between climatic and socio-economic factors may also occur: Mato Grosso's lack of insurance and weak governance may make cropping and sowing decisions more responsive to climate anomalies than regions with more codified agricultural practices (Cohn et al., 2016). The spatially and inter-annually variable sowing dates could translate into large fluctuations in yield. It is unclear at present which of these climatic, legal, and socio-economic constraints dominate contemporary sowing decisions, and whether any will impose hard limits on adaptation of sowing dates in response to climate change.

In the absence of either highly resolved datasets documenting sowing date, or a reasonable model to predict sowing data, agricultural predictions in Mato Grosso (and elsewhere) are forced to rely on untested assumptions about the timing of sowing. For example, many crop modeling studies assume that sowing occurs on the yield-maximizing date or at wet season onset (Grassini et al., 2015; Iizumi et al., 2019; Sacks et al., 2010; Waha et al., 2012). In South America, long growing seasons and intensive cropping practices mean that such assumptions could result in errors in sowing date as great as five months (Waha et al., 2012). Similarly, farmers who wish to double crop are sometimes presumed to plant immediately after the sanitary break to increase the likelihood of a successful second crop, but it is unclear whether the sanitary break truly imposes a hard limit on sowing (Pires et al., 2016). This gap in knowledge could be addressed with high-resolution sowing data that illuminates the limits of local adaptability, allowing targeted vulnerability assessments and policy evaluation.

Unfortunately, the high cost of ground surveys means sowing and harvest dates are not known at sufficient spatiotemporal resolution to isolate climatic and socio-economic controls on sowing behavior in Mato Grosso. The highest quality sowing data available over Mato Grosso is a weekly crop progress report produced by the Instituto Mato-Grossense de Economia Agropecuaria (IMEA) (IMEA, 2019), which reports the percent of soy planted over 7-day intervals. These reports are not spatially explicit and aggregate across important regional differences within the state (Picoli et al., 2018).

In other regions where sowing data are unavailable, time series analysis of remotely sensed vegetation indices have been used to extract phenological dates of natural vegetation and crops (Boschetti et al., 2009; Bolton et al., 2020; Gao et al., 2020; Jonsson and Edklundh, 2004; Sakamoto et al., 2005; Zhang et al., 2003, 2020). By calibrating and validating the phenological stages against ground-truth sowing and harvest dates, satellite-observed vegetation indices can be related to ground-level behavior. These approaches have been applied to agricultural production in the US, Europe and Asia: in the midwest US, 500 m sowing dates for maize, wheat and soybean achieved root mean squared errors (RMSE) of less than 10 days compared to state-level Crop Progress Reports (Ren et al., 2017; Urban et al., 2018); in Italy, 250 m rice planting and harvest dates have been estimated with error of 10 days (Boschetti et al., 2009); in Japan, 1 km planting and harvest dates were estimated with RMSE of 12.1 days and 10.6 days, respectively (Sakamoto et al., 2005); and phenological information extracted for wheat at 1 km scale in Punjab, India and for wheat and corn at 30 m in central China were both highly correlated to validation data collected at an aggregated scale (Lobell et al., 2013; Pan et al., 2015).

However, existing satellite-based approaches cannot be directly applied to Mato Grosso: the state's variety of cropping practices require an estimation method that avoids region-specific assumptions about crop timing (Zalles et al., 2019) and or that rely on local expert knowledge of agricultural practice (Begue et al., 2018). An absence of spatially explicit ground truth data requires the construction of proxies from high-resolution satellite data against which to test planting date estimation algorithms. Here, we propose an approach to address both of these issues, creating 500 m and 25 km sowing and harvest maps of soy, commonly the first of two crops planted within a growing season, across Mato Grosso for 2004 to 2014.

The resulting spatial maps of sowing date contain rich information about individual farmers' behavior within their climatological, logistical, and economic contexts - knowledge that will allow analysis of how yields were influenced by sowing timing, and extrapolation to farmer behavior and productivity under future scenarios. A preliminary analysis of the sowing date information is presented here, and identifies important spatial and temporal variability, trends, and constraints on sowing date, which can be immediately used to refine assumptions about soy agriculture in Mato Grosso.

2. Methods

2.1. Study site

Mato Grosso is a 900,000 km^2 state located in the southern Brazilian Amazon comprising three major biomes: Pantanal (tropical wetland, 62,000 km^2) in the south, Amazon (humid tropical forests, 481,000 km^2) in the north and Cerrado (tropical savannas, 360,000 km^2) in the center (Brown et al., 2013). The state experiences a hot, semi-humid to humid climate, with nearly constant and spatially homogeneous temperatures (22–26 °C) year round, and a strong north-south gradient in rainfall. At the north of the state, annual precipitation exceeds 2000 mm, with a 3 month dry season; in the south, annual rainfall is 1000 mm with a 5 month (May–October) dry season (Arvor et al., 2014).

Long rainy seasons and continued improvements in crop variety have contributed to rapid expansion of cultivated land in Mato Grosso (Abrahao and Costa, 2018). In 2014, the total row crop area was nearly 100,000 km^2 (Zalles et al., 2019), of which 85% was double cropped soy and 12% was single cropped soy, and agriculture comprised 72% of the state's gross domestic product (GDP) (Cohn et al., 2016) (Figure 1).

2.2. Approach

2.2.1. Proposed estimation method

While many remote sensing-based estimates of crop phenology exist, including in Brazil (Becker et al., 2020), they must be adjusted for use in Mato Grosso. First, the simplifying assumptions used in previous studies, including relatively rigid assumptions about crop phenology and timing, are not defensible on regional scales, particularly where multiple cropping intensities co-occur (Zalles et al., 2019). Second, ground data availability is a major constraint in Mato Grosso, creating difficulties for calibration and validation of remote sensing estimates. Without innovation in generating ground-truth datasets, spatially resolved satellite-based sowing and harvest dates can only be validated using aggregated data such as IMEA crop progress reports. To address these challenges, we used a scalable time series analysis method, adapted from a linearized harmonic time series fitting algorithm introduced by Clinton (2017), to extract phenological milestones from EVI derived from MODIS imagery. Requiring only linear regression, this algorithm was able to take advantage of Google Earth Engine (GEE)'s automatic parallelization and readily available satellite imagery. Clinton (2017)'s linearized harmonic time series fitting method, designed to fit periodic and stationary natural vegetation phenology, was expanded to account for non-periodic agricultural time series in which single, double and triple cropping can occur in succession, and in which the growing season may shift rapidly from year to year. The algorithm provided a scalable, rapid analysis that maintained accuracy and minimized assumptions regarding crop timing. The phenological milestones derived from this algorithm were related to reference ground sowing and harvest dates observed from high-resolution (3–5 m) Planet Labs imagery, producing a calibrated relationship between sowing/harvest and MODIS-derived phenological dates (Planet, 2017). We show that high-resolution imagery can provide validation data where survey information is unavailable.

2.3. Datasets and definitions

A range of remotely-sensed imagery, existing maps, ground-based datasets and climate products were used in the analysis. Remote sensing imagery consisted of the Enhanced Vegetation Index (EVI) calculated from cloud-masked MODIS 8-day composite products (MYD09A1 and MOD09A1) at 500 m resolution from 2004 to 2014 (NASA Goddard Space Flight Center, 2019a,b). High resolution optical images of soy fields in Mato Grosso were obtained from the PlanetScope (3 m) and RapidEye (5 m) satellites from August 1, 2016 to July 31, 2017 at three locations: 11 images of 32 km^2 at (-55.389, -11.868); 11 images of 55 km^2 at (-53.454, -15.396); and 16 images of 126 km^2 at (-57.731, -13.285) (Planet, 2017). The images from PlanetScope and RapidEye are referred to as "Planet Labs imagery".

We produced a crop cover map identifying areas of single and double cropped soy. A center pivot irrigation map for 2014 collected by Brazil's National Water Agency (ANA) (Brito, 2017) allowed us to identify rainfed fields for analysis. This irrigation map limited the study period to 2014 and before, but is the best data available in Mato Grosso.

Spera et al. (2014)
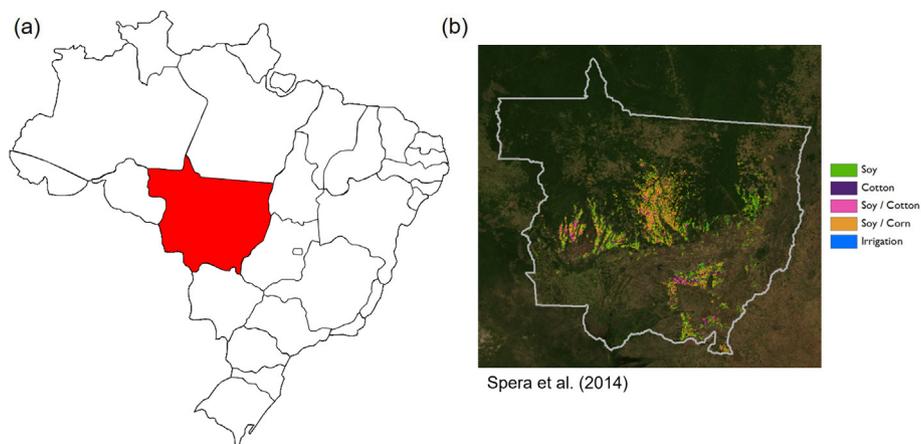
Figure 1. (a) Mato Grosso, Brazil and (b) its crop cover.

Wet season onset from 2004 to 2014 was calculated by Abrahao and Costa (2018). The anomalous accumulation (AA) method was applied to a gridded (0.25 × 0.25 deg) daily rainfall product produced through interpolation of 3625 rain gauges and 735 weather stations across Brazil (Liebmann et al., 2007; Xavier et al., 2016). In the anomalous accumulation (AA) method, the wet season onset date was defined based on the value of the anomalous accumulation [mm/day]:

$$AA(t) = \sum_{n=1}^{t} \left(R(n) - R_{ref} \right) \tag{1}$$

where $R(n)$ is the rainfall on day n and R_{ref} is a reference rainfall value, defined here as the agronomically significant threshold of 2.5 mm/day (Arvor et al., 2014). Here, $t = 1$ refers to July 1, a date within the dry season. The onset date was defined as the day at which the value of $AA(t)$ reaches its minimum (Liebmann et al., 2007).

Spatially aggregated validation was performed by comparing estimated sowing dates to (1) the University of Madison - Wisconsin's Sustainability and the Global Environment (SAGE) dataset, which compiles national and subnational planting and harvest date statistics circa 2000 (Sacks et al., 2010), and (2) weekly crop progress reports from IMEA (IMEA, 2019).

For time series analysis, we started the agricultural year on August 1, the middle of the dry season. To relate the calendar year to the agricultural year, we refer to "sowing year" and "harvest year". For example, for the agricultural year commencing August 1, 2013, the sowing year is 2013, and the harvest year is 2014. Because soy in Mato Grosso can be single or double cropped, we refer separately to "single cropped (SC) soy" and "double cropped (DC) soy", using "soy" to denote both SC and DC soy.

