ISPRS Journal of Photogrammetry and Remote Sensing 211 (2024) 83–103

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

# A satellite-field phenological bridging framework for characterizing community-level spring forest phenology using multi-scale satellite imagery

Chunyuan Diao [a,*], Carol K. Augspurger [b], Yilun Zhao [a], Carl F. Salk [c,d]

[a] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[b] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[c] Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences, Alnarp, Sweden
[d] Institute for Globally-Distributed Open Research and Education, Gothenburg, Sweden

ARTICLE INFO

*Keywords:*
Spring phenology
Forest community
Remote sensing
Field observations
Phenological indicator

ABSTRACT

Forest phenology, as a sensitive indicator of a forest's response to climate change and variability, has long been monitored using remote sensing, yet has seldom been interpreted or validated with spatially compatible, community-level field phenological observations. In temperate deciduous forests, multiple spring phenological phases are critical for modeling carbon storage and biogeochemical cycles. The simultaneous detection of all these critical phenological phases at the community level remains a long-standing challenge. To tackle these challenges, the objective of this study is to develop a novel satellite-field phenological bridging framework for characterizing all key spring phenological phases of a fragmented deciduous forest using multi-scale satellite time series. The framework consists of four key components: deep learning-based spatiotemporal image fusion, satellite-based forest phenology modeling, satellite-based forest phenological metric extraction, and field-based forest community phenological characterization. With the devised framework, we extract a total of 24 satellite phenological metrics of Trelease Woods, a forest fragment near Urbana, IL, USA. From weekly field phenological observations of 148 canopy trees of 15 common species of the forest community, we devise three summative field phenological indices to characterize community-level phenological states in spring. Under the satellite-field bridging framework, events during each key spring forest phenological phase (i.e., bud swell, budburst/leafing out, leaf expansion, and leaf maturity) are successfully detected using the fusion imagery (MAE from 1.1 to 2.9 days and bias from -2.4 to 1.5 days). However, the satellite detection of the earliest field events may be influenced by understory plants, soil background, and snow. The subsequent multi-scale, satellite phenological analysis underscores the importance of taking into account spatial scale and representation from both satellite and field phenological perspectives in building corresponding bridging relationships. Among the extracted pheno-metrics, bridging the threshold-based metrics to field phenological indices results in the highest accuracy (MAE less than 3 days and bias less than 2 days). The strong agreement among the field indices demonstrates the effectiveness of our field phenological surveying approach in generating community-wide forest phenological representation. Our study innovatively scales up the field phenological observations from the individual trees to the species to the community level, and the devised framework enables accurate retrieval of all key phenological events of community-wide, spring canopy development of the forest fragment.

## 1. Introduction

As a sensitive indicator of climate and environmental change, vegetation phenology has been widely studied to understand ecosystem functions and biosphere-atmosphere interactions under the changing environment (Chmielewski and Rötzer, 2001; Fitchett et al., 2015; Kramer et al., 2000; Parry, 2007; Piao et al., 2019; Xie et al., 2015). This seasonal dynamic of vegetation plays a critical role in regulating ecosystem carbon-energy-water cycling, predicting ecosystem changes at local to global scales, as well as assessing biosphere response and

* Corresponding author.
*E-mail address:* chunyuan@illinois.edu (C. Diao).

https://doi.org/10.1016/j.isprsjprs.2024.03.018
Received 22 November 2023; Received in revised form 24 February 2024; Accepted 28 March 2024
Available online 6 April 2024

feedback to the climate system (Caparros-Santiago et al., 2021; Gunderson et al., 2012; Peñuelas et al., 2009; Richardson et al., 2013; Tang et al., 2016; White et al., 2005). Over the past half century, the warming trends of climate have shifted vegetation phenological timing (e.g., earlier spring and extended growing season) in many temperate forest regions (Badeck et al., 2004; Cleland et al., 2007; Linderholm, 2006; Menzel and Fabian, 1999), with significant implications for ecosystem productivity (Keenan et al., 2014; Richardson et al., 2010), surface energy balance (Richardson et al., 2013; Schwartz and Crawford, 2001), biogeochemical cycling (Keenan et al., 2014; Piao et al., 2007; White et al., 1999), disturbance regimes (Dukes et al., 2009), trophic interactions (Both et al., 2009; Post et al., 2008), and species ranges (Chuine, 2010; Chuine and Beaubien, 2001). To characterize forest ecosystem response to a changing world, monitoring forest phenological dynamics over space and time is crucial.

Remote sensing has been increasingly utilized for spatially explicit monitoring of forest phenology at large scales, owing to its repeated viewing, synoptic coverage, and consistent measurement (Berra and Gaulton, 2021; Delbart et al., 2006; White et al., 2014). The satellites with high temporal resolutions (e.g., daily revisiting frequency), such as Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), and Visible Infrared Imaging Radiometer Suite (VIIRS), have been conventionally employed to characterize vegetation phenological transition dates (e.g., start of growing season) via the time series of vegetation index (VI) (Ahl et al., 2006; Diao, 2019a; Hmimina et al., 2013; Liu et al., 2015; Xiao et al., 2006; Zhang et al., 2003, 2017). The characterization of vegetation phenology usually involves the preprocessing of satellite VI time series to remove outliers and off-season vegetation cycles (Hird and McDermid, 2009; Zeng et al., 2020), fitting of preprocessed time series via mathematical functions (e.g., double logistic function, asymmetric Gaussian function, harmonic function, and wavelet function) (Atkinson et al., 2012; Beck et al., 2006; Jonsson and Eklundh, 2002; Sakamoto et al., 2005), and retrieval of phenological transition dates via phenometric extraction methods (e.g., threshold-based methods, curve feature-based methods, and pheno-network methods) (Diao, 2019a, 2019b; Tan et al., 2010; Zeng et al., 2020; Zhang et al., 2003; Gu et al., 2009; White et al., 2009). Under the phenological characterization scheme, several vegetation phenological products (e.g., VIPPHEN phenology product, MODIS land cover dynamics product, and VIIRS land surface phenology product) have been developed, further improving our understanding of vegetation phenological dynamics under climate change and environmental disturbance (Didan and Barreto, 2016; Ganguly et al., 2010; Zhang et al., 2018). Yet the characterized phenology from these satellites is usually at coarse spatial resolutions (e.g., 500 to 5600 m), with multiple land covers (e.g., forest, agriculture, and water body) as well as mixture of tree species of distinct phenology possibly in one pixel, which makes the characterization of forest phenology challenging, particularly in heterogenous and fragmented landscapes. The advent of higher spatial resolution (i.e., 30 m) Harmonized Landsat Sentinel-2 (HLS) opens up new opportunities to retrieve phenology of a forest community (Bolton et al., 2020; Moon et al., 2021), e.g., the integrated timing of phenological events among individuals and species within a forest community. Given frequent cloud cover and rainy conditions accompanying temperate forest green-up in spring, the unavailability of HLS images can be a significant limiting factor. The limited satellite images during spring, particularly in critical phenological transition windows, make it challenging to characterize community-level forest phenological dynamics and interpret satellite characteristic phenology in the context of field-observed forest phenological events.

The bridging of satellite- and field-based spring phenological measures of forest communities has been a long-standing challenge in remotely sensed phenological validations (Berra and Gaulton, 2021; Donnelly et al., 2022). The vegetation phenology estimated by remote sensing is fundamentally different from that observed at ground level. In contrast to field-observed phenological phases (e.g., bud burst and leaf unfolding phases) that have clear organismal changes with biophysical significance, the satellite-derived phenology is based mostly on the characterization of temporal variation in remotely sensed observations of surface greenness that may have equivocal biophysical meaning (Henebry and De Beurs, 2013; Schwartz, 2003). Also the timing of a phenological phase detected from different remote sensing data, preprocessing approaches, and pheno-metric extraction methods may vary extensively (Diao, 2020; Diao and Li, 2022; Schwartz et al., 2002; White et al., 2014, 2009). For instance, differences of up to 2 months between the timing of satellite-derived start of season (SOS) and that of ground phenology have been observed over broad regions of North America, due to the variety of pheno-metric extraction methods utilized (White et al., 2009). The intrinsic discrepancies between satellite- and field-phenological measures raise the challenge in interpreting and validating forest community phenology retrieved by remote sensing.

Studies seeking to validate remotely sensed forest spring phenological measures with field-based observations vary in terms of methodology and quality of community-based data of field measurements (Berra and Gaulton, 2021; Schwartz and Hanes, 2010). Field methods for studies of temperate forests include using 1) Phenocam near-surface camera observations (Klosterman et al., 2014; Richardson et al., 2018, 2009); 2) National Phenology Network (NPN) providing crowd-sourced information (Peng et al., 2017); and 3) direct field observations by one individual (Fisher and Mustard, 2007; Liang et al., 2011, 2014). Some field studies are less accurate because they sample a limited area and number of trees (or tree species) due to the scope of the camera or the labor-intensive nature of season-long observations. Studies are limited by the number of phenological phases (mostly budburst/leafing out in spring) compared, by a study site with a range in elevation and complex topography, and by including non-homogeneous forests (e.g., mixed deciduous and evergreen species). Others lack a forest census to enable weighting of a phenological phase of each tree species by its relative basal area or crown size. The lack of appropriate field representation of forest phenology at the community level creates a major barrier to understanding what satellite phenological measures represent. Intensive systematic and representative field phenological observations of tree species of known relative basal area or crown size within a forest community are necessary to compare satellite- and field-based phenological measures at the community level. Ideally, the entire span of spring forest phenological phases (e.g., bud swell, bud burst, leaf expansion, and leaf maturity) are incorporated, as these phases together are critical to quantify a forest's phenological development rate and have important implications for ecosystem functions, including carbon storage and biogeochemical cycling (Donnelly et al., 2017).

The overarching goal of this study is to develop a novel satellite-field phenological bridging framework for characterizing a diversity of spring phenological phases of a fragmented deciduous forest using multi-scale satellite time series. Using Trelease Woods, a deciduous forest fragment, as the case study, we specifically aim to 1) fuse multi-scale satellite imagery to retrieve a wide range of satellite phenological metrics of the fragmented forest, 2) quantify community-level field phenological progress of the forest with intensive field phenological observations, and 3) bridge the remotely sensed and field-based phenological measures at the community level for all key forest spring phenological phases.

## 2. Study area and data

### 2.1. Study area

The 12-ha study site is the North half of Trelease Woods, a 24.5 ha temperate deciduous mature forest fragment near Urbana, Illinois (IL), USA (40°09' N, 88°10' W) (Fig. 1). The fragment has been protected and managed as a natural area owned by the University of Illinois since 1915. It is embedded in an agricultural landscape dominated by corn and soybeans. Immediately adjacent to the site is restored prairie on N,

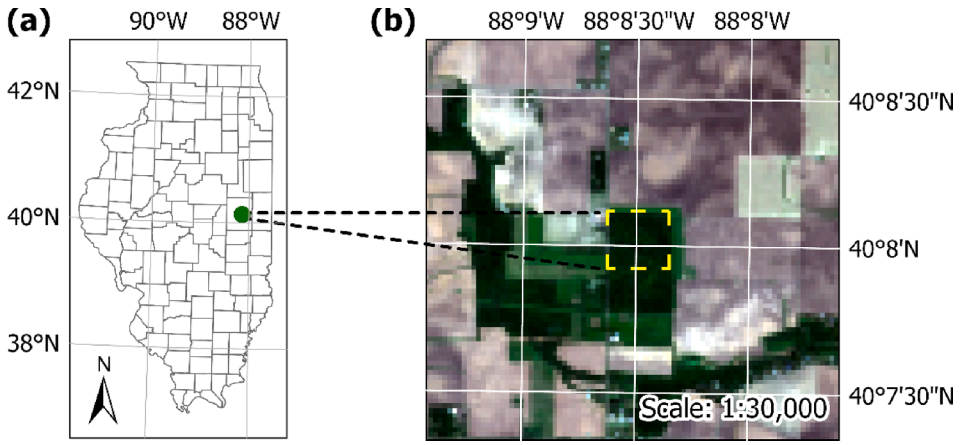


**Fig. 1.** (a) The location of study site in Illinois, USA. (b) The surrounding environment of the forest fragment study site. The green point in (a) and the dashed rectangle in (b) denote the study site.

E, and S aspects, and a road on the W. Topography is uniform with elevation varying by < 5 m. Among the 20 canopy tree species, the dominant species based on basal area are *Acer saccharum* Marsh., *Celtis occidentalis* L., and, until recently, *Fraxinus americana* L. The study period for monitoring spring phenology ranges from 2016 to 2020.

### 2.2. Remote sensing data

Taking into account the forest's fragmentation and size of our study site, we acquired the HLS images (tile 16TCK) covering the spring phenological period (i.e., images from day of year [DOY] 1 through 250) from 2016 to 2020. The HLS data are generated via the harmonization of Landsat 8 and Sentinel-2 data, which comprises imagery atmospheric correction, cloud masking, spatial co-registration, bidirectional reflectance distribution function (BRDF) adjustment, band pass adjustment, and temporal compositing (Claverie et al., 2018). The HLS data have a 30-m spatial resolution and 2-3 days temporal revisiting frequency. To ensure the quality of HLS image time series, the pixels contaminated by water, snow/ice, cloud, and cloud shadow in the study site are masked based on corresponding quality control layers. Only HLS images with more than 90% clear pixels of the study site are selected for further analysis. Due to the frequent cloud cover and rainy conditions in spring, the availability of HLS images during critical phenological transition periods is quite limited (Table S1), which makes it challenging to retrieve forest phenology with HLS data. As a result, we further fuse the HLS images (Table S1) with MODIS images to generate temporally dense fusion imagery with daily temporal revisiting frequency and 30-m spatial resolution.

