ISPRS Journal of Photogrammetry and Remote Sensing 216 (2024) 259–273

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

From satellite-based phenological metrics to crop planting dates: Deriving field-level planting dates for corn and soybean in the U.S. Midwest

Qu Zhou [a,b], Kaiyu Guan [a,b,c,d,*], Sheng Wang [a,b,*], James Hipple [e], Zhangliang Chen [a,b]

[a] Agroecosystem Sustainability Center, Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
[b] Department of Nature Resources and Environmental Sciences, College of Agricultural, Consumer and Environmental Sciences, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
[c] National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
[d] Department of Computer Science, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
[e] United States Department of Agriculture, Risk Management Agency, Washington, DC 20250, USA

ARTICLE INFO

Keywords:
Planting Dates
Phenology
Required Growing Degree Days
Corn and Soybean
U.S. Midwest
Harmonized Landsat and Sentinel-2

ABSTRACT

Information on planting dates is crucial for modeling crop development, analyzing crop yield, and evaluating the effectiveness of policy-driven planting windows. Despite their high importance, field-level planting date datasets are scarce. Satellite remote sensing provides accurate and cost-effective solutions for detecting crop phenology from moderate to high resolutions, but remote sensing-based crop planting date detection is rare. Here, we aimed to generate field-level crop planting date maps by taking advantage of satellite remote sensing-derived phenological metrics and proposed a two-step framework to predict crop planting dates from these metrics using required growing degree dates (RGDD) as a bridge. Specifically, we modeled RGDD from the planting date to the spring inflection date (derived from phenological metrics) and then predicted the crop planting dates based on phenological metrics, RGDD, and environmental variables. The ~3-day and 30-m Harmonized Landsat and Sentinel-2 (HLS) products were used to derive crop phenological metrics for corn and soybean fields in the U.S. Midwest from 2016 to 2021, and the ground truth of field-level planting dates from USDA Risk Management Agency (RMA) reports were used for the development and validation of our proposed two-step framework. The results indicated that our framework could accurately predict field-level planting dates from HLS-derived phenological metrics, capturing 77 % field-level variations for corn (mean absolute error, MAE=4.6 days) and 71 % for soybean (MAE=5.4 days). We also evaluated the predicted planting dates with USDA National Agricultural Statistics Service (NASS) state-level crop progress reports, achieving strong consistency with median planting dates for corn (R^2=0.90, MAE=2.7 days) and soybeans (R^2=0.87, MAE=2.5 days). The model's performance degraded slightly when predicting planting dates for fields with irrigation (MAE=5.4 days for corn, MAE=6.1 days for soybean) and cover cropping (MAE=5.4 days for corn, MAE=5.6 days for soybean). The USDA RMA Common Crop Insurance Policy (CCIP) provides county- or sub-county-level crop planting windows, which drive producers' decisions on when to plant. Within the CCIP-driven planting windows, higher prediction accuracies were achieved (MAE for corn: 4.5 days, soybean: 5.2 days). Our proposed two-step framework (phenological metrics-RGDD-planting dates) also outperformed the traditional one-step model (phenological metrics-planting dates). The proposed framework can be beneficial for deriving planting dates from current and future phenological products and contribute to studies related to planting dates such as the analysis of yield gaps, management practices, and government policies.

1. Introduction

Planting dates play an important role in crop development and growth, affecting crop yields, reflecting farmers' management practices under climate change, and evaluating the effectiveness of the government's policies on crop planting (Boyer et al., 2023; Cassman and Grassini, 2020; Khan et al., 2017b; Urban et al., 2018). Timely planting is crucial for root growth initiation and vegetative development (Khan

* Corresponding authors at: Agroecosystem Sustainability Center, Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA.
E-mail addresses: kaiyug@illinois.edu (K. Guan), shengwang12@gmail.com (S. Wang).

https://doi.org/10.1016/j.isprsjprs.2024.07.031
Received 16 February 2024; Received in revised form 23 July 2024; Accepted 31 July 2024
Available online 16 August 2024
0924-2716/© 2024 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

et al., 2017a) and to ensure favorable climate conditions during the critical growth stages such as flowering (Sacks et al., 2010). Delayed planting may lead to reduced photosynthesis and decreased duration of plant growth (Hu and Wiatrak, 2012), while early planting tends to have challenges in the seedling establishment because of low temperatures (Khan et al., 2017a). Inappropriate planting date is one of the major reasons for reduced crop yields (Bussmann et al., 2016). Choosing optimal planting dates can maximize beneficial conditions and minimize detrimental conditions during crop growth, ultimately improving crop yield potential (Hu and Wiatrak, 2012; Urban et al., 2018; Zhang et al., 2021). Optimal planting dates are typically within a "planting window" and farmers always face a decision of when to plant (Bussmann et al., 2016). Planting dates also reflect how farmers respond to weather variability, and shifting planting dates is an important adaptation strategy for farmers to maintain crop yield (Waha et al., 2013). Government policies related to planting dates should be able to guide farmers to plant crops within reasonable time windows to avoid yield deductions.

Accurate records of crop planting dates are crucial for assessing the agronomic and environmental impact of management activities on agricultural systems. Long-term planting date datasets are valuable for understanding farmers' planting decisions (Deines et al., 2023; Kucharik, 2006; Sacks et al., 2010). Adjusting crop growing seasons can reduce the impacts of weather variability on crop growth (Miller et al., 2021) and mitigate pest pressure (Pulakkatu-Thodi et al., 2014), but may also affect regional carbon cycling, water use, soil erosion, and surface energy balance (Deines et al., 2023; Ren et al., 2024). To maximize the ecological and economic benefits, optimal planting windows should be appropriately designed (Sacks et al., 2010). Determining yield-maximizing planting dates is economically important (Boyer et al., 2015; Egli and Cornelius, 2009; Hu and Wiatrak, 2012) and has been studied for decades (Boyer et al., 2023). Multiple factors may affect planting dates, such as economic factors, which may include the availability of agricultural machinery, farmer credit, and risk tolerance (Borchers et al., 2014; Johansen et al., 2012); and management factors, which may include irrigation, cover cropping, and tillage (Acharya et al., 2017; Teasdale and Mirsky, 2015), and planting intensity and crop varieties (Kucharik, 2006; Zhang et al., 2021). Each year the USDA Risk Management Agency (RMA) Common Crop Insurance Policy (CCIP) Special Provisions publish the earliest and final planting dates at the county or sub-county level for crop insurance in the U.S., serving as a guideline for farmers to choose economically advantageous planting dates (Schnitkey, 2013). Since employing crops with different planting dates is beneficial for sustainable agroecosystems (Cassman and Grassini, 2020; Isbell et al., 2017; Nicholls and Altieri, 2013), it is also important to determine optimal planting dates for maximizing both ecological and economic benefits. Thus, planting date records are crucial for policymakers to understand farmers' planting practices and design more flexible and beneficial planting windows for sustainable agriculture (Miller et al., 2021).

Despite their importance, planting date datasets are often limited to coarse resolutions or regional scales, while they are desired at fine resolutions or field level. When released to the public, planting dates are typically aggregated into large administrative regions (Urban et al., 2018), such as the state-level crop progress reports produced by the USDA National Agricultural Statistics Service (NASS). Due to the high spatial variability of soil properties, economic factors, and cropping conditions (Bussmann et al., 2016; Feola et al., 2015; Zhang et al., 2021), planting dates of adjacent fields can differ by as much as one month (Zhang et al., 2021). Thus, coarse-resolution planting date datasets could not meet the needs of many applications that rely on fine-resolution planting date information (Urban et al., 2018). Satellite remote sensing provides an opportunity to derive field-level planting date maps cost-effectively. However, few previous studies have been focused on mapping field-level planting dates at large scales (Deines et al., 2023), because most studies are focused on deriving crop phenological metrics (Gao et al., 2017; Piao et al., 2019; Zeng et al., 2020). Long-term crop phenology datasets are publicly available from moderate to fine resolutions at regional or global scales (Ganguly et al., 2010; Moon et al., 2022; Niu et al., 2022; Zhang et al., 2018a). Deriving field-level planting dates from crop phenological metrics will be beneficial for generating planting date maps with high accuracies and large spatiotemporal coverages.

Here, we primarily focused on developing a framework to accurately convert satellite remote sensing derived crop phenological metrics into field-level planting dates across a large region. Crop phenological metrics can be characterized from the satellite detected seasonality of crop growth (Zhang et al., 2018b), which can be further combined with the concept of crop growing degree-days (GDD) to determine crop status after planting (Delpierre et al., 2009; Lobell et al., 2011; Piao et al., 2019). GDD is a useful agriculture-climate indicator to predict crop growth stages, and required GDDs (RGDD) from planting dates to emergence, spring inflection date, or maturity are widely used to determine crop growth (Akyuz et al., 2017; Pathak and Stoddard, 2018; Sacks and Kucharik, 2011). RGDD should be dynamically modeled since it is affected by climate conditions (Pathak and Stoddard, 2018), soil properties (Trachsel et al., 2011), and growing regions (Sacks et al., 2010). Thus, we proposed a two-step framework (phenological metrics-RGDD-planting dates) to predict field-level planting dates from satellite-derived crop phenological metrics using RGDD as a bridge. Specifically, we used 30-m and ~3-day NASA's Harmonized Landsat and Sentinel-2 (HLS) products to derive field-level crop phenological metrics for corn and soybean in the U.S. Midwest from 2016 to 2021, and we used field-level planting dates from USDA RMA to develop models for RGDD and planting date predictions. An independent holdout of 20% USDA RMA field-level planting date reports and state-level USDA NASS crop progress reports were used for performance validation. This study can advance our capability to detect field-level planting dates across large regions for sustainable agriculture management.

2. Material

2.1. Study region

Our study region includes 12 states in the U.S. Midwest, including Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin (Fig. 1a). The corn-soybean system in this region is highly productive, accounting for over 85% of total U.S. (Deines et al., 2023) and over 30% of global corn and soybean production (USDA, 2023). The soil types in this region are mainly Alfisols, Mollisols, and Entisols (Clark et al., 2019; Potash et al., 2022), and the climate conditions can be characterized as semi-arid in the western part and warm-summer humid in the central and eastern parts (Peel et al., 2007). Since the 1980s, the U.S. Midwest has experienced increased daily temperature, increased precipitation from mid-spring to early summer, and decreased precipitation from mid-summer to early fall (Dai et al., 2016). Planting dates are generally earlier in the warmer states and later in the cooler states in the U.S. Midwest (Deines et al., 2023) with significant shifts over the past decades (Deines et al., 2023; Kucharik, 2006; Sacks and Kucharik, 2011).