2.4. Estimation of sowing and harvest dates

A four-step process was used to map sowing and harvest dates for Mato Grosso, and is illustrated in Figure 2. First, phenological milestones were estimated from MODIS-derived Enhanced Vegetation Index (EVI) (Step 1). Next, an equation relating the phenological milestones to
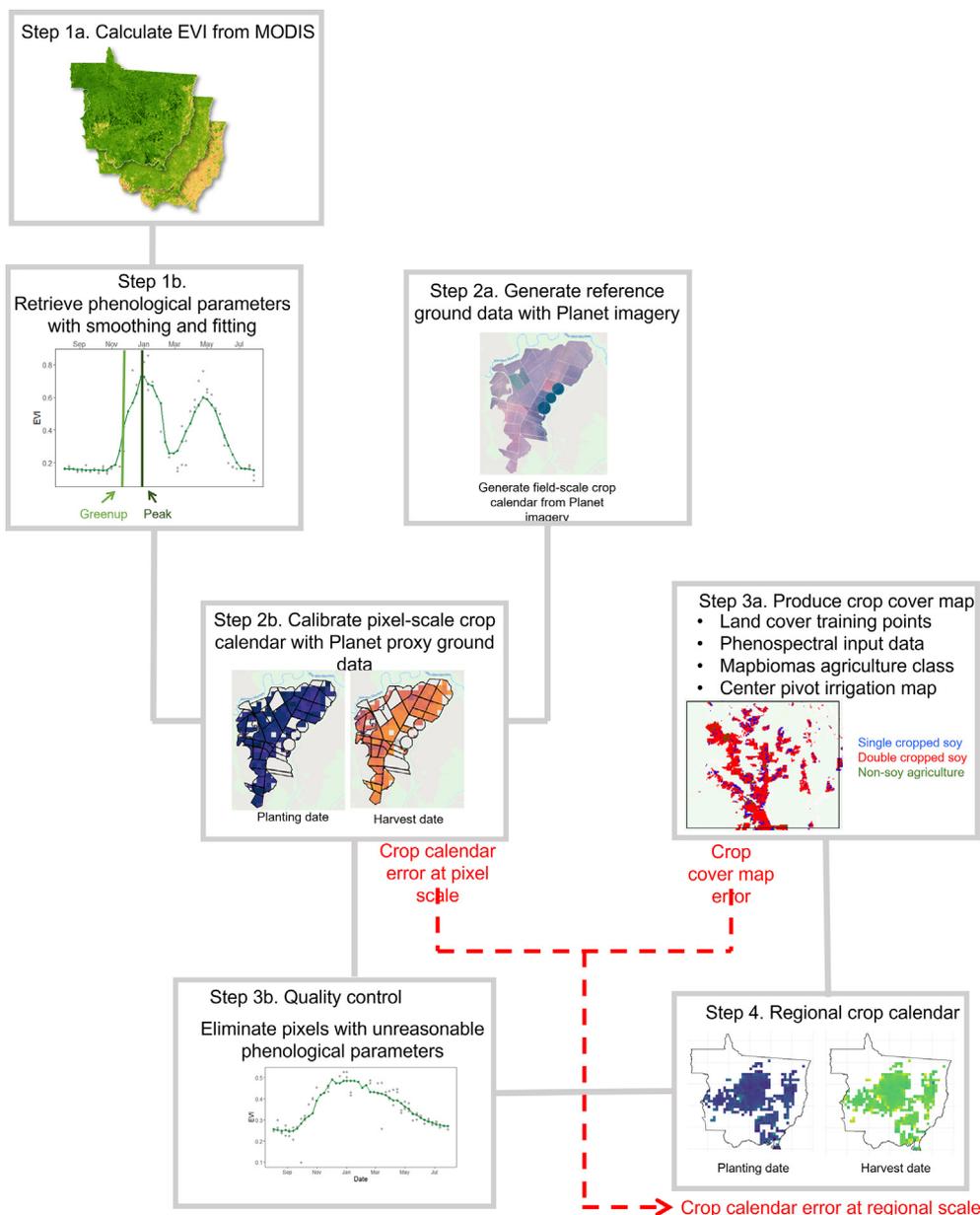
Figure 2. Method overview. Each labelled step is described in the text.

sowing and harvest dates was calibrated to create 500 m sowing and harvest maps (Step 2). In this step, reference ground data for sowing and harvest dates were also generated from Planet Labs imagery. Quality control and crop cover maps were applied as masks (Step 3), and a 25 km regional sowing and harvest date map was created (Step 4). Uncertainty in the resulting sowing and harvest dates at the 500 m and 25 km scale were quantified during Step 4. Each step is described in detail in the following sections.

2.4.1. Step 1: Extract phenological milestones from MODIS imagery

Soybean phenology was tracked with cloud-filtered enhanced vegetation index (EVI) from MODIS (MOD09A1 and MYD09A1) (Reed et al., 2009; Zhang et al., 2003) (Step 1a, Figure 2). The removal of the cloudy pixels by retaining only pixels labeled as "clear" in the "StateQA" band produced EVI time series containing idiosyncratic and irregular gaps, impacting both EVI (Figure 3a) and its time derivative, $dEVI/dt$ (Figure 3f, computed with a forward difference method).

The resulting gap-containing EVI and $dEVI/dt$ time series were smoothed with a series of moving average windows (Step 1b, Figure 2). The size and number of smoothing windows were chosen based on three criteria: (i) the percent of soy area for which reasonable sowing and harvest date estimates were obtained (as defined in Step 3); (ii) accuracy of sowing and harvest date estimates (as obtained by comparison to data described in Step 2); and (iii) robustness of estimates to the number of missing points in the EVI time series (addressed by degrading complete, cloud-filtered EVI time series: randomly eliminating 1 to 10 of the roughly 25 EVI points covering the growing window of the first crop). The best of the trialed methods was incorporated in the main algorithm. This involved smoothing each pixel's annual time series with two successive 20-day moving average windows (Figure 3b), and smoothing $dEVI/dt$ once with a 40-day moving average window (Figure 3f). While these windows are large compared to the roughly 120-day growing cycle of soy, they were necessary to eliminate high noise caused by cloud gaps and aerosols. The width of these windows is acceptable because the method aims to describe only broad features of the time series, such as the date of the peak and width of the soy curve, which are not degraded under the smoothing windows.

Next, phenological features were retrieved from the smoothed EVI time series using the proposed fitting algorithm (Step 1b, Figure 2). The "greenup" date for the first crop, $t_{greenup}^{first}$, was defined as the date of fastest increase in EVI. The "peak" day for first crop was similarly defined as the date of maximum EVI, $t_{peak,num}^{first}$. Peak and greenup dates are often used to estimate sowing and harvest dates in the literature (Sakamoto et al., 2005), and their high signal-to-noise ratio makes them relatively robust to data gaps (Hmimina et al., 2013; Pan et al., 2015).

We fitted the smoothed EVI time series to a 1st order harmonic function following Clinton (2017), which is similar to other functional forms used extensively in phenological studies (e.g. Geerken, 2009; Wagenseil and Samimi, 2016):

$$EVI = \beta_0 + \beta_1 \cdot t + A*cos(2\pi\omega \cdot t - \varphi) \qquad (2)$$

The terms β_0 and β_1 represent the mean EVI and the linear trend, respectively. The phase, φ, and amplitude, A, terms can be calculated by linear regression of EVI on time if frequency ω is known (Shumway and Stoffer, 2017). The remaining parameters can be found by linear regression. We estimated ω using the "quarter period", q, defined as the time difference between the $t_{peak,num}^{first}$ and the preceding $t_{greenup}^{first}$, where ω and q were then related as $q = \pi/2\omega$.

As shown in Table 1, the peak and greenup dates ($t_{peak,num}^{first}$, $t_{greenup}^{first}$) were identified by searching for maximum EVI or maximum $|dEVI/dt|$ in two time windows (see Figure 3 c, i).

We avoided computing dates in the middle of the wet season due to high levels of cloud cover. Therefore, the quarter period, q, for the first crop was estimated as $t_{peak,num}^{first} - t_{greenup}^{first}$ (Figure 3, c and h). Because this q was calculated for the first crop, the harmonic function was fitted only over the time window corresponding to the first crop (from $t_{peak,num}^{first} - 2q$ to $t_{peak,num}^{first} + q$, see Figure 3e). These windows maximized the number of EVI points available for fitting but avoid associating EVI observations with the wrong crop.

With ω for each crop cycle estimated from q, Eq. (2) was fitted to the EVI time series for each crop cycle, and A and φ obtained from the fitted values. The phase was then used to calculate the "analytic" peak date for the first crop as $t_{peak,fitted}^{first} = \frac{\varphi}{2\pi\omega}$. The analytic peak date is an estimate of peak greenness that is more robust to noise and data gaps than its numeric counterpart. We did not calculate an analytic greenup date.

During this step, the algorithm was applied to all pixels in Mato Grosso. Because the crop cover classification required the results of this algorithm, we were not able to mask soy pixels until after this step was completed. Fortunately, knowledge of whether a pixel is single or double cropped was not required to fit the algorithm. A sensitivity analysis showed that estimates of the first crop date were independent of the presence or absence of a second crop, so the same methods could be applied to single and double cropped pixels. Because crop cover data were not available to identify the second crop, we did not estimate sowing and harvest dates for the second crop.

The impact of crop management practices on our algorithm were also considered. Fertilization has been shown to cause a higher rate of phenology change (Thies et al., 1995) and higher yield in soy crop, potentially having a larger impact on crop development than climate change (Liu and Dai, 2020). Fertilization's impact on crop development may be manifest in the time series of EVI as a shorter growing period or
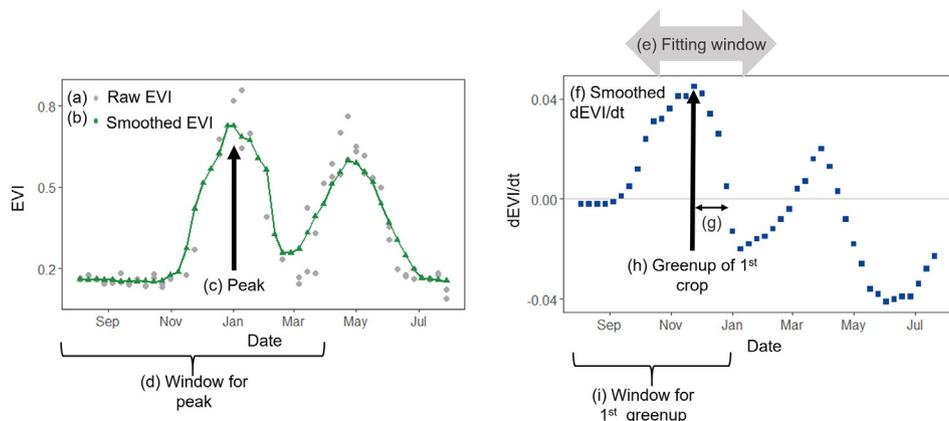
Figure 3. Time series analysis method.

Table 1. The date windows over which phenological stages were sought.

Description	Symbol	Date Window
Peak date of the first crop	$t_{peak,num}^{first}$	August 1 of the sowing year to March 31 of the harvest year
Greenup for first crop	$t_{greenup}^{first}$	August 1 of the sowing year until $t_{peak,num}^{first}$

higher peak. These shifts in crop timing and peak greenness are inherently captured in our time series analysis method because we make no assumptions about growing period or peak EVI value. Furthermore, as we ignore the last quarter of the crop cycle during fitting due to cloud cover, our method is robust to any potential acceleration of crop maturation that could cause the phenological curve to become asymmetric.

2.4.2. Step 2: Calibrate a relationship between phenological parameters and sowing/harvest reference ground data generated from Planet Labs imagery

The peak and greenup dates ($t_{peak,num}^{first}$, $t_{greenup}^{first}$) derived from EVI time series were used to estimate the sowing and harvest dates through equations calibrated to "ground-truth" data:

$$\text{Sowing date} = t_{peak,fitted}^{first} - p*q \tag{3}$$

$$\text{Harvest date} = t_{peak,fitted}^{first} + h*q \tag{4}$$

Where p and h represent the number of quarter periods, q, between the fitted peak date, $t_{peak,fitted}^{first}$, and the sowing and harvest dates, respectively.

Because ground survey data do not exist in Mato Grosso, we visually interpreted high-resolution Planet Labs imagery to identify sowing and harvest activity over soy fields in Mato Grosso. We refer to the dataset of sowing and harvest dates created from Planet Labs imagery as the "reference ground dataset" (Step 2a, Figure 2).

To create the reference ground dataset, we first identified three distinct soy regions from crop cover training data (Figure 4). Each studied area contained 40–80 soybean fields. The Planet Labs imagery obtained over these locations from 1 August 2016 to 31 July 2017 were

Figure 4. The location of MATOPIBA ground-truth survey and of Mato Grosso's Planet Labs image locations (denoted 1, 2 and 3).

manually delineated into fields based on the presence of clearly visible roads. These images were taken by the PlanetScope (3 m) and RapidEye (5 m) satellites, and offered enough spatial resolution to clearly delineate soy fields, observe greenup soon after sowing, and follow the progress of harvest equipment as bright, bare patches gradually replace mature, brown fields at the end of the growing season.

