

## CropPlantHarvest: A 500 m annual dataset of crop planting and harvesting dates (2001–2024) of the U.S. Midwest

Yin Liu<sup>1</sup>, Chunyuan Diao<sup>1</sup>

<sup>1</sup>Department of Geography and Geographic Information Science, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

5 *Correspondence to:* Chunyuan Diao (chunyuan@illinois.edu)

**Abstract.** As key components of agricultural management, planting and harvesting schedules have strongly influenced crop production by defining the length of the crop growing season and shaping the environmental conditions crops experience. Accurate knowledge of these management data is crucial for enhancing crop yield estimates by capturing the timing of crop development relative to weather and soil conditions, assessing climate adaptation by tracking shifts in farming practices over time, and supporting agricultural carbon accounting. Yet, existing planting and harvesting date datasets are largely based on state-level statistics or rule-based calendars that overlook intra-regional variability and the influence of human decision-making. The absence of long-term, high-resolution planting and harvesting date information hinders our ability to reconstruct historical agricultural practices and assess their agronomic and environmental consequences. In this study, we introduce CropPlantHarvest, the first dataset of annual corn and soybean planting and harvesting dates across the U.S. Midwest at 500 m resolution from 2001 to 2024. Planting dates are estimated using CropSow, an integrative remotely sensed crop modeling system that aligns simulated crop growth trajectories with satellite observations to retrieve field-level planting dates. Harvesting dates are retrieved using the Normalized Harvest Phenology Index (NHPI), a novel index that integrates Normalized Difference Vegetation Index (NDVI) and near-infrared (NIR) reflectance to detect harvesting events by capturing the distinct spectral transition from senescent crops to exposed crop residues. Validation against USDA crop progress reports and field-level dataset demonstrates high accuracy of CropPlantHarvest, with a mean absolute error of approximately 5 days for both crop species. This large spatial and temporal dataset captures management-driven variability in crop season timing and duration, supporting improved modeling of crop yields, greenhouse gas emissions, and resource use. It could also serve as a benchmark for refining remote-sensing phenology products and evaluating the agro-environmental impacts of evolving crop management decisions. CropPlantHarvest is available at <https://doi.org/10.5281/zenodo.16967482> (Liu and Diao, 2025).

25 **Keywords:** Agricultural management practices; Planting date; Harvesting date; Phenology; Remote Sensing

## 1 Introduction

30 Food security will be increasingly challenged in the coming years due to population growth, shrinking availability of arable  
land, shifting consumption patterns, and climate change (Beddington, 2010). By 2050, global food production will need to  
increase by 60% to feed an estimated 9.3 billion people. In light of these projected changes, adapting crop management  
practices is essential to improve production efficiency and meet future food demands (Challinor et al., 2014). Among all crop  
management practices, the timing of planting and harvesting are especially critical, as they define the start and end of the  
growing season and determine the environmental conditions crops experience during development (Baum et al., 2020; Liu et  
35 al., 2023, 2025). As human-determined decisions, planting and harvesting dates reflect adaptive responses to shifting climate  
conditions (Kusumastuti et al., 2016; Xu et al., 2019; Huang et al., 2019; Baum et al., 2020; Shew et al., 2020). Adjusting  
planting dates can reduce exposure to temperature extremes and mitigate yield losses, while harvesting at the optimal time  
preserves grain quality and minimizes losses from adverse weather, pests, and disease (Jain et al., 2016; Liu et al., 2023, 2025).  
Accurate information on planting and harvesting dates is essential for assessing climate adaptation. Such data also enhances  
40 crop yield estimation by aligning crop development stages with observed weather and soil conditions (Bolton and Friedl, 2013;  
Zhang and Diao, 2023; Guo et al., 2024; Pei et al., 2025), and they play a central role in agricultural carbon accounting by  
identifying the periods when key exchanges of carbon, water, and energy occur between the land surface and atmosphere  
(Delgrosso et al., 2005; Zhang et al., 2020). Reliable records of these management decisions are therefore indispensable for  
climate adaptation analysis, crop yield estimation, and greenhouse gas accounting, ultimately supporting more accurate  
45 modeling and sustainable agricultural management.

In response to this need, several global datasets (e.g., SAGE (Sacks et al., 2010), MIRCA2000 (Portmann et al., 2010),  
RiceAtlas (Laborte et al., 2017), and PSHW (Iizumi et al., 2019), GCPE (Mori et al., 2023)) have been generated to provide  
information on crop planting and harvesting. MIRCA2000 and RiceAtlas compile crop-specific planting and harvesting  
months at subnational scales by integrating regional databases, national reports, and expert knowledge. In contrast, SAGE,  
50 PSHW, and GCPE apply rule-based methods that relate long-term climate variables (e.g., temperature, precipitation, and  
growing degree days) to the likelihood of planting and/or harvesting events, producing global estimates at a 0.5° resolution.  
While these datasets have been widely used to support research and policy efforts in climate adaptation, crop modeling, and  
sustainable land use, the coarse spatial resolutions of these datasets limit their ability to capture fine-scale spatial heterogeneity  
in crop management practices. Furthermore, they provide rough crop calendars or climate-derived estimates rather than actual  
55 management dates, which limits their capacity to capture temporal variability in planting and harvesting activities across  
diverse farming systems (Liu et al., 2023). This underscores the need for long-term, high-resolution datasets on planting and  
harvesting dates to better represent spatiotemporal variability in agronomic practices and improve modeling accuracy.

To address the data gap, remote sensing has been increasingly employed to estimate planting and harvesting dates across large  
spatial and temporal scales by interpreting vegetation dynamics captured in Vegetation Index (VI) time series (Yang et al.,  
2026; Cao et al., 2024; Diao, 2020; Diao et al., 2021; Diao and Li, 2022; Gao et al., 2017; Liu et al., 2022a; Moulin et al.,  
60 1997; Sacks and Kucharik, 2011; Schwartz et al., 2002; Shen et al., 2023; Wu et al., 2017; Zhang et al., 2003). Several  
phenology datasets have been produced following this remote sensing-based strategies, where phenometrics derived from VI  
time series, such as the start of season (SOS) and end of season (EOS), are applied to approximate planting and harvesting  
timing (Bolton et al., 2020; Luo et al., 2020; Nieto et al., 2021; Niu et al., 2022). Yet, these datasets face important accuracy  
65 limitations in capturing planting and harvesting dates, as the phenometrics extracted typically using the feature points in VI  
curves do not consistently align with the actual timing of these field operations. Planting and harvesting practices are human-  
managed decisions that are primarily driven by external factors such as weather conditions, soil trafficability, and machinery  
availability, which can shift their feature positions on the VI curve (Liu et al., 2023, 2025). Moreover, planting is particularly  
challenging to detect directly, as it occurs before vegetation emerges and thus falls within an off-season flat portion of the VI

70 curve that provides little signal to indicate the event. Harvesting presents similar challenges when crops remain in the field long after maturity, in which case the event appears on a VI plateau rather than at a distinct turning point, as over-mature crops exhibit VI values comparable to bare soil. These limitations reduce the reliability of existing remote sensing-based datasets for accurately characterizing planting and harvesting dates (Gao and Zhang, 2021).

In this study, we introduce CropPlantHarvest, the first crop management practices dataset (i.e., planting and harvesting timing) for corn and soybean across the U.S. Midwest, mapped annually at 500 m resolution for 2001–2024. Planting dates are derived using CropSow, an integrative remotely sensed crop modeling system that assimilates remotely sensed VI time series with crop growth trajectories simulated by the Agricultural Production Systems sIMulator (APSIM) to estimate field-level planting dates (Liu et al., 2023). Harvesting dates are estimated with the Normalized Harvest Phenology Index (NHPI), a novel index that integrates Normalized Difference Vegetation Index (NDVI) and near-infrared (NIR) reflectance to detect the distinct spectral transition from senescent crops to exposed crop residues that marks harvesting events (Liu et al., 2025). Validation against field-level dataset and United States Department of Agriculture (USDA) crop progress reports demonstrates high accuracy of CropPlantHarvest, with a mean absolute error of approximately 5 days for both crop types. As one of the world’s most productive and intensively managed agricultural regions, the U.S. Midwest plays a pivotal role in global food supply, contributing roughly one-third of global corn exports and over one-quarter of global soybean exports (Kucharik and Ramankutty, 2005). Its vast and heterogeneous landscapes, diverse management practices, and high sensitivity to seasonal timing make it a natural laboratory for studying dynamic changes of crop planting and harvesting decisions over space and time. Understanding these dynamics here is particularly valuable because management strategies that succeed in such a diverse, climate-sensitive system can offer transferable insights for other agricultural regions worldwide. Moreover, the Midwest’s long history of consistent monitoring through yield statistics, crop progress reports, and environmental records provides a unique opportunity to link detailed planting and harvesting timings with production trends and climate adaptation strategies. Previous research at aggregated scales has shown that corn yield gains over the past half-century have been partially driven by earlier planting and longer growing seasons (Butler et al., 2018). By exploring these trends at fine spatial resolution, our 24-year dataset could enable more precise analyses of localized management adaptations, supporting improved crop models, yield estimates/forecasts, and agro-environmental assessments.

## 95 **2 Material and method**

### **2.1 Material**

To support the generation and validation of the planting and harvesting date products in U.S. Midwest, we utilize a comprehensive suite of datasets, including satellite imagery, crop type maps, meteorological data, soil property data, and ground reference observations. The U.S. Midwest states include Illinois (IL), Indiana (IN), Iowa (IA), Kansas (KS), Michigan (MI), Minnesota (MN), Missouri (MO), Nebraska (NE), North Dakota (ND), Ohio (OH), South Dakota (SD), and Wisconsin (WI) (Fig. 1). Satellite imagery provides the time series of spectral reflectance needed to detect planting and harvesting events. Crop type maps identify the locations of corn and soybean fields. Meteorological and soil datasets supply environmental inputs for the crop growth model used in planting date estimation. Ground reference observations are used both to calibrate the crop growth model used in planting date estimation and to evaluate the accuracy of the final planting and harvesting date estimates.

#### 105 **2.1.1 Satellite data**

We use the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD43A4 Version 6 product, a daily nadir Bidirectional Reflectance Distribution Function (BRDF)-adjusted reflectance dataset with a 500-meter spatial resolution, to detect corn and soybean planting and harvesting dates of U.S. Midwest from 2001-2024. From this dataset, we derive two key time series: the NDVI to characterize crop growth dynamics, and the Normalized Harvest Phenology Index (NHPI) to capture harvesting progression. The MODIS product offers an effective balance of spatial resolution, temporal frequency, and historical coverage, making it well-suited for long-term monitoring of planting and harvesting events. Its daily acquisition enables the

detection of subtle reflectance changes during both early and late crop development phases, which are critical periods for identifying planting and harvesting dates.

### 2.1.2 Crop type data

115 We use the USDA Cropland Data Layer (CDL), an annually updated crop-specific land cover product at 30 m spatial resolution,  
to extract spectrally pure crop signals from pure MODIS pixels of our study region (Boryan et al., 2011). MODIS pixels are  
classified as “pure” when at least 90% of the 30 m sub-pixels within the 500 m MODIS footprint are assigned to the same crop  
type. This at least 90% agreement threshold is selected based on a sensitivity analysis (Fig. S1), which shows that the threshold  
provides a conservative balance between minimizing mixed-pixel effects and retaining sufficient spatial coverage with pure  
120 pixels widely distributed for reliable downstream applications across the Midwest. Under this criterion, spectral mixing from  
other land covers (e.g., different crops, fallow land, or non-agricultural areas) is substantially reduced, such that the extracted  
reflectance time series primarily reflect the phenological characteristics of a single crop type. The CDL provides high  
classification accuracy for both corn and soybean, with producer’s and user’s accuracies typically exceeding 90%. For years  
or regions without CDL coverage prior to 2008, we supplement with the Corn–Soy Data Layer (CSDL), which offers historical  
125 corn and soybean classifications from 2001 onward.

### 2.1.3 Meteorological data

To support the simulation of crop growth trajectories in the planting date estimation process, we incorporate daily  
meteorological variables from the Daymet dataset (Thornton et al., 2020). These variables include minimum temperature,  
maximum temperature, precipitation, and shortwave radiation, which are required inputs for the crop growth model. Daymet  
130 provides continuous, gridded records at 1 km spatial resolution across North America from 1980 to present. For 2024, when  
Daymet data is not yet available at the time of our analysis, we substitute equivalent records from the Parameter-elevation  
Regressions on Independent Slopes Model (PRISM) dataset, which offers high-quality, gridded climate data at 4 km spatial  
resolution for the contiguous United States. In addition to driving the crop model simulation, these meteorological variables  
are also used to examine the influence of weather conditions on temporal variability in crop planting and harvesting dates.