2.2. Data

Four major categories of datasets were used in this study including ground truth data, satellite data, environmental data, and auxiliary data. The ground truth data included insurance unit-level RMA planting dates from 2016 to 2020 and state-level NASS crop progress reports from 2016 to 2021. The field polygons were derived from the crop units and covered the 12 states in the U.S. Midwest (Tables S1 and S2). These polygons were further compared with the NASS Cropland Data Layer (CDL) (Boryan et al., 2011). The crop type with the most CDL pixels in the field polygons was considered the CDL crop type. Only polygons

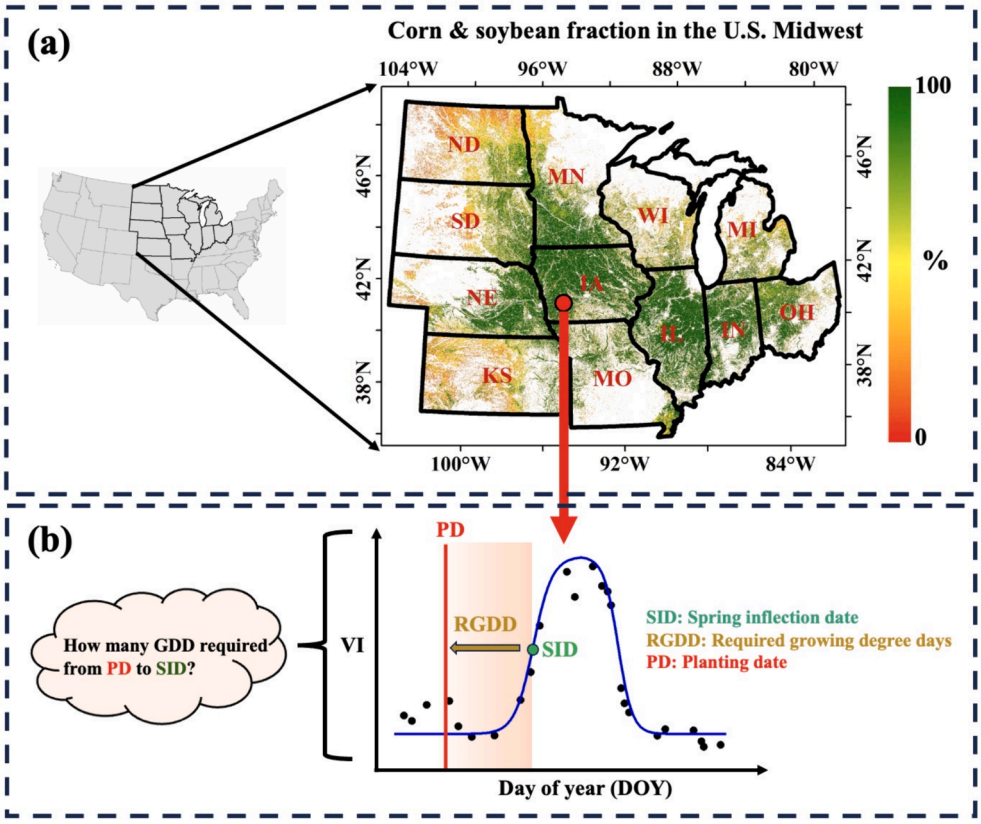

Fig. 1. (a) Study region and corn-soybean fractions in the U.S. Midwest derived from USDA National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL) from 2008 to 2022 and (b) the conceptual framework for predicting field-level planting dates from crop phenology using required growing degree days (RGDD) from the planting date (PD) to the spring inflection date (SID, derived from satellite vegetation index (VI) time series) as a bridge.

with matched CDL and RMA crop types were kept for model development and validation. State-level crop progress reports from the USDA NASS Quickstats (https://quickstats.nass.usda.gov) were used for additional independent validation. Since crop progress reports only provide week-ending statistics, the mid-week date was used for validation, calculated by subtracting 3.5 days from the week-ending date.

The satellite dataset was NASA's HLS observations (version 2) from 2016 to 2021 (Claverie et al., 2018), which are 30-m and ~3-day seamless products from both the Operational Land Imager (OLI) onboard Landsat-8 and Multi-Spectral Instrument (MSI) onboard Sentinel-2. The high-quality HLS time series after applying the quality assurance (QA) layer was used to derive field-level phenological metrics for corn and soybean fields in the U.S. Midwest. The QA layer in the HLS product was generated by the Fmask algorithm (Zhu et al., 2015), and pixels with a QA flag indicating cloud, cloud shadow, snow, and water were disregarded.

The environmental data included 10-m gSSURGO (Gridded Soil Survey Geographic Database) soil data in 2020 from the USDA-NRCS and 2.5-arc PRISM climatic data (Daly et al., 2015) from 2016 to 2021. The environmental datasets were used to develop models for RGDD and planting date predictions.

The auxiliary data included CDL and field boundary layer. The 30-m annual CDL data were used to identify corn and soybean pixels and to remove field polygons that CDL and RMA crop types were inconsistent. Our internal field boundary layer based on the common land unit (CLU) was refined by each year's CDL and was used to aggregate pixels into fields.

The HLS, gSSURGO, PRISM, and CDL data were matched at the field level based on the field boundary layer. Specifically, for fine-resolution data (e.g. HLS, gSSURGO, and CDL), multiple pixels within a field were averaged (for continuous numeric data such as HLS reflectance and gSSURGO clay content) or assigned the majority value (for categorical data such as CDL crop type and gSSURGO soil texture). For coarse-resolution data such as the 2.5-arc PRISM data, most fields were smaller than one PRISM pixel, and we used the pixel centered at the field to extract climatic information.

3. Methodology

To derive field-level planting dates from satellite-based phenological metrics, we proposed a two-step framework (phenological metrics-RGDD-planting date) with the following three major steps: (1) derive field-level crop phenological metrics from HLS vegetation index time series; (2) model RGDD using phenological metrics and environmental variables; and (3) predict field-level planting dates based on RGDD, phenological metrics, and environmental variables. A flowchart that summarizes these three steps is shown in Fig. 2.

3.1. Deriving phenological metrics

High-quality vegetation index time series is crucial to characterize crop growth and normalized difference vegetation index (NDVI) has been proven successful in detecting crop phenology (Beck et al., 2006; Liu et al., 2022; Zhang et al., 2003). Although the cloud mask (QA layer) was applied to the satellite products, the remaining residual noise caused by dust, cloud shadow, and snow in the time series may affect the usage of the NDVI time series. A revised algorithm from Chen et al., (2004) based on the Savitzky-Golay filter was used to remove outliers and reconstruct a high-quality NDVI time series, which included modifications for potential cover cropping and was more suitable for corn and soybean fields in the U.S. Midwest (Zhou et al., 2022).

The classic double logistic function was used to determine pheno-

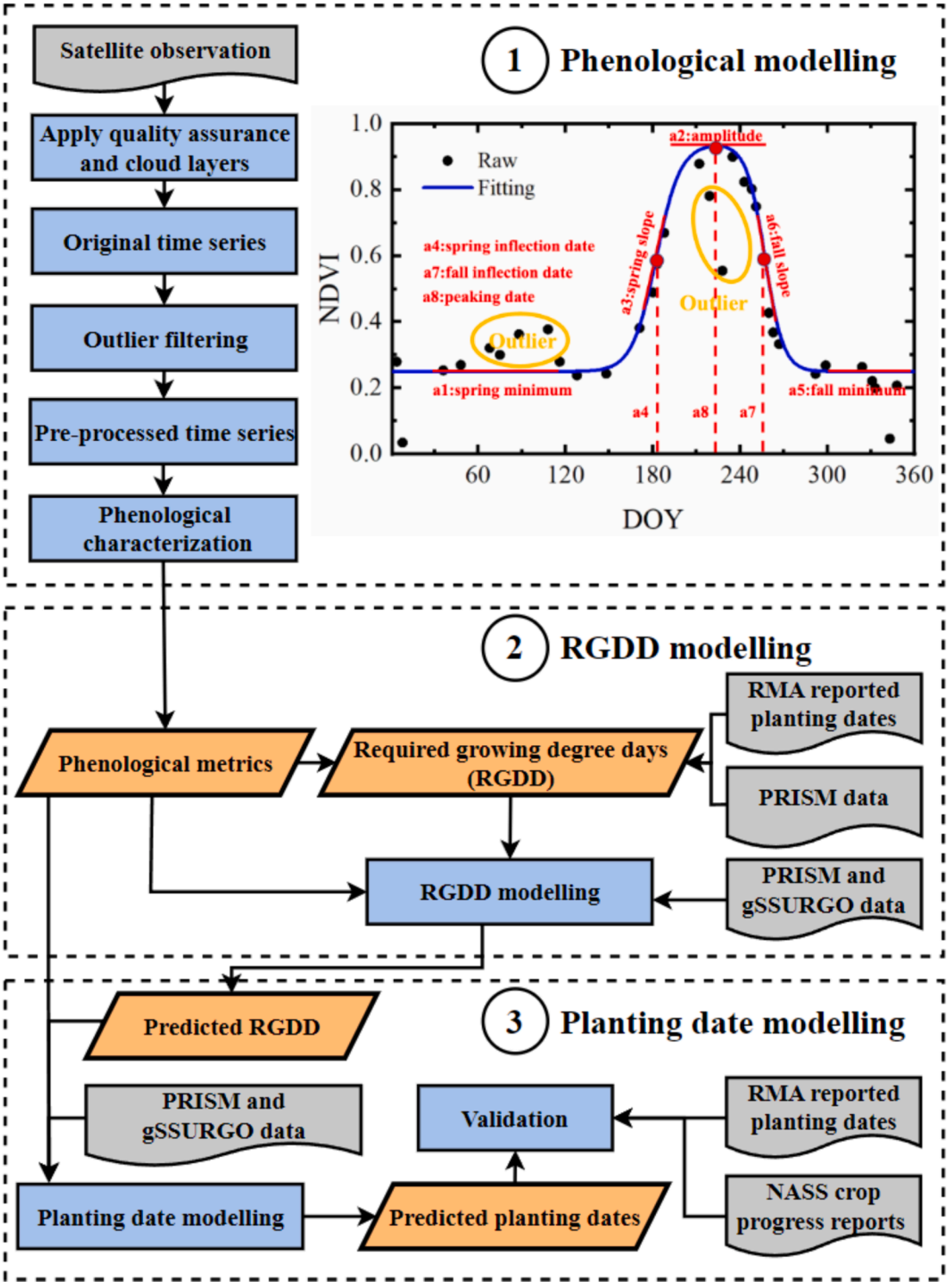

Fig. 2. Flowchart of deriving field-level planting dates from satellite remote sensing using a two-step framework. The framework includes three major modules: (1) deriving phenological metrics (parameters a1 to a8) from Harmonized Landsat and Sentinel-2 (HLS) normalized difference vegetation index (NDVI) time series, (2) modeling required growing degree days (RGDD) from planting date to spring inflection date, and (3) predicting field-level planting dates based on phenological metrics, RGDD, and environmental variables.

logical stages (Beck et al., 2006; Melaas et al., 2013; Zhang et al., 2003). For a given location, the time series ($G(t)$) can be modeled by two logistic functions for the spring ($S(t)$) and fall ($F(t)$) time series (Babcock et al., 2021). Thus, $G(t)$ can be written as Eq. (1).

$$G(t) = \begin{cases} S(t) = a_1 + \frac{a_2}{1 + e^{-a_3(t-a_4)}} 1 \le t \le a_8 \\ F(t) = a_5 + \frac{a_2}{1 + e^{-a_6(t-a_7)}} a_8 < t \le 365 \end{cases} \quad (1)$$

where a_1 is spring minimum greenness, a_2 is seasonal maximum greenness, a_3 controls spring green-up rate, a_4 is spring inflection date, a_5 is fall minimum greenness, a_6 controls fall green-up rate, a_7 is fall inflection date, and a_8 is the date when $S(t) = F(t)$. Since cover cropping may delay planting dates of main cash crops (Deines et al., 2023; Osipitan et al., 2019) as shown in Fig. S1, a Gaussian function (Eq. (2)) was included to model the growth of potential cover crops.

$$C(t) = a_9 exp\left(-\frac{(t - a_{10})^2}{2{a_{11}}^2} \right) 1 < t \le a_8 \quad (2)$$

where a_9 is the peak greenness of cover crops, a_{10} is the date of peaking cover crop greenness, and a_{11} is the standard deviation of cover crop greenness. To avoid the impacts of weeds, we assumed that $a_1 + a_9$ should be larger than 0.4 for cover cropping fields (Wang et al., 2023; Zhou et al., 2022). Thus, the NDVI time series with potential cover cropping can be written as Eq. (3).

$$G(t) = \begin{cases} S(t) + C(t) 1 \le t \le a_8 \\ F(t) a_8 < t \le 365 \end{cases} \quad (3)$$

where $C(t)$ equals zero when $a_1 + a_9 < 0.4$. The phenological metrics and their descriptions are summarized in Table S3.