The MODIS images for generating the fusion data are acquired from the MODIS MCD43A4 nadir BRDF-adjusted reflectance product (tile h11v04 and h11v05; version 6.1). This BRDF-adjusted product removes view angle effects using a semi-empirical BRDF model to retrieve the nadir surface reflectance, which can eliminate the influence of BRDF on leaf phenology detection (Galvão et al., 2011). It has 500-m spatial resolution with daily temporal resolution. For each date, the surface reflectance is temporally composited using all the images acquired from both Terra and Aqua during its 16-day compositing period in terms of image quality, temporal distance, and spatial coverage, which also diminishes the influence of atmospheric noise and cloud contamination. The MODIS MCD43A4 observations contaminated by snow are further removed using the snow layer of the MODIS MOD10A1 product (i.e., Terra Snow Cover Daily Global 500m product). The satellite observations that have significantly higher or lower values than neighboring observations (i.e., beyond three standard deviations of reflectance) are also masked. The pre-processed MODIS images are bilinearly resampled to the spatial resolution of HLS (30 m), and the resampled MODIS and HLS images are further spatially aligned. The blue, green, red, near-infrared (NIR), and two shortwave-infrared (SWIR) bands from both MODIS and HLS images are employed in the fusion process. Two vegetation indices, normalized difference vegetation index (NDVI) and 2-band enhanced vegetation index (EVI2), widely used in analyzing vegetation seasonal dynamics, are further calculated from the fusion imagery to build the time series for the following phenological analysis.

Given the size of the fragmented forest of our study site, we further assess multi-scale satellite time series (i.e., 500-m, 250-m, and 30-m scales) for characterizing spring phenology of the forest community. Specifically, we utilize the pre-processed MODIS MCD43A4 data for retrieving forest phenology at the 500-m scale; the study site is contained within one pixel of the MODIS MCD43A4 imagery. At the 250-m scale, we acquire the 8-day MODIS MOD09Q1 Terra surface reflectance product (version 6.1) for the study site. It is temporally composited using the corresponding images during the 8-day composite period in terms of cloud cover, solar zenith angle, and atmospheric interference. The observations contaminated by snow, cloud, and cloud shadow are eliminated using the corresponding quality control layers. The MODIS MOD09Q1 pixels totally within the study site are utilized; they occupy about two-thirds of the study area. For the 30-m scale, we further compare the forest community phenology characterized from the MODIS-HLS fusion data with that from the HLS-wise data, to assess the role of image fusion in phenological monitoring of the forest fragment. Given the limited availability of HLS data, the partially cloud-free HLS and Landsat 7 images of the study site are also utilized to construct time series for phenological monitoring under the HLS-wise scenario.

### 2.3. Field phenological observations

Phenological observations were made on 148 individual canopy trees located haphazardly and selected to include the 15 most common species, representing 96.3% of total basal area of trees > 22.9 cm diameter at breast height (DBH) (Table S2). Relative basal area for living species is adjusted as *Fraxinus americana* and *F. pennsylvanica* died during the study period. Sample size varies among species (mean = 9 individuals per species; range = 1-18).

The predominant phenological status of each canopy tree was determined by the same observer (CKA) using binoculars weekly from mid-February-June from 2016-2020. On a given census date, the phenological event recorded represented an appraisal of the tree crown as a whole (i.e., the predominant status of all developing vegetative units). Within each of three vegetative 'phases' (i.e., bud swell, budburst/leafing out, and leaf expansion), DOY was noted for three 'events' (i.e., when 1/3, 2/3, and 3/3 of the development phase was completed). The 10th event, DOY of leaf maturity, was also estimated as follows. The

mean number of days per species from DOY of end of leaf expansion to DOY of leaf maturity, based on observations in 1995-1999, was added to DOY of end of leaf expansion for each individual in each species in each study year of 2016-2020. DOY of these 10 phenological 'events' was used in analyses (Table 1). If not directly observed because of changes in the speed of development, an estimate of DOY, based on linear interpolation between DOYs with observations, was made for each non-observed event.

Daily temperature data for the study period at the Champaign, IL Weather Station (3S), 8 km SW of the study site at nearby Champaign, were collected by the National Weather Service Cooperative Observer Program (US-COOP, https://www.weather.gov/coop/) and data were obtained through the Midwestern Regional Climate Center (http://mrcc.purdue.edu/).

## 3. Methods

To characterize the community-level spring phenology of the fragmented deciduous forest, the satellite-field phenological bridging framework devised in this study consists of four key components: deep learning-based spatiotemporal image fusion (Section 3.1.1), satellite-based forest phenology modeling (Section 3.1.2), satellite-based forest phenological metric extraction (Section 3.1.3), and field-based forest community phenological characterization (Section 3.2) (Fig. 2).

### 3.1. Satellite-based forest community phenological characterization

#### 3.1.1. Deep learning-based spatiotemporal image fusion

In this study, we employ the hybrid deep learning-based image fusion model to create fusion imagery with daily temporal resolution and 30-m spatial resolution through blending MODIS and HLS images (Yang et al., 2021). As a spatiotemporal image fusion model, the hybrid deep learning model is chosen in this study due to its advantages in robustly modeling different levels of temporal reflectance changes among the images in fragmented landscapes. On each prediction date, the model generates the fusion image using the corresponding MODIS image on that date and two paired MODIS-HLS images of adjacent dates, with an image pair being the MODIS and HLS images obtained on the same date.

The hybrid deep learning model synthesizes two deep learning structures: super-resolution convolutional neural network (SRCNN) and long short-term memory (LSTM) (Fig. 3). The SRCNN includes three convolutional components: feature extraction, non-linear mapping, and reconstruction. Feature extraction of SRCNN learns salient spatial and texture features from MODIS images with convolution operations, generating the coarse MODIS feature maps (i.e., the first hidden layer). Non-linear mapping of SRCNN leverages convolution operations to map the spatial features of MODIS images to corresponding features of HLS images, and produces the fine HLS feature maps (i.e., the second hidden layer). Also, the satellite sensor-induced spectral reflectance differences are taken into account in non-linear mapping. Reconstruction of SRCNN restores spatially-degraded details in the MODIS images with convolution operations and constructs the super-resolution (SR) images at the HLS scale. Using MODIS-HLS image pairs as the training data, the SRCNN model can harmonize the spectral reflectance of MODIS and HLS images, while rebuilding degraded spatial structures and details of MODIS images in the resulting SR images. The LSTM includes two LSTM layers, each containing 100 LSTM units, to learn temporal and phenological changing patterns from a sequence of images for predicting the fusion image. Each LSTM unit includes a memory cell and three gates (i.e., input gates, forget gates, and output gates) for regulating the learning process. The gating structures help selectively retain or remove memory information in the LSTM unit over time. Following the two LSTM layers, a dropout layer is devised to overcome the overfitting of the model, and a fully connected layer is employed to predict the fusion imagery. During the training process, the LSTM model learns phenological changing patterns from the SR image sequence produced by the SRCNN. By leveraging these learned temporal phenological patterns, the LSTM model then constructs the final fusion image on the prediction date with corresponding HLS images during the prediction process. The description of the workflow of the model is in Supplementary Text S1, and the detailed settings of the model structure and parameters can be found in our previous study (Yang et al., 2021).

The hybrid deep learning model can effectively characterize the spatial relationship between MODIS and HLS images, as well as temporal evolving patterns in spectral reflectance of a sequence of images. With its capability in predicting varying levels of phenological dynamics in fragmented landscapes, the model is suited for constructing fusion images of our fragmented study area when the forest is experiencing rapid phenological changes in spring. Utilizing the hybrid deep learning model, we construct 30-m fusion images on a daily basis for our fragmented forest study site, spanning DOY 1-250 from 2016 to 2020. We then retain only the fusion images of relatively high quality in consideration of the source image quality (i.e., MODIS and HLS image), the temporal distance between image pair and prediction dates, and the spatial and temporal fusion coherency and representation. With the retained fusion images, we calculate the community-average NDVI and EVI2 of our forest study site. The community-average NDVI and EVI2 time series are utilized to model the fragmented forest spring phenology and retrieve corresponding phenological metrics from the satellite perspective.

#### 3.1.2. Forest phenology modeling

For both NDVI and EVI2, the satellite time series are first scrutinized visually to remove outliers whose values are drastically different from those of adjacent observations. After the visual assessment, we employ a night filter to eliminate the anomalous observations lower than a certain threshold (a threshold of 0.2 for NDVI and 0.1 for EVI2 based on corresponding off-season values) (Filippa et al., 2016). Those anomalously low values are usually caused by limited light conditions or residual atmospheric interferences. Then we utilize a spline filter to remove the spurious or abnormal observations outside the predetermined residual boundary via recursive spline smoothing and residual computation. Specifically, we fit the satellite time series with a cubic smoothing spline function and calculate the residuals between the fitted and observed values. The satellite observations with the residuals outside the boundary of $\mu \pm 3\sigma$ ($\mu$ and $\sigma$ denote the mean and standard deviation of the residuals, respectively) are flagged as spurious observations and thus are discarded. This spline filter is utilized iteratively to remove those spurious observations until no outlier is detected. Both the night and spline filters are devised to improve the quality of satellite time series by removing outliers, possibly due to source image quality (i.e., MODIS and HLS image), residual atmospheric interferences, and data fusion.

**Table 1**
The number, code, and description of 10 phenological events used in analyses.

| Phenology Number | Code | Description |
|---|---|---|
| 1 | B1 | Bud Swell (bud is 1/3 of final size before it bursts) |
| 2 | B2 | Bud Swell (bud is 2/3 of final size) |
| 3 | B3 | Bud Swell (bud is final size) |
| 4 | E1 | Bud Burst (bud has burst; first leaf tips visible beyond bud scales) |
| 5 | E2 | Leafing Out (individual leaf blade and stem visible) |
| 6 | E3 | Leafing Out (individual leaf blade and stem visible; leaf unfolding) |
| 7 | F1 | Leaf Expansion Begins (leaf fully unfolded and 1/3 of final size) |
| 8 | F2 | Leaf Expansion (leaf is 2/3 of final size) |
| 9 | F3 | Leaf Expansion Ends (leaf is final size but not final color) |
| 10 | M | Leaf Maturity (leaf is hardened with final deep green color) |

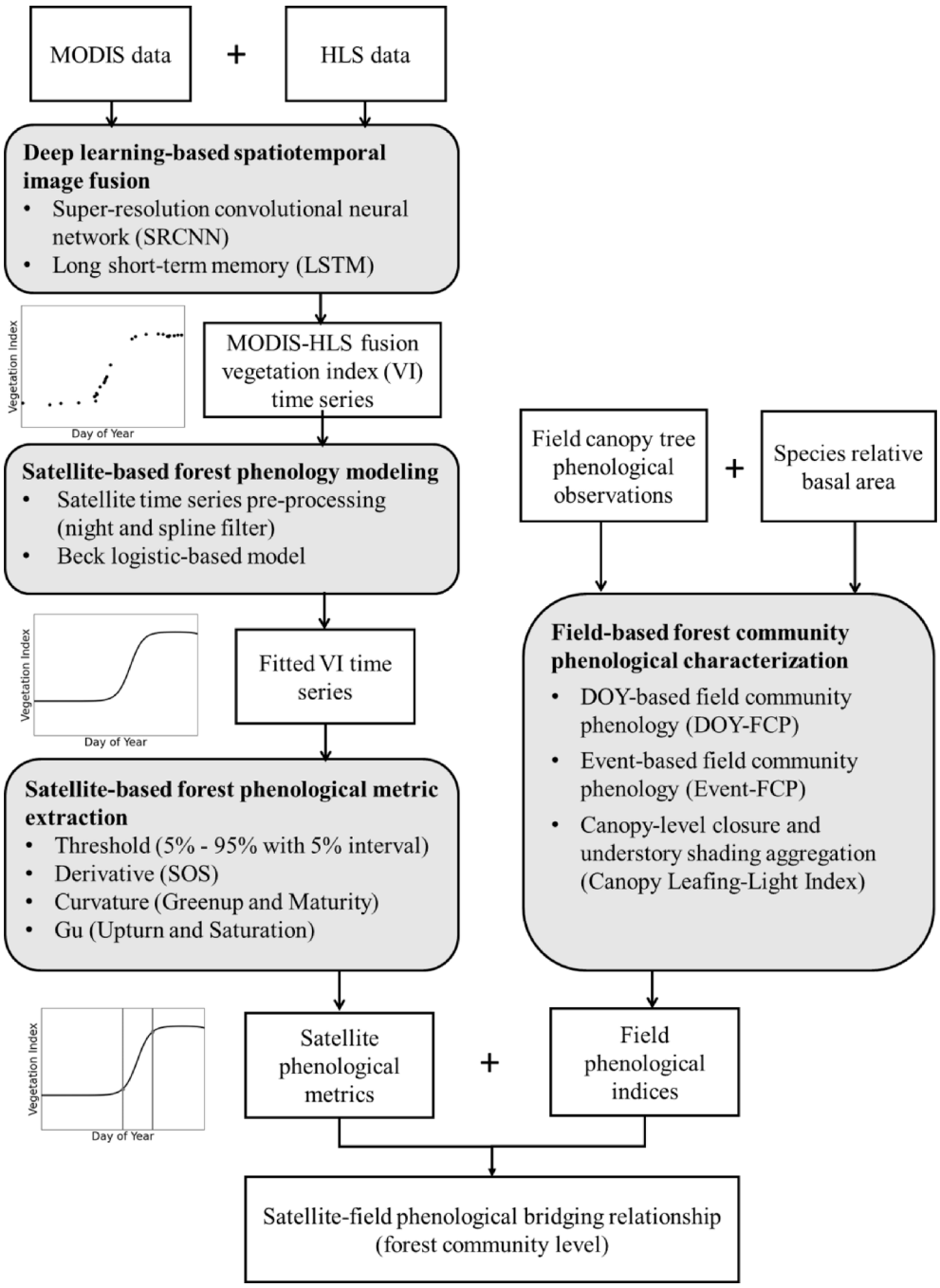


**Fig. 2.** The satellite-field phenological bridging framework for monitoring community-level spring phenological events of the fragmented deciduous forest.