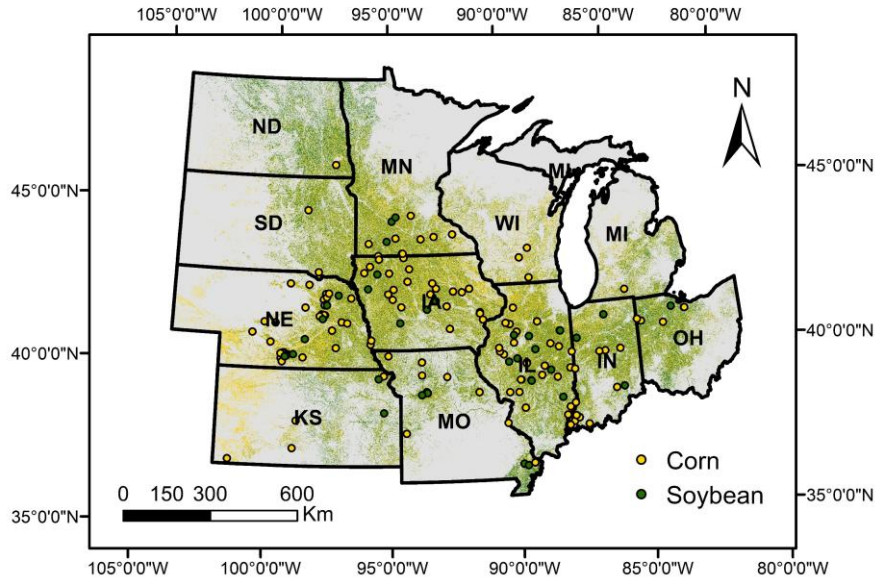
### 135 2.1.4 Soil data

In addition to meteorological variables, soil properties are incorporated to provide crop growth modeling inputs that control  
water and nutrient availability, root development, and other processes influencing crop development within the planting date  
estimation system. Soil variables, including hydraulic properties, pH, and organic matter content, are obtained from the  
Gridded Soil Survey Geographic (gSSURGO) database. gSSURGO is a 30 m resolution raster product derived from the vector-  
140 based SSURGO database, developed through extensive field sampling and laboratory analysis by the National Cooperative  
Soil Survey. These inputs are critical for accurately simulating crop development from planting through vegetative growth  
within the system.

### 2.1.5 Ground truth data

Two sources of ground truth datasets are used for model calibration and/or product validation. The first is the USDA Crop  
Progress Reports (CPRs), which provide weekly state-level cumulative percentages of crops reaching specific phenological  
145 stages (e.g., planting and harvesting). CPR data are used to calibrate the crop growth model in the planting date estimation  
system by aligning simulated planting date distributions with reported planting progress. CPRs also serve as a reference for  
evaluating the temporal consistency of planting and harvesting date estimates at the state level. The second source is Beck’s  
dataset, which contains field-level planting and harvesting records provided by Beck’s Hybrids and is publicly available at  
150 <https://www.beckshybrids.com/>. To account for the resolution difference between our 500 m product and Beck’s plot-level  
measurements, we select MODIS pixels with each falling entirely within the boundary of a crop field, filtering out those  
spanning multiple fields with PlanetScope imagery verified through visual interpretation. This yields a total of 187 field records

from 2016 to 2024 (Fig. 1), including 151 corn records and 36 soybean records. These records are used to validate the accuracy of both planting and harvesting estimates at the field scale.



155

Figure 1. Spatial distribution of corn and soybean planting and harvesting date records from Beck’s dataset (2016–2024) across 12 U.S. Midwestern states. Yellow points indicate corn fields, and green points indicate soybean fields.

## 2.2 Method

For each pure crop pixel, its MODIS NDVI time series serves as the foundation for estimating planting and harvesting dates in corn and soybean fields. Conventional remote sensing–based methods often fail to capture true planting and harvesting dates. This is because feature points on VI curves do not consistently correspond to the actual timing of these field operations. To overcome this limitation, we use CropSow for planting date estimation and the NHPI for harvesting date estimation. Specifically, the annual NDVI curve is divided at its maximum value into two segments. The pre-peak segment spans from the start of the year to the timing of maximum NDVI (typically June–July in the U.S. Midwest). It represents canopy development and supports planting date detection with the CropSow system. The post-peak segment extends from maximum NDVI through the end of the year. It represents canopy senescence and crop residue exposure and supports harvesting date detection with the NHPI-based method (Fig. 2). For planting date estimation, the CropSow system integrates a remote sensing–based phenological detection method with the APSIM crop growth model. The phenological detection method identifies Greenup, which indicates the onset of active crop growth near emergence. APSIM simulates the duration from planting to Greenup by accounting for spatially explicit soil properties, weather conditions, and crop characteristics within the soil–crop–atmosphere continuum. The planting date is then determined by subtracting the simulated planting–Greenup duration from the satellite-observed Greenup date. As a result, planting date estimation implicitly incorporates location information through the model inputs required by APSIM. For harvesting date estimation, the Normalized Harvest Phenology Index (NHPI), defined as the ratio of near-infrared (NIR) reflectance to NDVI, enhances spectral differences between senescent vegetation and exposed crop residues to signal harvest. A threshold-based approach is applied to the NHPI time series to detect the sharp transition that corresponds to harvesting date. This harvesting-date detection relies solely on temporal spectral dynamics and does not require explicit location-based inputs or ancillary environmental data. By combining these two event-specific methods across

175

180 all target pixels, we generate CropPlantHarvest, a consistent, large-scale dataset of corn and soybean planting and harvesting dates for the U.S. Midwest. The product is evaluated at field and state levels and used to assess long-term spatiotemporal trends and meteorological drivers.

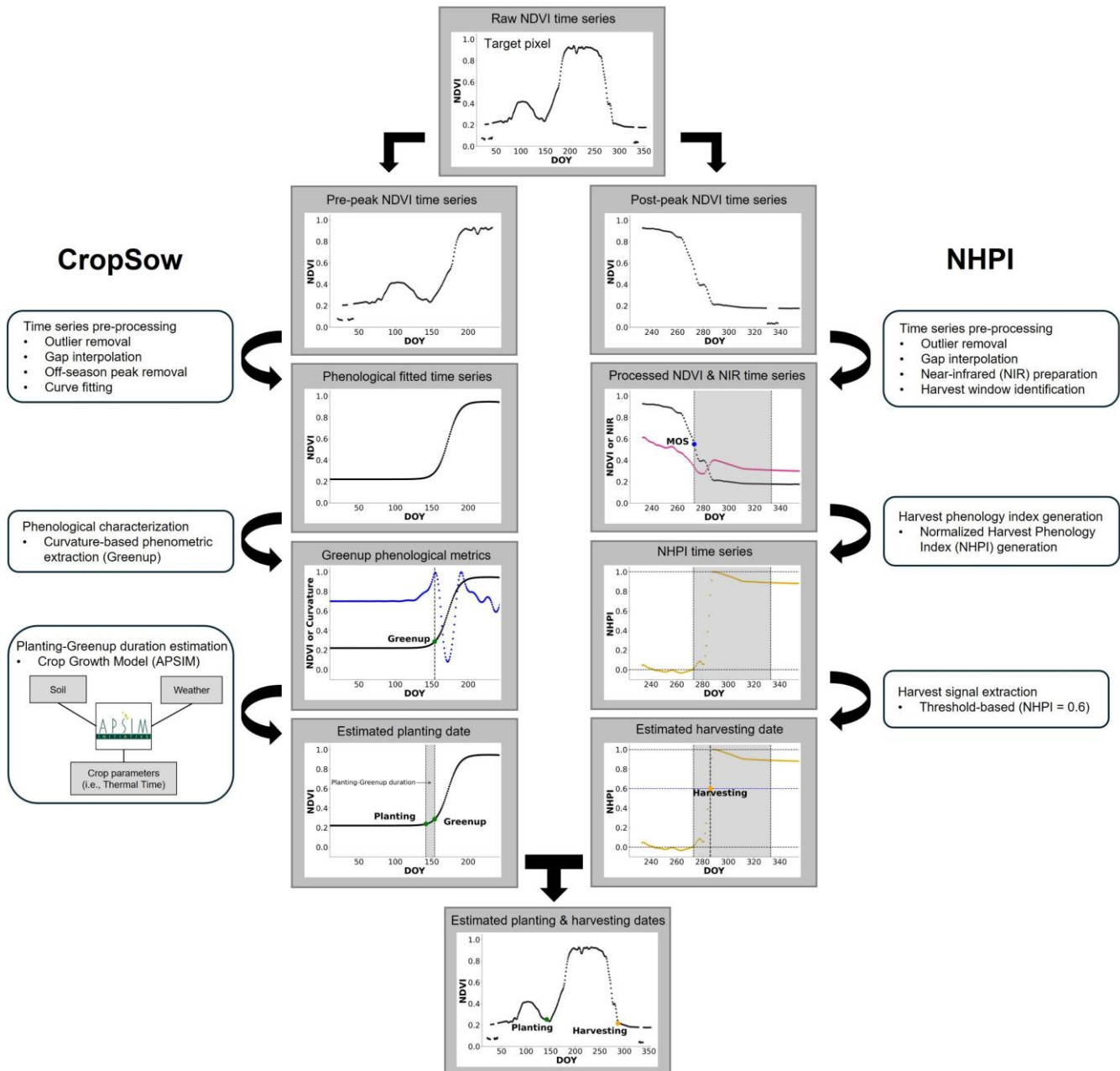


Figure 2. Workflow for estimating planting and harvesting dates from MODIS NDVI time series of pure crop pixels. The left branch shows the CropSow system for planting date estimation, while the right branch shows the NHPI-based method for harvesting date estimation.

### 185 2.2.1 Planting Date Estimation

With the pre-peak segment of the NDVI curve, the planting date is estimated using the CropSow system, which consists of three main components: time series pre-processing, phenological characterization, and estimation of the planting-to-Greenup duration (Liu et al., 2023). The first pre-processing step involves outlier removal, gap interpolation, off-season peak removal, and curve fitting. Outlier removal is performed sequentially using a quality assurance (QA) filter, a spline filter, a median absolute deviation (MAD) filter, and a snow filter. These filters eliminate implausible observations caused by snow cover, clouds, haze, or low illumination. The QA filter removes the observations flagged as poor quality in the satellite product, including those affected by cloud contamination and cloud shadows. As an additional quality-control measure, observations with NDVI values below zero are excluded because agricultural fields are generally expected to exhibit non-negative NDVI values under crop canopy, bare-soil, and crop-residue conditions. Negative NDVI values are therefore more likely to indicate contamination from water, snow, cloud shadows, or other anomalous surface conditions. After the initial quality screening, a spline filter is applied to smooth the NDVI curve and remove observations with residuals exceeding the mean plus/minus three standard deviations (Migliavacca et al., 2011). The MAD filter targets the removal of sharp spikes by applying the median absolute deviation criterion (Papale et al., 2006). Finally, the snow filter excludes observations with a Normalized Difference Snow Index (NDSI) above  $-0.2$ , removing snow-contaminated data that may not be fully identified by the previous steps. While NDSI values exceeding  $0.4$  are generally considered indicative of snow cover, we adopt a more conservative threshold of  $-0.2$  to minimize the risk of retaining snow-contaminated pixels in the time series (Liu et al., 2025). The filtered observations are then gap-filled using linear interpolation. Beyond outliers, the NDVI time series may contain off-season growth cycles unrelated to the target crop, typically resulting from weeds, cover crops, or double cropping. Such cycles can distort the characterization of phenological metrics for the target crop. To address this, a seasonality filter is applied to remove the off-season crop growth cycle (Diao, 2020). This filter employs a smoothing spline algorithm that fits a piecewise polynomial curve to the data, preserving the main seasonal growth trajectory while suppressing short-term fluctuations. Growth cycles are then delineated by turning points (i.e., peaks and troughs) in the smoothed curve, and each cycle is evaluated to determine whether it represents the main crop growth cycle. The evaluation considers both the NDVI magnitude of each peak and its timing relative to crop calendars from CPRs. This ensures consistency with the expected growth period of the target crop. The peak NDVI of the target crop cycle is required to occur between three weeks before the start of the corn silking stage (or the soybean flowering stage) and the start of the maturity stage reported in CPRs. Any cycles with localized peaks outside this temporal window are classified as off-season noise, and the segments between their preceding and subsequent troughs are removed.