3.2. Modeling RGDD

After obtaining phenological metrics, we derived RGDD from the planting date to the spring inflection date using Eq. (4). The planting date was derived from field-level data, the spring inflection date was derived from satellite-based phenological metrics, and GDD was derived from the PRISM climatic data. Since the field-level reports did not have direct measures of RGDD, the derived RGDD data in the unit of Celsius degree (°C) were used as "pseudo" ground truth for training and validating the RGDD model.

$$\mathrm{RGDD} = \sum_{i=\text{planting date}}^{\text{spring inflection date}} \mathrm{GDD_i} \quad (4)$$

where $\mathrm{GDD_i}$ is daily GDD, which can be calculated as $\mathrm{GDD} = (T_{max} + T_{min})/2\text{-}T_{base}$ (Liu et al., 2013). T_{max}, T_{min}, and T_{base} are the daily maximum temperature, daily minimum temperature, and base temperature, respectively. The growth of corn and soybean will stop when temperatures exceed a certain threshold temperature ($T_{threshold}$). Thus, two modifications in the calculation of GDD were applied: (1) any temperature below T_{base} is set to T_{base} and (2) any temperature above $T_{threshold}$ is cut off to $T_{threshold}$ (Curtis et al., 2023). The T_{base} is usually set as 10 ℃ for corn and soybean (MRCC, 2024), and the standard $T_{threshold}$ for corn is 30℃ and can also be used for soybean (NDAWN, 2024a, 2024b). The spring inflection date was selected to calculate RGDD for two reasons: (1) it is derived from the spring time series, which is relatively closer to the planting date than phenological metrics from the fall time series, and (2) it corresponds to the maximum change rate of the spring time series, which can reduce the uncertainty in determining the spring inflection date. Assuming a fitting error of $\Delta NDVI$ in the phenology curve, the uncertainty in determining a date in the phenology curve is $\frac{\Delta NDVI}{changerate}$. For example, a change of 0.01 in NDVI might result in a difference of several weeks in the early and peak NDVI time series but only several days near the spring inflection date (Fig. 2).

Table 1
Variables and their data source for model development of required growing degree days (RGDD) and planting dates in the U.S. Midwest from 2016 to 2021.

Category	Variable	Source	Resolution
Phenology	a_1 to a_8 [1]	HLS-derived	30 *m*
	Daily maximum and minimum temperature[2]		
	Monthly temperature from April to June	PRISM	2.5 arc
	Monthly precipitation from April to June		
Environment	Monthly VPD[3] from April to June		
	Clay concentration		
	Sand concentration		
	Silt concentration	gSSURGO	10 *m*
	Soil organic carbon concentration		
	Soil texture[4]		

[1] Parameters of double logistic functions.
[2] Daily maximum and minimum temperature: used for calculating growing degree days (GDD) and not used for the development of the RGDD model.
[3] VPD: vapor pressure deficit.
[4] Soil texture: derived based on the contents of clay, sand, and silt.

The model of RGDD can be written as:

$$\mathrm{RGDD} = \mathrm{Function}(\text{phenological metrics}, \text{environmental variables}) \quad (5)$$

The model assumed that RGDD is dynamically changed across space and time, which might be affected by crop growing status (phenological metrics) and environmental conditions (soil and climatic parameters). The environmental factors included climatic variables such as temperature, precipitation, and vapor pressure deficit (VPD), and soil variables such as clay, sand, silt, soil organic carbon (SOC) concentration, and soil texture. These variables and their data sources are listed in Table 1. The function in Eq. (5) was the eXtreme Gradient Boosting (XGBoost) machine learning algorithm (Chen and Guestrin, 2016). The 80% of "pseudo" field-level RGDD data were used for training and the remaining 20% were used for validation.

3.3. Predicting planting dates

The modeled RGDD was included in the prediction of planting dates based on our two-step model, which can be written as:

$$\text{Planting date} = \mathrm{Function}(\mathrm{RGDD}, \text{phenological metrics}, \text{environmental variables}) \quad (6)$$

In addition to the two-step planting date model (Eq. (6)), we also built two other planting date models for comparison: (1) a one-step model (Eq. (7)) that directly predicts planting dates from phenological metrics and environmental variables, which can help assess the benefits of using RGDD in our two-step model; (2) a model (Eq. (8)) that only utilizes phenological metrics to predict planting dates, which can allow us to compare the performance of our derived phenological metrics with other phenological metrics in predicting planting dates.

$$\text{Planting date} = \mathrm{Function}(\text{phenological metrics}, \text{environmental variables}) \quad (7)$$

$$\text{Planting date} = \text{Function}(\text{phenological metrics}) \quad (8)$$

The inputs for Eqs. (6) and (8) were listed in Table 1 and the RGDD in Eq. (6) was predicted using Eq. (5). The functions in Eqs. (6–8) were the XGBoost algorithm (Chen and Guestrin, 2016). The XGBoost provides many hyperparameters and we empirically set several parameters for both RGDD and planting date models: number of gradient-boosted trees (n_estimators=140), maximum tree depth for base learners (max_depth=10), learning rate (learning_rate=0.05), random number seed (random_state=42). Other hyperparameters were set as default. The training and validation of planting dates can refer to sections 3.4 and 3.5.

3.4. Spatiotemporal cross-validation

To assess the spatiotemporal transferability of our proposed model, we did spatiotemporal cross-validation by training and testing on different spatiotemporal subsets of the total reference data (Filippelli et al., 2024). We built the following three kinds of models: (1) trained and tested on the full dataset (full validation); (2) trained and tested on the leave-one-year-out dataset (temporal validation); and (3) trained and tested on the leave-20%-region-out dataset (spatial validation). Specifically, for full validation, about 80% of field-level planting date data were used for training and the remaining 20% were used for validation; for temporal validation, in the 80% of field-level data, one-year data were dropped for training and the dropped-year data in the remaining 20% were used for validation; for spatial validation, in the 80% of field-level data, 20%-county data were dropped for training and the dropped-county data in the remaining 20% were used for validation.

3.5. Accuracy assessment

To comprehensively assess our predicted planting dates, we conducted field-, county-, and state-level validation against multi-sourced planting date datasets. For the field-level validation, we used the holdout 20% field-level planting date data for independent validation. For the county-level validation, we compared our planting dates against the publicly available dataset from Deines et al. (2023). The dates that 25%, 50%, and 75% of acres planted within a county were compared. For the state-level validation, we compared our predicted planting dates with state-level NASS crop progress reports. The median predicted planting dates within a state were validated against the NASS-reported dates that 50% of acres in the state were planted.

Three widely used statistical indicators were used to evaluate the model's performances including the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). These indicators can be calculated as Eq. (9).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(P_i - \widehat{P}_i)^2}{\sum_{i=1}^{n}(P_i - \overline{P})^2} MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - R_i| RMSE = \sqrt{\sum_{i=1}^{n}\frac{(P_i - R_i)^2}{n}} \quad (9)$$

where P_i is the predicted value, R_i is the referenced value, $\overline{P}$ is the mean value of P_i, $\widehat{P}_i$ is estimated value from linear regression of P_i and R_i, and n

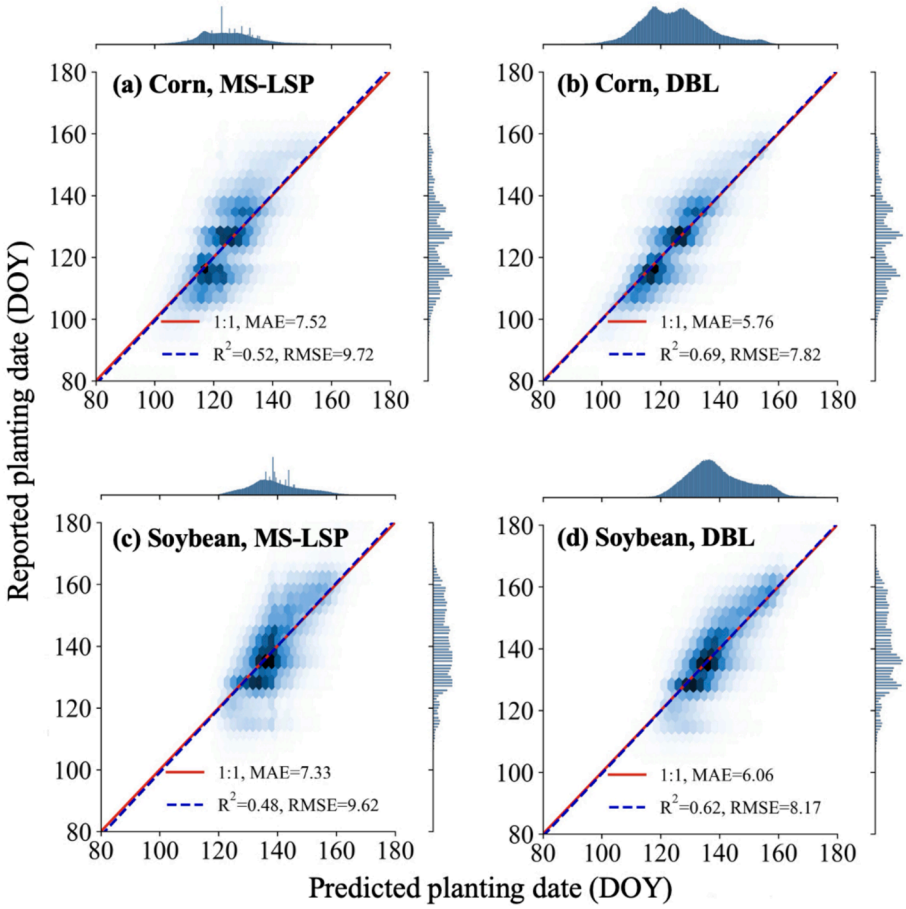

Fig. 3. Validation of predicted planting dates against independent holdout of 20% field-level planting dates in the U.S. Midwest from 2016 to 2019 based on Eq. (8). The model only uses phenological metrics to predict planting dates: (a) for corn using Multi-Source Land Surface Phenology (MS-LSP) metrics; (b) for corn using our double logistic (DBL) function-based phenology metrics; (c) for soybean using MS-LSP metrics; and (d) for soybean using DBL metrics. The histograms represent the density distributions of the reported and predicted planting dates. DOY refers to the day of the year.

is the number of matched samples.

4. Results and discussion

4.1. Performance of the two-step framework

Appropriate phenological modeling is crucial to derive planting dates from phenological metrics. We compared our double logistic (DBL) function-based phenological metrics (eight parameters) with the Multi-Source Land Surface Phenology (MS-LSP) products (20 parameters, excluding layers such as quality assurance, number of phenological cycles, and number of clear observations) (Bolton et al., 2020) in predicting field-level planting dates in the U.S. Midwest from 2016 to 2019. The performance of these phenological metrics to directly predict field-level planting dates (using Eq. (8)) for corn and soybean is shown in Fig. 3. Although MS-LSP has more parameters to characterize the phenological cycles, our DBL-based phenological metrics (Fig. 3b and d, R2 = 0.69, MAE = 5.76 days, and RMSE = 7.82 days for corn; R^2 = 0.62, MAE = 6.06 days, and RMSE=8.17 days for soybean) outperformed the MS-LSP phenological metrics (Fig. 3a and c, R2 = 0.52, MAE = 7.52 days, and RMSE = 9.72 days for corn; R^2 = 0.48, MAE = 7.33 days, and RMSE = 9.62 days for soybean). The higher performance of our DBL-based phenological metrics could be attributed to our phenological model being designed for corn and soybean, while the MS-LSP algorithm is for all land cover types. These results indicated that appropriate phenological modeling is crucial to improve field-level planting date estimation.