Following the night and spline filters, any gaps in the satellite time series are filled using linear interpolation.

After pre-processing satellite (NDVI or EVI2) time series, we utilize the Beck logistic-based model to fit the time series for tracking the spring leaf phenological trajectory of the forest community (Eq. (1)) (Beck et al., 2006). The Beck model is selected owing to its enhanced capability in capturing rapid spring phenological changes and diminishing the influence of off-season abnormal observations, as well as the phenological implications of its model parameters (Diao, 2020; Diao and Li, 2022; Zhang et al., 2003).

$$f(t) = m_{base} + (m_{max} - m_{base}) \bullet \frac{1}{1 + e^{(-n_1 * (t - n_2))}} \quad (1)$$

where $t$ is the date of year and $f(t)$ is the fitted VI value at time $t$. $m_{base}$ and $m_{max}$ denote the initial background and peak VI values, respectively. $n_1$ is the increase rate of the fitted curve at the curve inflection point ($n_2$). The initial background VI value is calculated as the median VI value of clear observations during the vegetative dormant phase before the beginning of season. To minimize the adverse influence of abnormal observations, particularly due to cloud, snow, and/or ice, the VI values lower than the initial background VI value are substituted with the background value. The other parameters of the Beck model are estimated with iterative non-linear least squares.

### 3.1.3. Forest phenological metric extraction

After fitting the satellite time series with the Beck model, we employ four pheno-metric extraction methods to retrieve a range of satellite phenological metrics. The pheno-metric extraction methods include

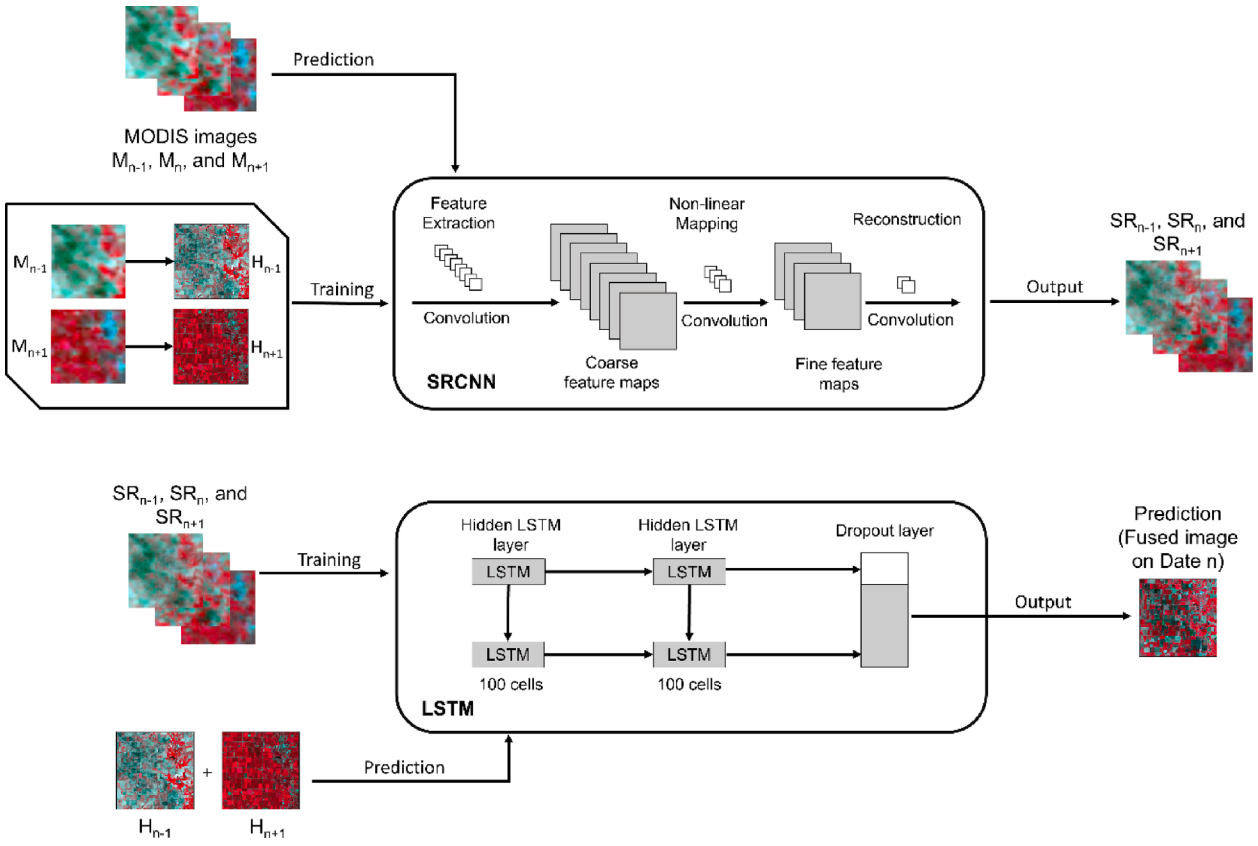


**Fig. 3.** Schematic of the hybrid deep learning model with structures of super-resolution convolutional neural network (SRCNN) and long short-term memory (LSTM) models. Detailed description of the workflow of the model is in Supplementary Text S1.

threshold-, derivative-, curvature-, and Gu-based methods (Fig. 4) (Filippa et al., 2016; Gu et al., 2009; Zeng et al., 2020; Zhang et al., 2003).

As one of the most widely used pheno-metric extraction methods, the threshold-based method extracts the phenological metrics using the absolute or relative thresholds of satellite time series. Compared to absolute thresholds (e.g., VI being 0.5), relative thresholds (e.g., VI being 50% of curve amplitude) have been found to be more robust and perform better across ecosystems and land cover types, and are thus utilized in this study (Berra and Gaulton, 2021; White et al., 1997; Zeng et al., 2020). Given a diversity of field phenological phases (e.g., bud swell, budburst/leafing out, leaf expansion, and maturity) considered in this study, we extract a variety of satellite phenological metrics using extensive relative thresholds, ranging from 5%-95% of curve amplitude with an interval of 5%. This range provides a total of 19 threshold-based metrics during the spring season.

Three curve feature-based pheno-metric extraction methods (i.e., curvature-, derivative-, and Gu-based methods) are also utilized for phenological metric retrieval via the distinct curve characteristics (Fig. 4). The curvature-based method is employed to retrieve two phenological metrics, namely Greenup and Maturity, according to the change rate of curvature of satellite time series. Specifically, Greenup and Maturity metrics are extracted based on the dates of the first and second local maxima of the change rate of curvature, respectively. The derivative-based method is used to retrieve the start of the season (SOS) metric, using the maximum of the first derivative of satellite time series in spring. The Gu-based method is employed to extract another two phenological metrics, namely Upturn and Saturation, based on the curve intersections. The Upturn metric is retrieved via the intersection of the recovery line and the baseline. The recovery line is tangent to the Beck-fitted curve at the maximum of the first derivative of the satellite time series, and the baseline is a horizonal line with its value being the initial background VI value. The Saturation metric is retrieved via the intersection of the recovery line and the max line, and the max line is a horizontal line with its value being the maximum VI value. These curve feature-based phenological metrics are characteristic of distinct curve properties caused by noticeable changes in leaf biochemical content (e.g., chlorophyll content) and/or biophysical structures, usually stemming from leaf phenological changes.

In total, we retrieve 19 threshold-based phenological metrics (i.e., 5%-95% thresholds with an interval of 5%) and five curve feature-based metrics (i.e., curvature-based Greenup and Maturity, derivative-based SOS, and Gu-based Upturn and Saturation). This diverse set of satellite metrics characterize the forest community phenological development remotely, and will be compared with the field community phenology indices (Section 3.2) for satellite-field phenological bridging of all the critical spring phenological phases (i.e., bud swell, budburst/leafing out, leaf expansion, and maturity events) at the community level.

### 3.2. *Field-based forest community phenological characterization*

Three summative field indices are calculated separately for each year to quantify community-level phenological states throughout canopy development in spring. These field indices take into account the phenological variations both within and between species, together with species composition and relative basal area, for scaling up the phenological observations from the individual trees to the species to the community level. In each index, the contribution of a given species is considered proportional to its relative basal area, based on a complete census of all trees >22.9 cm at 1.4 m height in the N half of Trelease Woods in 2005 (J. Edgington, unpublished data).

1. *DOY-based field community phenology (DOY-FCP).* Using the event status for each individual, we calculate the mean event status per species for each day with a change in event status, weight it by its relative basal area, and then sum for all species for each day with a change in event

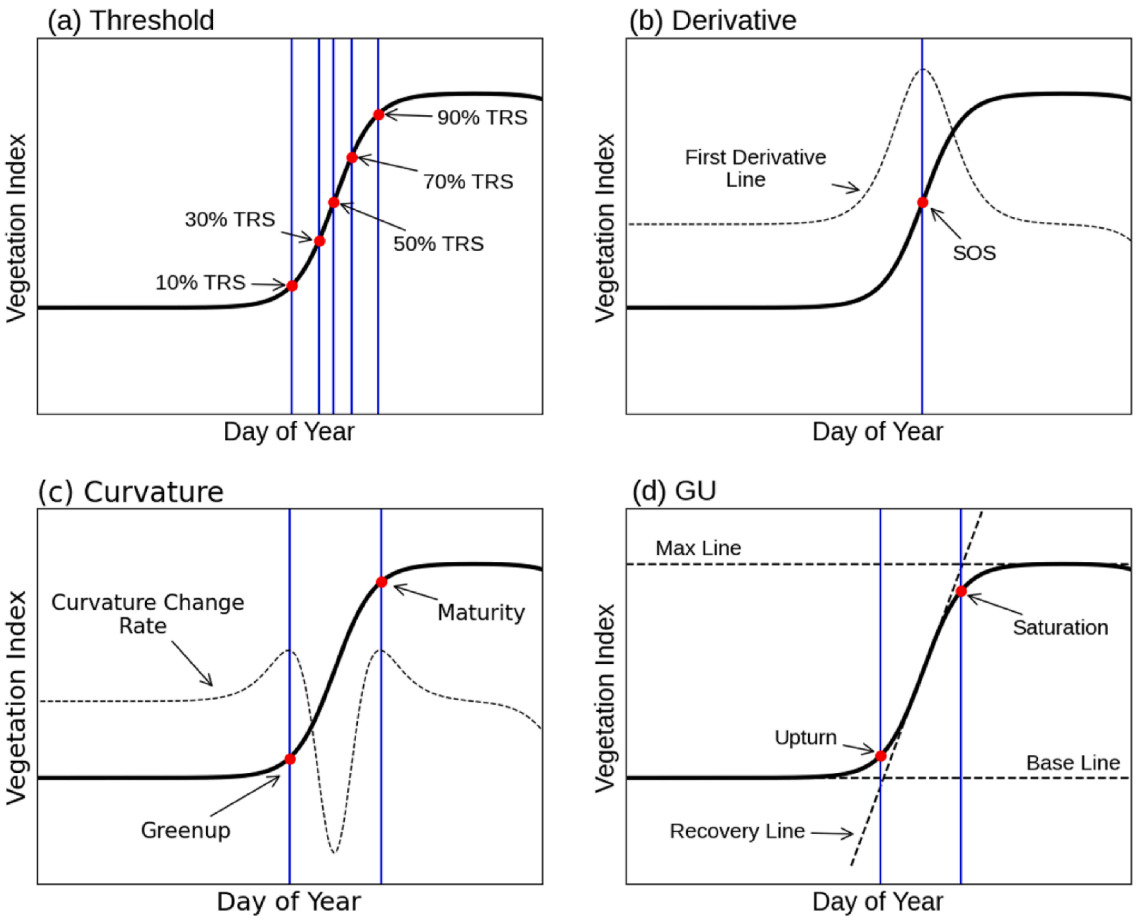


**Fig. 4.** Satellite phenological metric extraction using (a) threshold-based method, (b) derivative-based method, (c) curvature-based method, and (d) Gu-based method. For the threshold-based method, only five thresholds (i.e., 10%, 30%, 50%, 70%, and 90% threshold [TRS]) are shown for illustration purposes.

status (Eq. (2)).

$$ES(community, DOY) = \sum_{b=1}^{B}\left[\frac{\sum_{t_b=1}^{T_b} ES(t_b, DOY)}{T_b} * RBA(b)\right] \quad (2)$$

where $ES(community, DOY)$ is the event status of community on a given DOY. $B$ is the total number of species and $T_b$ is the total number of individuals of species $b$. $ES(t_b, DOY)$ is the event status of the individual $t_b$ of species $b$ on a given DOY. $RBA(b)$ is the relative basal area of species $b$.

This comprehensive method utilizes both observed and estimated event days to generate one community value per day (i.e., $ES(community, DOY)$) for the duration of spring phenology. DOY-FCP results from a sequential aggregation of data: 1) Event status for each individual tree of a given species is aggregated to a mean event status of that species on a given DOY. 2) Each species' mean event value is multiplied by its relative basal area to yield that species' relative event status; 3) all species' relative event values are added together to yield the event status of the forest community. With DOY-FCP, a logistic curve is further generated to represent phenological development of the community in detail throughout spring.