Following time series pre-processing, Beck’s double logistic method is applied to fit the pre-processed pre-peak segment to ensure a consistent representation of crop phenological dynamics (Beck et al., 2006). Compared to other fitting methods such as asymmetric Gaussian, Savitzky–Golay, quadratic, and nonlinear spherical methods, the double logistic approach can better characterize vegetation with rapid phenological transitions and relatively short growing seasons without overestimating season length. The Beck’s double logistic function (Eq. 1) includes six parameters:  $V_{base}$  (off-season NDVI),  $V_{max}$  (maximum NDVI),  $m_2$  and  $n_2$  (timing of inflection points for green-up and senescence, respectively), and  $m_1$  and  $n_1$  (rates of change at those inflection points). These parameters are estimated by minimizing the root mean square error (RMSE) between the fitted curve  $f(t)$  and the pre-processed NDVI time series. While only pre-peak observations are used for fitting at this stage, the full double logistic formulation is retained to preserve the functional continuity and physiological realism of the phenological profile. This approach primarily constrains parameters associated with the green-up phase (i.e.,  $V_{base}$ ,  $V_{max}$ ,  $m_1$  and  $m_2$ ), while senescence-related parameters (i.e.,  $n_1$  and  $n_2$ ) remain weakly constrained. The curvature-based phenometric extraction method is then applied to the fitted curve to characterize the Greenup phenometric, defined as the first local maximum in the curvature change rate (Beck et al., 2006).

$$f(t) = V_{base} + (V_{max} - V_{base}) * \left( \frac{1}{1+e^{(-m_1*(t-m_2))}} + \frac{1}{1+e^{(-n_1*(t-n_2))}} - 1 \right) \quad (1)$$

230 Lastly, the Agricultural Production Systems sIMulator (APSIM) is used to estimate the interval between planting and the remotely sensed Greenup phenometric, which informs planting date retrieval. APSIM simulates crop phenological development by accounting for plant dynamic interactions within the soil–crop–atmosphere continuum. Specifically, the crop  
235 undergoes germination and emergence after planting, eventually reaching the remotely sensed Greenup stage, which roughly corresponds to the V3 stage (three collared leaves) in corn or the V3 stage (third trifoliolate) in soybean. In APSIM, germination is triggered by soil moisture conditions, while subsequent phenological development is driven by the accumulation of thermal time, expressed as daily degree days. The rate of thermal time accumulation is driven by temperature exposure and regulated by water and nitrogen stresses simulated by the SoilWater and SoilNitrogen modules, respectively. The total thermal time  
240 requirement from emergence to the remotely sensed Greenup stage is denoted as *tt\_emerg\_to\_Greenup*. Given a specific *tt\_emerg\_to\_Greenup*, the planting date is inferred by iteratively adjusting the planting date until APSIM-estimated Greenup aligned with observed Greenup phenometric. Physiologically, *tt\_emerg\_to\_Greenup* is primarily determined by the number of leaves attained at Greenup and the thermal time required for the development of an individual leaf (i.e., *phyllochron*). Given that leaf number targets are relatively consistent (i.e., three-leaf stage) and that crop *phyllochron* is broadly comparable across  
245 cultivars and primarily regulated by environmental conditions, particularly temperature and irradiance (Birch et al., 1998; dos Santos et al., 2022), *tt\_emerg\_to\_Greenup* is calibrated at the state–year level to represent environmentally driven variability in *phyllochron*. The parameter *tt\_emerg\_to\_Greenup* is specifically calibrated annually using state-level CPRs and 1,000 randomly sampled pure MODIS pixels, by minimizing the RMSE between APSIM-simulated and CPR-reported planting date distributions. All APSIM input variables (e.g., fertilizer and planting density) follow settings established in previous planting  
250 date estimation study (Liu et al., 2023). Once calibrated, APSIM is run for each target crop pixel with varying planting dates and localized soil and weather inputs. Sub-state variability in soils and microclimate is represented through pixel-specific soil properties and daily weather inputs, which govern the effective rate of thermal time accumulation and phenological progression. The planting date that yields a simulated Greenup consistent with the observed Greenup phenometric is assigned as the estimated planting date for that field.

### 250 2.2.2 Harvesting Date Estimation

With the post-peak segment of the NDVI curve, harvesting dates are estimated using the NHPI-based method, which comprises three components: time series pre-processing, harvest phenology index generation, and harvest signal extraction (Liu et al., 2025). Time series pre-processing involves outlier removal, gap interpolation, preparation of the near-infrared (NIR) time series, and determination of the harvesting window. Outlier removal follows the same procedure applied in the pre-peak  
255 preprocessing, while gaps are filled through linear interpolation. In parallel, the NIR time series is processed by removing low-quality observations flagged during outlier detection and filling the resulting gaps with linear interpolation. This ensures consistency with the NDVI time series. The harvesting window is defined as the period starting from the middle of senescence (MOS), which is the date when NDVI declines to 50% of its senescence-phase amplitude (the difference between peak and minimum NDVI), and extending two months thereafter. This window isolates harvest-related changes, minimizes noise from  
260 other phenological phases, and improves the accuracy and efficiency of following harvesting date detection. This two-month window reflects the typical temporal separation between MOS and harvest observed in field-level datasets across the U.S. Midwest and is sufficiently flexible to capture interannual variability in harvest timing. The window length can be flexibly adjusted in regions with shorter or longer post-maturity field residence to accommodate different harvesting behaviors.

The second component focuses on generating the Normalized Harvest Phenology Index (NHPI), which is designed for  
265 capturing harvesting event signals. With pre-processed NDVI and NIR time series, the Harvest Phenology Index (HPI) is first calculated as the ratio of NIR to NDVI (Eq. 2). The NHPI is then obtained by normalizing the HPI within the harvesting window to a 0–1 scale using its local minimum and maximum values (Eq. 3). This normalization can enhance comparability across fields with varying environmental conditions.

$$HPI(t) = \frac{NIR(t)}{NDVI(t)} \quad (2)$$

$$NHPI(t) = \frac{HPI(t) - HPI_{min}(t_{start}, t_{end})}{HPI_{max}(t_{start}, t_{end}) - HPI_{min}(t_{start}, t_{end})} \quad (3)$$

where  $HPI_{min}(t_{start}, t_{end})$  and  $HPI_{max}(t_{start}, t_{end})$  represent the minimum and maximum HPI values observed during the harvesting window  $[t_{start}, t_{end}]$ , respectively. Here,  $t_{start}$  is defined as MOS, and  $t_{end}$  is two months after MOS.

In the final component, a threshold-based method is applied to extract the harvest signal from the NHPI time series. The harvesting date is identified as the first date when the NHPI exceeds the threshold 0.6. This threshold is determined through calibration using field-level ground truth data of our study region, where it yields the lowest error for both corn and soybean (Fig. S2). By amplifying the spectral contrast between senescent vegetation and crop residues, the NHPI-based method effectively characterizes harvesting transitions and enables robust, spatially and temporally generalizable detection of harvesting dates without the need for further calibration.

### 2.2.3 CropPlantHarvest Dataset Validation

We evaluate the final product at both the field and state levels. At the field level, planting and harvesting dates are validated against ground-truth data from Beck’s dataset, which covers the U.S. Midwest from 2016 to 2024. To account for the spatial resolution difference between our 500 m product and Beck’s field-level records, we select MODIS pixels with each falling entirely within the boundary of a crop field, filtering out those spanning multiple fields with PlanetScope imagery verified through visual interpretation. We quantify accuracy using Mean Absolute Error (MAE) and the coefficient of determination ( $R^2$ ) (Eqs. 4 and 5), based on comparisons between estimated and observed planting or harvesting dates for corn and soybean.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

where  $y_i$  is the observed planting or harvesting date of the sample  $i$ ,  $\bar{y}$  is the mean of all the observed planting or harvesting dates, and  $\hat{y}_i$  is the estimated planting or harvesting date of the sample  $i$ .  $n$  denotes the number of samples.  $\sum_{i=1}^n (y_i - \hat{y}_i)^2$  is the sum of squared errors, and  $\sum_{i=1}^n (y_i - \bar{y})^2$  is total sum of squares.

At the state level, we assess the planting and harvesting date products separately by comparing their aggregated temporal distributions with state-level statistics from USDA CPRs. For both planting and harvesting, the estimated dates of all corn or soybean pixels within each state and year are aggregated to generate corresponding cumulative distributions, formatted to match the structure of CPR data. To evaluate consistency, we sample each cumulative curve at 5% intervals between the 20th and 80th percentiles and compute the MAE and  $R^2$  to quantify agreement between the satellite-based estimates and CPR-reported values.

### 2.2.4 Trend and Meteorological Driver Analysis

To investigate long-term patterns in human-determined planting and harvesting dates, we conduct two complementary analyses: (1) a spatiotemporal trend analysis and (2) a meteorological driver analysis. For the spatiotemporal trends, county-level mean planting and harvesting dates are derived annually from the CropPlantHarvest product. Spatial patterns are examined by relating county-level values to latitude of the corresponding county centroids, given that latitude is a key determinant of temperature, which in turn influences agricultural management decisions. Temporal trends are estimated by fitting ordinary least squares (OLS) regressions of planting or harvesting dates against year at the county level. Temporal trends in growing season length are further assessed by calculating the difference between harvesting and planting dates and estimating its change

305 over time using the same OLS. Growing season length provides an integrative indicator of how crop management practices and climatic conditions jointly influence the duration of the cropping cycle. The resulting slopes, expressed in days per decade, are summarized at the state level to highlight regional patterns.

For the meteorological driver analysis, we evaluate the influence of climate conditions on spatiotemporal changes in planting and harvesting dates. County-level meteorological variables, including average of daily minimum temperature and maximum  
310 temperature, and total precipitation, are calculated from Daymet data for the early season (i.e., January to May) and harvesting season (i.e., August to October). These variables are selected because temperature and precipitation are the primary climate factors regulating planting and harvesting decisions. Minimum and maximum temperatures influence soil warming and crop development rates, which determine suitable windows for planting and harvesting, while precipitation affects soil trafficability and field accessibility for machinery operations. We apply county- and year-fixed effects panel regressions to isolate the effects  
315 of these variables, with county-fixed effects controlling for time-invariant county characteristics and year-fixed effects accounting for growing conditions common to all counties in a given year. The panel regression model is specified as follows:

$$Y_{it} = \beta_0 T_{min,it} + \beta_1 T_{max,it} + \beta_2 P_{it} + \alpha_i + \beta_t + \varepsilon_{it} \quad (6)$$

where  $Y_{it}$  denotes the planting or harvesting date for county  $i$  in year  $t$ .  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$  capture the effects of average daily minimum temperature ( $T_{min,it}$ ), maximum temperature ( $T_{max,it}$ ), and total precipitation ( $P_{it}$ ) during the planting or harvesting  
320 season for county  $i$  in year  $t$ , respectively.  $\alpha_i$  represents county-specific fixed effects.  $\beta_t$  denotes year-specific fixed effects.  $\varepsilon_{it}$  is the error term. We conduct separate regressions for corn and soybean, as well as for planting and harvesting timing, with variables expressed in both standardized and unstandardized forms to allow comparison of effect sizes.

### 3 Results

#### 3.1 Dataset Validation

##### 325 3.1.1 Field-level Validation

At the field level, CropPlantHarvest shows a strong agreement with reference planting and harvesting dates, demonstrating its accuracy in capturing the timing of crop management practices. For planting dates, the evaluation yields an  $R^2$  of 0.46 and a MAE of 6.40 days across 151 corn fields, and an  $R^2$  of 0.51 with a MAE of 6.31 days across 36 soybean fields (Fig. 3a). Most predictions fall within  $\pm 10$  days of observed planting dates, indicating reliable detection of planting timing. For harvesting  
330 dates, the results show greater alignment with reference data compared to planting dates, with an  $R^2$  of 0.749 and a MAE of 6.43 days for corn ( $N = 151$ ), and an  $R^2$  of 0.517 with a MAE of 5.39 days for soybean ( $N = 36$ ) (Fig. 3b). This is likely attributable to more distinct spectral transitions at harvesting compared to planting, as well as differences in the detection systems used for the two events. Despite efforts to select MODIS pixels fully contained within individual fields, some of the observed variations may be attributable to spatial mismatches between field-level observations and the 500 m resolution of the  
335 MODIS input data, as well as potential geolocation uncertainties.