We downloaded Planet Labs images over locations that (1) cover only soy fields (based on crop cover point data) and (2) are representative of all the potential sources of error in the estimation method. To ensure that the image locations contain soy, we chose combinations of fields and years that were reported as soy in the land cover data. In the interest of obtaining Planet Labs images that are representative of all potential error sources, we included both single and double cropped fields and chose images that were spatially far apart. A large, persistent storm system would generate the same gap in the vegetation index timeseries over adjacent fields, potentially missing a significant source of error that would have appeared if the storms occurred in a different pattern. Additionally, using images scattered over Mato Grosso ensured that the calibration was not biased toward a single producer or toward practices that are local to a specific region. Figure 4 shows the locations in Mato Grosso for which we downloaded images that span the whole growing season. We uploaded these Planet Labs images to Google Earth Engine, delineated individual fields by hand, and recorded the earliest and latest possible sowing and harvest dates for each field. Figure 5 displays some Planet Labs imagery samples, highlighting sowing and harvest detection and the influence of clouds on data creation.

Farming cycles were identified for each field using observations of bare soil, green vegetation, brown mature crops, and bare soil. This allowed estimation of harvest dates (illustrated in Figure 5) based on the distinct geometric patterns caused by harvesting equipment. Sowing dates were reported by visual assessment over a 2–5 week date range (depending on the quality and temporal resolution of the dataset) preceding the first observed increase in greenness in the soy fields. The final reference ground dataset consisted of the earliest and latest possible sowing and harvest dates for the date range 1 August 2016 to 31 July 2017. This dataset was used to evaluate the pixel-level accuracy of the estimation method over fields which are known to be soy, in the absence of additional error introduced by the land cover map at the regional level. We applied the proposed time series analysis method to the Planet Labs locations over the same date range in order to calculate estimation error; however, we were unable to generate sowing and harvest date estimates for all rainfed soy in Mato Grosso after 2014 due to lack of recent irrigation maps.

Our reference ground data was compared to a ground survey of 90 soy properties from 2010 to 2017, which reported sowing and harvest dates over the MATOPIBA region of Northeast Brazil (comprising the states of Maranhão, Tocantins, Piauí and Bahia) (Figure 4). While this "MATOPIBA survey" was the best property level ground truth data available to us in Brazil, it may have been affected by recall bias as it was conducted only during the 8th year. Therefore, we introduced large error bars to the original data if responders reported vague date ranges for sowing and harvest (such as "early October"). The width of these error bars was chosen based on the wording of the reported date range; for example, a report of "2nd week of October" received a range of October 10 +/- 7 days; "early October" received an error bar of October 10 +/- 15 days; and "October" received an error bar of October 15 +/- 20 days.

MATOPIBA survey data and Planet Labs imagery collections coincided for two properties, enabling a direct comparison between the two datasets. Existing survey data in MATOPIBA were found to be inadequate for ground-truthing soy in Brazil due to excessive spatial aggregation and concerns about farmers' recall bias. In the survey, a single set of sowing and harvest dates were reported for each property (on average, 52 km²). However, significant differences (up to 3 months) in the sowing and harvest dates across fields within each property were visually evident in the Plant Labs imagery. One property reported that their sowing and harvest activity occurred over a period of 1 and 1.5 months, respectively.
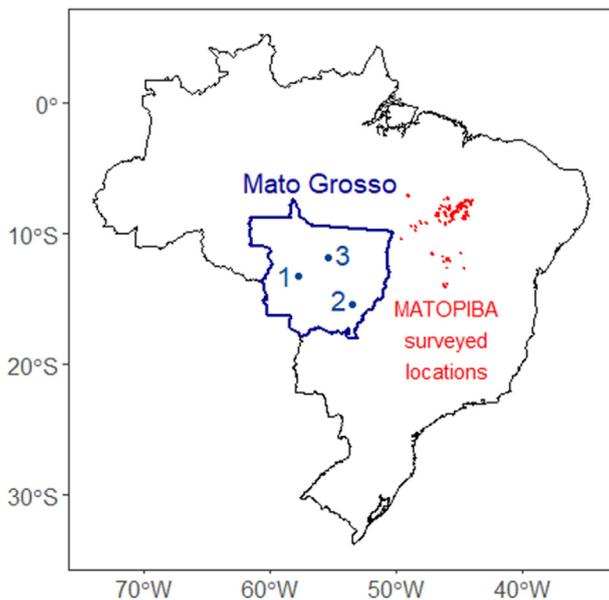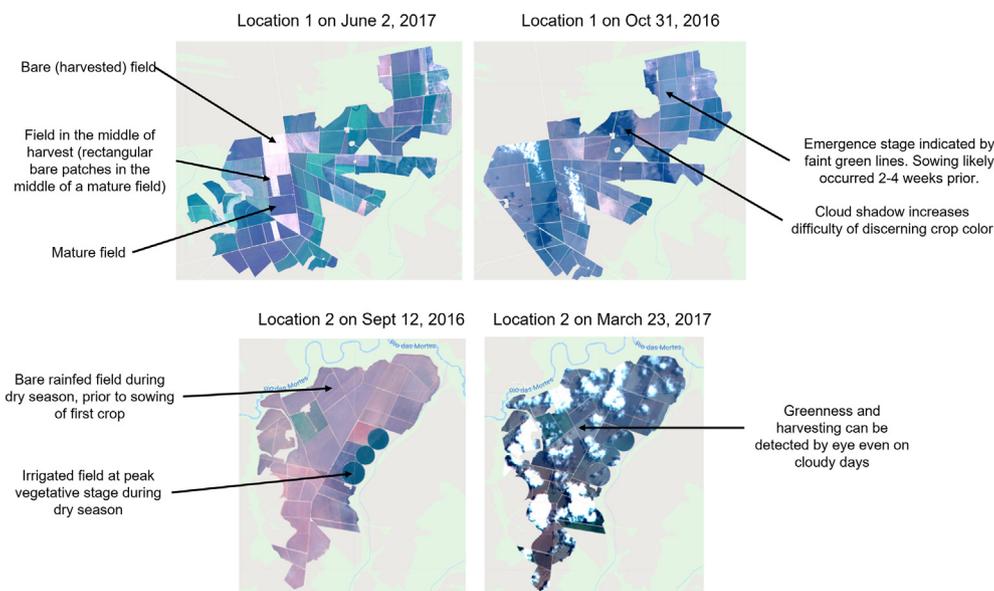
Figure 5. Planet Labs images from two locations in Mato Grosso, ranging from the start of the growing season (September) to the end (June) illustrate the visual cues that were used to estimate sowing and harvest dates for each field. Clouds and cloud shadows impacted the quality of the estimates.

For the same property, Planet Labs imagery suggested that sowing was spread over a 2 month period, and harvest activity over a 3 month period, an increase of 100%. This larger range indicated that at least some fields were not covered by the values reported in the MATOPIBA survey. For the other overlapping property, the reported sowing and harvest dates were not consistent with greenup and browndown observed in Planet Labs imagery in 3 out of 4 of its fields, with differences of up to 1.5 months.

Additionally, 5% of MATOPIBA survey responses reported identical dates across all years, suggesting that farmer recall may have impacted the quality of the results. Our examination of two of these properties suggested that spatial aggregation could lead to large errors in reported crop timing. Reference ground datasets may be a worthwhile alternative, even where survey data are available, because it avoids the greater imperfections in some survey datasets.

Given the uncertainties in the sowing and harvest date information, we computed the RMSE as the difference (in days) between the estimate and the nearest date in the reported date range if the estimated date fell outside of the range, or zero when the estimate fell inside the range.

The best calibration RMSE of 2.5 and 1.6 days for sowing and harvest date across all three Planet Labs imagery sites (Figure 4) was achieved by setting the sowing date as $t_{peak,fitted}^{first} - 1.75q$, and harvest as $t_{peak,fitted}^{first} + 1.1q$ (Step 2b, Figure 2). Cross-validation, in which one Planet Labs imagery site was removed per calibration, estimated out-of-sample prediction RMSE as 2.92 and 1.61 days for sowing and harvest date. We tested the sensitivity of the estimated sowing and harvest dates to variations in time series analysis parameters that were settled by trial and error by comparing each version of the estimated dates to the reference ground dataset. The parameters included in the sensitivity analysis were p and h (Eqs. (3) and (4)), moving window sizes for smoothing (Step 1), and the window over which $t_{peak,num}^{first}$ was sought (Step 1).

The parameters used in the method were found through sensitivity analysis. In Step 1, we used smoothing windows to compensate for cloud-induced noise in the EVI timeseries, and calibrated parameters that related two phenological indicators, peak date and quarter period, to sowing and harvest dates. The size of the smoothing windows and value of the calibrated parameters were selected to minimize error as compared to the reference ground dataset obtained from Planet Labs imagery. As shown in Table 2, of the six parameters chosen through sensitivity analysis, five had a significant impact on the estimated sowing date. As expected, the "peak cutoff date" had no effect on the estimate for the first crop. Here, the peak cutoff date refers to the date after which EVI values were ignored for first crop estimates, and the date before which EVI values were ignored for second crop estimates. In other words, EVI values after the peak cutoff date (second crop) did not make any impact on estimates based on EVI values before the peak cutoff date (first crop). EVI observations after the chosen peak cutoff date of April 1 were not considered for estimating the sowing/harvest dates of the first crop, *regardless* of the cropping intensity.

Notably, we found that two smoothing windows were necessary for EVI. An overly wide smoothing window may merge two separate peaks into one, falsely increasing the estimated quarter period. An increased quarter period means that the crop growing season length is over-estimated, which would result in sowing date estimates that are too early,

Table 2. Sensitivity of sowing/harvest estimate error on timeseries algorithm parameters, and the error-minimizing parameter values. The estimates were most sensitive to the sowing and harvest calibration parameters, followed by dEVI/dt's smoothing window size.

Timeseries analysis parameter	Effect on sowing date error [days of error per unit change in parameter]	Effect on harvest date error [days of error per unit change in parameters]	Error-minimizing value
Peak cutoff date [days]	0	0	April 1
First moving window for EVI smoothing [days]	0.01	0.1	20
Second moving window for EVI smoothing [days]	0.02	0.14	20
Smoothing window for dEVI [days]	0.1	0.1	40
Sowing calibration parameter (p) [-]	21	n/a	1.75
Harvest calibration parameter	n/a	13	1.1

and harvest date estimates that are too late. However, an overly narrow smoothing window would distort the location of the peak and greenup in unpredictable ways: noise from images impacted by clouds or aerosol would cause the estimated peak and greenup date to be too late or too early, causing sowing dates that are either too early (if peak and greenup are too close together, or if peak is too early) or too late (if peak and greenup are too far apart, or if peak is too late). The use of two successive smoothing windows balanced the need for increased smoothing without blending a double cropped field into a single peak.

2.4.3. Step 3: Apply crop cover and quality control masks

To produce soy sowing and harvest dates over Mato Grosso required knowing the location and cropping intensity of soy agriculture across the state (Step 3a, Figure 2). To map soy agriculture and its intensity, we adapted an existing crop classification technique tested for soy and corn in Parana State, Brazil (Zhong et al., 2016). We trained a Cartesian classifier in GEE using topographic, phenological and spectral information derived from Landsat and MODIS. All phenological and spectral input data in our study were derived from the EVI time series calculated in Step 1b (cloud-filtered and smoothed), while topographic data (elevation, slope, aspect and hillshade) were derived from the Shuttle Radar Topography Mission (Farr et al., 2007). While we only report sowing and harvest dates for the first crop (soy), phenological information for the second crop was used as input to the crop cover classifier.

Training data constraints meant that only three land cover types were classified: single cropped soy, double cropped soy, and non-soy agriculture. After the classifier labeled all pixels in Mato Grosso as one of these three classes, we masked out non-agricultural lands using the Mapbiomas land cover dataset. The phenological and spectral information used to classify pixels as "single cropped soy", "double cropped soy", and "non-soy agriculture" is shown in Figure 6. The Shuttle Radar Topography Mission (SRTM) provided topographic information on elevation, slope, aspect, and hillshade at 30 m resolution that was used in land cover classification (Farr et al., 2007). Crop cover training data (9,000 points, 2003–2017) was formed by intersecting a Landsat-based crop classification produced by Agrosatelite (2017) with a roadside survey of Mato Grosso's agricultural areas conducted by Embrapa and the Kansas Biological Survey (Kastens et al., 2017). The classes in this training dataset were "double cropped soy", "single cropped soy", and "non-soy agriculture", and did not include points separating agriculture from non-agricultural cover. Therefore we supplemented our classified map with Mapbiomas v3, a 30 m resolution land cover map of Brazil from 1985 to 2017 that identified row crop agriculture (Mapbiomas, 2019).