2. *Event-based field community phenology (Event-FCP)*. Using the first day for each of 10 events for each individual, we calculate the mean day per species for each event, weight it by a species' relative basal area, and then sum for all species for each of 10 events (Eq. (3)). Values between weekly census days arise because of estimation of days for events not directly observed. This method generates one simplified, summative value for the community for each event.

$$DOY(community, ES_X) = \sum_{b=1}^{B}\left[\frac{\sum_{t_b=1}^{T_b} DOY(t_b, ES_X)}{T_b} * RBA(b)\right] \quad (3)$$

where $DOY(community, ES_X)$ is the community DOY of event $X$ ($X$ ranges from 1 to 10). $B$ is the total number of species and $T_b$ is the total number of individuals of species $b$. $DOY(t_b, ES_X)$ is the DOY of the individual $t_b$ of species $b$ reaching the event $X$. $RBA(b)$ is the relative basal area of species $b$.

Comparable to DOY-FCP, Event-FCP also results from a sequential aggregation of data: 1) DOY for each individual tree of a given species is aggregated to a mean DOY of that species for a given event. 2) Each species' mean DOY value is multiplied by its relative basal area to yield that species' relative DOY; 3) all species' relative DOY values are added together to yield community-level DOY for that event. Both DOY-FCP and Event-FCP are community-level field phenological indices to characterize the variability in canopy phenology events of the forest community in the study area. These indices incorporate both intra- and interspecific phenological variation, along with species composition and relative basal area, for generating community-wide forest phenological representation. Each of these two types of field community phenology indices then is compared to the corresponding satellite-derived phenological metrics.

3. *Canopy-level closure and understory shading estimates (Canopy Leafing-Light Index)*. We further characterize community-wide canopy development based on canopy light transmission to the understory, dominated by *Asimina triloba* and saplings of *Acer saccharum* and *Aesculus glabra* in the study site. Under-canopy light levels are related to visible phenology of canopy trees (Schwartz et al., 2013). This Canopy

Leafing-Light Index has daily values from the perspective of an individual sapling in the understory, based, in part, on events from the beginning of bud burst (E1) through end of full expansion (F3) of canopy trees. It starts with Begin Bud Burst (E1), not Begin Bud Swell (B1), as initial calculations showed that the bud swell phase of canopy trees does not change canopy light transmission. This index is estimated from ground-based observations of canopy tree phenology which are aggregated to the community level and matched to observations at sapling heights of the proportion of above-canopy irradiance transmitted to the understory (I/Io, where I = irradiance reaching understory individuals and Io = irradiance in the open) (Augspurger et al., 2005). The Canopy Leafing-Light Index adds a new perspective to the exploration of what biophysical characteristics of phenology can be captured by remote sensing. It represents the perspective of seasonal changes in understory light availability due to canopy closure above.

Calculating the index is a four-step process. First, ground-level solar radiation measurements at 1- or 3-minute intervals are summed to total daily solar radiation from a nearby NOAA Surface Radiation Budget (SURFRAD) site. Second, a baseline estimate is developed that is the same every year to account for light interception by branches and trunks. Wintertime I/Io values are fit to a function with a maximum at the summer solstice and minimum at the winter solstice, and symmetrical around the summer solstice to account for changes in trunk and branch light interception due to seasonal changes in sun angle. Third, contributions of different leafing events (i.e., E1 through F3) to shading are estimated by optimizing the fit of aggregated canopy phenology (I/Io values). Event-specific estimates are iteratively improved to reach a near-optimal estimate for each event. Fourth, the interception of light by trunks and branches is combined with leafing phenology to determine daily total percentage interception of incoming solar radiation. The relative basal area-weighted functions of each species are then summed on each census day to create a community-wide canopy function that summarizes the state of the entire forest's canopy from the first day of bud burst (E1) to the last day of leaf expansion (F3). Interpolation is used to fill in days lacking observations. Full details of this process are described in Augspurger and Salk (in review).

### 3.3. Accuracy assessment

To evaluate the performance of the hybrid deep learning model in generating the fusion imagery, we compare the fusion results with corresponding HLS images on the HLS image acquisition dates. Specifically, root mean square error (RMSE), erreur relative global adimensionnelle de synthese (ERGAS), and spectral angle mapper (SAM) are used as the accuracy metrics, as these spectral-based metrics have high correlation to the accuracy of reconstructed phenology (Yang et al., 2021; Zhu et al., 2022). RMSE quantifies the band-specific mean spectral reflectance difference between the fusion and corresponding HLS images. ERGAS accommodates the spectral reflectance difference across bands, as well as the resolution difference between MODIS and HLS images to quantify the mean spectral reflectance difference between the fusion and corresponding HLS images across bands, with a lower value indicating more comparable fusion results (Eq. (4)).

$$ERGAS = 100\frac{R_h}{R_l}\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left[\frac{RMSE(\widehat{L}_n)^2}{\mu_n{}^2}\right]} \tag{4}$$

where $R_h$ and $R_l$ are the spatial resolutions of HLS and MODIS images, respectively. $N$ is the number of bands. $\widehat{L}_n$ is the estimated spectral reflectance of band $n$, and $\mu_n$ is the mean reflectance of band $n$.

SAM assesses the mean spectral reflectance similarity between the fused and corresponding HLS images across bands via spectral angels (Eq. (5)). The spectral reflectance of each pixel of imagery is represented as an N-dimensional spectral vector (N is the band number of imagery), and the mean spectral angle between the fusion and HLS reflectance of all the pixels in the N-dimensional space is then calculated as SAM. Similar to RMSE and ERGAS, a lower SAM value denotes more spectrally similar fusion results with less spectral distortion. As red and NIR bands are used for deriving NDVI and EVI2, the fusion accuracy metrics are calculated using these two bands.

$$SAM = \frac{1}{M}\sum_{m=1}^{M} arccos\frac{\sum_{n=1}^{N}\left(\widehat{L}_m^n L_m^n\right)}{\sqrt{\sum_{n=1}^{N}\left(\widehat{L}_m^n\right)^2\sum_{n=1}^{N}\left(L_m^n\right)^2}} \tag{5}$$

where $M$ is the number of pixels in the fusion image and $N$ is the number of bands. $\widehat{L}_m^n$ and $L_m^n$ denote the reflectance of pixel $m$ in band $n$ of the fusion and HLS images, respectively.

To assess the performance of the Beck model in fitting the VI satellite time series derived from the fusion imagery, we calculate the difference between fitted and observed VI values using RMSE. The community-level satellite-field phenological bridging relationship of the fragmented forest is assessed using mean absolute error (MAE) and bias. Specifically, the timing of each field phenological event of the forest community is compared with that of all the satellite metrics using MAE and bias during the study period; the satellite metric with the smallest MAE and bias is then bridged to the corresponding field event. MAE quantifies the mean absolute difference between the satellite phenometrics and field community phenology metrics during the study period, while bias measures the mean difference of the satellite-field metric pairs accordingly.

To further evaluate the potential of extending the community-level satellite-field phenological bridging relationship to the tree crown level, we locate seven large individual trees, with each occupying a large portion of a pixel in the fusion imagery. These trees include four *Quercus macrocarpa* (bur oak) trees and three *Quercus rubra* (red oak) trees, randomly distributed across the study site; they are among the largest trees with observed phenology in the field. We evaluate the accuracy of the satellite-field phenological bridging relationship derived at the community level for those seven large individual trees of all study years, using MAE and bias.

To assess the impact of multi-scale satellite imagery (i.e., 500-m, 250-m, and 30-m scales) on community-level phenological bridging, we also calculate the MAE and bias between field and corresponding satellite phenological metrics derived under 500-m MODIS, 250-m MODIS, MODIS-HLS fusion, and HLS-wise scenarios, and compare the accuracy metrics across scenarios accordingly.

## 4. Results

### 4.1. Field phenological characterization of forest community

Based on DOY-FCP, mean Begin Bud Burst (E1) for the five years was DOY 115 (April 25); the range among years was DOY 107 (April 17) to DOY 124 (May 4). Mean Full Leaf Expansion (F3) was DOY 140 (May 20); the range was DOY 133 (May 13) to DOY 147 (May 27). Timing varied among years as spring development, from Begin Bud Burst (E1) through Full Leaf Expansion (F3), ranged from 15 - 30 days. Speed of development between phenological events was faster from Begin Bud Burst (E1) to End Leafing Out (E3) (mean 6.3 days) than from Begin Leaf Expansion (F1) to Full Leaf Expansion (F3) (mean 14.0 days).

Much intra- and inter-specific variation in phenology exists in the field observations (Fig. 5). This variation is obscured when mean values are calculated, then weighted by basal area and summed for all species to generate the indices. For example, variation among all individuals, regardless of species, is noticeable, as a mean of 50 days among the five years exists between the DOY of first and last trees to reach Full Leaf Expansion (F3); the range is from 38 days in 2018 to 56 days in both 2016 and 2017. Also, e.g., based on Event-FCP, on the DOY in 2016 when the typical tree reaches Full Leaf Expansion (F3), only 29.1% of individuals are at F3; 39.9% are at earlier events, and 31.0% are at Full

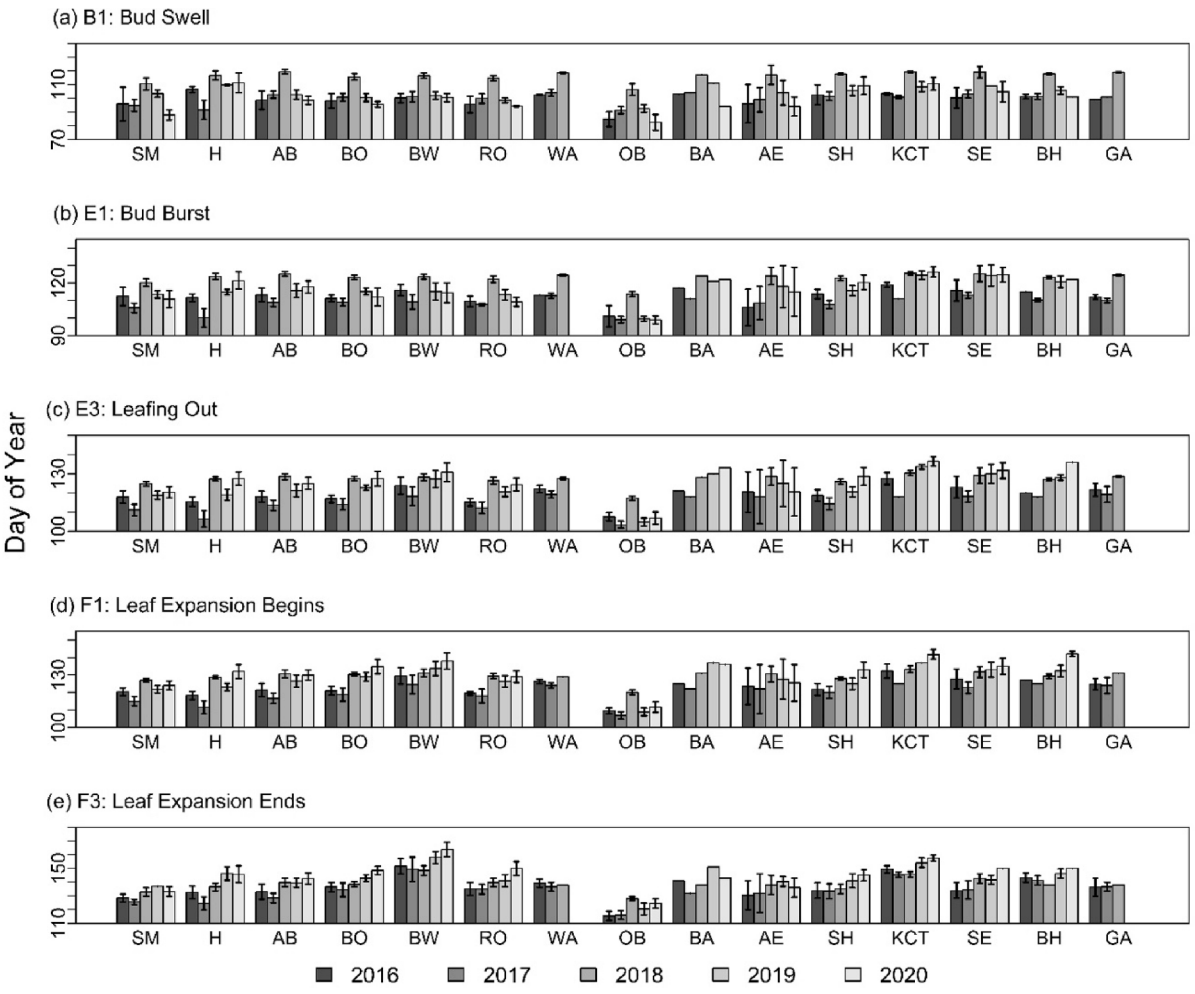


**Fig. 5.** Variation among field observations in species' mean (+/- 1 standard deviation) day of year (DOY) of five key spring phenological events for 2016-2020. The standard deviation indicates the intra-specific variation in DOY for each species. The order of species is based on their relative basal area (highest at far left; lowest at far right). See Table S2 for common name, scientific nomenclature, abbreviation, and relative basal area for each species.

Leaf Maturity (M). Therefore, even when the community is considered to have reached Full Leaf Expansion (F3), more than a third of individuals still have not reached that event, while nearly a third have previously completed this development. This variation contributes to the discrepancy between ground- and satellite-observations.

The community phenology development from the field perspective, based on the DOY- FCP index, can be modeled well via a logistic curve (Fig. 6). The logistic curve holds the biophysical inference as a 'growth curve' (Liang et al., 2011). The RMSE between the observed and fitted FCP values ranges from 0.159 to 0.351 for the five years. This strong fit provides support for using the same type of curve for the satellite time series fitting, as is widely used in previous remotely sensed phenological studies (Klosterman et al., 2014; Liang et al., 2011; White et al., 2014; Zhang et al., 2003).