After deriving phenological metrics, the second step in our framework was to predict RGDD, which can be accurately modeled using phenological metrics and environmental variables. As shown in Fig. 4, our RGDD model could explain 81% of field-level RGDD variation for corn with an MAE of 25.21 degrees and 75% of RGDD variation for soybean with an MAE of 38.66 degrees (Fig. 4a and c) in the U.S. Midwest from 2016 to 2020. More GDD was needed for soybean (average RGDD of about 460 degrees) than for corn (about 355 degrees), which was consistent with previous studies (Zhong et al., 2014). Higher performance for corn was achieved potentially because corn is a C4 plant, while soybean is a C3 plant. C4 plants have a more efficient photosynthetic pathway and higher light use efficiency, especially in warm and sunny conditions (Ciampitti and Vyn, 2012; Sage, 2004; Yin and Struik, 2009). Thus, the growth of corn has a faster response to GDD which may lead to a better performance in RGDD modeling. Spatially, the MAEs of predicted RGDD for both corn (Fig. 4b) and soybean (Fig. 4d) were generally smaller in the northern U.S. Midwest than those in the southern U.S. Midwest. The lower performance of RGDD modeling in the southern part might be due to the increase in vegetative growth duration and decrease in reproductive growth duration from lower to higher latitudes (Liu et al., 2013). Since the period from the planting date to the spring inflection date mainly covered the crop's vegetative growth, the longer vegetative growth duration made it more difficult to predict RGDD. In most counties in the U.S. Midwest, MAE for RGDD modeling was less than 30 degrees for corn and 50 degrees for soybean, which might result in 3-day and 5-day uncertainties in planting date predictions if the temperature was constantly 20 degrees. Overall, RGDD can be accurately modeled from phenological metrics and environmental variables.

The two-step framework (phenological metrics-RGDD-planting date) can accurately predict field-level planting dates for corn and soybean in the U.S. Midwest from 2016 to 2020 based on satellite-derived phenological metrics, modeled RGDD, and environmental variables. When validated against independently holdout of 20% field-level planting date data, the two-step model could explain 77% of field-level planting date variation for corn with an MAE of 4.6 days and an RMSE of 6.6 days (Fig. 5a) and 71% of planting date variation for soybean with an MAE of 5.4 days and an RMSE of 7.5 days (Fig. 5c). The performance of corn was slightly better than soybean, which was consistent with a previous one-step model (Deines et al., 2023). Spatially, the MAEs of predicted planting dates for both corn (Fig. 5b) and soybean (Fig. 5d) were smaller

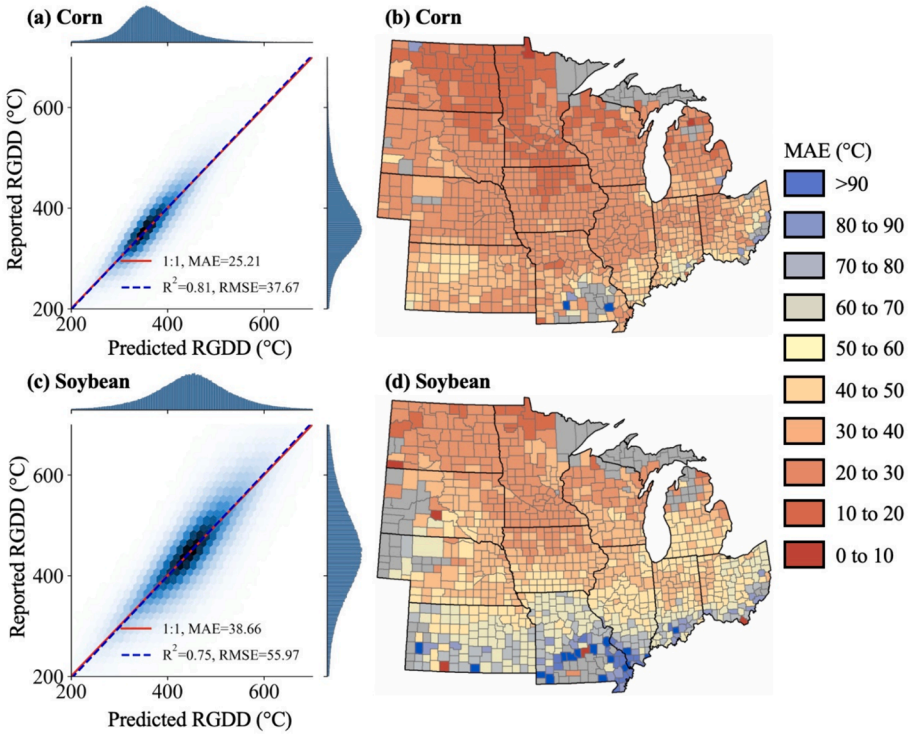

Fig. 4. Validation of predicted field-level required growing degree days (RGDD) against independent holdout of 20% field-level data for (a) corn and (c) soybean in the U.S. Midwest from 2016 to 2020. The histograms represent the density distributions of the reported and predicted RGDD. The spatial distributions of county-level mean absolute errors (MAE) of predicted RGDD in the unit of Celsius degree (°C) for (b) corn and (d) soybean are also plotted.

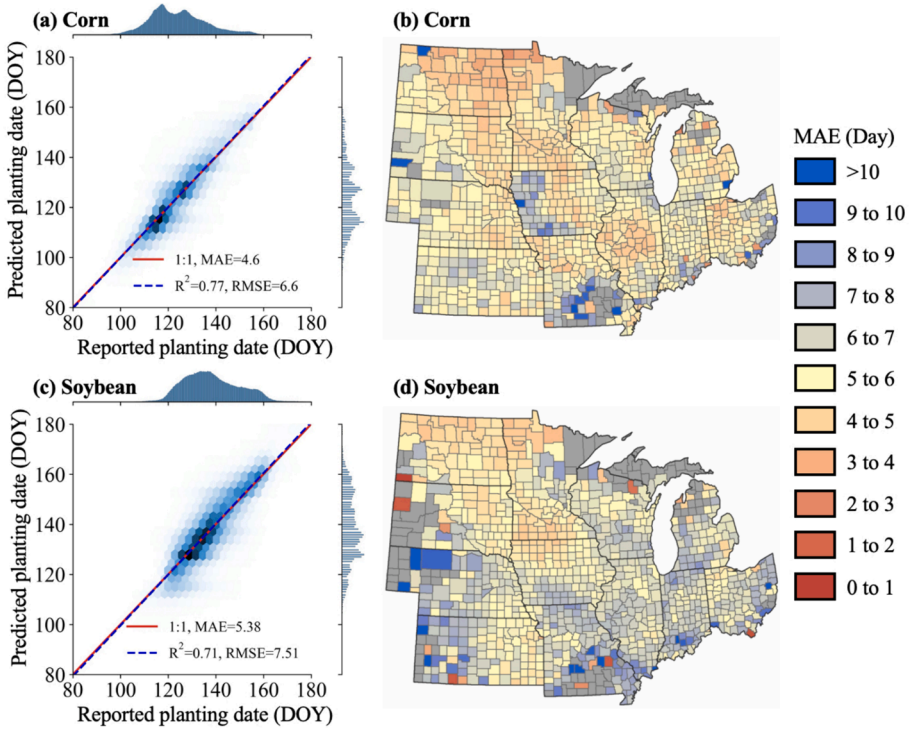

Fig. 5. Validation of predicted planting dates against independent holdout of 20% field-level planting dates in the U.S. Midwest from 2016 to 2019 based on Eq. (6). The model uses phenological metrics, environmental variables, and RGDD to predict planting dates for (a) corn and (c) soybean in the U.S. Midwest from 2016 to 2020. The histograms represent the density distributions of the reported and predicted planting dates. The spatial distributions of county-level mean absolute errors (MAE) of predicted planting dates for (b) corn and (d) soybean are also plotted.

in the central and northern U.S. Midwest than those in the western and southern U.S. Midwest. The western parts of South Dakota, Nebraska, Kansas, and the southern part of Missouri had the largest MAEs, while North Dakota, Minnesota, and Iowa had relatively lower MAEs.

At the county level, our predicted planting dates were consistent with those from a previous study. Our predicted dates when 25%, 50%, and 75% acres were planted within a county agreed well with the dates provided by Deines et al. (2023) from 2016 to 2020 in the U.S. Midwest (Fig. S2). In the comparison of our and Deines's planting date data for dates when 25%, 50%, and 75% acres were planted within a county, the R^2 ranged from 0.81 to 0.84 for corn and 0.70 to 0.79 for soybean; the MAE ranged from 3.06 to 3.79 days for corn and 3.41 to 3.69 days for soybean; and the RMSE ranged from 4.59 to 5.50 days for corn and 5.08 to 5.65 days for soybean. The predicted planting dates had better consistency for corn than for soybean, likely because both our and Deines's predictions of planting dates had higher accuracies for corn. Thus, more efforts should be made to the field-level planting date detection of soybean in the future.

Table 2
Model performance with different spatiotemporal cross-validation strategies including (1) trained and tested on the full dataset (full validation); (2) trained and tested on the leave-one-year-out dataset (temporal validation); and (3) trained and tested on the leave-20%-region-out dataset (spatial validation).

Crop	Validation	R^2	MAE (day)	RMSE (day)
Corn	full	0.77	4.60	6.60
	temporal	0.65	6.15	8.03
	spatial	0.75	4.77	6.79
Soybean	full	0.71	5.38	7.51
	temporal	0.61	6.62	8.76
	spatial	0.69	5.56	7.70

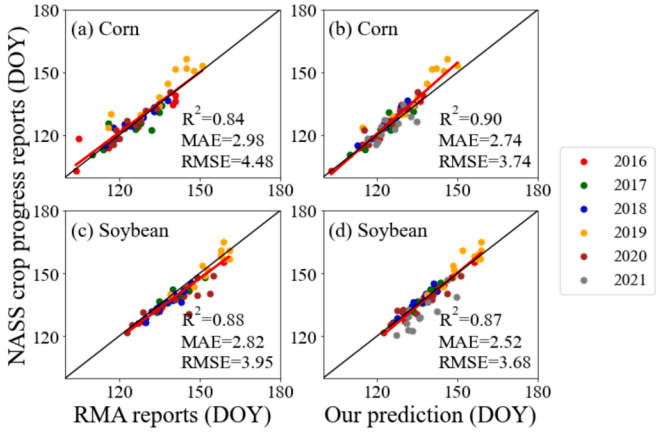

Fig. 6. Validation of predicted state-level median planting dates against the NASS crop progress reports in the U.S. Midwest from 2016 to 2021. (a) RMA vs NASS for corn; (b) our prediction vs NASS for corn; (c) RMA vs NASS for soybean; and (d) our prediction vs NASS for soybean. The colors of the dots refer to different years.

When aggregated to the state level and compared with state median planting dates from the NASS crop progress reports from 2016 to 2021 in the 12 states in the U.S. Midwest, our planting date prediction can achieve an R^2 of 0.90 and MAE of 2.74 days for corn (Fig. 6b), and an R^2 of 0.87 and MAE of 2.52 days for soybean (Fig. 6d), which is more favorable compared with previous studies (Deines et al., 2023; Ren et al., 2017; Sakamoto et al., 2011; Urban et al., 2018), whose R^2 ranges from 0.58 to 0.76 and MAEs >4 days. Considering the biases between field-level and state-level reports were 2.98 days for corn (Fig. 6a) and 2.82 days for soybean (Fig. 6c), our model trained on field-level reports was highly accurate at the state level.