Because cloud filtering of the MODIS data in Step 1a introduced spatial gaps in the EVI images that, if not filled, would be propagated as gaps in the classifier's training data, the EVI images were gapfilled over space with a mean square kernel. This ensured that each point in the land cover dataset contained a full set of input data.

The classifier was trained on spectral information benchmarked on phenological stages, rather than on calendar dates, to allow the classifier

to align input data across years and locations. This alignment produced a classifier that was robust to interannual and inter-regional variations in the sowing and harvest dates, sensitive to physiological and seasonality differences among crop types and crop intensity levels, and relevant in extrapolated contexts (Zhong et al., 2016).

In addition to first crop phenology, whose estimation is detailed in the main text, we also used second crop phenology during classification. The method to derive second crop phenology was similar to that for the first crop, with the following changes to Step 1b:

1. Instead of extracting greenup ($t_{greenup}$) from the smoothed EVI timeseries, we extracted the "browndown" date, $t_{browndown}^{second}$, defined as the date of fastest decrease in EVI. The "quarter period", q, was then defined as the time difference between the $t_{peak,num}^{second}$ and $t_{browndown}^{second}$.
2. The time windows within which the peak and browndown dates ($t_{peak,num}^{second}$, $t_{browndown}^{second}$) were identified by searching for maximum EVI or maximum $|dEVI/dt|$ are listed in Table 3.
3. We fitted the timeseries to the 1st order linearized harmonic using only EVI values corresponding to the second crop (from $t_{peak,num}^{second} - q$ to $t_{peak,num}^{second} + 2q$).

Our crop cover map achieved an overall accuracy of 82.2 +/- 0.5%, calculated via cross-validation with the crop cover dataset in which 10% of points were eliminated per calibration. Consumer's and producer's accuracy are displayed in Table 4. The 25 km scale sowing and harvest date ranges reported for single and double cropped soy in Table 5 account for this classification error, and as discussed in Section 3.3, differences among cropping intensities can be discerned despite crop cover classification error.

Next, we used a map of center pivot irrigation locations in 2014 to mask out potentially irrigated pixels from 2004 to 2014 (Step 3a, Figure 2). Irrigated fields do not conform to the assumptions made in the time series analysis, and correspond to fundamentally different sowing dates and adaptation options than rainfed fields. We made the assumption that non-irrigated locations in 2014 were also non-irrigated in previous years. This was justified because center pivot is a permanent structure that, once installed, will not be dismantled for several years. This means that a field that was rainfed in 2014 is unlikely to have been irrigated in past years. In the absence of center pivot data for all years in our study period, we created a conservative mask for center pivot by eliminating all pixels in previous years (2004–2013) that were irrigated in 2014.

In a final quality control step, we tested the predicted sowing and harvest dates against a series of rules intended to screen out implausible results (Step 3b, Figure 2), requiring that:

1. The peaks of the first and second crop were more than 20 days apart.
2. Sowing occurred between August 1 of the sowing year and May 31 of the harvest year. This rule recognizes that no soy is planted during the height of the dry season.
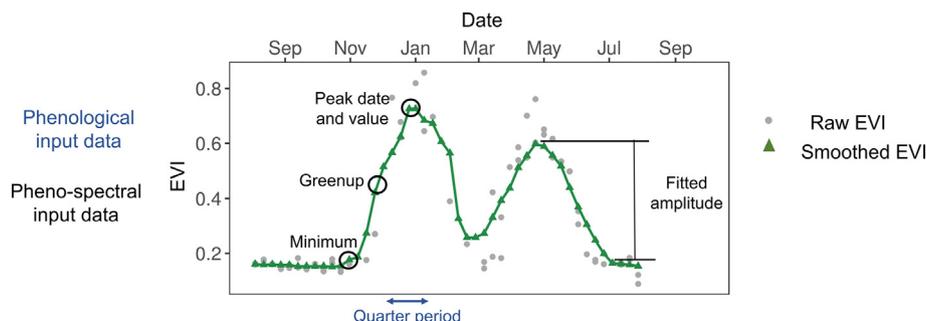
Figure 6. Input data for the crop classification for a sample pixels and year.

Table 3. The date windows over which phenological stages were sought.

Description	Symbol	Date Window
Peak date of the second crop	$t_{peak,num}^{second}$	April 1 to July 31 of the harvest year
Browndown for the second crop	$t_{browndown}^{second}$	$t_{peak,num}^{2}$ until July 31 of the harvest year

Table 4. Crop cover accuracy and confusion matrix.

	Single cropped soy	Double cropped soy	Non soy agriculture	Producer's accuracy
Single cropped soy	440	548	3	44%
Double cropped soy	301	4367	53	92%
Non soy agriculture	9	128	107	43%
Consumer's accuracy	59%	86%	66%	82.5%

Table 5. Sowing and harvest date error at aggregated scales. This combines pixel level errors in sowing and harvest date estimates and land cover classification errors. Pixel scale errors dominate.

Total error [days]	Single cropped soy	Double cropped soy	All soy (single + double cropped)
Sowing	6.9 ± 18.7	6.9 ± 17.5	6.9 ± 17.4
Harvest	1.9 ± 21.3	1.8 ± 19.9	1.8 ± 19.8

3. The crop cycle of soy (sowing to harvest) was between 60 and 150 days. The average soy crop cycle is 120 days long, 90 days for short cycle varieties.
4. A soy pixel must have a raw peak EVI of at least 0.8 during the growing season and a fitted EVI amplitude of at least 0.15. This rule filters out natural vegetation pixels that are misclassified as soy using EVI properties of soy that are established in literature. Soy pixels display a much higher seasonal change in EVI than forest cover and have a larger peak EVI value than savannah, so pixels can be filtered by peak EVI and the amplitude of the fitted EVI curve.

2.4.4. Step 4: Create 25 km sowing and harvest dates and quantify estimation errors

Steps 1 to 3 allowed us to produce 500 m sowing and harvest date maps for single and double cropped soy from 2004 to 2014. We then aggregated our estimates to 25 km cells to match the scale of available wet season onset data and to visualize broader spatial trends. Our estimates contain errors from two sources: (1) pixel-scale errors associated with time series analysis and the calibrated equation; and (2) regional-scale (25 km) errors associated with the crop cover map. Pixel-scale errors apply at field scales, but when these are aggregated to regional scales, errors associated with the crop cover map must be included.

Errors at the pixel level were quantified with the reference ground data (Figure 7). However, because the reference ground data were reported as a range of plausible sowing/harvest dates, the calibration data themselves contain uncertainty. We used the law of total probability to aggregate the error and its uncertainty at individual pixel scales into an error distribution that describes all pixels.

We modeled the uncertainty in pixel-level error as a uniform distribution spanning from the difference between the estimate and the lower bound, and the difference between the estimate and the upper bound of the reported range. Each pixel in the dataset contained its own unique pixel-level errors, which were then aggregated into a single pixel-level error distribution, p(x), describing all pixels, using the law of total probability (Eq. (5)). The distributions of the error bounds, *a* and *b*, were found by fitting a normal distribution to *a* and *b* values found at individual pixels. Eq. (5) was solved numerically in R. Sowing and harvest error distributions were treated separately.

$$p(x) = \int_0^{+\infty} \int_{-\infty}^{+\infty} \frac{H(x-a) - H(x-y-a)}{y} *p(a)*p(y)\, da\, dy \quad (5)$$

Where :
a = estimate − upper bound of reported range
b = estimate − lower bound of reported range
$y = b - a$
H = Heaviside function
$p(a)$ = probability density distribution of a
$p(y)$ = probability density distribution of b
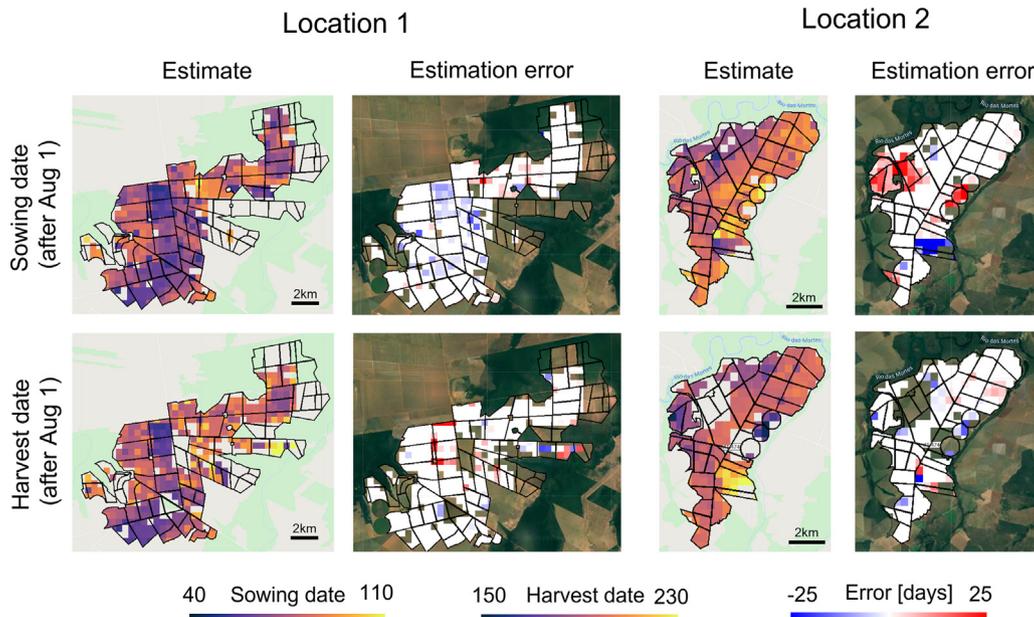
Figure 7. Estimated pixel-scale sowing and harvest dates and pixel-scale estimation errors for Planet Labs data locations (as labeled in Figure 4). The pixels in these maps were quality masked as described in Step 3. Some fields do not contain reported data because there were not enough Planet Labs images to construct a range of possible sowing/harvest dates less than 1.5 months long. The errors shown in this figure are defined following Section 2.4.2.

At regional scale, our estimates also contain errors due to misclassified soy cover. Because the sowing and harvest date estimation method was independent of land cover classification, the misclassification in land cover contributed to sowing and harvest date as a mislabeling of pixels when aggregating. For example, the sowing date of double cropped soy within a 25 km region has uncertainties associated with both the sowing date error at individual pixels and the misclassification of double cropped soy in the region.

We simulated the error introduced by misclassification through bootstrapping. We generated a "true" land cover map for a 25 km region containing the average proportion of single cropped, double cropped soy and non-soy agriculture found in Mato Grosso. From this, we generated many "erroneous" land cover maps in agreement with the confusion matrix. For each erroneous land cover map, we calculated the median sowing and harvest dates for single cropped soy, double cropped soy, and all soy pixels. The difference between this median and the corresponding median for the "true" land cover map represented the error introduced to the aggregated sowing and harvest dates by erroneous land cover classification. The total error at the 25 km aggregated scale was a simple sum of the pixel level estimate error and the error introduced by the land cover classification.

3. Results

3.1. Spatial pattern and variation in sowing and harvest dates

Figure 7 shows quality-controlled estimates of the sowing and harvest dates for two Planet Labs imagery locations in Mato Grosso. At pixel scale, the data reveal large differences in the timing of soy agriculture across adjacent fields, showing that neighboring fields of 1–2 km in size can differ in their sowing dates by more than one month, and in their harvest dates by more than two months.

Figure 8 shows median sowing and harvest dates over 25 km cells for single and double cropped soy for selected years between 2004 and 2014. The maps display interannual and regional variation in the sowing and harvest dates.