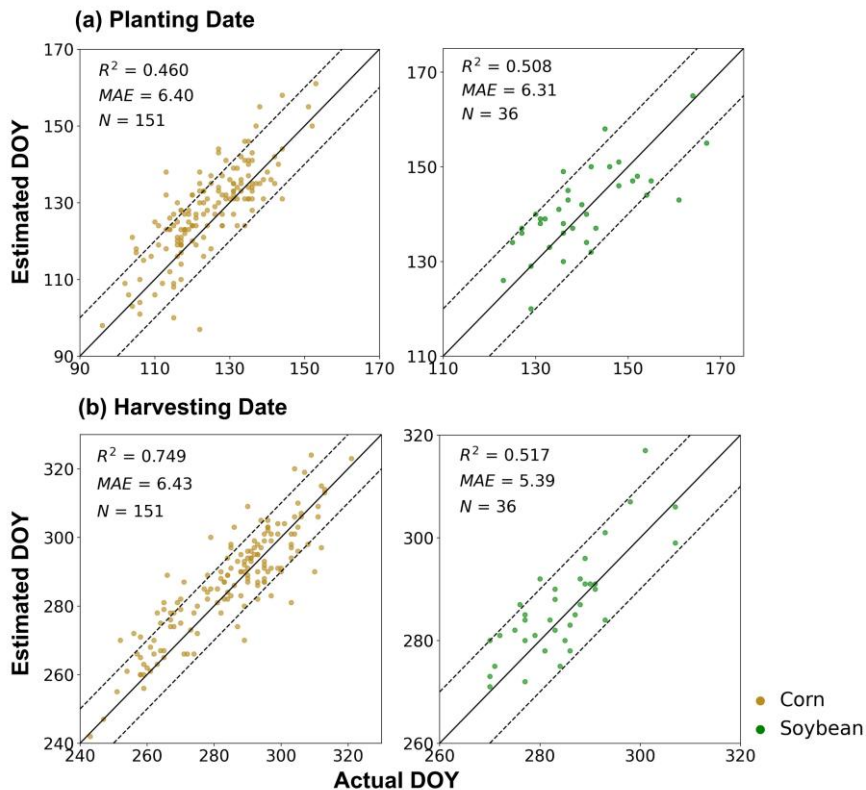
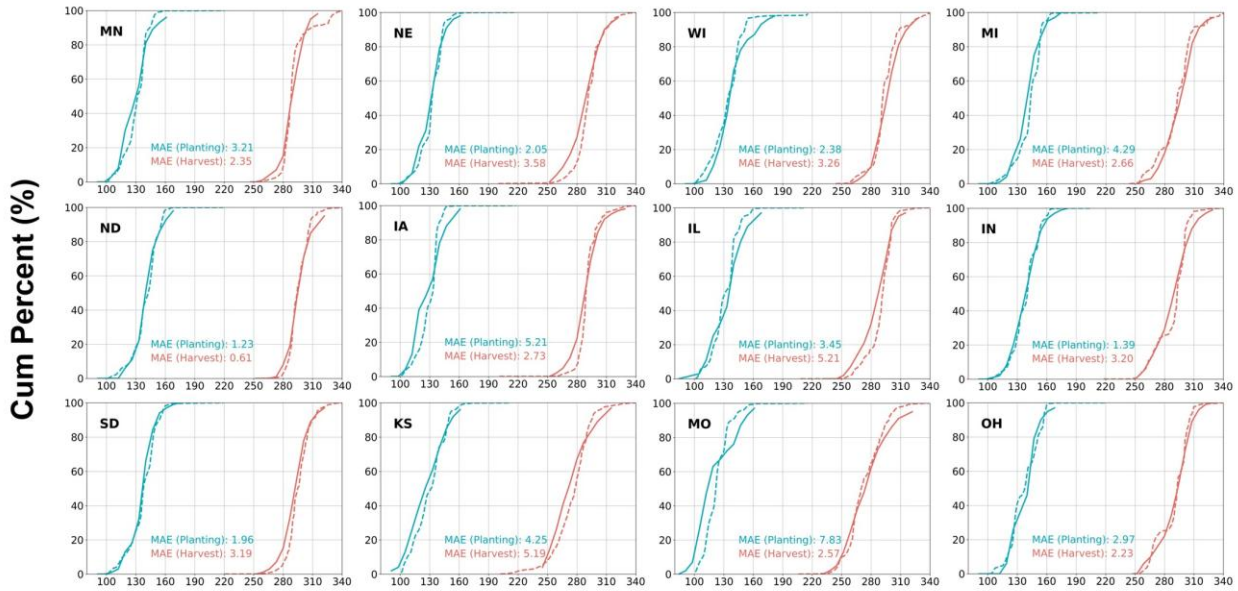


Figure 3. Field-level validation of estimated planting dates (a) and harvesting dates (b) for corn and soybean in the U.S. Midwest from 2016 to 2024. The solid line represents the 1:1 relationship between estimated and observed dates, and the dashed lines indicate  $\pm 10$  days difference between estimated and observed ones.

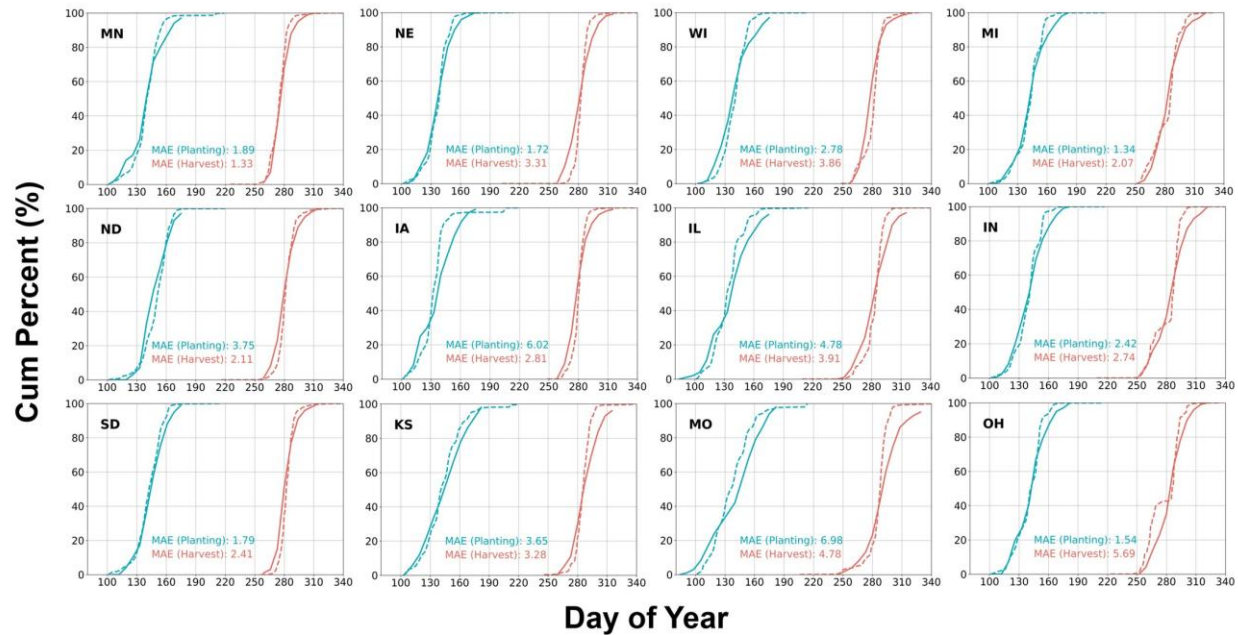
### 340 3.1.2 State-level Validation

Using 2024 as an example, the temporal cumulative distributions of the estimated planting and harvesting dates are compared with those derived from CPR data for corn and soybean across 12 states in the U.S. Midwest (Fig. 4). The comparison shows strong consistency, indicating that CropPlantHarvest effectively reproduces the temporal progression of planting and harvesting for both crops and across all states. The MAEs between the estimated and reported cumulative curves are consistently within one week. For corn, planting date MAEs are less than 4 days in 8 states, and harvesting date MAEs are less than 4 days in 10 states. For soybean, planting date MAEs are below 4 days in 9 states, and harvesting date MAEs are below 4 days in 10 states. Spatial variations in the cumulative distributions across states reflect heterogeneity in planting and harvesting schedules, which are shaped by differences in environmental and climatic conditions, as well as management practices. Additionally, the temporal gap between the planting and harvesting curves exhibits substantial differences across states, reflecting differences in cultivar selection (e.g., maturity length), shaped by local agronomic practices and climate.

### (a) Corn



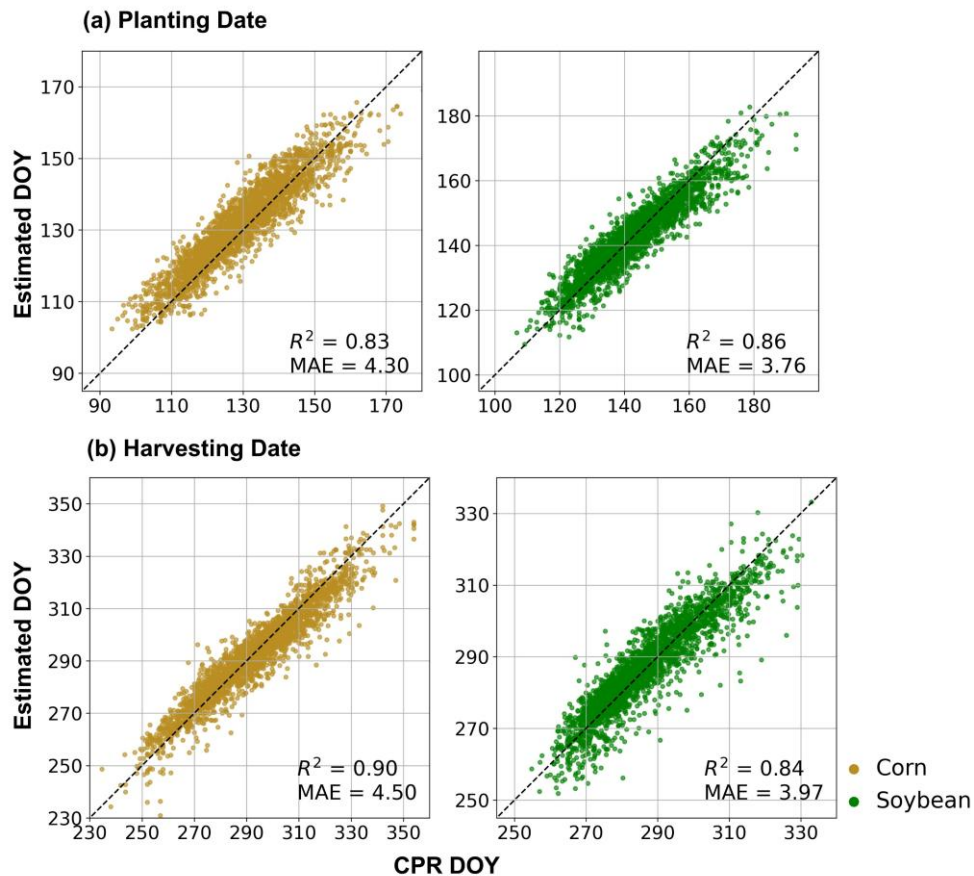
### (b) Soybean



--- Estimated Planting Date    — Ground Truth Planting Date    - - - Estimated Harvest Date    — Ground Truth Harvest Date

Figure 4. Temporal cumulative distributions of estimated planting and harvesting dates compared with corresponding state-level CPR data for corn (a) and soybean (b) in 2024.