The spatiotemporal transferability should be identified before applying our framework outside the spatiotemporal range of our training datasets (Filippelli et al., 2024). Compared with full validation

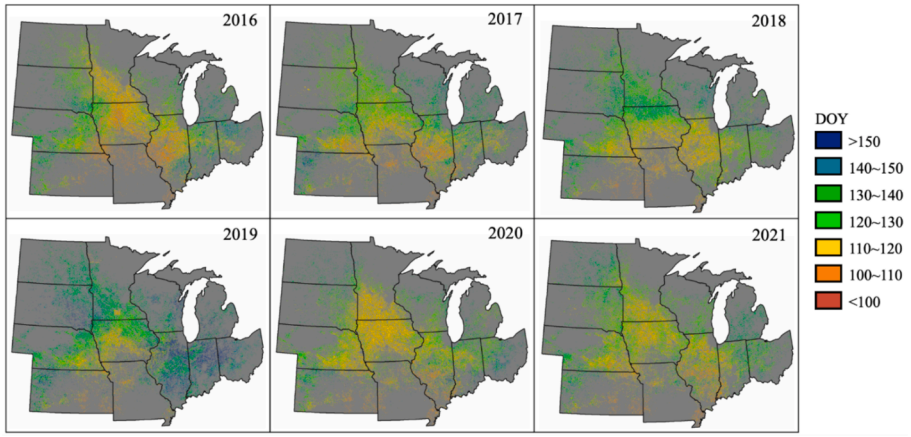

Fig. 7. Field-level planting dates of corn in the U.S. Midwest from 2016 to 2021.

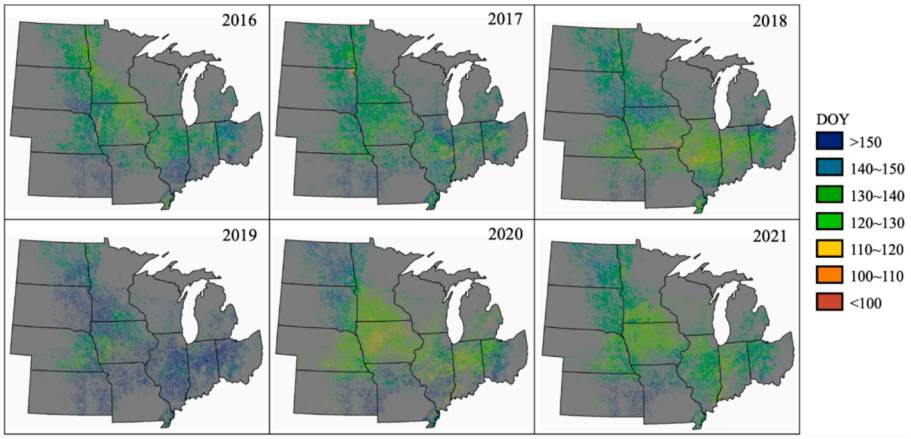

Fig. 8. Field-level planting dates of soybean in the U.S. Midwest from 2016 to 2021.

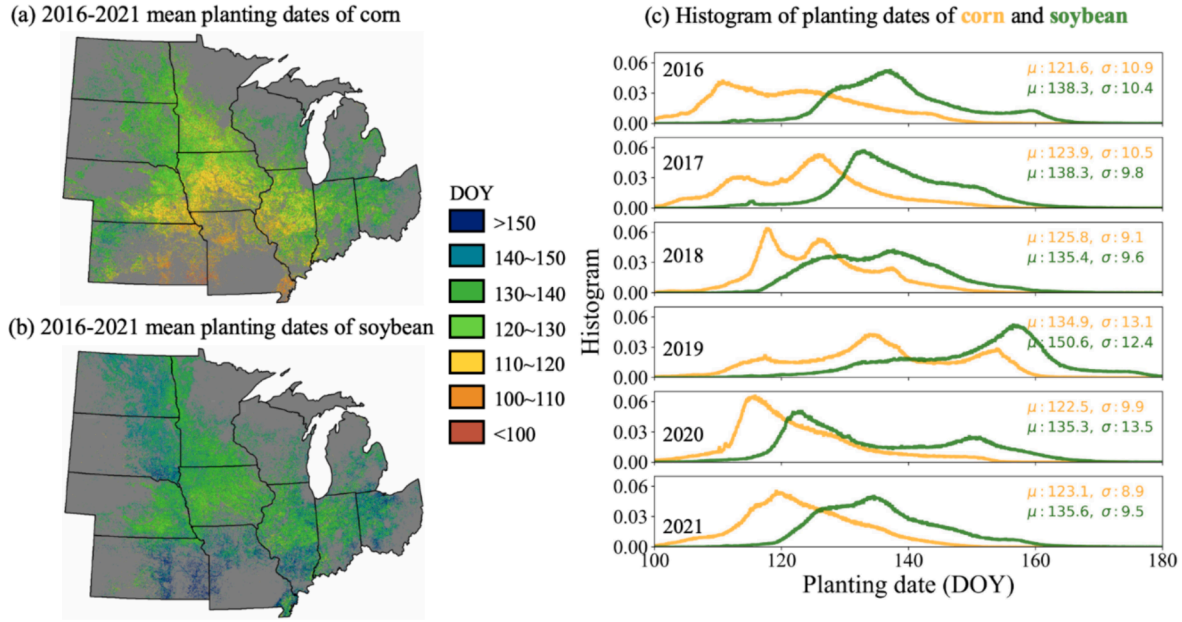

Fig. 9. Mean field-level planting dates of (a) corn and (b) soybean in the U.S. Midwest from 2016 to 2021 and (c) the histograms of planting dates in each year. The μ and σ parameters are the mean value and standard deviation of the samples.

(Table 2), temporal validation (leave-one-year-out) had a lower performance in field-level planting date detection with a RMSE increased by 1.23 days for corn (from 6.60 to 8.03 days) and 1.25 days for soybean (from 7.51 to 8.76 days), while the spatial validation (leave-20%-region-out) resulted in a smaller decrease in performance, with a RMSE increased by only 0.19 days for both corn and soybean (corn: from 6.60 to 6.79 days; soybean: from 7.51 to 7.70 days). The spatiotemporal transferability of corn and soybean was similar, with the decreased performance caused by temporal transfer being about six times that caused by spatial transfer. The results indicated that the two-step framework had better spatial transferability than temporal transferability in predicting field-level planting dates for corn and soybean fields.

4.2. Spatiotemporal pattern of planting dates

The field-level planting date maps for corn and soybean from 2016 to 2021 in the U.S. Midwest provided detailed information on spatial and temporal variations of planting dates in recent years. Annual field-level planting dates for corn and soybean are separately plotted in Figs. 7 and 8, and the mean planting dates of corn and soybean in the U.S. Midwest from 2016 to 2021 are shown in Fig. 9. The planting dates in the central U.S. Midwest were generally earlier than those in the other parts of the U.S. Midwest, with large variations within states (Fig. 9). For example, the south vs. the north of Iowa and Illinois and the west vs. east Nebraska and Kansas had considerably different planting dates, which cannot be captured by state-level NASS crop progress reports. The field-level planting date maps demonstrated that (1) soybean is typically planted two weeks after corn in the U.S. Midwest, (2) planting dates in the U.S. Midwest may vary more than two months even for the same crop, and (3) 30-m and ~3-day HLS products are useful for field-level predictions of planting dates.

Although large shifts in planting dates in the U.S. Midwest were reported in previous studies (Deines et al., 2023; Kucharik, 2006; Sacks and Kucharik, 2011), no significant changes were observed in planting dates of corn and soybean in this region from 2016 to 2021 (Fig. S3). However, we detected abnormally late plating in 2019 for both corn and soybean in the U.S. Midwest (Figs. 7 and 8). In 2019, corn and soybean were planted 9–13 days later and 12–15 days later, respectively (Fig. 9c), due to the heavy spring rainfall in 2019 (Gao et al., 2020). In 2020, corn and soybean in Iowa were planted slightly earlier than the other years (Figs. 7 and 8). The R^2 between field-reported and satellite-predicted planting dates in 2020 in Iowa was 0.41 for corn and 0.44 for soybean, while during 2016–2020 was 0.66 and 0.67 for corn and soybean, respectively (Table S4). The lower model performance in Iowa in 2020 might be attributed to the intense and fast-moving windstorms known as a derecho on 10 August 2020 across Iowa's agricultural regions (Hosseini et al., 2020), which caused damage to corn and soybean and affected the crop NDVI time series. This further led to unreliable phenological modeling and resulted in inaccurate predictions of planting dates in 2020 in Iowa (Fig. S4). Thus, field-level planting date records may help understand the impacts of natural disasters and inform appropriate management practices for future adaptations.

Accurate planting date records for almost all corn and soybean fields in the U.S. Midwest allow for the analysis of the impacts of agricultural management such as irrigations and cover cropping. The performance of field-level planting date prediction for corn and soybean with irrigation/non-irrigation and cover cropping/non-cover cropping is shown in Fig. 10. The irrigation information was extracted from field reports and cover cropping information was derived from the satellite-based phenological modeling (Fig. S1). For cover cropping fields, the MAE for corn and soybean in the U.S. Midwest was 5.39 days and 5.63 days, respectively, while for the non-cover cropping fields, the performance was improved by 0.8 days for corn and 0.9 days for soybean. This indicated that cover cropping had similar impacts on the model's performance for both corn and soybean. Generally, cover crops are planted in the non-growing seasons, which may not have much influence on the growth curve of cash crops (Zhou et al., 2022). Thus, cover cropping has no different impacts on the satellite time series for corn and soybean, leading to similar impacts on predicting planting dates for corn and soybean. For irrigation fields, the MAE for corn and soybean in the U.S. Midwest was 5.36 days and 6.11 days, respectively, while for the non-irrigation fields, the performance was improved by 0.8 days for corn and 1.4 days for soybean. This indicated that irrigation reduced the model's performance more for soybean fields. The two-step model was based on RGDD, an accumulation of thermal heat (Cayton et al., 2015), and irrigation directly affects the temperatures through evapotranspiration and further influences the model's performance. The cooling effects of irrigation are typically more pronounced and last longer for corn than for soybean (Chen et al., 2018). Thus, irrigation should lead to larger impacts on the model's performance for corn from the aspects of cooling effects. However, the opposite results were obtained in this study, which may be due to the difference in corn and soybean structure. Since PRISM-modeled temperature rather than the land surface temperature was used, soybean structure may induce larger discrepancies between the two kinds of temperature. Overall, both cover cropping and irrigation decreased the model's performance in predicting field-level planting dates for corn and soybean. The impacts of cover cropping were similar for corn and soybean fields, while irrigation had a larger impact on soybean fields than on corn fields. Further research is needed

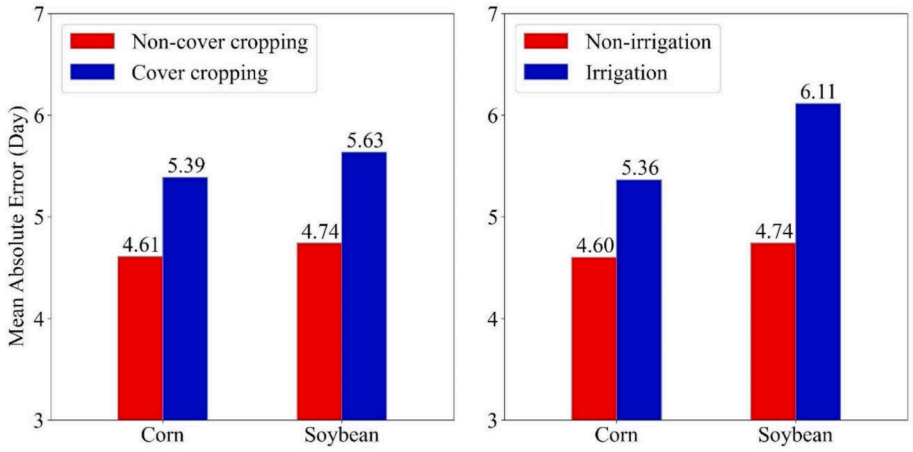

Fig. 10. Performance of predicted field-level planting dates for corn and soybean fields with cover cropping/non-cover cropping and irrigation/non-irrigation in the U.S. Midwest from 2016 to 2020.