Phenological development, as measured by field observations, varies considerably among the five study years (Fig. 7a). Development is fastest for all events in 2017, slowest for early events in 2018, and slowest for late events in 2020. These differences among years in phenological development are related to differences among years in February-May temperatures (Fig. 7b). Year 2017 has a very warm February and the earliest start of bud swell. Years 2016, 2017, and 2019 have different March temperatures, the same quite warm April temperatures, and different May temperatures. Although they have differing starting and ending phenology, these three years have an equivalent rate of development through April. The laggard Year 2018 is the second coolest March and coolest April, but is very warm in May when it very quickly finishes leaf development. Year 2020 has the second warmest March, but then the second coolest April and coolest May. That year has a rapid start, but very slow development thereafter.

The three field-based indices of canopy phenological development are in strong agreement overall (Fig. 8). One exception is with the final event 10, Full Leaf Maturity (M). A mean day of 10 in Event-FCP is always earlier than the day based on DOY-FCP that requires that all individuals reach full maturity in order to have a value of 10. DOY-FCP is a data-intensive index and reflects actual observations made. Event-FCP is, in a sense, a 'summary' of DOY-FCP and collapses the data observed and reduces the 'noise' of the DOY-FCP index. The excellent alignment between the DOY-FCP and Event-FCP makes it reasonable to use the simplified Event-FCP (orange squares in Fig. 8) for making the comparisons below with satellite-based phenological metrics. Specifically, six key phenological events during spring, namely Begin Bud Swell (B1), Begin Bud Burst (E1), End Leafing Out (E3), Begin Leaf Expansion (F1), End Leaf Expansion (F3), and Leaf Maturity (M), are selected to be compared with satellite-based measures.

In general, temporal patterns for the Canopy Leafing-Light Index, reflecting shade in the understory, are similar to the canopy development curves, except they are slower to increase in the early days, especially in years 2017 and 2020, and stabilize at slightly lower values than the canopy development readings. These discrepancies are explained by the assumptions of the Canopy-Leafing-Light Index. This index starts with Begin Bud Burst (E1), not Begin Bud Swell (B1) as do the other two aggregation methods, because it is assumed that the three

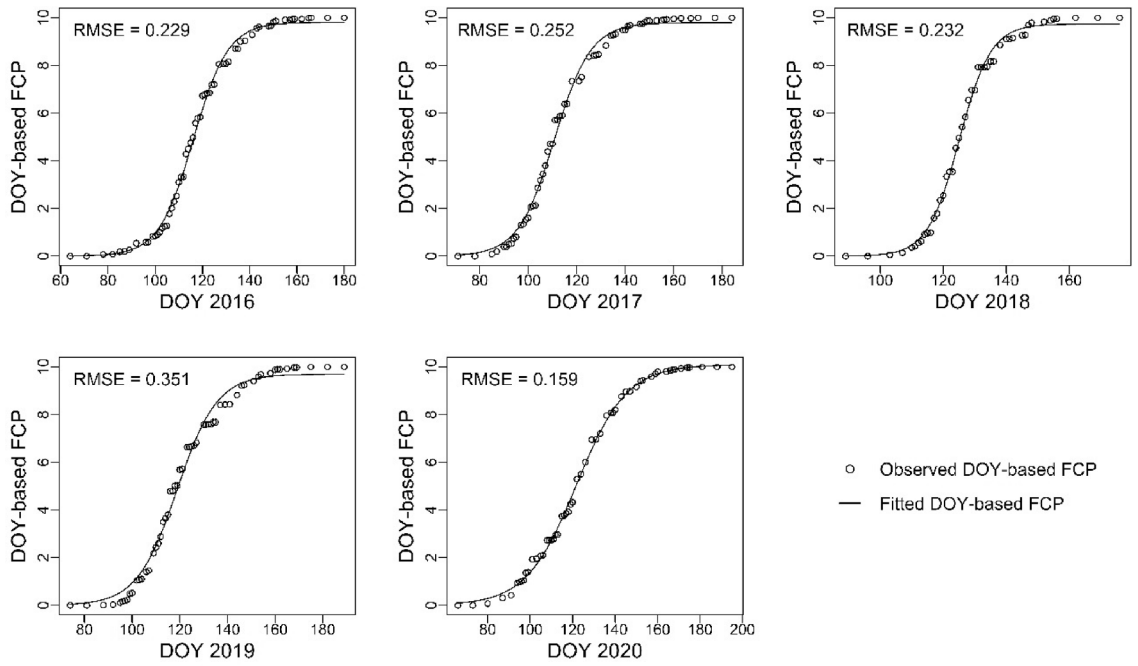


**Fig. 6.** Temporal pattern of phenological development during spring in each of five study years, based on field observations, calculated as the index DOY-based field community phenology (DOY-FCP). A dot represents the mean value on each day with an event change. The index ranges from 0 – 10; each number represents a unique phenological event (see Table 1). The observations are fit via a logistic curve; the root mean square error (RMSE) of the fit is in the upper left.

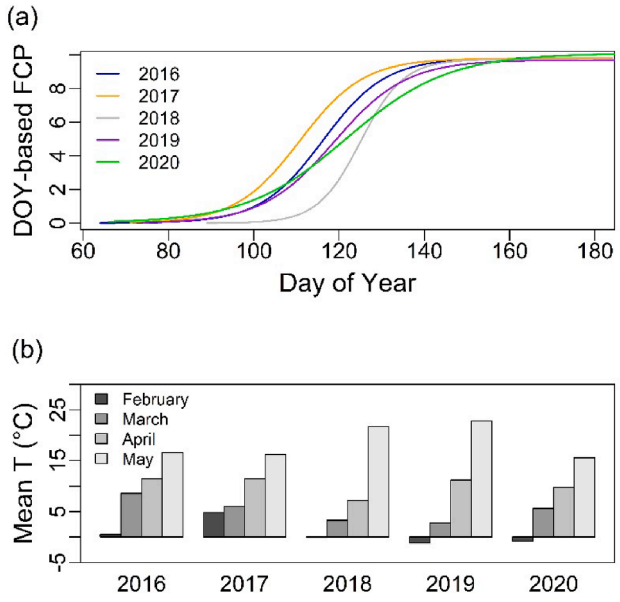


**Fig. 7.** (a) A comparison among the five study years of the temporal pattern of phenological development during spring. Each curve represents the logistic fit to the observations in Fig. 6. (b) Mean of mean daily temperature (°C) for each month, February through May, in each year of 2016-2020.

events observed during Bud Swell (B1-B3) by canopy trees do not change light interception below. Likewise, this index never reaches a value of 10.0 because it is assumed that even at Full Leaf Expansion (F3), 100% of light is not intercepted. Otherwise, the curve is in reasonable alignment with the other two indices and is a good summation of seasonal changes in shading in the understory due to vegetation above. Given the similar temporal patterns among Event-FCP, DOY-FCP, and Canopy Leafing-Light Index, we will use only Event-FCP as the community-level field phenological measure to be compared with corresponding satellite phenological metrics in subsequent analyses.

Considerable differences occur among years in the mean DOY of six key phenological events (Fig. 9). The earliest phenology for all events is in 2017 with its early spring warmth, while events through Begin Leaf Expansion (F1) are the latest in 2018 with a cool spring. Differences among years decrease as later events occur. The pattern is quite similar among years for Full Leaf Expansion (F3) and Leaf Maturity (M) as the rate of late leaf development is relatively consistent among years. The natural inter-annual variation provides a strong test of the validity of satellite observations over years varying in phenological development.

### 4.2. *Satellite-based phenological characterization of forest community*

#### 4.2.1. *Spatiotemporal image fusion*

With the hybrid deep learning model, the fusion imagery is generated for the study site throughout the spring season from 2016 to 2020. The fusion images are visually comparable to the corresponding HLS images on the HLS image acquisition dates, including the color tone (spectral similarity), texture, and spatial patterns of different land cover types (e.g., forest, agriculture, water, and urban land cover types) (Fig. 10). By contrast, the MODIS images are more blurred and spatial patterns of land covers can hardly be captured. Our study site, with its relatively small size, cannot be observed in the MODIS imagery. Through the spatiotemporal image fusion, the spectral details and spatial patterns of our study site can be recovered in the fusion imagery with distinct boundaries and color tones from surrounding prairie and agricultural fields, demonstrating the capability of the fusion model in capturing the rapid spring phenological change of the forest community among the imagery in this fragmented and heterogeneous landscape.

We further evaluate the performance of the fusion results using RMSE, ERGAS, and SAM (Table 2). These metrics are calculated based on the valid pixels of cloud-free HLS images (Table S1) and corresponding fusion images before DOY 200 for all study years. DOY 200 is selected as the cut-off date as the forest community went into the mature phenological phase before this date during the study period. For the study years, the average RMSEs for both red and NIR bands are below 0.026, with corresponding standard deviations mostly less than 0.01.

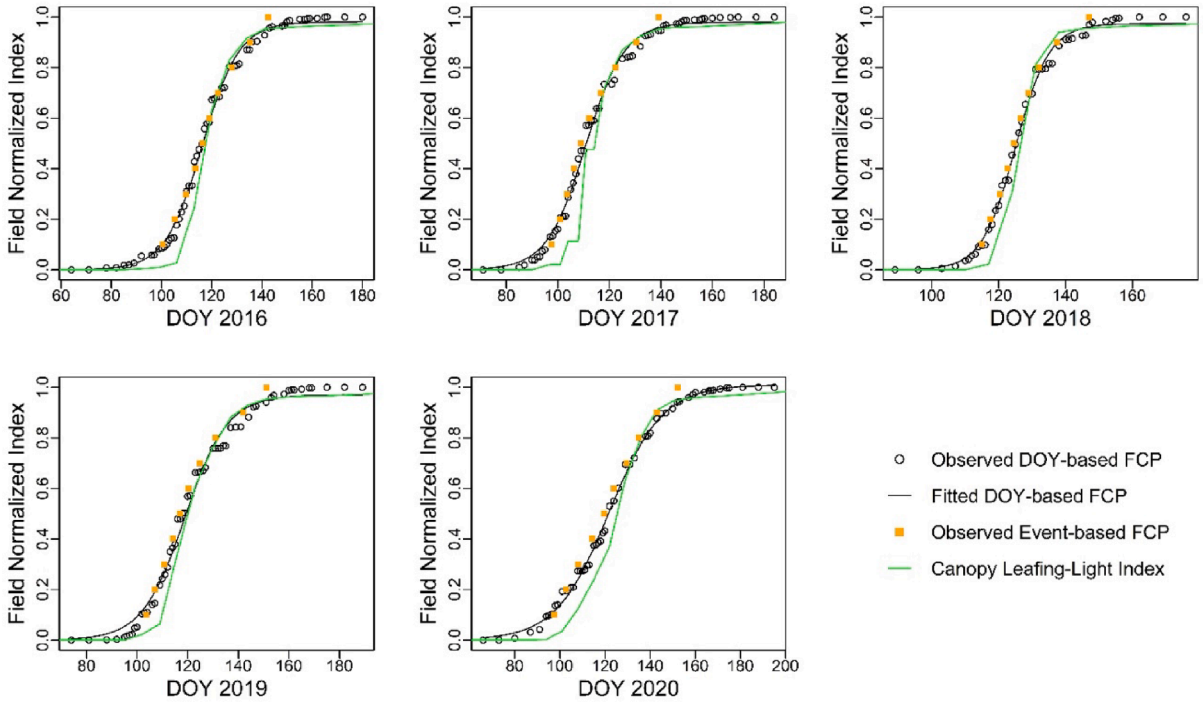


**Fig. 8.** A comparison using normalized values of the three field-based indices for phenological development of the community during spring in each year 2016-2020. Also shown is the logistic curve fit to the dots representing the observed DOY-FCP.

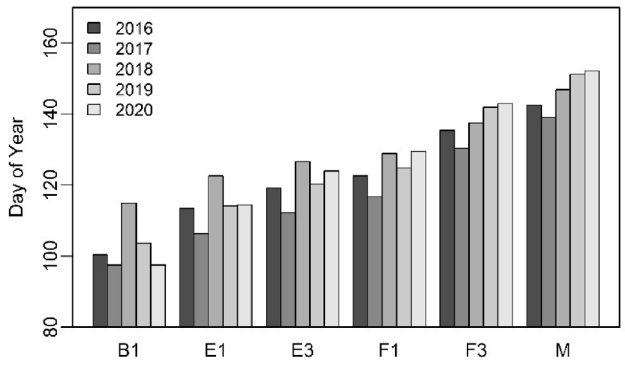


**Fig. 9.** Based on the Event-based field community phenology (Event-FCP) index, a comparison among the five study years in the mean DOY for each of six key phenological events: Begin Bud Swell (B1), Begin Bud Burst (E1), End Leafing Out (E3), Begin Leaf Expansion (F1), End Leaf Expansion (F3), and Leaf Maturity (M) (see Table 1).

The average ERGAS values are between 0.4 and 0.875, and the average SAM values are about 0.032. These evaluation metrics demonstrate the good performance of the hybrid deep learning model in generating the fusion imagery in the heterogenous landscapes, in the context of relevant existing fusion studies in the region (Yang et al., 2021). The fusion images are further scrutinized to retain only the images of relatively high quality in consideration of the source image quality (i.e., MODIS and HLS image), the temporal distance between image pair and prediction dates, and the spatial and temporal fusion coherency and representation. The community-average NDVI and EVI2 time series are then generated from the retained fusion images for forest phenology modeling.

#### 4.2.2. Satellite phenological metric

The Beck logistic curves demonstrate a strong fit to both fused community-average NDVI and EVI2 observations of phenological development in all years (Fig. 11). The RMSE values between the fused observed NDVI and fitted NDVI range from 0.022 to 0.048 for the five years. Similarly, the RMSE values between the fused observed EVI2 and fitted EVI2 range from 0.017 to 0.041 for the five years. The strong fit for both NDVI and EVI2 indicates the potential of these indices for monitoring forest phenology and capturing its seasonal variation. Comparable to field phenological observations, the satellite-derived fitting curves show fastest phenological development of forest community in 2017, and slowest development for early phenological events in 2018. Overall, these results demonstrate that the Beck model is appropriate for use in tracking phenological development at the study site.