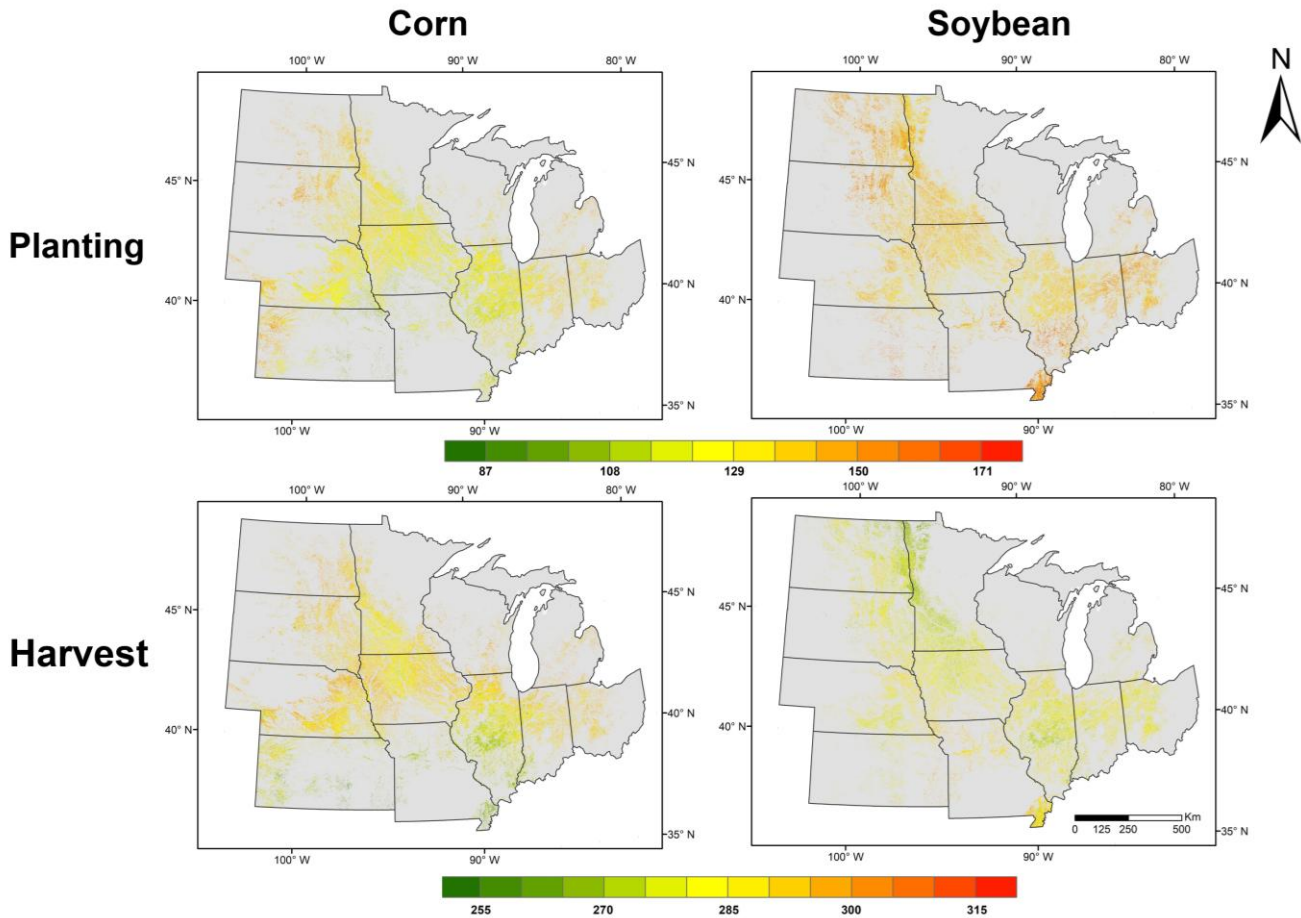
355 Across all states and years, the estimated planting dates show strong agreement with CPR statistics, with an  $R^2$  of 0.83 and MAE of 4.30 days for corn, and an  $R^2$  of 0.86 and MAE of 3.76 days for soybean (Fig. 5a). Estimated harvesting dates also align well with CPR data, achieving an  $R^2$  of 0.90 and MAE of 4.50 days for corn, and an  $R^2$  of 0.84 and MAE of 3.97 days for soybean (Fig. 5b). These results show improved accuracy over previous state-level planting and harvesting date estimates from remotely sensed phenology characterization studies using SOS and EOS phenometrics (Yang et al., 2020; Shen et al., 2022), which report MAEs exceeding one week. The improvement is likely due to the tailored design of the proposed systems to specifically capture human-managed planting and harvesting activities. For planting, the CropSow system detects the emergence of a visible signal in the VI curve and estimates the time lag between planting and that detected signal, overcoming the limitation that planting typically occurs during the flat portion of the VI curve. For harvesting, the NHPI-based system identifies spectral feature points that exhibit stronger and more consistent correspondence with actual harvesting dates than those obtained directly from VI time series, addressing the inconsistent relationships between VI-based metrics and harvesting activity. Overall, the field-level and state-level evaluations confirm the reliability of the CropPlantHarvest product in capturing planting and harvesting dynamics for both corn and soybean across spatial and temporal scales.



370 Figure 5. State-level validation of estimated planting dates (a) and harvesting dates (b) for corn and soybean in the U.S. Midwest from 2001 to 2024. For each state–year, validation points are sampled from the estimated cumulative distribution percentiles (20%–80%) at 5% intervals and compared with corresponding CPR data, yielding one point per sampled percentile for each state–year–crop combination.

### 3.2 Spatial-temporal Trends Analysis

The averaged planting and harvesting date maps from 2001 to 2024 reveal distinct spatial patterns for both corn and soybean fields across the U.S. Midwest (Fig. 6). Corn planting generally begins earlier in southern regions and progressively later at higher latitudes, reflecting temperature-related constraints that influence planting decisions. Soybean planting dates show greater uniformity, likely due to staggered planting relative to corn and weaker temperature dependence. Corn harvesting follows a similar north–south gradient, occurring earlier in southern areas and later in northern regions, largely shaped by its planting pattern. In contrast, soybean harvesting exhibits weaker latitudinal patterns. Interannual variability patterns (Fig. S3) further differ between crops, with corn exhibiting greater variability in harvesting dates than in planting dates, while soybean shows more comparable variability between planting and harvesting. These high-resolution maps also capture substantial within-state variation not reflected in state-level CPRs. For instance, northern Illinois consistently shows later corn harvests than southern Illinois despite similar planting windows. Such fine-scale detail underscores the value of this CropPlantHarvest product for understanding both local and regional cropping dynamics.



385 Figure 6. Spatial patterns of annual averages of planting and harvesting dates during 2001–2024 for corn and soybean.

To characterize spatial variability in corn and soybean planting and harvesting dates across the U.S. Midwest, we further analyze the relationship between planting or harvesting dates and geographic location. As shown in Fig. 7, corn planting dates exhibit a nonlinear increasing trend with latitude across the U.S. Midwest. This pattern is likely related to the influence of

390 temperature on planting decisions. In lower-latitude areas (e.g., ~36–38°N), planting typically occurs between day of year  
 (DOY) 90 and 115, while in higher-latitude regions (above 42°N), planting shifts later, often occurring between DOY 120 and  
 140. Warmer spring conditions in the south accelerate soil warming and enable earlier attainment of thermal thresholds  
 required for seed germination and crop establishment. These favorable conditions provide greater planting flexibility, allowing  
 farmers to begin planting corn earlier in the season. In contrast, cooler spring temperatures at higher latitudes delay planting  
 395 readiness, resulting in progressively later planting dates toward the north. Corn harvesting dates follow a segmented latitudinal  
 pattern, characterized by a steeper increase in lower-latitude areas (south of 41.11°N) and a more gradual increase in higher-  
 latitude regions. Harvesting dates range from approximately DOY 250 in southern counties to around DOY 310 in the north.  
 This trend generally corresponds to the planting pattern, as earlier planting often leads to earlier harvesting. However, minor  
 deviations between the two are likely driven by hybrid selection and management strategies. Variations in corn cultivars  
 influence the timing of physiological maturity, which in turn affects harvesting dates across different latitudes.

400 Soybean planting dates display a more diverse latitudinal pattern compared to corn. In southern regions (below 40.34°N),  
 soybean is often planted later than that in some northern areas, generally between DOY 120 and 160. This delay is likely due  
 to double cropping systems involving winter wheat, where soybean planting is postponed until winter wheat harvest is  
 completed (Borchers et al., 2014). In contrast, in central and northern regions soybean planting dates remain relatively  
 consistent across latitudes, typically within DOY 130–145, reflecting the limited adoption of double cropping and the flexibility  
 405 of later planting relative to corn once low-temperature risks have passed. Soybean harvesting dates exhibit a relatively stable  
 latitudinal pattern, with harvest typically occurring around DOY 290. This pattern may be attributed to cultivar selection  
 adapted to local conditions, enabling crops to reach maturity at a similar time despite differences in planting dates. In addition,  
 soybean development is further regulated by photoperiod besides temperature, which contributes to a more compact and  
 uniform growth cycle across latitudes. These factors together likely contribute to the observed uniformity in soybean harvesting  
 410 timing.

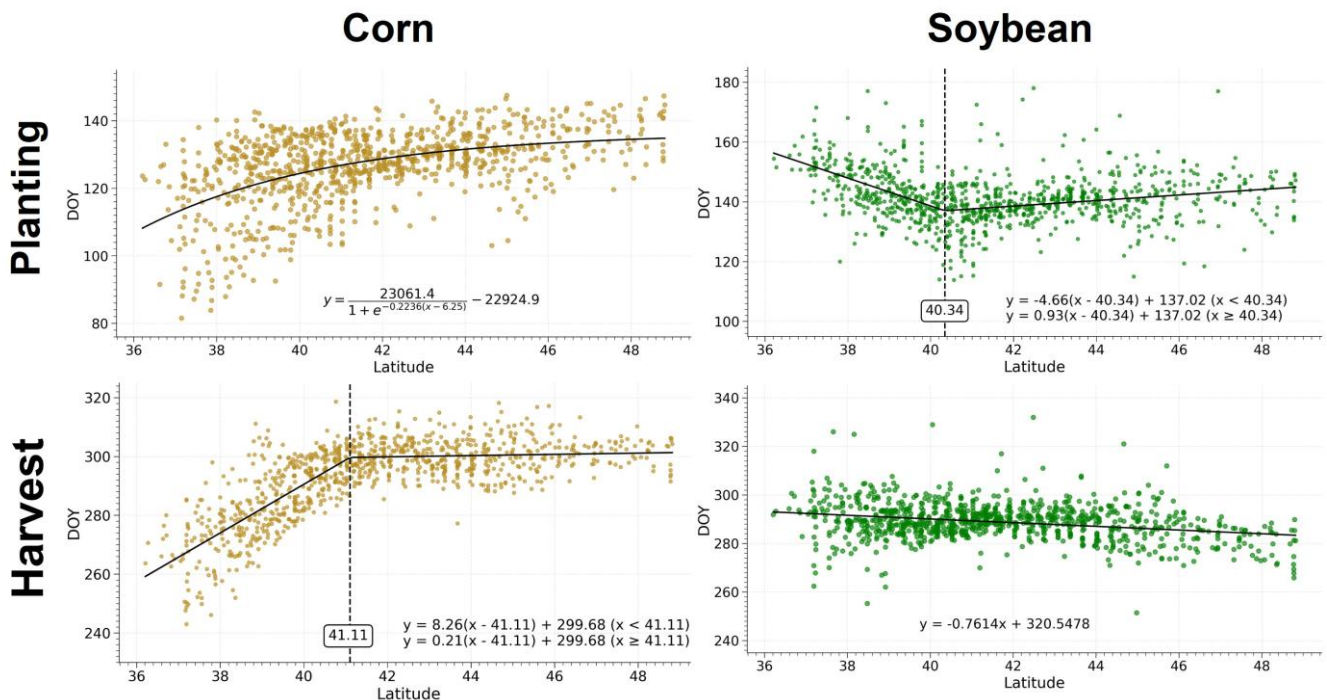


Figure 7. Relationship between mean county-level planting or harvesting dates and latitude in the U.S. Midwest for corn and soybean from 2001 to 2024, with a fitted trend line illustrating the latitude-dependent variation in planting or harvesting timing.

415 Fig. 8 presents trends in planting dates, harvesting dates, and growing season lengths for corn and soybean across 12 U.S.  
Midwestern states from 2001 to 2024. Positive values indicate delays, while negative values represent earlier timing. In general,  
corn planting dates show delayed trends across all states, though the magnitude varies. These delays may be driven by  
increasingly wet spring conditions, greater weather variability, or more cautious planting decisions in response to early-season  
risks. In contrast, soybean planting dates exhibit more heterogeneous patterns. Delays in states such as Michigan and Missouri  
420 may be due to double cropping with winter wheat or shifts in local management practices, while earlier planting trends in states  
like Illinois and Nebraska may result from increased planting flexibility or a shift from strictly sequential planting (soybean  
after corn) to partially overlapping planting windows. Harvesting date trends are generally smaller in magnitude than planting  
trends. Corn harvesting dates have been delayed in several states, including Illinois, Kansas, Missouri, and Ohio, likely due to  
the adoption of longer-season hybrids. Other states show relatively stable harvesting timing. Soybean harvesting dates remain  
425 relatively stable across the region. As a result, the corn growing season appears shortened in most states, suggesting that  
delayed planting has been the dominant driver. In contrast, the soybean growing season remains relatively stable across most  
states, likely because synchronized shifts in planting and harvesting have offset one another. These results suggest that corn  
and soybean systems are both adapting dynamically to changing climatic and agronomic conditions. Additionally, these  
differences reflect site-specific management strategies shaped by human decisions, which are influenced by varying  
430 understandings of local climate, cropping systems, and operational constraints.