The long-term spatial pattern of sowing and onset dates, averaged over 2004 to 2014, is shown in Figure 9. For both cropping intensities, central Mato Grosso is planted earlier, while areas closer to the state border are planted later. As we explore in greater depth in a subsequent paper, Zhang et al. (2021), onset likely plays a large role in determining these the regional differences in sowing dates, as the spatial patterns in sowing roughly follow patterns of onset. However, onset cannot explain
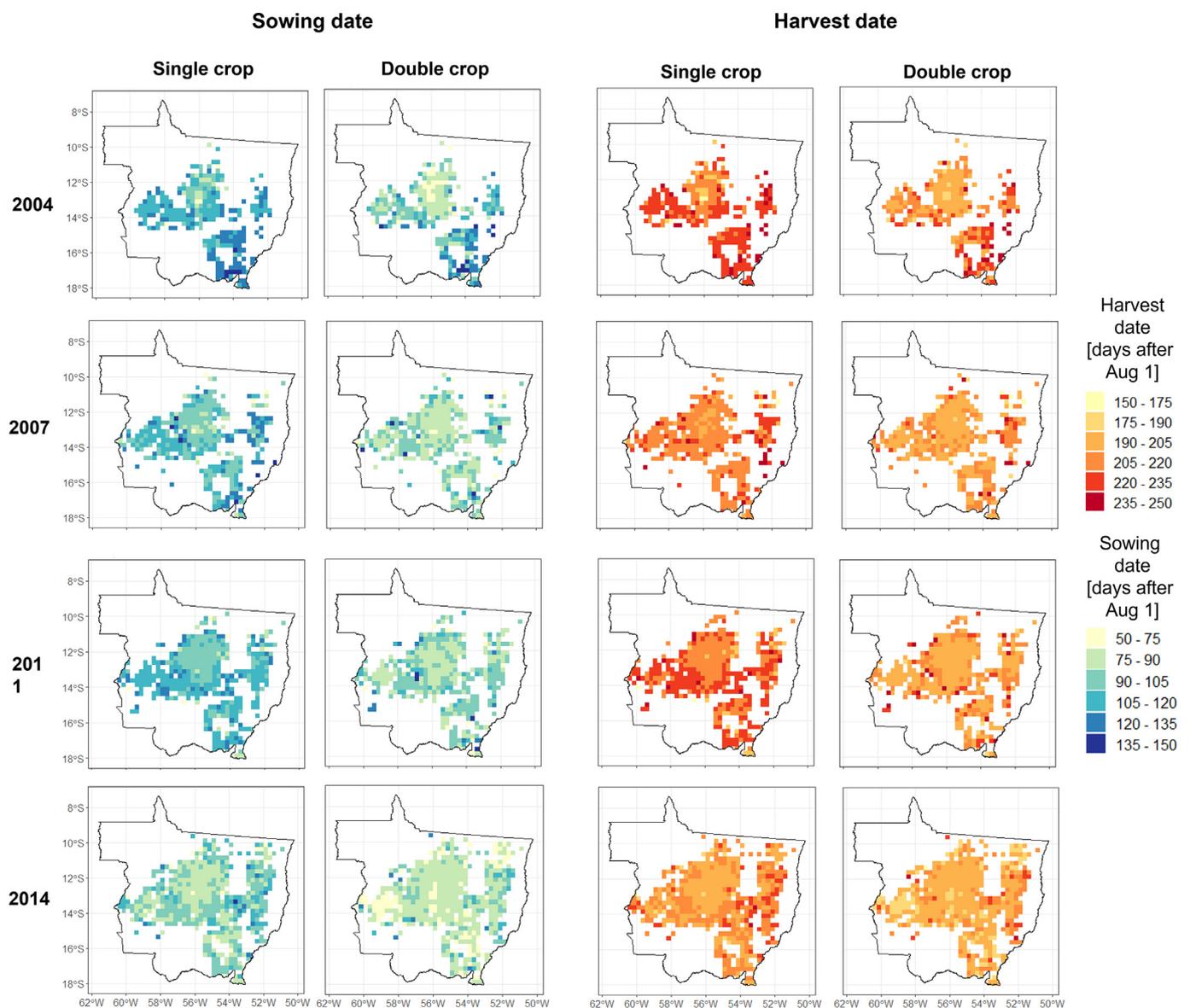
Figure 8. Estimated median sowing and harvest dates, in days after August 1 of the sowing year, over 25 km cells.

all of the regional variability: time-averaged onset is more spatially homogeneous than time-averaged sowing dates. The difference in spatial pattern of sowing dates (and of sowing dates after the influence of onset is removed, Figure 11) between the two cropping intensities also hints at non-climatic controls on sowing decisions. Single-cropped soy appears to have a stronger spatial pattern in sowing dates than double cropped soy, though they experience the same onset. Single cropped soy is both more spatially autocorrelated and more variable than double cropped soy: Moran's I is 0.577 and 0.537 for single and double cropped soy, respectively, and standard deviation is 8.24 and 7.77 days. Additionally, in the central part of Mato Grosso, single and double cropped soy are planted at very similar times, but sowing dates diverge at the edges of the state.

3.2. Interannual trends and constraints in sowing dates

3.2.1. Sanitary break is not a hard limit for early sowing, but wet season onset may be influential

Figure 10a shows histograms of sowing and harvest dates for single and double cropped soy from 2004 to 2014, overlaid on Mato Grosso's median wet season onset for the corresponding year in blue; Figure 10b shows the delay between sowing date and onset for single and double cropped soy from 2004 to 2014. These data reveal that the delay consistently (with the exception of 2010) decreased from 2004 to 2014 for both single and double cropped soy. During 2010, there was an anomalously early onset of the wet season, but sowing dates did not shift to a correspondingly early time. By 2014, the delay for both double and single cropped soy were both at an all-time low of 19 and 30 days, respectively.

For most soy agriculture in Mato Grosso, sowing dates occur much later than either the wet season onset or the sanitary break, indicating that neither of these constraints is a hard limit on early sowing. Our results indicate that the mean sowing date for double cropped fields is 89 days after August 1 (late October), while the mean sowing date for single

cropped fields is 98 days after August 1 (early November) (Figure 10a). Thus, the average sowing date for double cropped soy is over a month after the end of the sanitary break. Similarly, Figure 10a reveals a delay between sowing and wet season onset of at least 19 days, up to three months for late-planted single cropped soy, and up to two months for late-planted double cropped soy. This delay in sowing may indicate that the onset date used here is estimated too early. However, the onset definition used here produces mean yearly onset estimates close to alternative onset definitions. The anomalous accumulation definition produces onset estimates no more than 2 days later than a definition based on frequency of rainy days within a 4-week period, and no more than 5 days later than a definition based on accumulated depth of precipitation over a 50-day period. The anomalous accumulation definition more often produces onset estimates that are *earlier* than the other definitions. An alternative explanation may stem from the large size of farms: properties of 10,000 ha may require up to 4 weeks to complete sowing (Pires et al., 2016). Therefore, observed sowing dates may not reflect intended sowing dates, especially for larger properties.

Additionally, it is clear that the onset and sanitary break *do* have impact on crop timing: sowing dates almost never occur before these two constraints (Figure 10a). Among the two limits, the wet season onset appears to exert the more consistent pull on behavior. Except for the anomalously early-onset year of 2010, onset occurred after the end of the sanitary break, suggesting that the sanitary break is rarely relevant in driving sowing behavior compared to onset (Figure 10a). Additionally, as sowing dates moved closer to the onset, the probability density distribution of sowing dates became more concentrated. This "piling up" effect is most obvious in 2014 and suggests that farmers collectively pay closer attention to onset as they attempt to plant earlier each year (Figure 10b).

3.2.2. Double cropped soy is planted earlier than single cropped soy

Double cropped soy was consistently planted earlier than single cropped soy, although this gap shrank over time - the difference between
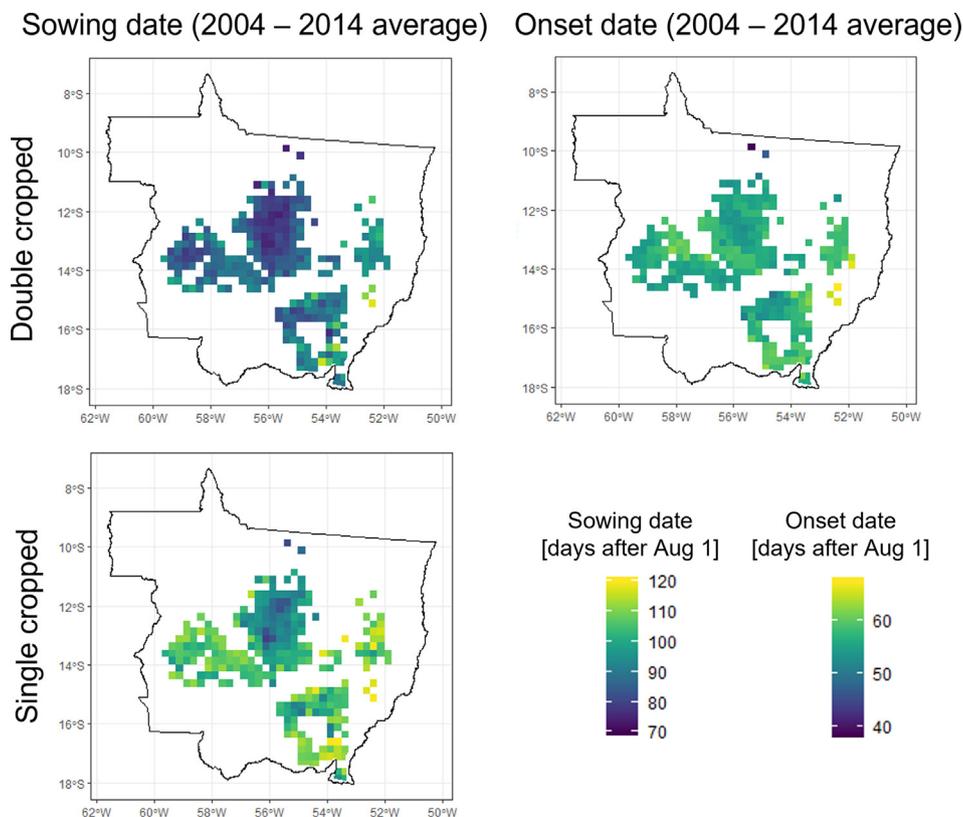
Figure 9. Estimated mean sowing and onset dates, averaged from 2004 to 2014, in days after August 1 of the sowing year. The areas shown represent only soy planted in all years of the study period.
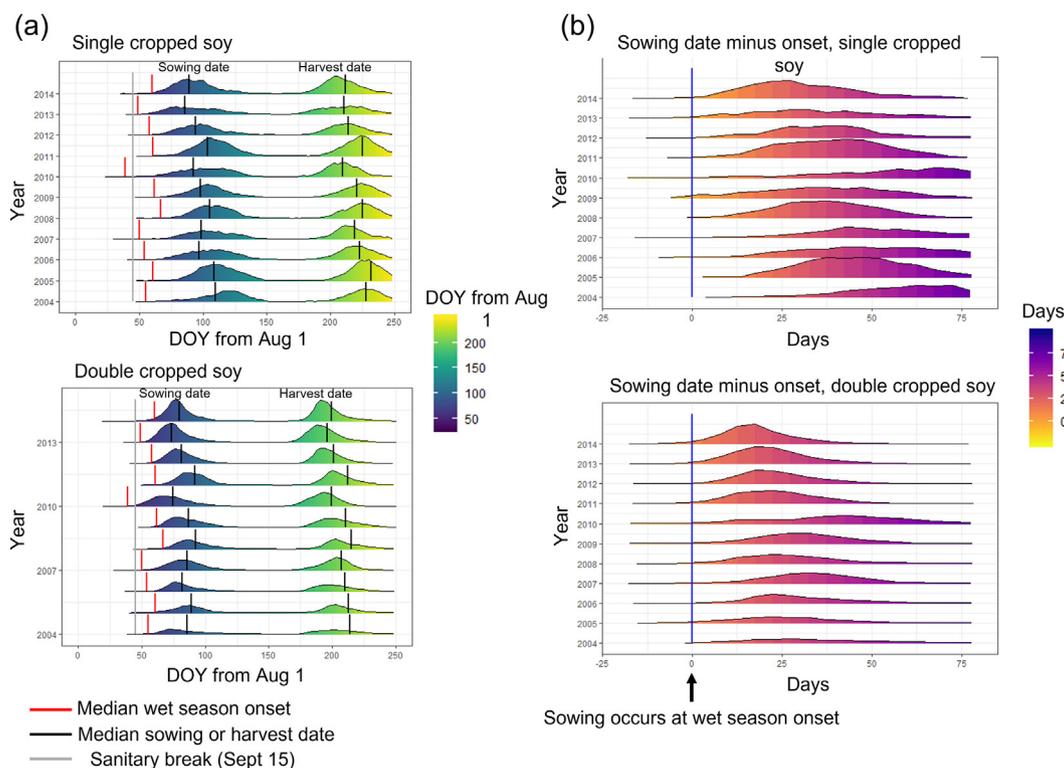
Figure 10. (a) Histogram of estimated sowing and harvest dates for single and double cropped soy across Mato Grosso from 2004 to 2014. The median wet season onset across the state is shown in red vertical lines, median sowing and harvest dates in black vertical lines, and the sanitary break in a gray vertical line. (b) The delay between sowing and onset dates across Mato Grosso. Positive value indicates that sowing occurs after onset. A delay of zero is drawn in the blue vertical line.

median sowing date for double and single cropped soy across Mato Grosso decreased from 24 days in 2004 to 10 days in 2014 (Figure 10). This is mostly associated with earlier sowing of single cropped soy, rather than later sowing of double cropped soy. The stronger trend for single cropped soy could indicate farmers sowing early to allow for double cropping but making a subsequent decision not to plant a second crop. Double cropped soy sowing dates also appear more sensitive to wet season onset than single cropped soy. For example, in 2010, double cropped soy was planted earlier to match the unusually earlier onset, while single cropped soy did not adjust as strongly to the earlier onset. The left side of double cropped sowing date histogram in Figure 10 pushes against the wet season onset, while the right side of the histogram tapers off around 110 days after August 1 - a cutoff consistent with the need to harvest soy in time for the second crop to be planted. Single cropped soy, in contrast, is less constricted on both ends and its probability density distributions are much wider.