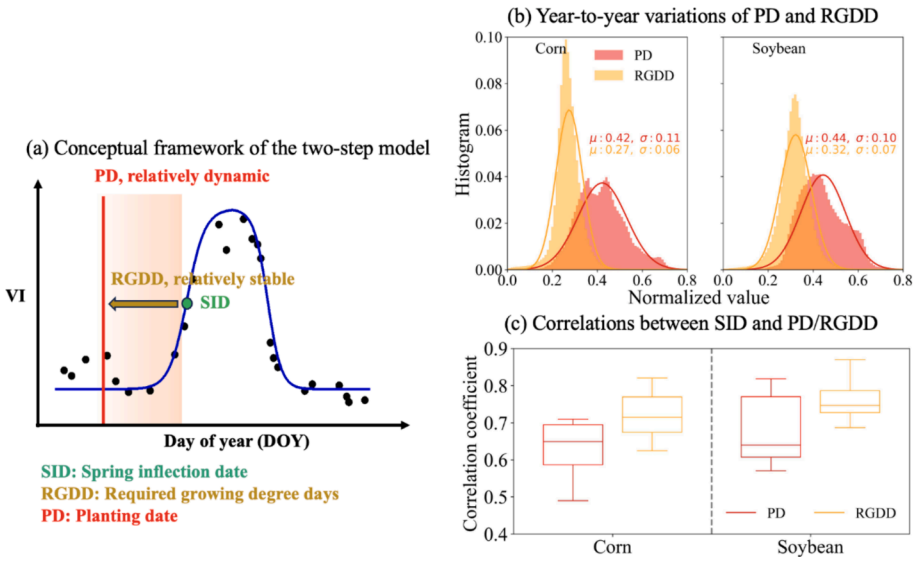

Fig. 11. Advantages of the two-step model in predicting field-level planting dates from phenological metrics: (a) conceptual scheme of deriving planting dates from phenological metrics using satellite vegetation index (VI) time series; (b) year-to-year variations of planting dates and required growing degree days (RGDD) in the U. S. Midwest from 2016 to 2020. The μ and σ parameters are the mean and standard deviation of the samples; (c) correlations between spring inflection dates and planting dates/RGDD.

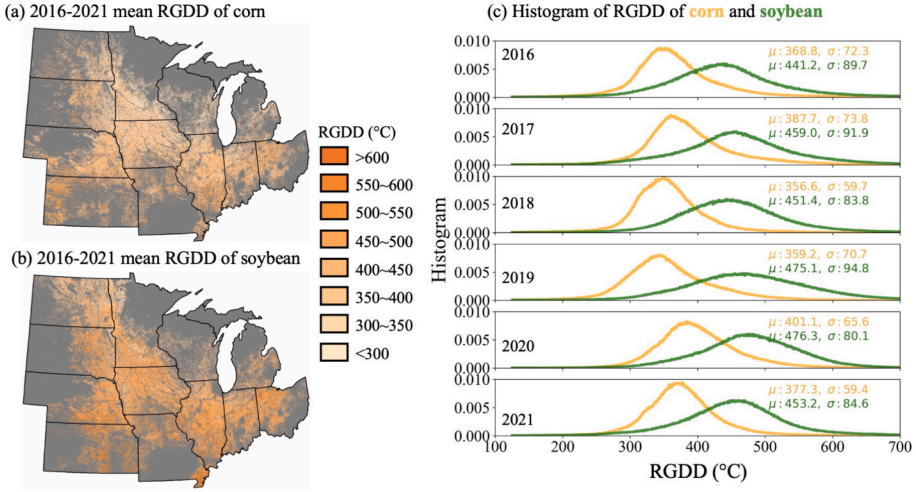

Fig. 12. Mean field-level required growing degree days (RGDD) in the unit of Celsius degree (°C) of (a) corn and (b) soybean in the U.S. Midwest from 2016 to 2021 and (c) the histograms of RGDD in each year. The μ and σ parameters are the mean value and standard deviation of the samples.

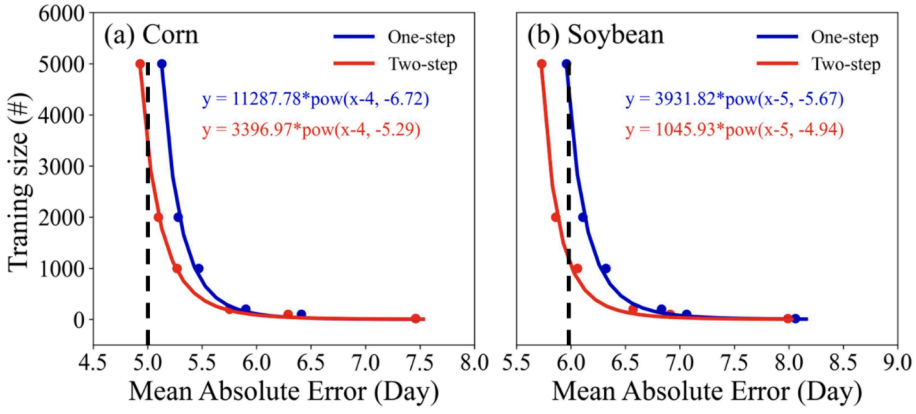

Fig. 13. Performance of predicted field-level planting dates of (a) corn and (b) soybean using one-step (phenological metrics-planting dates) and two-step (phenological metrics-RGDD-planting dates) models with varied training samples in the U.S. Midwest from 2016 to 2020.

to understand these effects better.

4.3. Benefits of the two-step framework

The two-step framework can provide more flexibility and robustness in model development, which has been widely adopted in remote sensing communities (Sakamoto et al., 2010; Santini et al., 2010; Zeng et al., 2018). In our proposed two-step framework (phenological metrics-RGDD-planting date), RGDD served as a bridge to link phenological metrics to planting dates (Fig. 11a). The benefits of using RGDD can be summarized as follows: (1) The correlations between RGDD and spring inflection date were stronger than those between planting dates and spring inflection date (Fig. 11c), which can lead to higher modeling performance. In other words, the importance of spring phenological metrics on planting dates is enhanced by RGDD, making it more suitable for years with natural disasters. For example, the strong windstorms in Iowa in 2020 occurred on 10 August (Hosseini et al., 2020), which only affected the fall phenology, so models relying more on spring phenology are theoretically less affected. (2) Compared with planting dates, the spatial and temporal variations of RGDD were much smaller (Fig. 11b, 12, S5, and S6). Considering the phenological metrics are robust from the satellite time series and less affected by single satellite observations, the relatively stable RGDD is easier to predict (Fig. 4). Because of its stability, empirical RGDD values are widely used to estimate crop growth. For example, the seed company Beck's Hybrids uses 120 RGDD for corn and 130 RGDD for soybean to determine the required time for seeds to emerge (Gauck, 2019). (3) RGDD can account for spatial and temporal variations in temperature and the impacts of weather variability (Cayton et al., 2015), which is highly consistent with growing-season temperature and a typical measure of crop development (Lobell et al., 2011). The RGDD in the U.S. Midwest during 2016–2021 for corn and soybean was relatively stable even with a delayed planting of about two weeks in 2019 (Figs. S5 and S6). (4) RGDD can potentially provide insights into how planting dates and phenological metrics are linked. Even if the planting date can vary more than two months in the U.S. Midwest (Figs. 7 and 8), RGDD for crop growth remained relatively stable (Figs. S5 and S6), which was beneficial to understand crop growth under variable conditions.

The advantages of using RGDD also led to higher accuracies in predicting field-level planting dates for corn and soybean in the U.S. Midwest. Fig. 13 shows the MAEs between satellite-predicted and reported field-level planting dates using the two-step (phenological metrics-RGDD-planting dates) and traditional one-step (phenological metrics-planting dates) models for corn and soybean fields with varying numbers of training samples. With the same number of training samples, the two-step model outperformed the one-step model for both corn and soybean, significantly reducing the need for ground "truth" samples. For example, to achieve an MAE of 5 days for corn fields, the two-step model requires ~3400 fields per state per year while the one-step model requires ~11300 fields (about three times) per state per year. To achieve an MAE of 6 days for soybean fields, the two-step model requires ~1000 fields per state per year while the one-step model requires ~3900 fields (about four times) per state per year. It is extremely beneficial because field-level ground truth is typically labor- and cost-intensive to obtain (Deines et al., 2023).

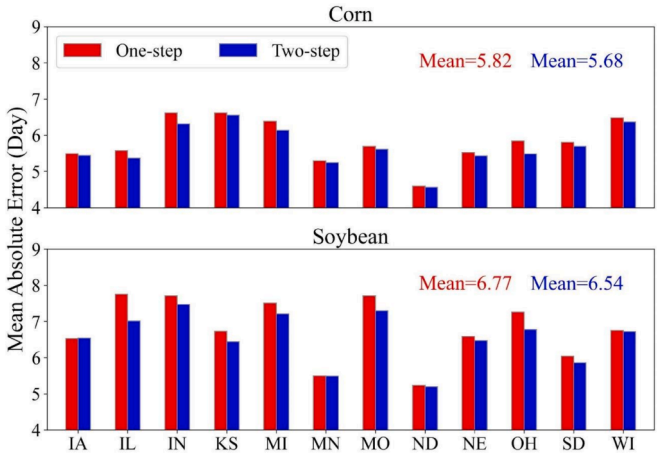

Fig. 14. Performance of predicted field-level planting dates of corn and soybean using one-step (phenological metrics-planting dates) and two-step (phenological metrics-RGDD-planting dates) models in the U.S. Midwest from 2016 to 2020 with 1000 samples per state per year.

The two-step model outperformed the one-step model for all states in the U.S. Midwest. For example, with 1000 samples per state per year, a relatively large number of ground truth samples (Deines et al., 2023), the two-step model achieved better performance than the one-step model for all states in the U.S. Midwest with improved MAEs for corn from 5.82 to 5.68 days and for soybean from 6.77 to 6.54 days (Fig. 14). The improvements of the two-step model were more obvious for extreme events like the flooding in the U.S. Midwest in 2019 (Fig. S7). These results indicated that the two-step model is spatially and temporally stable and less sensitive to extreme events.

The two-step framework for field-level detection of planting dates demonstrated high accuracies for corn and soybean fields in the U.S. Midwest from 2016 to 2020. The framework has great flexibility for wider applications with some necessary modifications. First, the framework can be used for planting date detection for other crops including summer and winter crops. For summer crops such as sorghum, which have similar phenological cycles to corn and soybean (Masialeti et al., 2010), the framework can be directly applied. However, for winter crops such as winter wheat, which has different phenological cycles from corn and soybean (Masialeti et al., 2010), modifications in the phenological modeling are required. The NDVI time series for current phenological modeling is from spring to winter, but for winter crops, it should start from the summer to next year's summer (Lu et al., 2014). For other crops like alfalfa, which experience "grow and cut" cycles in summer and fall (Masialeti et al., 2010), the double logistic function may fail to capture the crop growth, and more advanced phenological modeling algorithms are required. Second, the framework can be adapted to detect other crop stages such as silking, maturity, and harvest dates by changing the bridge (RGDD in this study). For example, the cooling degree days (CDD, similar to the concept of GDD) are useful for determining crop leaf senescence date (Piao et al., 2019). Replacing GDD with CDD, the framework might be able to detect field-level leaf senescence dates. Harvest detection may be more complex because it is not only controlled by crop phenological cycles but also affected by farmers' management (Kusumastuti et al., 2016). A certain period after phenological-derived heading dates might be suitable for harvesting (Sakamoto et al., 2005). Thus, the framework may provide a time window for harvesting, and the detection of extract harvest dates may require other information such as synthetic aperture radar (SAR) observations (Meroni et al., 2021). Third, the framework can be extended to larger regions and longer periods. Current phenological modeling is based on the 30-m and ~3-day HLS NDVI time series, which started in 2015. For larger-region (e.g. global) and longer-period (e.g. back to 2000) applications, Landsat archives may be a good alternative (Deines et al., 2023). However, the frequency of high-quality Landsat observations is too sparse to capture short-term changes during crop growth (Zhou et al., 2022). Thus, phenological datasets from multi-sensor fused products are more favorable (Liao et al., 2019; Luo et al., 2018; Sadeh et al., 2021).