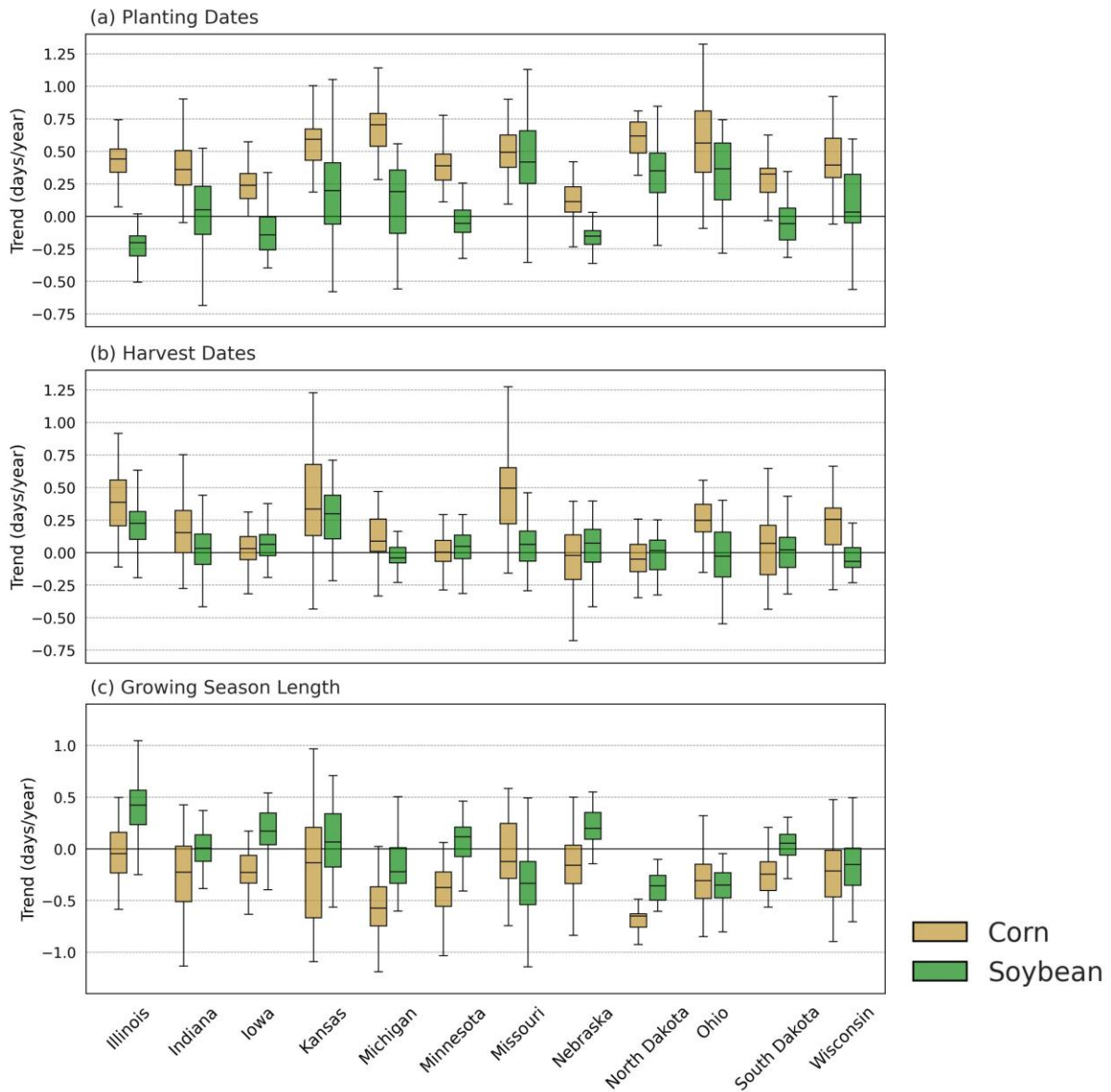


Figure 8. The trends of planting dates, harvesting dates, and growing season lengths during 2001-2024 by crops and states.

### 3.3 Meteorological Drivers of Planting and Harvesting

Regarding planting timing, Table 1 summarizes the regression results from county- and year-fixed effects panel models relating early season (January–May) weather conditions to the timing of planting dates for corn and soybean. For corn, higher mean minimum temperatures during spring are significantly associated with earlier planting dates, with each 1 °C increase in minimum temperature advancing planting by about 0.85 days ( $p < 0.001$ ). Higher total precipitation is significantly associated with delayed planting, with each additional millimeter of spring precipitation delaying planting by about 0.0232 days ( $p < 0.001$ ). In contrast, mean maximum temperature during spring shows no statistically significant association with planting dates ( $p = 0.632$ ). This pattern is likely because, in the U.S. Midwest, corn is planted as early as possible once the frost-free period

begins and soil conditions are sufficiently dry, so minimum temperature and total precipitation serve as the primary climatic constraints. The standardized coefficient for precipitation (3.35) is slightly greater in magnitude than that for minimum temperature (-2.84), suggesting that excess soil moisture may be a more limiting factor than temperature for timely corn planting. For soybean, higher mean minimum and maximum temperatures have comparable effects on planting dates. Each 1 °C increase in minimum temperature advances planting by about 0.38 days, while higher maximum temperatures are associated with changes of similar magnitude. Precipitation is significantly linked to delayed planting, with each additional millimeter in spring delaying planting by about 0.0318 days ( $p < 0.001$ ). The standardized coefficient for precipitation (4.6) is substantially greater than those for temperature (around -1.3), indicating that excess soil moisture is a more important constraint than temperature for soybean planting. This pattern reflects the fact that soybean is typically planted after corn, when the frost-free period has already begun, so minimum temperature exerts less influence on planting decisions and soil moisture becomes the dominant factor. In general, soybean planting, similar to corn, tends to occur earlier in warmer conditions and later in wetter conditions. Additional panel regressions using detrended planting dates are conducted to isolate interannual weather effects from long-term management and technological trends, yielding comparable results (Table S1).

Table 1. Summary of fitted coefficients for early season weather variables in county- and year-fixed effects panel regression models estimating county-level planting day for corn and soybean. The weather variables include mean minimum temperature ( $T_{\min}$ ), mean maximum temperature ( $T_{\max}$ ), and total precipitation (Precip) during January–May. Coefficient type indicates whether coefficients are based on non-standardized or standardized weather variables. Standard errors are shown in parentheses. Statistical significance is indicated using multiple p-value thresholds (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ) to distinguish different levels of statistical significance. Coefficients without symbols are not statistically significant.

Crop type	Corn		Soybean	
	Non-standardized	Standardized	Non-standardized	Standardized
$T_{\min}$	-0.8515 (0.1482) ***	-2.8396 (0.4941) ***	-0.3834 (0.1583) *	-1.3116 (0.5418) *
$T_{\max}$	0.0540 (0.1127)	0.1940 (0.4050)	-0.3483 (0.1303) **	-1.2580 (0.4708) **
Precip	0.0232 (0.0009) ***	3.3492 (0.1264) ***	0.0318 (0.0009) ***	4.6135 (0.1375) ***
Adjusted R <sup>2</sup>	0.721	0.721	0.546	0.546

\* for  $p < 0.05$ , \*\* for  $p < 0.01$ , \*\*\* for  $p < 0.001$

Regarding harvesting timing, Table 2 summarizes the regression results from county- and year-fixed effects panel models relating late-season (August–November) weather conditions to the timing of harvesting dates for corn and soybean. A 1 °C increase in minimum temperature is associated with a delay in harvesting of 2.14 days for corn ( $p < 0.001$ ) and 2.01 days for soybean ( $p < 0.001$ ). This is because, in the U.S. Midwest, higher minimum temperatures often indicate more humid conditions (Kunkel et al., 2013). Factors such as soil moisture, evapotranspiration from crops, and the insulating effect of humid air contribute to warmer nights, and this elevated humidity slows the dry-down process, ultimately delaying harvesting decisions. In contrast, a 1 °C increase in maximum temperature is associated with an advance in harvesting of 4.00 days for corn ( $p < 0.001$ ) and 3.35 days for soybean ( $p < 0.001$ ). This is likely due to faster crop dry-down under warmer daytime conditions. Each additional millimeter of precipitation during the harvest window is associated with a delay in harvesting of 0.01–0.02 days ( $p < 0.001$ ) for both crops, as wet weather can impede field access and machine operation. Based on standardized coefficients, maximum temperature has a much stronger influence on harvest timing than minimum temperature or total precipitation. Overall, these results show the asymmetric influence of diurnal temperature patterns and seasonal rainfall on planting and harvesting timing, with distinct crop-specific responses across the two management phases. Additional panel regressions using detrended harvesting dates are conducted to isolate interannual weather effects from long-term management and technological trends, yielding comparable results (Table S2).

475 Table 2. Summary of fitted coefficients for harvesting season weather variables in county- and year-fixed effects panel  
 regression models estimating county-level harvesting day for corn and soybean. The weather variables include mean minimum  
 temperature ( $T_{\min}$ ), mean maximum temperature ( $T_{\max}$ ), and total precipitation (Precip) during August–October. Coefficient  
 type indicates whether coefficients are based on non-standardized or standardized weather variables. Standard errors are shown  
 in parentheses. Statistical significance is indicated using multiple p-value thresholds (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ )  
 480 to distinguish different levels of statistical significance. Coefficients without symbols are not statistically significant.

Crop type	Corn		Soybean	
	Non-standardized	Standardized	Non-standardized	Standardized
$T_{\min}$	2.1408 (0.1998) ***	5.0634 (0.4726) ***	2.0141 (0.1652) ***	4.7819 (0.3922) ***
$T_{\max}$	-3.9971 (0.1803) ***	-11.2444 (0.5072) ***	-3.3498 (0.1717) ***	-9.4188 (0.4828) ***
Precip	0.0160 (0.0014) ***	1.4946 (0.1354) ***	0.0143 (0.0012) ***	1.3148 (0.1119) ***
Adjusted $R^2$	0.711	0.711	0.536	0.536

\* for  $p < 0.05$ , \*\* for  $p < 0.01$ , \*\*\* for  $p < 0.001$

#### 4 Discussion

We introduce CropPlantHarvest, the first long-term, 500 m spatial resolution dataset of planting and harvesting dates for corn  
 and soybean across the U.S. Midwest, covering 2001–2024. By pinpointing the onset and end of vegetation activity, it delivers  
 temporally precise crop growth season boundaries that are essential for understanding management adaptations to climate  
 485 change, improving yield estimation, and guiding greenhouse gas accounting, ultimately enabling more accurate modeling and  
 sustainable agricultural management.

CropPlantHarvest offers two key advantages over existing planting and harvesting datasets. First, it provides detailed spatial  
 and temporal varying estimates of planting and harvesting dates for corn and soybean at 500 m spatial resolution and daily  
 temporal resolution (e.g., DOY 102). Developed using daily MODIS observations, it detects subtle vegetation changes that  
 490 mark the start and end of the growing season, enabling precise identification of these management-driven events. While global  
 crop calendar products (e.g., SAGE, MIRCA2000, RiceAtlas, PSHW, GCPE) have been invaluable for mapping broad-scale  
 crop seasonality, their coarse spatial (e.g.,  $0.5^\circ$  grid or regional) and temporal (e.g., monthly or half-monthly) resolutions limit  
 their ability to capture interannual variability or localized management practices. CropPlantHarvest overcomes these  
 constraints with fine-scale, annually varying estimates that support detailed monitoring and analysis of spatial and temporal  
 495 heterogeneity in planting and harvesting behaviors. Second, CropPlantHarvest achieves high accuracy, with a MAE of around  
 4 days at the state level when validated against independent CPRs. This performance surpasses previous phenology-based  
 datasets derived from SOS- or EOS-based phenometrics (Yang et al., 2020; Shen et al., 2022) and phenology matching methods  
 (Sakamoto, 2018), all of which report MAEs exceeding 10 days. The improved accuracy of CropPlantHarvest stems from its  
 integration of two tailored systems: CropSow, which estimates planting dates by assimilating satellite-derived phenological  
 500 signals with crop growth trajectories simulated by a process-based crop growth model, and the NHPI-based method, which  
 identifies harvesting dates by capturing the distinct spectral transition from senescent vegetation to exposed crop residues using  
 a combination of NDVI and NIR reflectance. By integrating additional modeling and spectral signatures beyond VI time series  
 alone, the CropSow system and the NHPI-based method more directly link satellite signals to operational events in the field,  
 addressing a long-standing challenge in remote sensing–based estimation of planting and harvesting dates, where  
 505 inconsistencies between VI trajectories and actual management actions often lead to substantial errors.