3.3. Error analysis and validation

To quantify error at the pixel scale, we compare our estimates to the sowing and harvest dates observed directly from the high resolution Planet Labs imagery. These pixel-level errors quantify the accuracy of the sowing and harvest date estimation method itself, independently of errors in the land cover map. Following Eq. (5), the pixel-level bias and its confidence interval is 6.9 ± 16.5 days for sowing and 1.8 ± 18.7 days for harvest. This error includes the differences between the estimated and reported sowing and harvest dates, and the uncertainties in the reported sowing and harvest dates themselves.

Errors at regional scale represent a of combination errors in land cover classification and uncertainties in pixel-scale sowing and harvest date estimates. They are shown in Table 5.

Though we are unable to validate our estimates to a spatially resolved regional dataset, we ensure that our estimates agree with spatially

aggregated regional statistics. The estimated sowing and harvest dates indicate a mean crop cycle length of 112 days, consistent with reported values (90–120 days) (Abrahao and Costa, 2018). The SAGE dataset indicates that in Brazil, soy is planted in late November and harvested in late March, equivalent to 120 and 240 days after August 1, respectively (Sacks et al., 2010). This estimate is closest to sowing and harvest date estimates for single cropped soy in 2004 (109 and 227 days after August 1) - close to the year during which the SAGE data were collected (circa 2000), and during a time when single cropped soy was the dominant sowing intensity. Although SAGE data represent both irrigated and rainfed cropland globally, only 2.5% of Mato Grosso's row crop was irrigated as recently as 2017 (Hampf et al., 2020). Almost all soy in Mato Grosso was rainfed during the SAGE period. Finally, a weekly crop progress report for Mato Grosso's soy is available from the Instituto Mato-Grossense de Economia Agropecuária (IMEA) agency (IMEA, 2019). The date at which 50% of Mato Grosso is planted is comparable to the reported values, as shown in Table 6. The SAGE and IMEA datasets address potential errors due to differences in seed quality. Differences in seed quality may impact our estimate of sowing date: poor seeds may germinate more slowly, or not at all, creating noisy or even absent information in the time series. This source of error would not otherwise be captured from Planet Labs as it depends on satellite-visible greening.

The findings must be placed in context of estimation error. This is especially important because the lack of ground truth crop timing data introduced large uncertainties in validation, generating imprecise estimates. Our pixel scale errors (RMSE of 6.9 +/- 16.5 days for sowing and 1.8 +/- 18.7 days for harvest) are comparable to those from other satellite-based estimates of soy phenology in the central US (RMSEs between 3.2 and 6.9 days) that are validated against county-level USDA NASS Crop Progress Reports (Ren et al., 2017; Urban et al., 2018; Zeng et al., 2016). The lower precision of the errors in this analysis is due to the uncertain nature of sowing and harvest dates derived from Planet Labs imagery. Uncertainties due to error in validation data are not generally

included in other soy phenology studies, and it is possible that the level of precision achieved by these studies would be lower if this uncertainty were included in error analysis.

Fortunately, important features of sowing behavior can be detected despite error and imprecision in both the pixel scale estimates and the land cover map. Differences between median sowing dates of single and double cropped soy (10–24 days, depending on the year), the trend toward earlier sowing (13–20 days over the study period, for double and single cropped soy), and the delay between sowing date and onset (29–41 days) all appear clearly against the magnitude of RMSE. While imprecision at *pixel* scale (16.5 days) is sizeable compared to the detected differences, the imprecision of *aggregated* differences is much lower.

4. Discussion

The results suggest that the common assumption that the onset controls the sowing date of tropical rainfed crops is too simplistic, and often incorrect in Mato Grosso. Instead, Figure 10 reveals a delay between sowing and wet season onset of up to 3 months for single cropped soy and up to 2 months for double cropped soy, a window first posited by Abrahao and Costa (2018) based on soy's photoperiod and climatological requirements. Averaged across all years, the delay between onset and sowing date was 41 days for single cropped soy and 29 days for double cropped soy. The smaller delay over double cropped fields is important because much of Mato Grosso's agricultural revenue depends on the feasibility of double cropping (Arvor et al., 2012a). If onset is delayed sufficiently to make the earlier sowing dates of double cropping impossible, the state may suffer a loss of profit. Further, the magnitude of delay changed over time for all soy: the delay was larger in earlier years, suggesting factors that allow farmers to plant closer to the onset each year. The presence of a delay, and its change over time, means that sowing and harvest date estimates in this region should not be based upon precipitation data alone. This also means that global datasets depicting sowing and harvest dates and which rely on the assumption that sowing date occurs at the onset of the wet season may be up to 3 months in error.

Importantly, the presence of this delay does not mean that farmers are insensitive to the wet season onset. Rather, the delay between sowing dates and the wet season (and occasionally the sanitary break, when it occurs later than the onset) could be attributed to a variety of non-climatic factors acting in conjunction with the wet season onset: (1) the time (up to 4 weeks) needed to complete sowing operations (Pires et al., 2016), (2) the result of logistical and economic constraints that delay farmers from sowing at a desired time (Waha et al., 2012), or (3) a deliberate choice to improve soy yields: simulations suggest that moving sowing date from Sept 25 to Oct 5 increases soy yields by increasing the precipitation received by crops (Pires et al., 2016). In addition, a host of economic, logistical, and social factors may cause farmers to accelerate or delay sowing depending on the availability of agricultural credit (Abrahao and Costa, 2018), crop prices (Borchers et al., 2014), risk aversion (Bussmann et al., 2016; Feola et al., 2015; Kala, 2015), use of irrigation (Feola et al., 2015), memory of planting dates in recent years and other subjective beliefs (Kala, 2015), and/or recommendations from agricultural extension services (Bussmann et al., 2016). Sowing dates therefore arise as complex decisions made by farmers, based both on climate and culture. The causes of the deviation between the onset and the observed sowing dates have significant implications for the future of double cropping in Mato Grosso, and should be examined further.

More evidence of the non-climatic constraints on sowing can be found in the spatial patterns of sowing dates, emphasizing the need for high-quality sowing data that can isolate these effects. At pixel scale, the maps of soy sowing and harvest dates reveal that differences in crop timing between nearby fields are comparable to or greater than inter-annual and regional differences in the sowing and harvest dates. This suggests that field-scale knowledge is essential to characterizing sowing and harvest dates across a property, an insight that should inform future survey design. For example, soil treatments, such as mineral fertilization and the application of lime to address soil pH, are a staple of soybean agronomy in Mato Grosso (Neill et al., 2017). It is possible that delays due to transportation, equipment and labor required to apply soil treatments may impact available sowing dates, contributing to the large field-scale variation in sowing dates.

At regional scales, a distinct spatial pattern of sowing dates emerges: the center of the state is planted earlier than most other areas. This could be a result of the uniquely favorable climate, soil and topography surrounding the BR163 highway running north and south across the state (Picoli et al., 2018), or due to differences in average farm size across the state. While the spatial patterns of sowing for both double and single cropped soy roughly match the spatial pattern of onset, it is clear from the spatial patterns remaining in the delay (in which the effect of onset is taken out) that sowing dates follow a spatial pattern that is independent of onset. In Figure 11, the presence of a spatial pattern in delay indicates that onset is not the only factor influencing sowing dates. Importantly, this pattern differs by cropping intensity: sowing dates and delays for double cropped soy appear more spatially homogeneous. This discrepancy may stem from differences in resource access or optimal sowing time. Single cropping practitioners may lack the resources to complete the more expensive (but more profitable) double cropping operations, and farmers located away from the central highway network may be unable to plant as early as they desire due to logistic constraints (Picoli et al., 2018). On the other hand, the spatial homogeneity in double cropping may result from the necessity of sowing as early as possible to allow the second crop to mature before the dry season begins (Abrahao and Costa, 2018; Pires et al., 2016). Single cropping practitioners have more flexibility in choosing a yield-optimizing time, which may contribute to the spatial heterogeneity. In the future, the state's growing transportation network (including the expansion of the road network north into the Amazon), improvements in technology that allow farmers to plant closer to the wet season onset, and shifting cropping practices may create a changing spatial pattern in sowing dates, highlighting the importance of spatiotemporally resolved sowing information (Cohn et al., 2016; Picoli et al., 2018).

Though Mato Grosso is the focus of this study, the insights and techniques introduced here can be applied to understand climate risks to rainfed agriculture worldwide. The magnitude of change in the wet season is expected to be concerning in many rainfed regions: in Malawi, the RCP8.5 climate scenario will shorten the growing season by 20–55 days by midcentury (Vizy et al., 2015). El Nino events in Indonesia are expected to increase in the probability of a highly disruptive 30-day delay in monsoon onset from 9 - 18% in 2007 to 30–40% in 2050 (Naylor et al., 2007). In Burkina Faso, the rainy season onset will be delayed by an average of one week in 2021–2050 compared to the 1971–2000 baseline under the A1B scenario (Ibrahim et al., 2014), and in West Africa the combined effect of delayed onset and earlier demise will cause a 20% reduction in the length of the growing season by 2050 (Sarr, 2012).

Table 6. Comparing the estimated sowing and harvest dates to IMEA's weekly crop progress reports. IMEA does not report crop progress separately for single and double cropped soy.