4.4. Performance for "within" and "outside" policy-driven planting windows

The fine-resolution planting date records allow for analysis of crop planting practices that are "within" or "outside" the CCIP-driven earliest

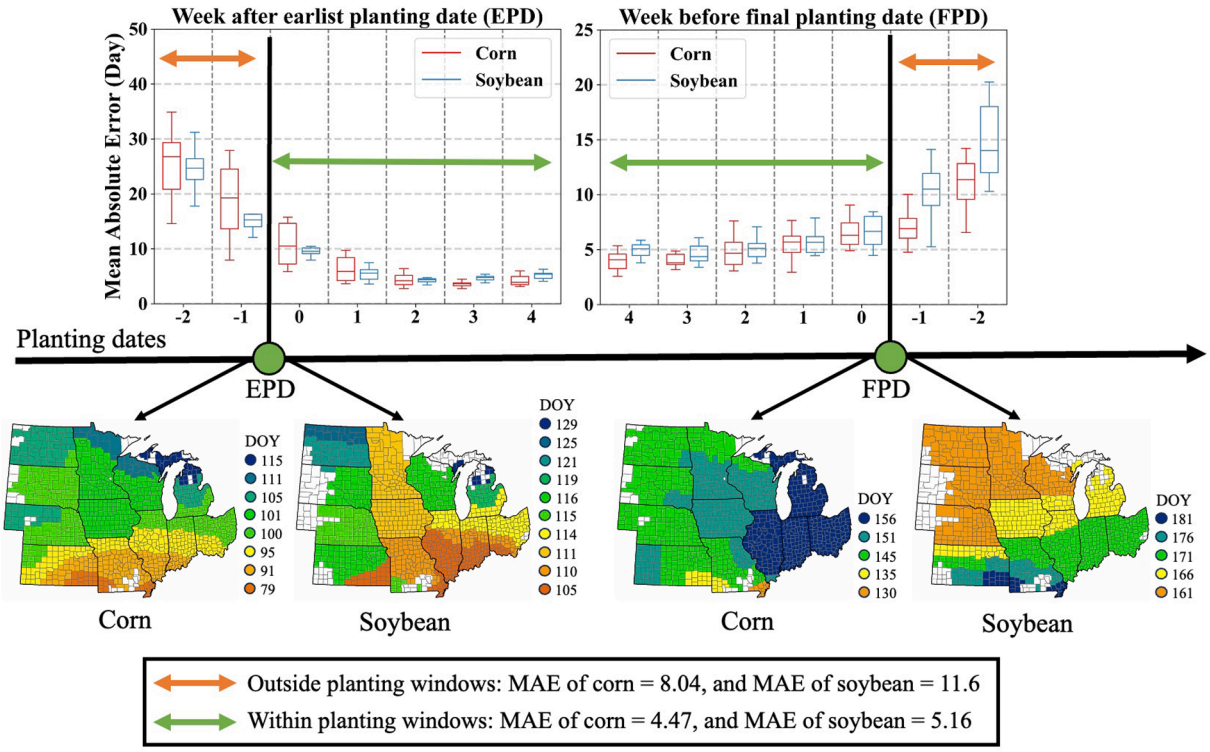

Fig. 15. Performance of predicted field-level planting dates of corn and soybean in the U.S. Midwest from 2016 to 2020 for different time windows defined by the USDA RMA CCIP-driven earliest planting dates (EPD) and final planting dates (FPD). The bottom four maps are the spatial distributions of the USDA-suggested EPD for corn and soybean and FPD for corn and soybean in the unit of day of the year (DOY).

and final planting dates. The CCIP-driven planting windows are used for crop insurance products including yield protection, revenue protection with harvest price exclusion, and revenue protection plans (Schnitkey, 2013). Fields planted before the earliest planting date are not eligible for replant payments and for fields planted after the final planting date, the guarantee will be reduced by 1% per day up to 25 days and will be 60% after 25 days (Schnitkey, 2013). The CCIP-driven planting windows are designed to guide farmers to ensure crop yield production, while both planting too early and too late could induce yield deduction, thus the suggested planting windows should be close to optimal planting windows (Khan et al., 2017b).

We used the CCIP-driven earliest and final planting dates to evaluate the performance of our model-derived field-level planting dates. The model's performance of field-level planting dates "within" and "outside" the CCIP-driven earliest and final planting dates is shown in Fig. 15. Planting dates "outside" the CCIP-driven planting windows were more difficult to predict, especially for those planted earlier than the CCIP-driven earliest planting dates. The MAEs of predicted field-level planting dates for corn and soybean fields "within" the CCIP-driven planting windows in the U.S. Midwest were 4.47, and 5.16 days, respectively. For corn and soybean fields with reported planting dates "outside" the CCIP-driven planting windows, MAEs of model-predicted planting dates were 8.04 and 11.6 days, respectively, which was about two times those "within" the CCIP-driven planting windows. In addition, for fields planted two weeks before the CCIP-driven earliest planting dates and fields planted two weeks after the CCIP-driven final planting dates, the MAEs were around 30 days and 20 days, respectively.

The framework, based on crop phenology, has limitations for fields planted too early or too late. For example, crops planted in February or March may have similar vegetation time series because seeds may be dormant until favorable weather conditions. The higher model performance indicates higher accordance with the crop phenological growth and a higher possibility of non-reduced crop yield. Better performance of predicted field-level planting dates within the CCIP-driven planting windows indicated that the time windows are relatively suitable for crop growth at least from the aspect of crop phenology. To predict planting dates for fields planted too early or too late, other types of algorithms rather than phenology-based approaches are needed.

5. Conclusion

This study proposed a two-step framework to convert phenological metrics to planting dates using RGDD as a bridge. The 30-m and ~3-day HLS remote sensing time series were used to derive field-level phenological metrics, the field-level reports of planting dates were used for model training and validation, and the state-level crop progress reports were used for additional validation. The framework can accurately predict field-level planting dates with an MAE of 4.6 days for corn and 5.4 days for soybean. After aggregating to the state level, the predicted planting dates were well consistent with the state-level crop progress reports with MAEs <3 days for both corn and soybean. The performance in cover cropping and irrigation systems degraded slightly and the MAEs were still less than one week. Higher accuracies were achieved for planting dates within policy-driven planting windows. The proposed two-step framework (phenological metrics-RGDD-planting dates) also outperformed the one-step model (phenological metrics-planting dates), which can significantly reduce the required number of ground truth samples and can be beneficial for deriving planting dates from current and future phenological products. This work can further contribute to the analysis of yield gaps, management practices, and government policies on crop planting.

CRediT authorship contribution statement

Qu Zhou: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Kaiyu Guan:** Writing – review & editing, Resources, Funding acquisition, Conceptualization. **Sheng Wang:** Writing – review

& editing, Supervision, Conceptualization. **James Hipple:** Writing – review & editing, Funding acquisition. **Zhangliang Chen:** Writing – review & editing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Planting date maps will be aggregated in time at the county level and made publicly available through a public repository. The field-level data are not publicly available and were supplied by RMA under contract for the purpose of this work. These data are protected under the disclosure agreements between USDA RMA and the University of Illinois and Section 1619 of the Food, Conservation, and Energy Act of 2008 (PL 110- 246), 7 U.S.C. 8791. All other input data can be acquired from the sources as noted in this paper.

Acknowledgments

We acknowledge the supports from the USDA Risk Management Agency under agreement number RMA22CPT0012747 and the National Science Foundation (NSF) CAREER Award by the Environmental Sustainability Program. K.G. and Q.Z. acknowledge the support from the NASA FINESST award. S.W. is supported by the NASA Early Career Investigator Program in Earth Science (80NSSC24K1057). This work is also supported by the USDA National Institute of Food and Agriculture (NIFA) Rapid Response to Extreme Weather Events Across Food and Agriculture Systems (A1712) program, project award no. 2023-68016-41371, and USDA NIFA Artificial Intelligence for Future Agricultural Resilience, Management, and Sustainability (AI-FARMS) project. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2024.07.031.