The CropPlantHarvest dataset supports diverse downstream applications. It enables assessments of long-term trends in farmer  
 management decisions and offers valuable insights into their adaptation to climate change. When combined with  
 meteorological data, it reveals that for corn, higher minimum temperatures during the early season tend to advance planting  
 dates, whereas increased precipitation during the same period delays them. Additional analyses using weather variables

510 restricted to the pre-planting period yield comparable patterns and further strengthen the causal interpretation of temperature and precipitation effects on planting timing (Table S3 and S4). During the harvesting season, higher maximum temperatures generally lead to earlier harvests by accelerating grain dry-down, while higher minimum temperatures are linked to later harvests due to elevated humidity, which slows dry-down. These findings help explain the shift from the continued advancement of planting dates in the latter half of the 20th century to delayed planting around the early 21st century, linked to wetter springs, as well as the persistent delay in harvesting associated with the “warming hole” in the U.S. Midwest—  
515 characterized by rising minimum temperatures but relatively stable maximum temperatures (Sacks and Kucharik, 2011; Basso et al., 2021). Further analysis could examine how these climate–management relationships vary across crop cultivars and regions, and assess their impacts on yield to inform policy interventions aimed at enhancing food security and system resilience. In addition, as critical inputs for process-based and remote sensing–based models, high-resolution planting and harvesting  
520 dates enable more accurate crop yield estimation, carbon accounting, and monitoring of other management practices. For crop yield estimation, they improve calibration by aligning simulated phenological stages with observed management events, thereby reducing prediction uncertainty (Zhang and Diao, 2023; Luo et al., 2023; Pei et al., 2025). For carbon accounting, they define the active growing and off-season periods, enabling earth system models to more accurately estimate biomass accumulation and seasonal carbon exchange, thereby improving assessments of carbon dynamics in Agric. Syst. (Delgrosso et al., 2005; Zhang et al., 2020). For crop management monitoring, they constrain remote sensing analyses to specific time  
525 windows, thereby improving the accuracy of detecting practices such as irrigation scheduling (Xie and Lark, 2021), cover-crop timing (Zhou et al., 2022; Wang et al., 2023), and tillage detection (Liu et al., 2022b; Zhang et al., 2024; Ma et al., 2026). This temporal detail provides a robust basis for yield forecasting, carbon accounting, and management monitoring of the U.S. agriculture, and establishes a foundation for extending applications to other regions.

530 Despite its spatial and temporal specificity and high accuracy in capturing human-managed planting and harvesting dates, CropPlantHarvest has several limitations. In order to provide long-term coverage back to 2001, the dataset generation framework relies on MODIS imagery at 500 m resolution. To mitigate errors associated with sub-pixel heterogeneity at this scale, a strict pixel-purity criterion ( $\geq 90\%$ ) is applied. This conservative design improves the reliability of the final dataset but reduces spatial coverage in highly fragmented agricultural regions (Fig. S1). Future work could leverage spatiotemporal image  
535 fusion (Yang et al., 2021; Lyu et al., 2025), integrating fine-spatial, coarse-temporal-resolution data (e.g., Landsat) with fine-temporal, coarse-spatial-resolution data (e.g., MODIS), to generate long-term satellite observations with both fine spatial detail and frequent temporal coverage, thereby enabling more precise estimation of planting and harvesting dates. This approach would preserve the historical record while adding detailed spatial information, thereby improving the resolution and reliability of detected cropland management practices. Future work could also leverage multispectral and hyperspectral foundation  
540 models trained on large-scale Earth observation datasets to learn transferable spatiotemporal representations of crop development, thereby improving the characterization of crop phenology and the retrieval of planting and harvesting dates from multi-source satellite observations (Yang et al., 2024; Li et al., 2026). In addition, estimating planting dates with CropSow requires the calibration of the thermal time threshold in the crop growth model, which is used to simulate the duration between planting and the remotely sensed Greenup phenometric. The accuracy of this calibration can be influenced by the quality and  
545 representativeness of the calibration data. Biases may occur when regional calibration data (i.e., CPR) are derived from limited or spatially imbalanced samples, which may not adequately represent regional planting conditions. Future work that calibrates CropSow using denser, field-level planting observations would help reduce these biases and further improve estimation accuracy.

## 5 Conclusion

550 In this study, we present CropPlantHarvest, the first annual, 500 m resolution dataset of corn and soybean planting and harvesting dates across the U.S. Midwest from 2001 to 2024. CropPlantHarvest provides essential data for climate adaptation

analysis, crop yield estimation, and greenhouse gas accounting, filling the long-standing gap in consistent, fine-resolution datasets for large-scale monitoring of crop management practices. It leverages an integrative remotely sensed crop modeling system (i.e., CropSow) to estimate planting dates and a novel spectral index (i.e., NHPI) to detect harvesting dates with MODIS imagery, enabling the characterization of fine-scale spatiotemporal dynamics of these management practices. Validation against Beck's field-level observations and CPR state-level statistics demonstrate high accuracy, with mean absolute errors generally around 5 days for both crop species. The high spatial and temporal resolution of CropPlantHarvest enables tracking of shifts in planting and harvesting strategies over two decades, offering new opportunities to assess the impacts of climate variability, management adaptation, and policy on agricultural systems. By providing high-resolution planting and harvesting dates, CropPlantHarvest serves as a benchmark for refining remote-sensing phenology products and evaluating the agro-environmental impacts of evolving crop management decisions. It also enables more accurate yield and carbon-cycle modeling and improved monitoring of management practices (e.g., tillage, cover cropping, and irrigation) by clearly delineating growing and off-season periods, thereby supporting food system resilience across U.S. agriculture and beyond.

### **Data Availability**

Our CropPlantHarvest dataset, which provides planting and harvesting dates for corn and soybean fields at 500 m spatial resolution across the U.S. Midwest from 2001 to 2024, can be accessed via Zenodo: <https://doi.org/10.5281/zenodo.16967482> (Liu and Diao, 2025).

### **Author Contributions**

YL and CD conceptualized the study and developed the methodology. YL conducted data curation and formal analysis, implemented the software, and generated the visualizations. CD secured funding, managed the project, provided supervision and resources, and contributed to writing – review and editing. YL prepared the original draft with contributions and revisions from all co-authors.

### **Competing Interests**

The authors declared that none of the authors has any competing interests.

### **Acknowledgements**

This study is supported partly by the National Science Foundation (2048068 and 2518299) and partly by the National Aeronautics and Space Administration (80NSSC21K0946). It is also supported by the Taylor Geospatial Institute GISCoR (Geospatial Institute Seed Grant Program to stimulate Collaborative Research) (113356).

### **Reference**

- Basso, B., Martinez-Feria, R. A., Rill, L., and Ritchie, J. T.: Contrasting long-term temperature trends reveal minor changes in projected potential evapotranspiration in the US Midwest, *Nat Commun*, 12, 1476, <https://doi.org/10.1038/s41467-021-21763-7>, 2021.
- Baum, M. E., Licht, M. A., Huber, I., and Archontoulis, S. V.: Impacts of climate change on the optimum planting date of different maize cultivars in the central US Corn Belt, *Eur. J. Agron.*, 119, 126101, <https://doi.org/10.1016/j.eja.2020.126101>, 2020.
- Beck, P. S. A., Atzberger, C., Høgda, K. A., Johansen, B., and Skidmore, A. K.: Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI, *Remote Sens. Environ.*, 100, 321–334, <https://doi.org/10.1016/j.rse.2005.10.021>, 2006.

- 590 Beddington, J.: Food security: contributions from science to a new and greener revolution, *Phil. Trans. R. Soc. B*, 365, 61–71, <https://doi.org/10.1098/rstb.2009.0201>, 2010.
- Birch, C. J., Vos, J., Kiniry, J., Bos, H. J., and Elings, A.: Phyllochron responds to acclimation to temperature and irradiance in maize, *Field Crops Res.*, 59, 187–200, [https://doi.org/10.1016/S0378-4290\(98\)00120-8](https://doi.org/10.1016/S0378-4290(98)00120-8), 1998.
- 595 Bolton, D. K. and Friedl, M. A.: Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics, *Agric. For. Meteorol.*, 173, 74–84, <https://doi.org/10.1016/j.agrformet.2013.01.007>, 2013.
- Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., and Friedl, M. A.: Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery, *Remote Sens. Environ.*, 240, 111685, <https://doi.org/10.1016/j.rse.2020.111685>, 2020.
- 600 Borchers, A., Truex-Powell, E., Wallander, S., and Nickerson, C.: Multi-cropping practices: Recent trends in double-cropping, U.S. Department of Agriculture, Economic Research Service, <https://doi.org/10.22004/AG.ECON.262122>, 2014.
- Boryan, C., Yang, Z., Mueller, R., and Craig, M.: Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program, Geocarto Int., 26, 341–358, <https://doi.org/10.1080/10106049.2011.562309>, 2011.
- 605 Butler, E. E., Mueller, N. D., and Huybers, P.: Peculiarly pleasant weather for US maize, *Proceedings of the National Academy of Sciences*, 115, 11935–11940, <https://doi.org/10.1073/pnas.1808035115>, 2018.
- Cao, R., Li, L., Liu, L., Liang, H., Zhu, X., Shen, M., Zhou, J., Li, Y., and Chen, J.: A spatiotemporal shape model fitting method for within-season crop phenology detection, *ISPRS J. Photogramm. Remote Sens.*, 217, 179–198, <https://doi.org/10.1016/j.isprsjprs.2024.08.009>, 2024.
- 610 Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., and Chhetri, N.: A meta-analysis of crop yield under climate change and adaptation, *Nature Clim Change*, 4, 287–291, <https://doi.org/10.1038/nclimate2153>, 2014.
- Delgrosso, S., Mosier, A., Parton, W., and Ojima, D.: DAYCENT model analysis of past and contemporary soil NO and net greenhouse gas flux for major crops in the USA, *Soil Tillage Res.*, 83, 9–24, <https://doi.org/10.1016/j.still.2005.02.007>, 2005.
- Diao, C.: Remote sensing phenological monitoring framework to characterize corn and soybean physiological growing stages, *Remote Sens. Environ.*, 248, 111960, <https://doi.org/10.1016/j.rse.2020.111960>, 2020.
- 615 Diao, C. and Li, G.: Near-Surface and High-Resolution Satellite Time Series for Detecting Crop Phenology, *Remote Sensing*, 14, 1957, <https://doi.org/10.3390/rs14091957>, 2022.
- Diao, C., Yang, Z., Gao, F., Zhang, X., and Yang, Z.: Hybrid phenology matching model for robust crop phenological retrieval, *ISPRS J. Photogramm. Remote Sens.*, 181, 308–326, <https://doi.org/10.1016/j.isprsjprs.2021.09.011>, 2021.
- 620 Gao, F. and Zhang, X.: Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities, *J Remote Sens*, 2021, 2021/8379391, <https://doi.org/10.34133/2021/8379391>, 2021.
- Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., Mueller, R., Johnson, D. M., and Prueger, J. H.: Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery, *Remote Sens. Environ.*, 188, 9–25, <https://doi.org/10.1016/j.rse.2016.11.004>, 2017.

- 625 Guo, T., Diao, C., Yang, Z., Liu, Y., and Zhang, C.: Towards scalable crop yield estimation through integration of crop model and deep learning, in: AGU fall meeting abstracts, tex.adsnote: Provided by the SAO/NASA Astrophysics Data System, GC41M-0104, 2024.
- Huang, J., Gómez-Dans, J. L., Huang, H., Ma, H., Wu, Q., Lewis, P. E., Liang, S., Chen, Z., Xue, J.-H., Wu, Y., Zhao, F., Wang, J., and Xie, X.: Assimilation of remote sensing into crop growth models: Current status and perspectives, *Agric. For. Meteorol.*, 276–277, 107609, <https://doi.org/10.1016/j.agrformet.2019.06.008>, 2019.
- 630 Iizumi, T., Kim, W., and Nishimori, M.: Modeling the Global Sowing and Harvesting Windows of Major Crops Around the Year 2000, *J. Adv. Model. Earth Syst.*, 11, 99–112, <https://doi.org/10.1029/2018MS001477>, 2019.
- Jain, M., Srivastava, A., Balwinder-Singh, Joon, R., McDonald, A., Royal, K., Lisaius, M., and Lobell, D.: Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data, *Remote Sensing*, 8, 860, <https://doi.org/10.3390/rs8100860>, 2016.
- 635 Kucharik, C. J. and Ramankutty, N.: Trends and Variability in U.S. Corn Yields Over the Twentieth Century, *Earth Interact.*, 9, 1–29, <https://doi.org/10.1175/EI098.1>, 2005.
- Kunkel, K. E., Stevens, L. E., Stevens, S. E., Sun, L., Janssen, E., Wuebbles, D., Hilberg, S. D., Timlin, M. S., Stoecker, L., Westcott, N. E., and Dobson, J. G.: Part 3. Climate of the Midwest U.S., 2013.
- Kusumastuti, R. D., Donk, D. P. van, and Teunter, R.: Crop-related harvesting and processing planning: a review, *Int. J. Prod. Econ.*, 174, 76–92, <https://doi.org/10.1016/j.ijpe.2016.01.010>, 2016.
- 640 Laborte, A. G., Gutierrez, M. A., Balanza, J. G., Saito, K., Zwart, S. J., Boschetti, M., Murty, M. V. R., Villano, L., Aunario, J. K., Reinke, R., Koo, J., Hijmans, R. J., and Nelson, A.: RiceAtlas, a spatial database of global rice calendars and production, *Sci Data*, 4, 170074, <https://doi.org/10.1038/sdata.2017.74>, 2017.
- Li, W., Liang, S., Chen, K., Chen, Y., Ma, H., Xu, J., Ma, Y., Zhang, Y., Guan, S., Fang, H., and Shi, Z.: AgriFM: A multi-source temporal remote sensing foundation model for Agriculture mapping, *Remote Sens. Environ.*, 334, 115234, <https://doi.org/10.1016/j.rse.2026.115234>, 2026.
- Liu, L., Cao, R., Chen, J., Shen, M., Wang, S., Zhou, J., and He, B.: Detecting crop phenology from vegetation index time-series data by improved shape model fitting in each phenological stage, *Remote Sens. Environ.*, 277, 113060, <https://doi.org/10.1016/j.rse.2022.113060>, 2022a.
- 650 Liu, Y. and Diao, C.: CropPlantHarvest: A 500 m annual dataset of crop planting and harvesting dates (2001–2024) of the U.S. Midwest, <https://doi.org/10.5281/zenodo.16967482>, 2025.
- Liu, Y., Rao, P., Zhou, W., Singh, B., Srivastava, A. K., Poonia, S. P., Van Berkel, D., and Jain, M.: Using Sentinel-1, Sentinel-2, and Planet satellite data to map field-level tillage practices in smallholder systems, *PLoS ONE*, 17, e0277425, <https://doi.org/10.1371/journal.pone.0277425>, 2022b.
- 655 Liu, Y., Diao, C., and Yang, Z.: CropSow: An integrative remotely sensed crop modeling framework for field-level crop planting date estimation, *ISPRS J. Photogramm. Remote Sens.*, 202, 334–355, <https://doi.org/10.1016/j.isprsjprs.2023.06.012>, 2023.
- Liu, Y., Diao, C., Yang, Z., Mei, W., and Guo, T.: A novel normalized harvest phenology index (NHPI) for corn and soybean harvesting date detection using landsat and sentinel-2 imagery on google earth engine, *Remote Sens. Environ.*, 331, 115016, <https://doi.org/10.1016/j.rse.2025.115016>, 2025.
- 660