Year	IMEA-reported date of 50% planted	Estimated date of 50% planted (SC, DC)	IMEA-reported date of 50% harvested	Estimated date of 50% harvested (SC, DC)
2013	October 25	October 25, October 24	February 25	February 27, February 12
2014	October 24	October 29, October 20	February 20	February 28, February 16

Delay (2004 – 2014 average)
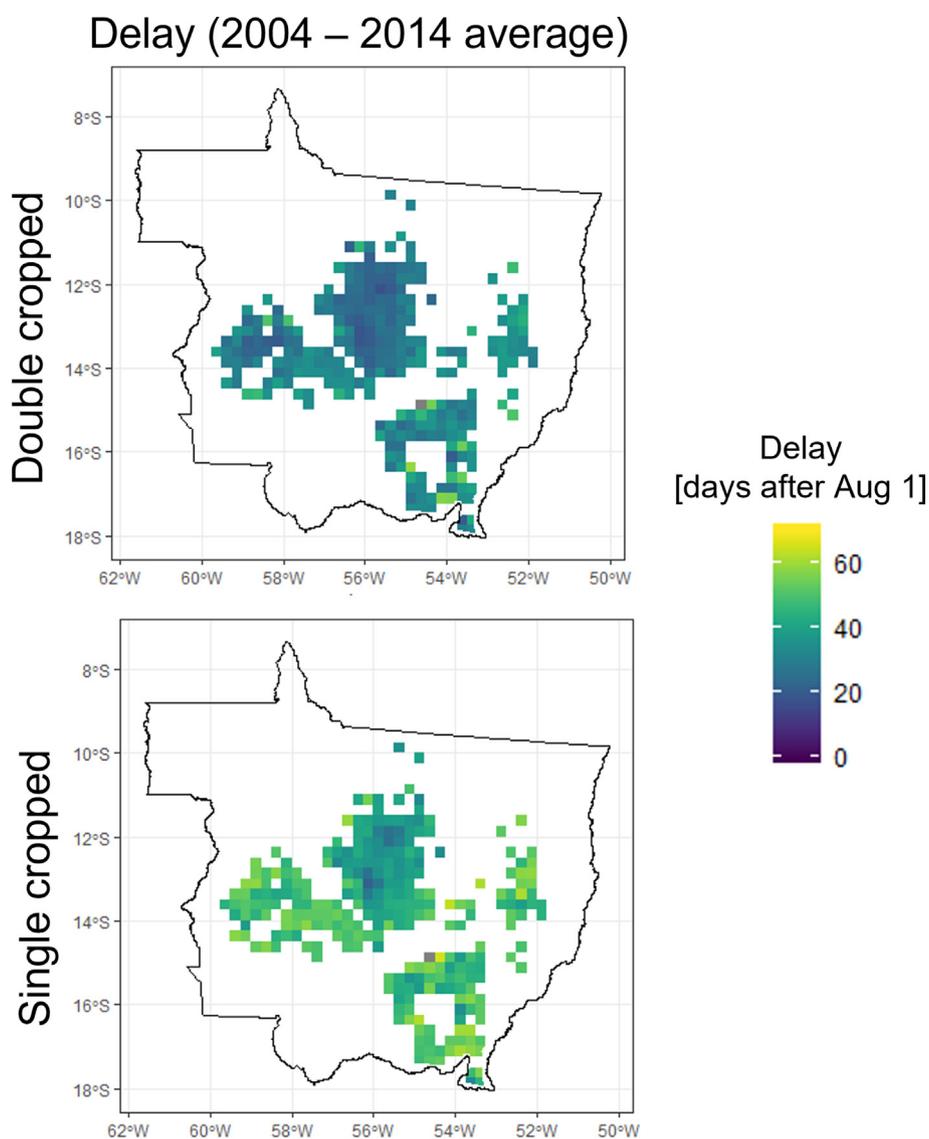
Figure 11. Estimated delay between median sowing dates and onset, averaged from 2004 to 2014, in days after August 1 of the sowing year. The areas shown represent only soy planted in all years of the study period.

While many studies have highlighted planting dates' potential as an adaptation strategy to a shortening wet season, there is still a need for a realistic understanding of planting behavior and its future trajectory. Studies that simulate optimal, yield-maximizing planting dates under climate change scenarios may recommend unrealistic dates, or dates that prevent intensive cropping systems. In Cameroon, optimized planting dates were three months later than traditionally observed planting dates (Laux et al., 2010); in Sudan, the recommended planting date was two to four weeks earlier than actual practice (Bussmann et al., 2016). The large gap between actual and optimal dates suggests that the optimized scenario may not be implemented without systematic (and possibly unlikely) changes in agricultural practice. Additionally, it is implausible that all farmers, in their various socio-economic contexts, will adapt to an equal degree. Generation of a similar high-quality sowing date dataset may be crucial to quantify risk to agricultural productivity, especially in vulnerable tropical and developing regions.

5. Conclusions

In dynamic areas like Mato Grosso, where climatic, technological, and economic factors cause sowing dates to shift quickly in time and space,

spatiotemporally resolved sowing and harvest date information is essential to forming an effective response to climate change. The dataset generated in this study demonstrates that the relationship between sowing date and wet season onset for rainfed crops varies with space, time, and cropping intensity. Sowing estimates that assume rainfed crops in tropical regions will be planted at wet season onset, without considering location, year and cropping intensity, could therefore introduce errors of up to three months. We also observe that in Mato Grosso, double cropped soy is planted closer to the start of the wet season onset with each passing year. Likely the result of improving equipment availability or crop varieties, this trend suggests that a delay in onset could force sowing to less favorable dates, reversing at least a decade of logistic and technological advances. Given the region's heavy reliance on agricultural revenue, the resulting decline in soy yield across Mato Grosso could be a significant shock.

These insights could aid efforts to assess the impact of climate change on Brazilian agriculture in a way that reflects the diverse non-climatic factors that go into sowing decisions (Hertel et al., 2010; Reidsma et al., 2010). Crop modeling efforts often resort to approximations in which sowing is triggered based on precipitation, temperature, or soil moisture (Hampf et al., 2020; Jones et al., 2003; Stockle et al., 2003;

Elliott et al., 2015; Schlenker and Lobell, 2010). These rules perform well on a global or continental scale, but are less appropriate on smaller scales (Dobor et al., 2016). It would be beneficial to incorporate into these predictions what we now know about the sizeable and shifting delay between onset and sowing dates in Mato Grosso.

A realistic understanding of sowing behavior, and consequently an accurate depiction of future yields, is only possible with updated, highly resolved sowing data. By introducing an estimation method that closes the information gap on sowing dates without the need for expensive ground survey data, we gain insights into sowing behavior for a data-scarce but agriculturally important area. Similar high-quality data will be valuable for risk assessment in regions such as southern Asia and southern Africa, which face not only the most severe consequences of warming, but also data scarcity and limited adaptive capacity (Laux et al., 2010; Lobell et al., 2008).

Declarations

Author contribution statement

Minghui Zhang: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Gabriel Abrahao; Avery Cohn; Sally Thompson: Conceived and designed the experiments.

Jake Campolo: Performed the experiments.

Funding statement

Data availability statement

Data associated with this study has been deposited at Moderate Resolution Imaging Spectroradiometer (MODIS).

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Abrahao, G., Costa, M., 2018. Evolution of rain and photoperiod limitations on the soybean growing season in Brazil: the rise (and possible fall) of double-cropping systems. Agric. For. Meteorol. 245, 32–45.

Agrosatelite, 2017. Planted area and crop yield. https://agrosa telite.com.br/en/products/safras/#planted-area.

Alexander, P., Rabin, S., Anthoni, P., Henry, R., Pugh, T.A., Rounsevell, M.D., Arneth, A., 2018. Adaptation of global land use and management intensity to changes in climate and atmospheric carbon dioxide. Global Change Biol. 24, 1–10.

Alexandrov, V., Eitzinger, J., Cajic, V., Oberforster, M., 2002. Potential impact of climate change on selected agricultural crops in north-eastern Austria. Global Change Biol. 8, 372–389.

Anwar, M.R., Liu, D.L., Macadam, I., Kelley, G., 2013. Adapting agriculture to climate change: a review. Theor. Appl. Climatol. 113, 225–245.

Arvor, D., Dubreuil, V., Simoes, M., Begue, A., 2012a. Mapping and spatial analysis of the soybean agricultural frontier in mato grosso, Brazil, using remote sensing data. Geo J. 78, 833–850.

Arvor, D., Meirelles, M., Dubreuil, V., Begue, A., Shimabukuro, Y.E., 2012b. Analyzing the agricultural transition in mato grosso, Brazil, using satellite-derived indices. Appl. Geogr. 32, 702–713.

Arvor, D., Dubreuil, V., Ronchail, J., Simoes, M., Funatsu, B.M., 2014. Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in mato grosso (Brazil). Int. J. Climatol. 34, 2622–2633.

Baldwin, B.S., Cossar, R.D., 2009. Caster yield in response to planting date at four locations in the south-central United States. Indust. Crops Prod. 29, 316–319.

Becker, W.R., Silva, L.C.d.A., Richetti, J., Lo, T.B., Johann, J.A., 2020. Harvest date forecast for soybeans from maximum vegetative development using satellite images. Int. J. Rem. Sens. 42, 1121–1138.

Begue, A., Lebourgeois, V., Bappel, E., Todoroff, P., Pellegrino, A., Baillarin, F., Siegmund, B., 2010. Spatio-temporal variability of sugarcane fields and recommendations for yield forecast using ndvi. Int. J. Rem. Sens. 31, 5391–5407.

Begue, A., Arvor, D., Bellon, B., Betbeder, J., de Abelleyra, D., Ferraz, R., Lebourgeois, V., Lelong, C., Simoes, M., Veron, S., 2018. Remote sensing and cropping practices: a review. Rem. Sens. 10, 1–32.

Bolton, D., Gray, J., Melaas, E., Moon, M., Eklundh, L., Friedl, M., 2020. Continental-scale land surface phenologyfrom harmonized landsat 8 and sentinel-2 imagery. Rem. Sens. Environ. 240, 1–16.

Borchers, A., Truex-Powell, E., Wallander, S., Nickerson, C., 2014. Multi-cropping practices: recent trends in double-cropping. U. S. Dep. Agric. Econom. Informat. Bull. 125, 1–22.

Boschetti, M., Stroppiana, D., Brivio, P., Bocchi, S., 2009. Multi-year monitoring of rice crop phenology through time series analysis of modis images. Int. J. Rem. Sens. 30, 4643–4662.

Brito, S., 2017. Atlas da irriga ção mostra estudo do uso daˊagua na agricultura nacional. Agencia Nacional deAguas, Embrapa.

Brown, J.C., Kastens, J.H., Coutinho, A.C., Victoria, D.d.C., Bishop, C.R., 2013. Classifying multiyear agricultural land use data from mato grosso using time-series modis vegetation index data. Rem. Sens. Environ. 130, 39–50.

Bussmann, A., Elagib, N.A., Fayyad, M., Ribbe, L., 2016. Sowing date determinants for sahelian rainfed agriculture in the context of agricultural policies and water management. Land Use Pol. 52, 316–328.

Challinor, A., Watson, J., Lobell, D., Howden, S., Smith, D., Chhetri, N., 2014. A meta-analysis of crop yield under climate change and adaptation. Nat. Clim. Change 4, 287–291.

Clinton, N., 2017. Time Series Analysis in Earth engine.goo.gl/lMwd2Y. (Accessed 24 June 2018).

Cohn, A.S., VanWey, L.K., Spera, S.S., Mustard, J.F., 2016. Cropping frequency and area response to climate variability can exceed yield response. Nat. Clim. Change 6, 601–604.

Correa, P., Schmidt, C., 2014. Public research organizations and agricultural development in Brazil: how did embrapa get it right? The World Bank 145, 1–10.

Costa, M., Pires, G., 2010. Effects of amazon and central Brazil deforestation scenarios on the duration of the dry season in the arc of deforestation. Int. J. Climatol. 30, 1970–1979.

Dharmarathna, W., Herath, S., Weerakoon, S., 2014. Changing the planting date as a climate change adaptation strategy for rice production in kurunegala district, Sri Lanka. Sustain. Sci. 9, 103–107.

Dobor, L., Barcza, Z., Hlasny, T., Arendas, T., Spitko, T., Fodor, N., 2016. Crop planting date matters: estimation methods and effect on future yields. Agric. For. Meteorol. 223, 103–115.

Dounias, I., Aubry, C., Capillon, A., 2003. Decision-making processes for crop management on african farms.modelling from a case study of cotton crops in northern Cameroon. Agric. Syst. 73, 233–260.

Dubreuil, V., Debortoli, N., Funatsu, B., Nedelec, V., Durieux, L., 2012. Impact of land-cover change in the southern amazonia climate: a case study for the region of alta floresta, mato grosso, Brazil. Environ. Monitor. Assess. 184, 877–891.

Elliott, J., Muller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K., Buchner, M., Foster, I., et al., 2015. The global gridded crop model intercomparison: data and modeling protocols for phase 1 (v1.0).Geoscientific Model Development, 8, 261–277.

Farr, T.G., Rosen, P.A., Caro, E., et al., 2007. The shuttle radar topography mission. Re. Geophys. 45, 1–33.

Feola, G., Lerner, A.M., Jain, M., Montefrio, M.J.F., Nicholas, K.A., 2015. Researching farmer behaviour in climate change adaptation and sustainable agriculture: lessons learned from five case studies. J. Rural Stud. 39, 78–84.

Fu, R., Yin, L., Li, W., et al., 2013. Increased dry-season length over southern amazonia in recent decades and its implication for future climate projection. Proc. Nat. Acad. Sci. 1073, 1–6.

Gao, F., Anderson, M., Daughtry, C., Karnieli, A., Hively, D., Kustas, W., 2020. A within-season approach for detecting early growth stages in corn and soybean using high temporal and spatial resolution imagery. Rem. Sens. Environ. 242, 1–19.