The Beck logistic curves for fused community-average NDVI and EVI2 correspond well to the patterns of the Event-FCP during spring (Fig. 12). Given both NDVI and EVI2 fitted with logistic curves, as well as the similar temporal patterns between Event-FCP and DOY-FCP, the logistic fitted DOY-FCP curves are also shown in Fig. 12 for easier comparison. In each year, the field fitted DOY-FCP curve is a tighter fit to the corresponding NDVI curve than to the EVI2 curve. The EVI2 curve slightly delays the NDVI curve in each year. Its delay is partly because EVI2 is less sensitive to the early bud/leaf phenology development with its design to suppress the canopy and soil background signal. Therefore, in the following analyses, NDVI is used to compare phenological development with Event-FCP.

Several phenological metrics are extracted from the community-average NDVI fitted curves and compared with Event-FCP (Fig. 13). The satellite phenological metrics include threshold-based metrics (10%, 30%, 50%, 70%, and 90% shown in Fig. 13), curvature-based metrics (Greenup and Maturity), Gu-based metrics (Upturn and Saturation), and derivative-based metric (Der-SOS). The derived phenological metrics from the NDVI fitted curves overlap with specific field-based phases. The 10% threshold, Gu-Upturn, and curvature-Greenup approximate the field B1-B3 phase; 30% and 50% thresholds and Der-SOS approximate the field E1-E3 phase; 70% threshold approximates the field F1-F3 phase; the Gu-Saturation, curvature-Maturity and 90% threshold approximate the field Leaf Maturity DOY.

Besides the satellite phenological metrics calculated from the

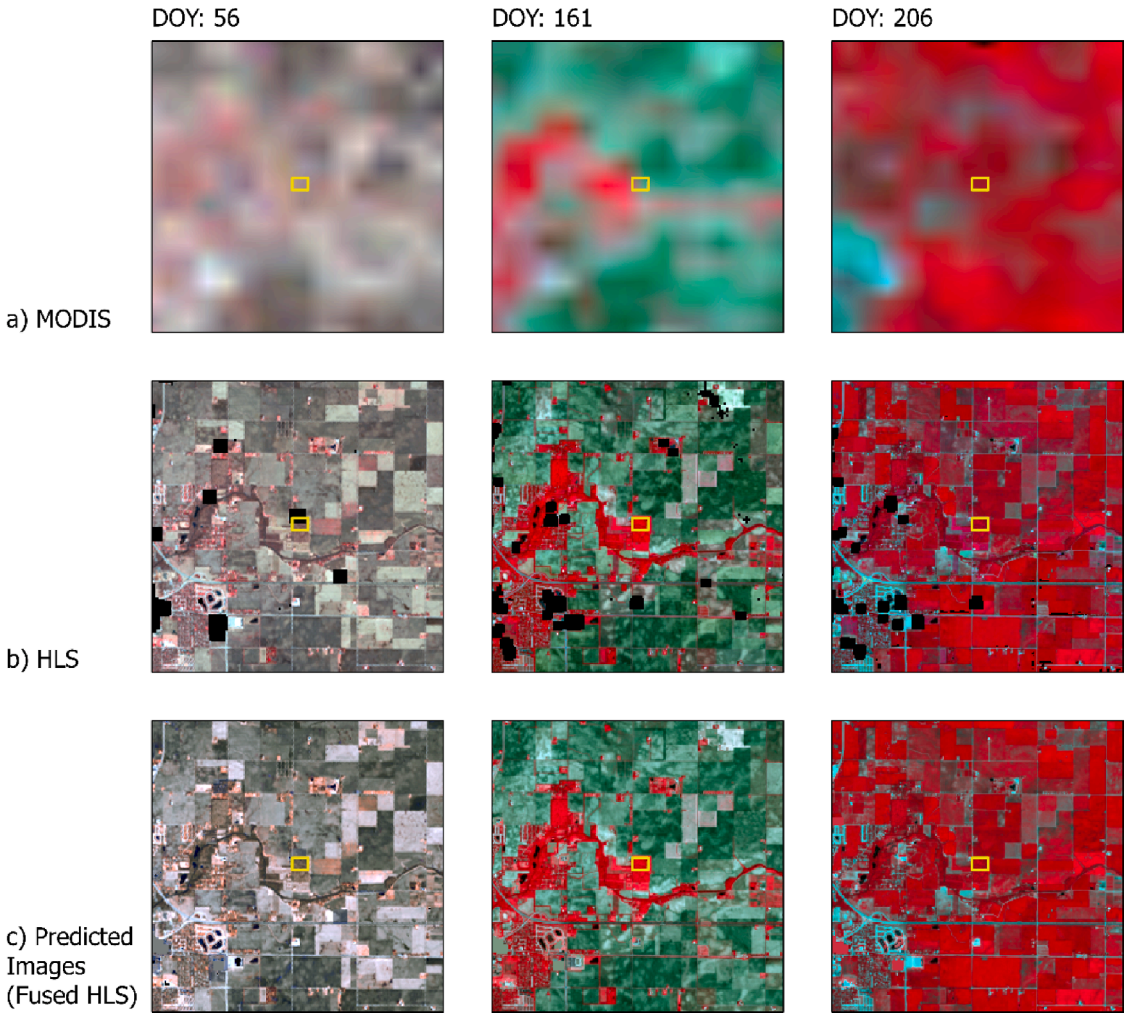


**Fig. 10.** For selected DOYs in 2019, the images from Moderate Resolution Imaging Spectroradiometer (MODIS) and Harmonized Landsat Sentinel-2 (HLS) satellites and corresponding fusion image examples from the hybrid deep learning model. The 12-ha study site is indicated by the yellow rectangle. The contaminated pixels in HLS images are masked in black.

**Table 2**
Accuracy of fusion results (mean ± standard deviation) by the hybrid deep learning model before DOY 200 from 2016-2020.

| Year | RMSE of Red | RMSE of NIR | ERGAS (2 band) | SAM (2 Band) |
|---|---|---|---|---|
| 2016 | 0.014 ± 0.007 | 0.022 ± 0.011 | 0.875 ± 0.451 | 0.034 ± 0.011 |
| 2017 | 0.014 ± 0.008 | 0.026 ± 0.014 | 0.846 ± 0.570 | 0.031 ± 0.005 |
| 2018 | 0.007 ± 0.001 | 0.013 ± 0.002 | 0.400 ± 0.049 | 0.032 ± 0.001 |
| 2019 | 0.009 ± 0.002 | 0.012 ± 0.004 | 0.454 ± 0.061 | 0.032 ± 0.007 |
| 2020 | 0.012 ± 0.008 | 0.022 ± 0.015 | 0.763 ± 0.482 | 0.030 ± 0.011 |

community-average NDVI fitted curves (diamond points in Fig. 14), we also calculate the mean and standard deviation of each satellite phenological metric derived from all the pixels in the study site using the fusion data (Fig. 14). Like community-wide field phenological measures, the satellite phenological metrics also exhibit large inter-annual variation in terms of both average NDVI-derived DOY and mean DOY of all the pixels in the study site. Among the years, the satellite-derived metrics are the earliest in 2017 with its early warm spring. The satellite metrics in early spring (e.g., curvature-Greenup, 10% threshold, Gu-Upturn, and 30% threshold) are much delayed in 2018 with its relatively low spring temperatures. Similar inter-annual patterns are found for satellite metrics corresponding to the same field phenological phase (e.g., 10% threshold, Gu-Upturn, and curvature-Greenup for the field budswell phase; 30% and 50% thresholds and Der-SOS for the field budburst/leafing out phase). For most satellite metrics, the community-average NDVI-derived DOY is close to the corresponding mean DOY of all the pixels in the study region, and is thus utilized in the subsequent satellite-field phenological comparison analysis.

### 4.3. Bridge satellite- and field-based phenological measures

Each of six key field phenological events (B1, E1, E3, F1, F3, and M) of the forest community during spring is compared to a range of satellite-derived phenological metrics derived from four pheno-metric extraction methods. Each of these six field events is best matched by a specific satellite-derived metric in terms of MAE and bias (Fig. 15; Table 3). Specifically, the field events Begin Bud Swell (B1), Begin Bud Burst (E1), End Leafing Out (E3), Begin Leaf Expansion (F1), and End Leaf Expansion (F3) are best approximated by the 10%, 30%, 45%, 60%, and 85% thresholds of NDVI time series, respectively. Leaf Maturity (M) corresponds well to the curvature-Maturity metric. With the extensive thresholds (5%-95% thresholds with an interval of 5%) tested in this study, most field events are best aligned with specific thresholds. The inter-annual variation in field phenological events is also captured well by the corresponding satellite phenological metrics, such as the earlier start of almost all the events in 2017 and the delayed start of several events (B1, E1, E3, and F1) in 2018.

We further assess the accuracy of field-observed events estimated by the best matching satellite-derived phenological metrics for the study

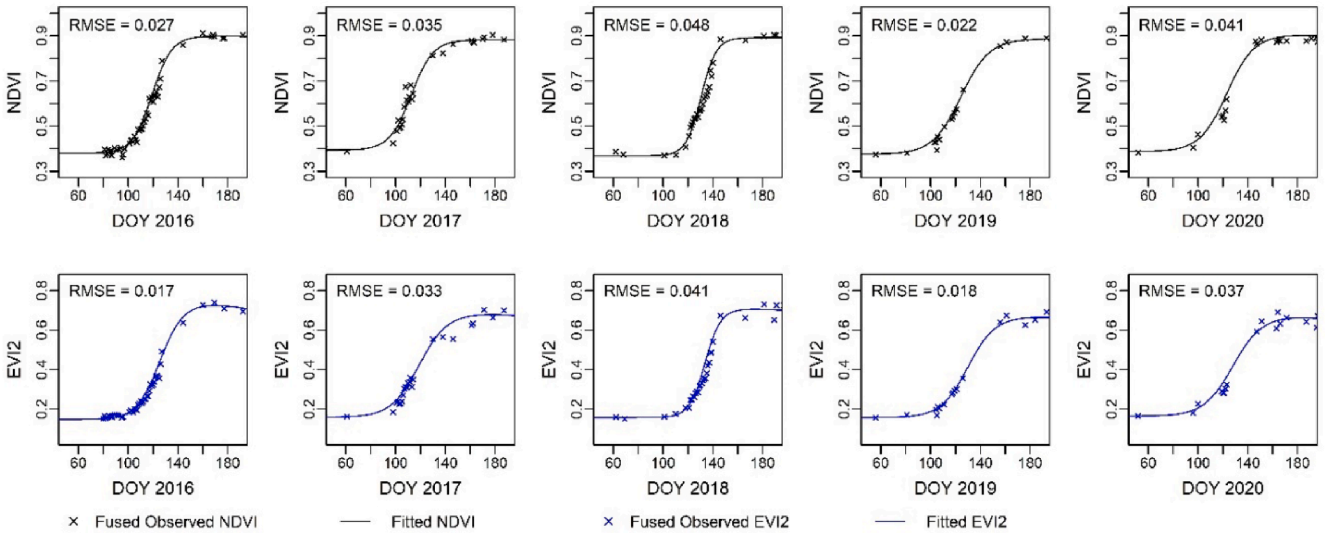


**Fig. 11.** The Beck logistic fitting to both fused community-average normalized difference vegetation index (NDVI) and 2-band enhanced vegetation index (EVI2) observations with RMSE calculated for each study year.

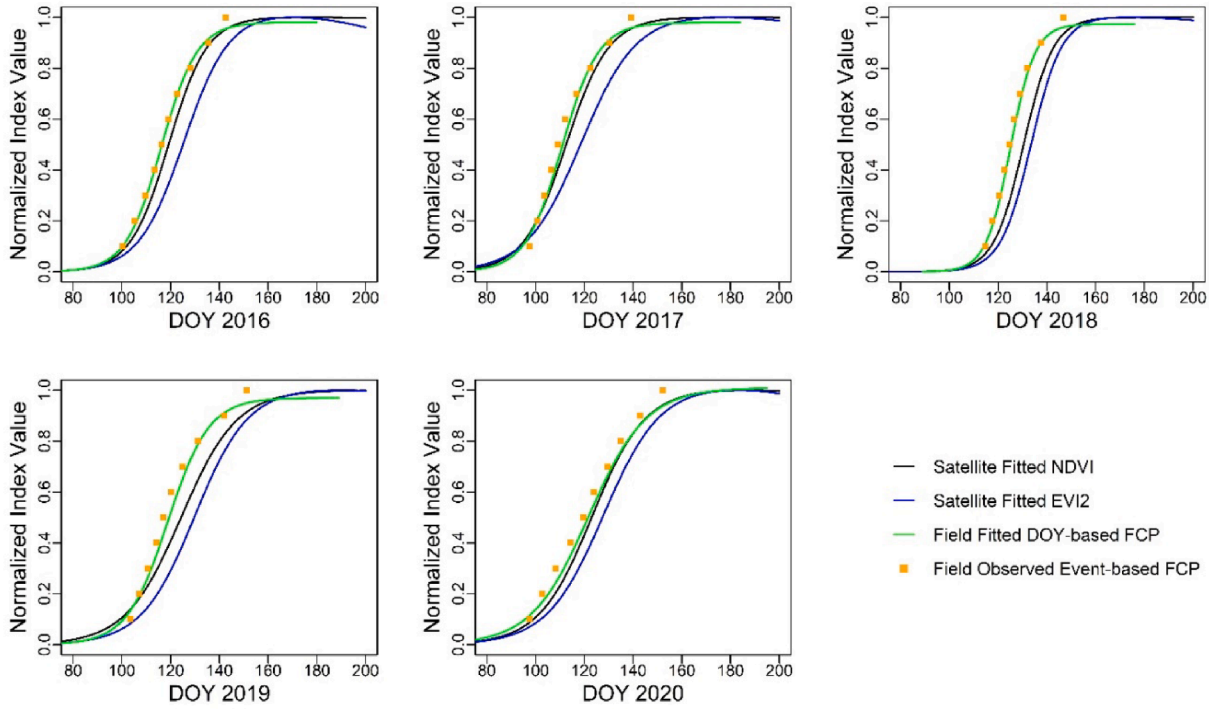


**Fig. 12.** A comparison of normalized values of the satellite-fitted community-average NDVI, EVI2, and field-based DOY-FCP and Event-FCP for each study year.

years using MAE and bias (Table 3). Comparable to Fig. 15, the field-based phenological measures are well bridged to corresponding satellite metrics. The MAE ranges among events from 1.1 - 2.9 days, with standard error ranging from 0.4 - 1 day. Field B1, E1, F1, and F3 are earlier than their corresponding satellite metrics by 0.02 - 1.5 days, while Field E3 and Leaf Maturity are later by 0.06 and 2.4 days, respectively. The standard errors of bias for all the events are less than 1.5 days. As curve feature-based metrics (curvature-based Greenup and Maturity, Gu-based Upturn and Saturation, and derivative-based SOS) have been widely used in satellite-based phenological studies, we further estimate their corresponding best matching field events from Table 1 (Fig. 16).