References

Acharya, J., Bakker, M.G., Moorman, T.B., Kaspar, T.C., Lenssen, A.W., Robertson, A.E., 2017. Time interval between cover crop termination and planting influences corn seedling disease, plant growth, and yield. Plant Dis 101, 591–600.
Akyuz, F.A., Kandel, H., Morlock, D., 2017. Developing a growing degree day model for north dakota and northern minnesota soybean. Agric for Meteorol 239, 134–140.
Babcock, C., Finley, A.O., Looker, N., 2021. A Bayesian model to estimate land surface phenology parameters with harmonized Landsat 8 and Sentinel-2 images. Remote Sens Environ 261, 112471.
Beck, P.S.A., Atzberger, C., Høgda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. Remote Sens Environ 100, 321–334. https://doi.org/10.1016/j.rse.2005.10.021.
Bolton, D.K., Gray, J.M., Melaas, E.K., Moon, M., Eklundh, L., Friedl, M.A., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. Remote Sens Environ 240, 111685.
Borchers, A., Truex-Powell, E., Wallander, S., Nickerson, C., 2014. Multi-cropping practices: recent trends in double cropping. USDA Econ. Res. Serv. 22, 15.
Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int 26, 341–358.
Boyer, C.N., Stefanini, M., Larson, J.A., Smith, S.A., Mengistu, A., Bellaloui, N., 2015. Profitability and risk analysis of soybean planting date by maturity group. Agron J 107, 2253–2262.
Boyer, C.N., Park, E., Yun, S.D., 2023. Corn and soybean prevented planting acres response to weather. Appl Econ Perspect Policy 45, 970–983.
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 Policy 52, 316–328.
Cassman, K.G., Grassini, P., 2020. A global perspective on sustainable intensification research. Nat Sustain 3, 262–268.
Cayton, H.L., Haddad, N.M., Gross, K., Diamond, S.E., Ries, L., 2015. Do growing degree days predict phenology across butterfly species? Ecology 96, 1473–1479.
Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.
Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. Remote Sens Environ 91, 332–344.
Chen, F., Xu, X., Barlage, M., Rasmussen, R., Shen, S., Miao, S., Zhou, G., 2018. Memory of irrigation effects on hydroclimate and its modeling challenge. Environ. Res. Lett. 13, 064009.
Ciampitti, I.A., Vyn, T.J., 2012. Physiological perspectives of changes over time in maize yield dependency on nitrogen uptake and associated nitrogen efficiencies: A review. Field Crops Res 133, 48–67.
Clark, J.D., Veum, K.S., Fernández, F.G., Camberato, J.J., Carter, P.R., Ferguson, R.B., Franzen, D.W., Kaiser, D.E., Kitchen, N.R., Laboski, C.A.M., 2019. United states midwest soil and weather conditions influence anaerobic potentially mineralizable nitrogen. Soil Sci. Soc. Am. J. 83, 1137–1147.
Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S.V., Justice, C., 2018. The harmonized landsat and sentinel-2 surface reflectance data set. Remote Sens Environ 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002.
Curtis, Z., Clark, B., Larson, Z., 2023. Understanding Growing Degree Days [WWW Document]. https://extension.psu.edu/understanding-growing-degree-days.
Dai, S., Shulski, M.D., Hubbard, K.G., Takle, E.S., 2016. A spatiotemporal analysis of Midwest US temperature and precipitation trends during the growing season from 1980 to 2013. Int. J. Climatol. 36, 517–525.
Daly, C., Smith, J.I., Olson, K.V., 2015. Mapping atmospheric moisture climatologies across the conterminous United States. PLoS One 10, e0141140.
Deines, J.M., Guan, K., Lopez, B., Zhou, Q., White, C.S., Wang, S., Lobell, D.B., 2023. Recent cover crop adoption is associated with small maize and soybean yield losses in the United States. Global change biology 29 (3), 794–807.
Deines, J.M., Swatantran, A., Ye, D., Myers, B., Archontoulis, S., Lobell, D.B., 2023. Field-scale dynamics of planting dates in the US Corn Belt from 2000 to 2020. Remote Sens Environ 291, 113551.
Delpierre, N., Dufrêne, E., Soudani, K., Ulrich, E., Cecchini, S., Boé, J., François, C., 2009. Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. Agric for Meteorol 149, 938–948.
Egli, D.B., Cornelius, P.L., 2009. A regional analysis of the response of soybean yield to planting date. Agron J 101, 330–335.
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, 74–84.
Filippelli, S.K., Schleeweis, K., Nelson, M.D., Fekety, P.A., Vogeler, J.C., 2024. Testing temporal transferability of remote sensing models for large area monitoring. Science of Remote Sensing 9, 100119. https://doi.org/10.1016/j.srs.2024.100119.
Ganguly, S., Friedl, M.A., Tan, B., Zhang, X., Verma, M., 2010. Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. Remote Sens Environ 114, 1805–1816.
Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., Mueller, R., Johnson, D.M., Prueger, J.H., 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sens Environ 188, 9–25.
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. Remote Sens Environ 242, 111752.
Gauck, S., 2019. Agronomy Talk: Corn and Soybean Early-Season Emergence [WWW Document]. https://www.beckshybrids.com/resources/agronomy-talk/agronomy-talk-corn-and-soybean-early-season-emergence#.
Hosseini, M., Kerner, H.R., Sahajpal, R., Puricelli, E., Lu, Y.-H., Lawal, A.F., Humber, M. L., Mitkish, M., Meyer, S., Becker-Reshef, I., 2020. Evaluating the impact of the 2020 Iowa derecho on corn and soybean fields using synthetic aperture radar. Remote Sens (Basel) 12, 3878.
Hu, M., Wiatrak, P., 2012. Effect of planting date on soybean growth, yield, and grain quality. Agron J 104, 785–790.
Isbell, F., Adler, P.R., Eisenhauer, N., Fornara, D., Kimmel, K., Kremen, C., Letourneau, D. K., Liebman, M., Polley, H.W., Quijas, S., 2017. Benefits of increasing plant diversity in sustainable agroecosystems. J. Ecol. 105, 871–879.
Johansen, C., Haque, M.E., Bell, R.W., Thierfelder, C., Esdaile, R.J., 2012. Conservation agriculture for small holder rainfed farming: Opportunities and constraints of new mechanized seeding systems. Field Crops Res 132, 18–32.
Khan, A., Najeeb, U., Wang, L., Tan, D.K.Y., Yang, G., Munsif, F., Ali, S., Hafeez, A., 2017a. Planting density and sowing date strongly influence growth and lint yield of cotton crops. Field Crops Res 209, 129–135.
Khan, A., Wang, L., Ali, S., Tung, S.A., Hafeez, A., Yang, G., 2017b. Optimal planting density and sowing date can improve cotton yield by maintaining reproductive organ biomass and enhancing potassium uptake. Field Crops Res 214, 164–174.
Kucharik, C.J., 2006. A multidecadal trend of earlier corn planting in the central USA. Agron J 98, 1544–1550.
Kusumastuti, R.D., Van Donk, D.P., Teunter, R., 2016. Crop-related harvesting and processing planning: a review. Int J Prod Econ 174, 76–92.

Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., Song, Y., 2019. Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. Sci. Total Environ. 650, 1707–1721.

Liu, L., Cao, R., Chen, J., Shen, M., Wang, S., Zhou, J., He, B., 2022. 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.

Liu, Y., Xie, R., Hou, P., Li, S., Zhang, H., Ming, B., Long, H., Liang, S., 2013. Phenological responses of maize to changes in environment when grown at different latitudes in China. Field Crops Res 144, 192–199.

Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat Clim Chang 1, 42–45.

Lu, L., Wang, C., Guo, H., Li, Q., 2014. Detecting winter wheat phenology with SPOT-VEGETATION data in the north china plain. Geocarto Int 29, 244–255.

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

Masialeti, I., Egbert, S., Wardlow, B.D., 2010. A comparative analysis of phenological curves for major crops in Kansas. Gisci Remote Sens 47, 241–259.

Melaas, E.K., Friedl, M.A., Zhu, Z., 2013. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. Remote Sens Environ 132, 176–185.

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

Miller, N., Tack, J., Bergtold, J., 2021. The impacts of warming temperatures on US Sorghum yields and the potential for adaptation. Am J Agric Econ 103, 1742–1758.

Moon, M., Richardson, A.D., Milliman, T., Friedl, M.A., 2022. A high spatial resolution land surface phenology dataset for AmeriFlux and NEON sites. Sci Data 9, 448.

MRCC, 2024. Growing Degree Days [WWW Document]. https://mrcc.purdue.edu/gismaps/gddinfo#:~:text=Modified%20Growing%20Degree%20Days%3A&text=If%20the%20low%20is%20below,is%20usually%2050°%20F.

NDAWN, 2024a. Soybean Growing Degree Days [WWW Document]. https://ndawn.ndsu.nodak.edu/help-soybean-growing-degree-days.html.

NDAWN, 2024b. Corn Growing Degree Days (GDD) [WWW Document]. https://ndawn.ndsu.nodak.edu/help-corn-growing-degree-days.html.

Nicholls, C.I., Altieri, M.A., 2013. Plant biodiversity enhances bees and other insect pollinators in agroecosystems. A Review. Agron Sustain Dev 33, 257–274.

Niu, Q., Li, X., Huang, J., Huang, H., Huang, X., Su, W., Yuan, W., 2022. A 30 m annual maize phenology dataset from 1985 to 2020 in China. Earth Syst Sci Data 14, 2851–2864.

Osipitan, O.A., Dille, J.A., Assefa, Y., Radicetti, E., Ayeni, A., Knezevic, S.Z., 2019. Impact of cover crop management on level of weed suppression: a meta-analysis. Crop Sci 59, 833–842.

Pathak, T.B., Stoddard, C.S., 2018. Climate change effects on the processing tomato growing season in California using growing degree day model. Model Earth Syst Environ 4, 765–775.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol Earth Syst Sci 11, 1633–1644.

Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X.U., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: Current progresses and challenges. Glob Chang Biol 25, 1922–1940.

Potash, E., Guan, K., Margenot, A., Lee, D., DeLucia, E., Wang, S., Jang, C., 2022. How to estimate soil organic carbon stocks of agricultural fields? Perspectives using ex-ante evaluation. Geoderma 411, 115693.

Pulakkatu-Thodi, I., Shurley, D., Toews, M.D., 2014. Influence of planting date on stink bug injury, yield, fiber quality, and economic returns in Georgia cotton. J Econ Entomol 107, 646–653.

Ren, J., Campbell, J.B., Shao, Y., 2017. Estimation of sos and eos for midwestern us corn and soybean crops. Remote Sens (basel) 9, 722.

Ren, Y., Qiu, J., Zeng, Z., Liu, X., Sitch, S., Pilegaard, K., Yang, T., Wang, S., Yuan, W., Jain, A.K., 2024. Earlier spring greening in Northern Hemisphere terrestrial biomes enhanced net ecosystem productivity in summer. Communications Earth & Environment 5 (1), 122.

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

Sacks, W.J., Kucharik, C.J., 2011. Crop management and phenology trends in the US Corn Belt: Impacts on yields, evapotranspiration and energy balance. Agric for Meteorol 151, 882–894.

Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J.P., Zhang, Y., Rozenstein, O., Manivasagam, V.S., Chenu, K., 2021. Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. Int. J. Appl. Earth Obs. Geoinf. 96, 102260 https://doi.org/10.1016/j.jag.2020.102260.

Sage, R.F., 2004. The evolution of C4 photosynthesis. New Phytol. 161, 341–370.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., Ohno, H., 2005. A crop phenology detection method using time-series MODIS data. Remote Sens Environ 96, 366–374.

Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E., Arkebauer, T.J., 2010. A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. Remote Sens Environ 114, 2146–2159.

Sakamoto, T., Wardlow, B.D., Gitelson, A.A., 2011. Detecting spatiotemporal changes of corn developmental stages in the US corn belt using MODIS WDRVI data. IEEE Trans. Geosci. Remote Sens. 49, 1926–1936.

Santini, F., Alberotanza, L., Cavalli, R.M., Pignatti, S., 2010. A two-step optimization procedure for assessing water constituent concentrations by hyperspectral remote sensing techniques: An application to the highly turbid Venice lagoon waters. Remote Sens Environ 114, 887–898.

Schnitkey, G., 2013. Early planting and final planting dates for crop insurance. Farmdoc Daily 3.

Teasdale, J.R., Mirsky, S.B., 2015. Tillage and planting date effects on weed dormancy, emergence, and early growth in organic corn. Weed Sci 63, 477–490.

Trachsel, S., Kaeppler, S.M., Brown, K.M., Lynch, J.P., 2011. Shovelomics: high throughput phenotyping of maize (Zea mays L.) root architecture in the field. Plant Soil 341, 75–87.

Urban, D., Guan, K., Jain, M., 2018. Estimating sowing dates from satellite data over the US Midwest: A comparison of multiple sensors and metrics. Remote Sens Environ 211, 400–412.

USDA, 2023. World Agricultural Production, https://apps.fas.usda.gov/psdonline/circulars/production.pdf.

Waha, K., Müller, C., Bondeau, A., Dietrich, J.P., 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. Glob. Environ. Chang. 23, 130–143.

Wang, S., Guan, K., Zhang, C., Jiang, C., Zhou, Q., Li, K., Qin, Z., Ainsworth, E.A., He, J., Wu, J., Schaefer, D., 2023. Airborne hyperspectral imaging of cover crops through radiative transfer process-guided machine learning. Remote Sensing of Environment 285, 113386.

Yin, X., Struik, P.C., 2009. Theoretical reconsiderations when estimating the mesophyll conductance to CO2 diffusion in leaves of C3 plants by analysis of combined gas exchange and chlorophyll fluorescence measurements. Plant Cell Environ 32, 1513–1524.

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., Hong, Y., 2018. A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. ISPRS J. Photogramm. Remote Sens. 141, 30–45.

Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. Remote Sens Environ 237, 111511.

Zhang, M., Abrahao, G., Cohn, A., Campolo, J., Thompson, S., 2021. A MODIS-based scalable remote sensing method to estimate sowing and harvest dates of soybean crops in Mato Grosso. Heliyon, Brazil, p. 7.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens Environ 84, 471–475. https://doi.org/10.1016/S0034-4257(02)00135-9.

Zhang, X., Jayavelu, S., Liu, L., Friedl, M.A., Henebry, G.M., Liu, Y., Schaaf, C.B., Richardson, A.D., Gray, J., 2018a. Evaluation of land surface phenology from VIIRS data using time series of PhenoCam imagery. Agric for Meteorol 256, 137–149.

Zhang, X., Liu, L., Liu, Y., Jayavelu, S., Wang, J., Moon, M., Henebry, G.M., Friedl, M.A., Schaaf, C.B., 2018b. Generation and evaluation of the VIIRS land surface phenology product. Remote Sens Environ 216, 212–229.

Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. Remote Sens Environ 140, 1–13.

Zhou, Q., Guan, K., Wang, S., Jiang, C., Huang, Y., Peng, B., Chen, Z., Wang, S., Hipple, J., Schaefer, D., 2022. Recent rapid increase of cover crop adoption across the US Midwest detected by fusing multi-source satellite data. Geophys Res Lett 49.

Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. Remote Sens Environ 159, 269–277.