- Luo, L., Sun, S., Xue, J., Gao, Z., Zhao, J., Yin, Y., Gao, F., and Luan, X.: Crop yield estimation based on assimilation of crop models and remote sensing data: A systematic evaluation, *Agric. Syst.*, 210, 103711, <https://doi.org/10.1016/j.agsy.2023.103711>, 2023.
- 665 Luo, Y., Zhang, Z., Chen, Y., Li, Z., and Tao, F.: ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products, *Earth Syst. Sci. Data*, 12, 197–214, <https://doi.org/10.5194/essd-12-197-2020>, 2020.
- Lyu, F., Yang, Z., Diao, C., and Wang, S.: Multistream STGAN: A Spatiotemporal Image Fusion Model With Improved Temporal Transferability, *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, 18, 1562–1576, <https://doi.org/10.1109/JSTARS.2024.3506879>, 2025.
- 670 Ma, Y., Shen, Y., Swatantran, A., Kelly, C., and Lobell, D. B.: STaPL: Scale Transfer with Pseudo-Labeling for satellite-based mapping of agricultural practices, *Remote Sens. Environ.*, 343, 115500, <https://doi.org/10.1016/j.rse.2026.115500>, 2026.
- Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., Cogliati, S., Manca, G., Diotri, F., Busetto, L., Cescatti, A., Colombo, R., Fava, F., Morra di Cella, U., Pari, E., Siniscalco, C., and Richardson, A. D.: Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO<sub>2</sub> uptake, *Agric. For. Meteorol.*, 151, 1325–1337, <https://doi.org/10.1016/j.agrformet.2011.05.012>, 2011.
- 675 Mori, A., Doi, Y., and Iizumi, T.: GCPE: The global dataset of crop phenological events for agricultural and earth system modeling, *J. Agric. Meteorol.*, 79, 120–129, <https://doi.org/10.2480/agrmet.D-23-00004>, 2023.
- Moulin, S., Kergoat, L., Viovy, N., and Dedieu, G.: Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements, *J. Climate*, 10, 1154–1170, [https://doi.org/10.1175/1520-0442\(1997\)010%3C1154:GSAOVP%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010%3C1154:GSAOVP%3E2.0.CO;2), 1997.
- 680 Nieto, L., Schwalbert, R., Prasad, P. V. V., Olson, B. J. S. C., and Ciampitti, I. A.: An integrated approach of field, weather, and satellite data for monitoring maize phenology, *Sci Rep*, 11, 15711, <https://doi.org/10.1038/s41598-021-95253-7>, 2021.
- Niu, Q., Li, X., Huang, J., Huang, H., Huang, X., Su, W., and Yuan, W.: A 30 m annual maize phenology dataset from 1985 to 2020 in China, *Earth Syst. Sci. Data*, 14, 2851–2864, <https://doi.org/10.5194/essd-14-2851-2022>, 2022.
- 685 Papale, D., Reichstein, M., Canfora, E., Aubinet, M., Bernhofer, C., Longdoz, B., Kutsch, W., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a more harmonized processing of eddy covariance CO<sub>2</sub> fluxes: algorithms and uncertainty estimation., <https://doi.org/10.5194/bgd-3-961-2006>, 2006.
- 690 Pei, J., Tan, S., Zou, Y., Liao, C., He, Y., Wang, J., Huang, H., Wang, T., Tian, H., Fang, H., Wang, L., and Huang, J.: The role of phenology in crop yield prediction: Comparison of ground-based phenology and remotely sensed phenology, *Agric. For. Meteorol.*, 361, 110340, <https://doi.org/10.1016/j.agrformet.2024.110340>, 2025.
- Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, *Glob. Biogeochem. Cycles*, 24, <https://doi.org/10.1029/2008GB003435>, 2010.
- 695 Sacks, W. J. and Kucharik, C. J.: Crop management and phenology trends in the U.S. Corn Belt: Impacts on yields, evapotranspiration and energy balance, *Agric. For. Meteorol.*, 151, 882–894, <https://doi.org/10.1016/j.agrformet.2011.02.010>, 2011.

- Sacks, W. J., Deryng, D., Foley, J. A., and Ramankutty, N.: Crop planting dates: an analysis of global patterns, *Glob. Ecol. Biogeogr.*, 19, 607–620, <https://doi.org/10.1111/j.1466-8238.2010.00551.x>, 2010.
- 700 dos Santos, C. L., Abendroth, L. J., Coulter, J. A., Nafziger, E. D., Suyker, A., Yu, J., Schnable, P. S., and Archontoulis, S. V.: Maize Leaf Appearance Rates: A Synthesis From the United States Corn Belt, *Front. Plant Sci.*, 13, 2022.
- Schwartz, M. D., Reed, B. C., and White, M. A.: Assessing satellite-derived start-of-season measures in the conterminous USA, *Int. J. Climatol.*, 22, 1793–1805, <https://doi.org/10.1002/joc.819>, 2002.
- 705 Shen, Y., Zhang, X., and Yang, Z.: Mapping corn and soybean phenometrics at field scales over the United States Corn Belt by fusing time series of Landsat 8 and Sentinel-2 data with VIIRS data, *ISPRS J. Photogramm. Remote Sens.*, 186, 55–69, <https://doi.org/10.1016/j.isprsjprs.2022.01.023>, 2022.
- Shen, Y., Zhang, X., Yang, Z., Ye, Y., Wang, J., Gao, S., Liu, Y., Wang, W., Tran, K. H., and Ju, J.: Developing an operational algorithm for near-real-time monitoring of crop progress at field scales by fusing harmonized Landsat and Sentinel-2 time series with geostationary satellite observations, *Remote Sens. Environ.*, 296, 113729, <https://doi.org/10.1016/j.rse.2023.113729>, 2023.
- 710 Shew, A. M., Tack, J. B., Nalley, L. L., and Chaminuka, P.: Yield reduction under climate warming varies among wheat cultivars in South Africa, *Nat Commun*, 11, 4408, <https://doi.org/10.1038/s41467-020-18317-8>, 2020.
- Thornton, M. M., Shrestha, R., Wei, Y., Thornton, P. E., Kao, S., and Wilson, B. E.: Daymet: Daily surface weather data on a 1-km grid for north america, version 4, , <https://doi.org/10.3334/ORNLDAAAC/1840>, 2020.
- 715 Wang, X., Blesh, J., Rao, P., Paliwal, A., Umashaanker, M., and Jain, M.: Mapping cover crop species in southeastern Michigan using Sentinel-2 satellite data and Google Earth Engine, *Front. Artif. Intell.*, 6, 1035502, <https://doi.org/10.3389/frai.2023.1035502>, 2023.
- Wu, C., Peng, D., Soudani, K., Siebicke, L., Gough, C. M., Arain, M. A., Bohrer, G., Lafleur, P. M., Peichl, M., Gonsamo, A., Xu, S., Fang, B., and Ge, Q.: Land surface phenology derived from normalized difference vegetation index (NDVI) at global FLUXNET sites, *Agric. For. Meteorol.*, 233, 171–182, <https://doi.org/10.1016/j.agrformet.2016.11.193>, 2017.
- 720 Xie, Y. and Lark, T. J.: Mapping annual irrigation from Landsat imagery and environmental variables across the conterminous United States, *Remote Sens. Environ.*, 260, 112445, <https://doi.org/10.1016/j.rse.2021.112445>, 2021.
- Xu, J., Han, H., Ning, T., Li, Z., and Lal, R.: Long-term effects of tillage and straw management on soil organic carbon, crop yield, and yield stability in a wheat-maize system, *Field Crops Res.*, 233, 33–40, <https://doi.org/10.1016/j.fcr.2018.12.016>, 2019.
- 725 Yang, Y., Ren, W., Tao, B., Ji, L., Liang, L., Ruane, A. C., Fisher, J. B., Liu, J., Sama, M., Li, Z., and Tian, Q.: Characterizing spatiotemporal patterns of crop phenology across North America during 2000–2016 using satellite imagery and agricultural survey data, *ISPRS J. Photogramm. Remote Sens.*, 170, 156–173, <https://doi.org/10.1016/j.isprsjprs.2020.10.005>, 2020.
- Yang, Z., Diao, C., and Li, B.: A Robust Hybrid Deep Learning Model for Spatiotemporal Image Fusion, *Remote Sensing*, 13, 5005, <https://doi.org/10.3390/rs13245005>, 2021.
- 730 Yang, Z., Diao, C., Zhao, Y., Liu, Y., and Gao, F.: Near real-time crop phenology characterization at field scale through weakly supervised deep learning, in: AGU fall meeting abstracts, citation-key: 2024AGUFMB52C...05Ytex.adsnote: Provided by the SAO/NASA Astrophysics Data System, B52C-05, 2024.

- Yang, Z., Diao, C., and Liu, Y.: Chapter 5 - Monitoring crop phenology with remote sensing, in: *Agricultural applications of earth observation*, edited by: Fu, P., Elsevier, 117–143, <https://doi.org/10.1016/B978-0-443-38345-8.00002-8>, 2026.
- 735 Zhang, C. and Diao, C.: A Phenology-guided Bayesian-CNN (PB-CNN) framework for soybean yield estimation and uncertainty analysis, *ISPRS J. Photogramm. Remote Sens.*, 205, 50–73, <https://doi.org/10.1016/j.isprsjprs.2023.09.025>, 2023.
- Zhang, W., Yu, Q., Tang, H., Liu, J., and Wu, W.: Conservation tillage mapping and monitoring using remote sensing, *Comput. Electron. Agric.*, 218, 108705, <https://doi.org/10.1016/j.compag.2024.108705>, 2024.
- 740 Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., Reed, B. C., and Huete, A.: Monitoring vegetation phenology using MODIS, *Remote Sens. Environ.*, 84, 471–475, [https://doi.org/10.1016/S0034-4257\(02\)00135-9](https://doi.org/10.1016/S0034-4257(02)00135-9), 2003.
- Zhang, Y., Marx, E., Williams, S., Gurung, R., Ogle, S., Horton, R., Bader, D., and Paustian, K.: Adaptation in U.S. Corn Belt increases resistance to soil carbon loss with climate change, *Sci Rep*, 10, 13799, <https://doi.org/10.1038/s41598-020-70819-z>, 2020.
- 745 Zhou, Q., Guan, K., Wang, S., Jiang, C., Huang, Y., Peng, B., Chen, Z., Wang, S., Hipple, J., Schaefer, D., Qin, Z., Stroebel, S., Coppess, J., Khanna, M., and Cai, Y.: Recent Rapid Increase of Cover Crop Adoption Across the U.S. Midwest Detected by Fusing Multi-Source Satellite Data, *Geophys. Res. Lett.*, 49, e2022GL100249, <https://doi.org/10.1029/2022GL100249>, 2022.