Specifically, curvature-Greenup, Gu-Upturn, derivative-SOS, Gu-Saturation, and curvature-Maturity best correspond to the field events Begin Bud Swell (B1), Middle Bud Swell (B2), End Leafing Out (E3), End Leaf Expansion (F3), and Leaf Maturity (M), respectively. The MAE ranges among events from 2.1 - 3.1 days, with standard error around 1 day (Table S3). The satellite metrics Greenup, Upturn, and Maturity are earlier than corresponding field events by 2.4 - 3.0 days. Der-SOS and Saturation are later than corresponding field ones by about 2 days. Similar to Fig. 15, the inter-annual phenological patterns of field events of the forest community can be captured by corresponding satellite curve-feature metrics.

Using the best matching satellite phenological metric for each key field phenological event, we also assess the satellite-field phenological bridging relationship for seven large individual trees for the study years;

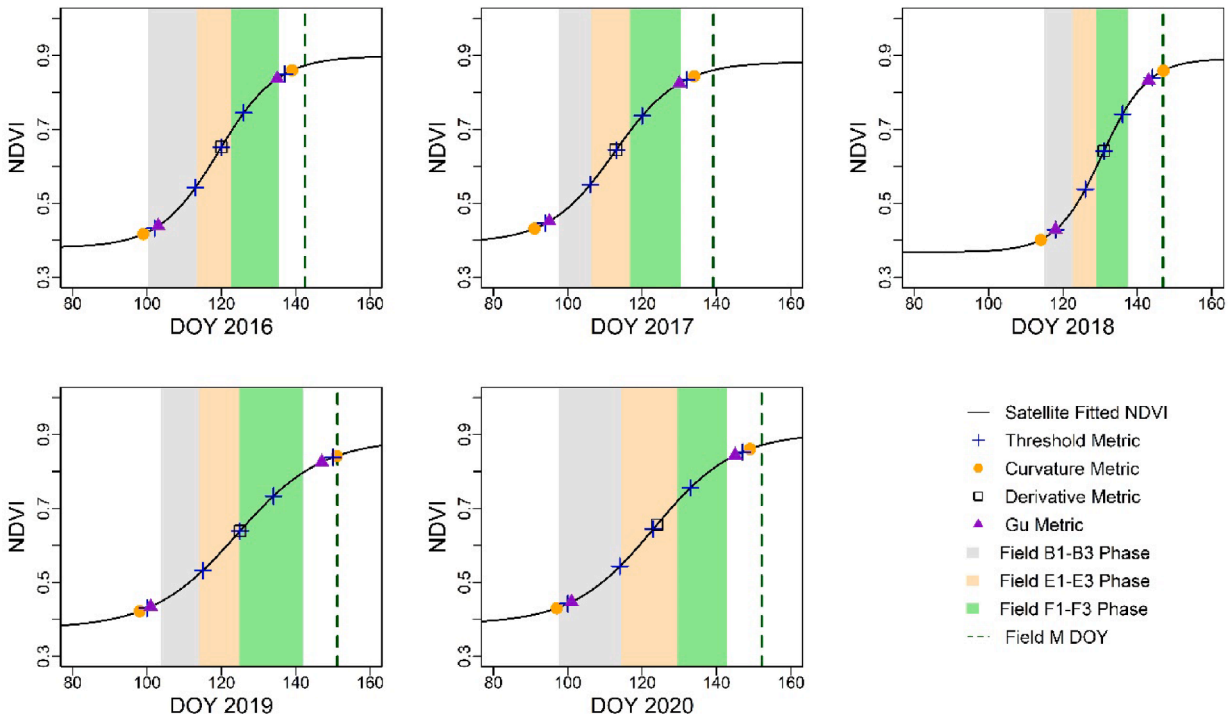


**Fig. 13.** The position of derived phenological metrics from NDVI fitted curves, overlapped with field Event-FCP phenological phases for each study year. The satellite phenological metrics include threshold-, curvature-, derivative-, and Gu-based metrics. Vertical bars represent field observed Event-FCP phases.

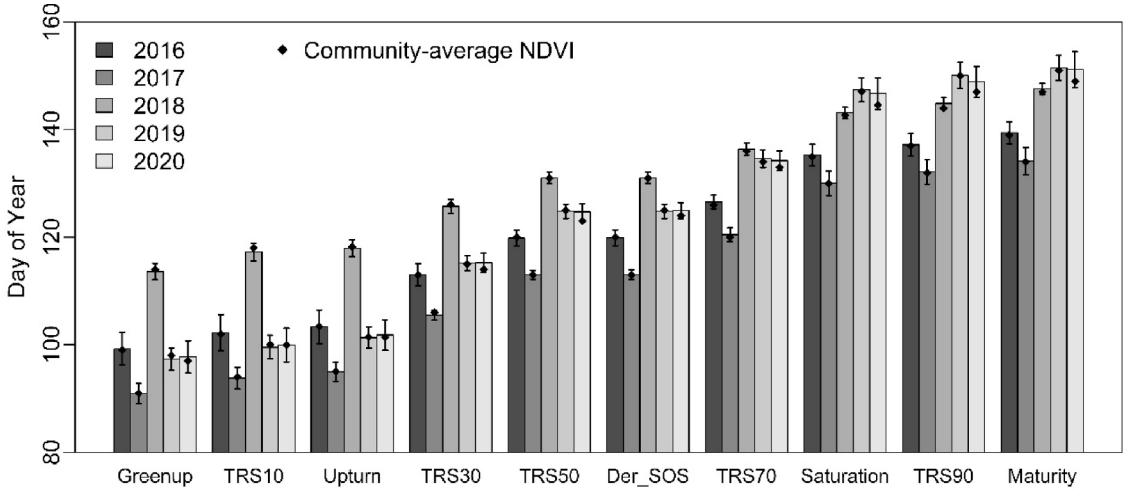


**Fig. 14.** The inter-annual variation in satellite-derived phenological metrics. Each bar indicates the mean value of phenological metrics derived from all the pixels of the study site using the fusion data; the error bar is the corresponding standard deviation. The diamond point denotes the phenological metric generated from the community-average NDVI fitted curve, which is close to the mean of the corresponding pixel-based phenological metric.

each tree occupies a large portion of a pixel. In general, field phenology of large individual trees can be estimated well from the community-derived satellite-field bridging relationship (Fig. 17). The field event and corresponding satellite metric dates are mostly aligned along the 1:1 line, with correlation coefficient r about 0.56 - 0.81. For large trees, the MAE among events ranges from 3.2 to 7.0 days (Table 4). Specifically, the field events E1, E3, and F1 can be estimated with relatively high accuracy (MAE around 3 - 4 days, bias around -0.7 - 1 day, and r around 0.68 - 0.81). The early field event B1 is more subject to the influence of understory vegetation (i.e., herb species and tree saplings) and soil background. The weak signal of budswell is also more difficult to be detected with remote sensing. At the end of spring, the transition from leaf expansion (F3) to maturity (M) also becomes harder to capture by satellite imagery compared to the earlier phases (e.g., leaf out and early leaf expansion phases), partly due to more subtle changes in tree leaf status, canopy structure, and spectral response. The variation in DOY of the field events F3 and M among selected individual trees are more difficult to characterize, with larger estimation errors (MAE around 5 - 7 days, negatively biased by 2 – 6 days, and r around 0.56 – 0.58).

### 4.4. *Impact of multi-scale satellite imagery on phenological bridging*

With our study site being a fragmented forest, we assess the impact of satellite imagery of varying spatial and temporal resolutions on the satellite-field phenological bridging relationship. For each study year, the NDVI time series of the forest community derived from different

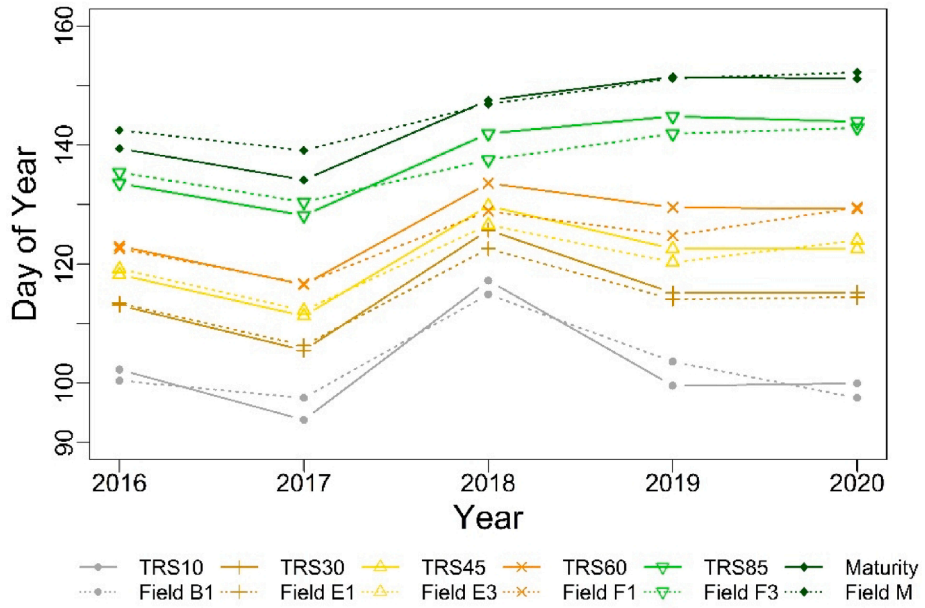


**Fig. 15.** The difference in DOY for the best match between six field-observed events versus corresponding satellite-derived phenological metrics.

**Table 3**
The accuracy of field-observed events estimated by the best matching satellite-derived phenological metrics (from Fig. 15), as measured by MAE and bias. A negative value for bias indicates the field measure of the event is later than the satellite estimate.

| Field Measure | Corresponding Satellite Metric | MAE (days) | Standard Error (MAE; days) | Bias (days) | Standard Error (Bias; days) |
|---|---|---|---|---|---|
| Field B1 | TRS10 | 2.86 | 0.37 | 0.02 | 1.48 |
| Field E1 | TRS30 | 1.08 | 0.59 | 0.64 | 0.73 |
| Field E3 | TRS45 | 2.1 | 0.46 | -0.06 | 1.15 |
| Field F1 | TRS60 | 2.1 | 0.86 | 1.5 | 1.13 |
| Field F3 | TRS85 | 2.66 | 0.58 | 0.38 | 1.44 |
| Field M | Maturity | 2.42 | 0.98 | -2.38 | 1.01 |

satellite products vary greatly (Fig. 18). The NDVI temporal pattern from 500m MODIS data is consistently delayed compared to that of other satellite data (i.e., 250m MODIS, HLS-wise, and fused data). This delayed pattern is mostly attributable to the mixture of land covers within the pixel, which contains not only the forest community, but also the surrounding prairie and agricultural fields. Compared to the forest community, the phenological development of both prairie and agricultural fields is much delayed, with emergence dates of agricultural fields around middle-late May for the study years. With only the forest contained in the pixel of 250m MODIS, its phenological patterns are different from those of 500m MODIS and are more comparable to those of fused data. Yet the pixel of 250m MODIS represents only a portion of the forest community and the NDVI phenological patterns of 250m MODIS exhibit varying degrees of congruence with those of fused data across years. This varying relationship is caused mostly by the difference

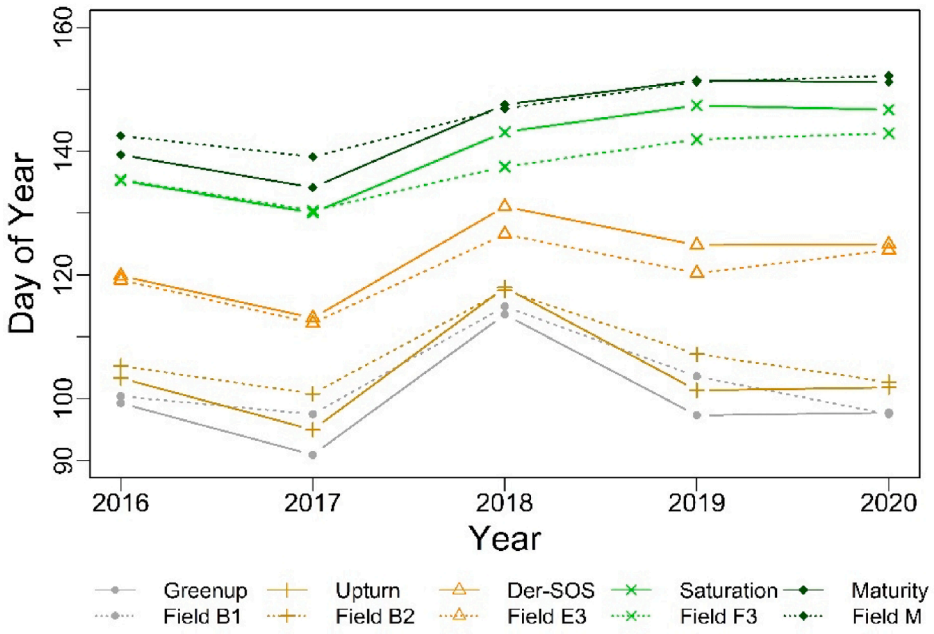


**Fig. 16.** The difference in DOY for the best match between satellite curve feature-based metrics versus corresponding field-observed events (from Table 1).