Geerken, R.A., 2009. An algorithm to classify and monitor seasonal variations in vegetation phenologies and the inter-annual change. ISPRS J. Photogram. Rem. Sens. 64, 422–431.

Gourdji, S.M., Sibley, A.M., Lobell, D.B., 2013. Global crop exposure to critical high temperatures in the reproductive period: historical trends and future project. Environ. Res. Lett. 8, 1–10.

Grassini, P., van Bussel, L.G., et al., 2015. How good is good enough? data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Res. 177, 49–63.

Hampf, A.C., Stella, T., Berg-Mohnicke, M., Kawohl, T., Kilian, M., Nendel, C., 2020. Future yields of double-cropping systems in the southern amazon, Brazil, under climate change and technological development. Agric. Syst. 177, 1–18.

Hassan, R.M., Nhemachena, C., 2008. Determinants of african farmers' strategies for adapting to climate change: multinomial choice analysis. Res. Agric. Appl. Econom. 3, 83–104.

Hertel, T.W., Burke, M.B., Lobell, D.B., 2010. The poverty implications of climate-induced crop yield changes by 2030. Glob. Environ. Change 20, 577–585.

Hmimina, G., Dufrene, E., Pontailler, J., Delpierre, N., et al., 2013. Evaluation of the potential of modis satellite data to predict vegetation phenology in different biomes: an investigation using ground-based ndvi measurements. Rem. Sens. Environ. 132, 145–158.

Howden, S.M., Soussana, J.-F., Tubiello, F.N., Chhetri, N., Dunlop, M., Meinke, H., 2007. Adapting agriculture to climate change. Proc. Nat. Acad. Sci. 104, 19691–19696.

Ibrahim, B., Karambiri, H., Polcher, J., Yacouba, H., Ribstein, P., 2014. Changes in rainfall regime over Burkina Faso under the climate change conditions simulated by 5 regional climate models. Clim. Dyn. 42, 1363–1381.

Iizumi, T., Kim, W., Nishimori, M., 2019. Modeling the global sowing and harvesting windows of major crops around the year 2000. J. Adv. Model. Earth Syst. 11, 99–112.

IMEA, 2019. Crop progress report.imea.com.br/imea-site/relatorios-mercado.

Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., Ritchie, J., 2003. The dssat cropping system model. Eur. J. Agronomy 18, 235–265.

Jonsson, P., Edklundh, L., 2004. Timesat – a program for analyzing time-series of satellite sensor data. Comp. Geosci. 30, 833–845.

Kala, N., 2015. Ambiguity aversion and learning in a changing world: the potential effects of climate change fro India agriculture. Working Paper.

Kastens, J.H., Brown, J.C., Coutinho, A.C., Bishop, C.R., Esquerdo, J.C.D., 2017. Soy moratorium impacts on soybean and deforestation dynamics in mato grosso, Brazil. PLoS ONE 12, 1–21.

Khanal, U., Wilson, C., Hoang, V.-N., Lee, B., 2018. Farmers' adaptation to climate change, its determinants and impacts on rice yield in Nepal. Ecol. Econ. 144, 139–147.

Knox, J., Hess, T., Daccache, A., Wheeler, T., 2012. Climate change impacts on crop productivity in african and south asia. Environ. Res. Lett. 7, 1–8.

Kucharik, C.J., 2006. A multidecadal trend of earlier corn planting in the central USA. Agronomy J. 98, 1544–1550.

Kucharik, C.J., Serbin, S.P., 2008. Impacts of recent climate change on Wisconsin corn and soybean yieldtrends. Environ. Res. Lett. 3, 1–10.

Laux, P., Jackel, G., Tingem, R.M., Kunstmann, H., 2010. Impact of climate change on agricultural productivity under rainfed conditions in Cameroon - a method to improve attainable crop yields by planting dateadaptations. Agric. For. Meteorol. 150, 1258–1271.

Liebmann, B., Camargo, S.J., Seth, A., Marengo, J.A., Carvalho, L.M., Allured, D., Fu, R., Vera, C.S., 2007. Onset and end of the rainy season in south America in observations and the echam 4.5 atmospheric general circulation model. Am. Meteorol. Soc. J. Clim. 20, 2037–2050.

Liu, Y., Dai, L., 2020. Modelling the impacts of climate change and crop management measures on soybean phenology in China. J. Cleaner Prod. 262, 1–10.

Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P., Naylor, R.L., 2008. Prioritizing climate change adaptation needs for food security in 2030. Science 319, 607–610.

Lobell, D.B., Ortiz-Monasterio, J.I., Sibley, A.M., Sohu, V., 2013. Satellite detection of earlier wheat sowing in India and implications for yield trends. Agric. Syst. 115, 137–143.

Ma, S., Churkina, G., Trusilova, K., 2012. Investigating the impact of climate change on crop phenological events in europe with a phenology model. Int. J. Biometeorol. 56, 749–763.

Mapbiomas, P., 2019. Mapbiomas - collection 3.1 of brazilian land cover & use map series. http://mapbiomas.org/.

O. B., P. G., NASA Goddard Space Flight Center, Ocean Ecology Laboratory, 2019a. Moderate Resolution Imaging Spectroradiometer (Modis) Aqua myd09a1.006.

O. B., P. G., NASA Goddard Space Flight Center, Ocean Ecology Laboratory, 2019b. Moderate Resolution Imaging Spectroradiometer (Modis) Terra mod09a1.006.

Naylor, R.L., Battisti, D.S., Vimont, D.J., Falcon, W.P., Burke, M.B., 2007. Assessing risks of climate variability and climate change for Indonesian rice agriculture. Proc. Nat. Acad. Sci. 104, 7752–7757.

Neill, C., Jankowski, K., Brando, P.M., et al., 2017. Surprisingly modest water quality impacts from expansion and intensification of large-scale commercial agriculture in the brazilian amazon-cerrado region. Trop. Conserv. Sci. 10, 1–5.

Ortiz-Monasterio, J., Dhillon, S., Fischer, R., 1994. Date of sowing effects on grain yield and yield components of irrigated spring wheat cultivars and relationships with radiation and temperature in ludhiana, India. Field Crops Res. 37, 169–184.

Pan, Z., Huang, J., Zhou, Q., Wang, L., et al., 2015. Mapping crop phenology using ndvi time-series derived from hj-1a/b data. Int. J. Appl. Earth Observ. Geoinformat. 34, 188–197.

Picoli, M., Camara, G., Sanches, I., Simoes, R., Carvalho, A., et al., 2018. Big earth observation time series analysis for monitoring brazilian agriculture. ISPRS J. Photogramm. Rem. Sens. 145, 328–339.

Pires, G.F., Abrahao, G.M., Brumatti, L.M., Oliveira, L.J., H, C.M., Liddicoat, E., Kato, Spencerand, Ladle, R.J., 2016. Increased climate risk in brazilian double cropping agriculture systems: Implications for land use in northern Brazil. Agric. For. Meteorol. 228, 286–298.

Planet, 2017. Planet application program interface: in space for life on earth. https://api.planet.com.

Reed, B., Schwartz, M., Xiao, X., 2009. Phenology of Ecosystem Processes. Springer, New York, New York.

Reidsma, P., Ewert, F., Lansink, A.O., Leemans, R., 2010. Adaptation to climate change and climate variability in european agriculture: the importance of farm level responses. Eur. J. Agronom. 32, 91–102.

Ren, J., Campbell, J.B., Shao, Y., 2017. Estimation of sos and eos for midwestern us corn and soybean crops. Rem. Sens. 9, 1–14.

Rosenzweig, C., Elliott, J., et al., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model comparison. Proc. Nat. Acad. Sci. 111, 3268–3273.

Sacks, W., Deryng, D., Foley, J., Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. Glob. Ecol. Biogeogr. 19, 607–620.

Sakamoto, T., Van Nguyen, N., Ohno, H., Ishitsuka, N., Yokozawa, M., 2005. A crop phenology detection method using time-series modis data. Rem. Sens. Environ. 96, 366–374.

Sarr, B., 2012. Present and future climate change in the semi-arid region of west africa: a crucial input for practical adaptation in agriculture. Atmospher. Sci. Lett. 13, 108–112.

Schlenker, W., Lobell, D.B., 2010. Robust negative impacts of climate change on african agriculture. Environ. Res. Lett. 5, 1–20.

Schmidhuber, J., Tubiello, F.N., 2007. Global food security under climate change. Proc. Nat. Acad. Sci. 104, 19703–19708.

Shumway, R., Stoffer, D., 2017. Time series analysis and its applications: with R examples.

Soltani, A., Sinclair, T., 2012. Modeling Physiology of Crop Development, Growth and Yield. CAB International.

Spera, S.A., Winter, J.M., Partridge, T.F., 2020. Brazilian maize yields negatively affected by climate after land clearing. Nature Sustain. 3, 845–852.

Stockle, C.O., Donatelli, M., Nelson, R., 2003. Cropsyst, a cropping systems simulation model. Eur. J. Agronomy 18, 289–307.

Thies, J., Singleton, P., Bohlool, B., 1995. Phenology, growth, and yield of field-grown soybean and bush bean as a function of varying modes of n nutrition. Soil Biol. Biochem. 27, 575–583.

Thornton, P.K., Kristjanson, P., Forch, W., Barahona, C., Cramer, L., Pradhan, S., 2018. Is agricultural adaptation to global change in lower-income countries on track to meet the future food production challenge? Global Environ. Change 52, 37–48.

Urban, D., Guan, K., Jain, M., 2018. Estimating sowing dates from satellite data over the u.s. midwest: a comparison of multiple sensors and metrics. Rem. Sens. Environ. 211, 400–412.

Victoria, D.d.C., da Pax, A.R., Coutinho, A.C., Kastens, J., Brown, J.C., 2012. Cropland area estimates using modis ndvi time series in the state of mato grosso, Brazil. Pesquisa Agropecuaria Brasileira 47, 1270–1278.

Vizy, E.K., Cook, K.H., Chimphamba, J., McCusker, B., 2015. Projected changes in Malawi's growing season. Clim. Dynam. 45, 1673–1698.

Wagenseil, H., Samimi, C., 2006. Assessing spatio-temporal variations in plant phenology using fourier analysis on ndvi time series: results from a dry savannah environment in Namibia. Int. J. Rem. Sens. 27, 3455–3471.

Waha, K., van Bussel, L., Muller, C., Bondeau, A., 2012. Climate-driven simulation of global crop sowing dates. Global Ecol. Biogeogr. 21, 247–259.

Waha, K., Muller, C., Bondeau, A., Dietrich, J., Kurukulasuriya, P., Heinke, J., Lotze-Campen, H., 2013. Adaptation to climate change through the choice of cropping system and sowing date in sub-saharan africa. Global Environ. Change 23, 130–143.

Waongo, M., Laux, P., Traore, S., Sanon, M., Kunstmann, H., 2014. A crop model and fuzzy rule based approach for optimizing maize planting dates in Burkina Faso, west africa. J. Appl. Meteorol. Climatol. 53, 598–613.

Xavier, A.C., King, C.W., Scanlon, B.R., 2016. Daily gridded meteorological variables in Brazil (1980–2013). Int. J. Climatol. 36, 2644–2659.

Zalles, V., Hansen, M.C., Potapov, P.V., et al., 2019. Near doubling of Brazil's intensive row crop area since 2000. Proc. Nat. Acad. Sci. 116, 428–435.

Zeng, L., Wardlow, B.D., Wang, R., Shan, J., Tadesse, T., Hayes, M.J., Li, D., 2016. A hybrid approach for detecting corn and soybean phenology with time-series modis data. Rem. Sens. Environ. 181, 237–250.

Zhang, M., Abrahao, G., Thompson, S., 2021. Sensitivity of soybean planting date to wet season onset in mato grosso, Brazil and implications under climate change (accepted). Clim. Change.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Feed, B.C., Huete, A., 2003. Monitoring vegetation phenology using modis. Rem. Sens. Environ. 84, 471–475.

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. Photogram. Rem. Sens. 161, 37–51.

Zhong, L., Hu, L., Gong, P., Biging, G.S., 2016. Automated mapping of soybean and corn using phenology. ISPRS J. Photogram. Rem. Sens. 119, 151–164.