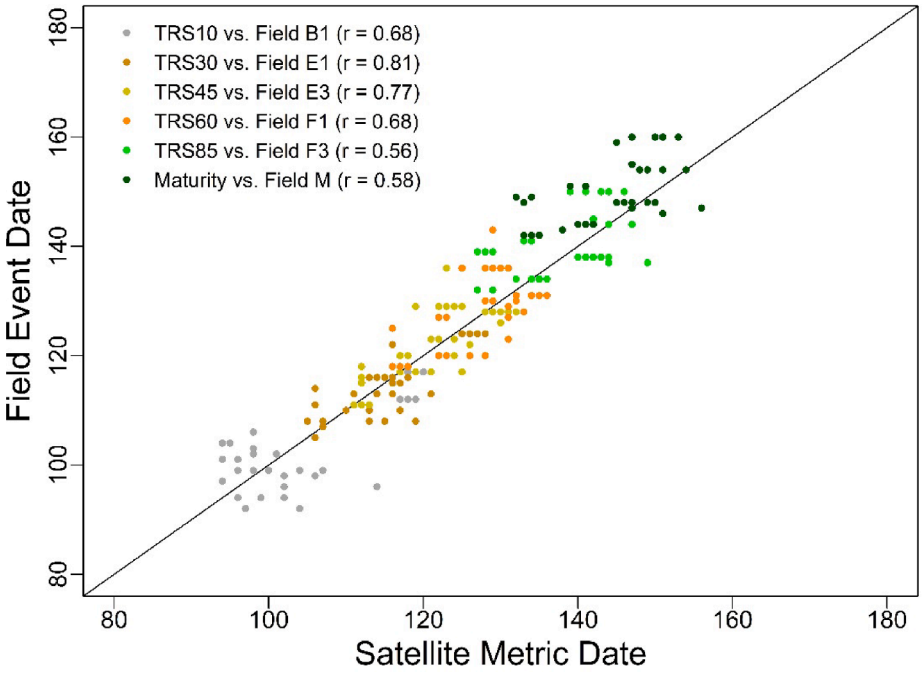


**Fig. 17.** For six phenological events, the relationship between DOY for field-data vs. satellite-derived metrics of all study years for selected pixels of seven large trees, each occupying a large portion of a pixel.

**Table 4**
The accuracy of satellite-field phenological bridging relationship for seven large individual trees of all study years (from Fig. 17), as measured by MAE and bias. A negative value for bias indicates the field measure of the event is later than the satellite estimate.

| Field Measure | Corresponding Satellite Metric | MAE (days) | Standard Error (MAE; days) | Bias (days) | Standard Error (Bias; days) |
|---|---|---|---|---|---|
| Field B1 | TRS10 | 5.63 | 0.62 | 1.86 | 1.10 |
| Field E1 | TRS30 | 3.17 | 0.44 | 1.00 | 0.68 |
| Field E3 | TRS45 | 3.23 | 0.50 | -0.71 | 0.73 |
| Field F1 | TRS60 | 4.03 | 0.55 | -0.60 | 0.88 |
| Field F3 | TRS85 | 5.11 | 0.62 | -2.31 | 1.00 |
| Field M | Maturity | 7.00 | 0.82 | -6.03 | 1.02 |

in species composition and abundance with the change of pixel coverage. Compared to the fused data, the NDVI phenological patterns of HLS-wise data are more fluctuating, possibly due to contamination from residual atmospheric interference, cloud, cloud shadow, and snow.

The satellite-field phenological bridging relationship across scales is assessed further using MAE and bias (Table S4 and Fig. 19). For all field events, the fused data achieve the smallest MAE (1.1 to 2.9 days) and bias (-2.4 to 1.5 days), followed by HLS-wise data (MAE: 2.8 to 6.5 days; bias: -4.6 to 4.0 days). The 500m MODIS data yield the largest error and bias due to the heterogeneous landscape covered by the pixel. For most events, the 500m MODIS satellite phenological metrics are delayed by over 3 weeks compared to corresponding field events, indicating the negative effect of the mixed pixel on satellite phenological extraction accuracy. For the 250m MODIS data, the MAE ranges from 4.4 to 19.3 days and bias from -11.4 to 19.2 days. The field events E1, E3 and F1 can be detected with smaller errors compared to the early (B1) and late (F3 and M) events in spring, yet the overall relatively large errors of 250m MODIS data indicate that the satellite phenological detection accuracy can be degraded when the forest community is not entirely covered by the pixels.

## 5. Discussion

The remote sensing-based phenology studies have largely expanded the scope of traditional phenology observations by enabling high-frequency and large spatial coverage phenological monitoring which may not be feasible by direct field observations (Berra and Gaulton, 2021; Ganguly et al., 2010; Zhang et al., 2018). The phenological monitoring using different sensors at multiple scales facilitates the characterization of spatio-temporal phenological patterns and dynamics across the globe, improving our understanding of large-scale phenological responses to climate change and environmental disturbance, feedbacks to the climate system, and impacts on ecosystem structures and functions (Piao et al., 2019; Tang et al., 2016; White et al., 2009). The fusion of satellite imagery of varying resolution characteristics further enhances the phenological characterization towards finer spatial and temporal scales (Liang et al., 2014; Tian et al., 2024). However, the existing remotely sensed phenological studies are mostly conducted at the landscape scale with a pixel possibly mixed by different land covers or tree species of distinct phenology, and may have large uncertainties in detecting phenological events at forest community levels, particularly in heterogenous and fragmented landscapes (Donnelly et al., 2022; Fisher and Mustard, 2007). The differences among satellite sensors, spectral indices, and algorithms utilized to monitor phenology can produce very different results (Berra and Gaulton, 2021; White et al., 2014). Also remotely sensed phenology may have equivocal biophysical meaning and is challenging to be validated by ground phenological observations due to the scale mismatch in measurement (Donnelly et al., 2022). In temperate deciduous forests, multiple phenological phases (e.g., bud swell, bud burst/leafing out, leaf expansion, and leaf maturity) are critical for understanding the forest phenological development timing and duration, and their response to climate change in spring. These phenological phases, affected by a variety of environmental factors, have important implications for carbon storage and biogeochemical cycling, yet have seldom been studied simultaneously in satellite and field phenological comparison studies (Donnelly et al., 2017). Through developing a novel satellite-field phenological bridging framework, this study eases the challenges of previous studies and contributes to a more explicit understanding of the relationship between satellite-derived

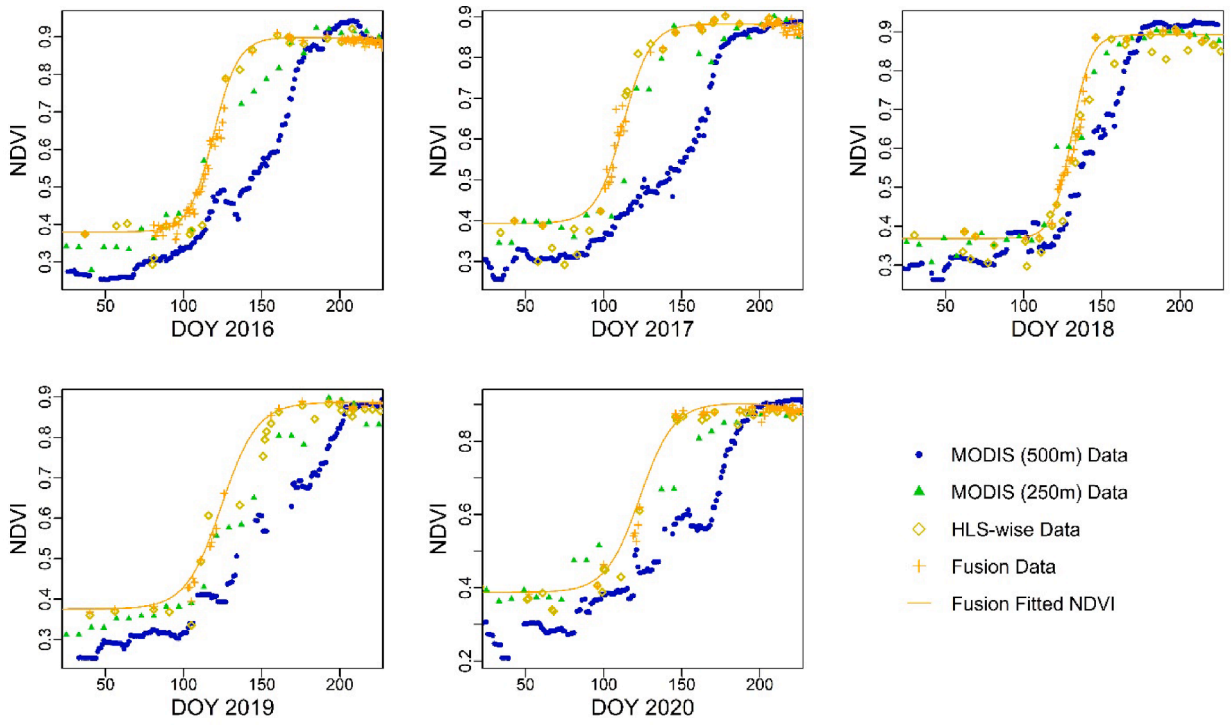


**Fig. 18.** NDVI time series of the study site using MODIS products at 500m (MODIS MCD43A4) and 250m (MOD09Q1) resolutions, original HLS-wise data, and fused data for each of the study years. The Beck-fitted NDVI curves from fused data are also shown.

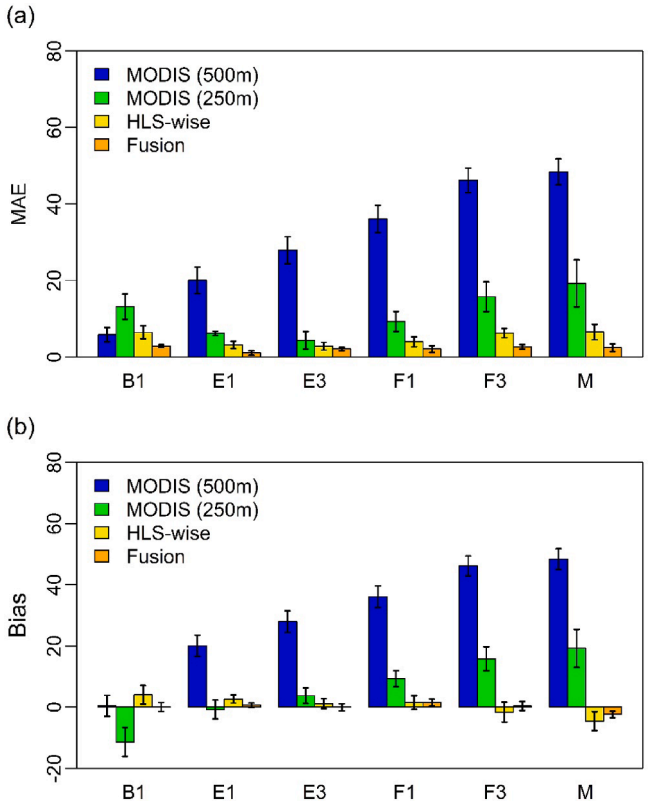


**Fig. 19.** The mean absolute error (MAE) and bias of satellite-field phenological bridging relationship using satellites of various spatial and temporal resolutions (i.e., 500m MODIS, 250m MODIS, HLS-wise, and fused data).

phenological metrics and field phenological observations of a fragmented temperate deciduous forest for a diversity of phenological phases during the spring season. It also innovatively scales up the field phenological observations from the individual trees to the species to the community level, and underscores the importance of taking into account spatial scale and representation from both satellite and field phenological perspectives in building corresponding bridging relationships.

To remotely retrieve spring phenology of the fragmented forest, we devise several key components in the phenological bridging framework, including deep learning-based spatiotemporal image fusion, satellite-based forest phenology modeling, satellite-based forest phenological metric extraction, and field-based forest community phenological characterization. The monitoring of fragmented forest phenology requires satellite observations with both high spatial and temporal resolutions, the acquisition of which remains a significant hurdle in remotely sensed phenological studies due to the sensor resolution tradeoff. The frequent cloud cover and rainy conditions during spring phenological development largely reduce the availability of HLS observations of the study site. A paucity of HLS images during critical phenological transition periods led to both larger MAE and bias for all phenological events in this study, despite phenological curves being generated. The uncertainty of phenological detection has been over two weeks in previous studies due to the limited availability of temporal observations around transition periods, indicating the importance of data quality as well as temporal distributions of satellite observations (Gao et al., 2020; Tian et al., 2024; Zhang et al., 2017, 2009). Despite being temporally more frequent, the 500m MODIS data yield much larger detection errors (e.g., over 3 weeks) for most phenological stages, due to the forest fragmentation and mixture with surrounding prairie and agricultural fields. The landscape heterogeneity, with a patchwork of the forest and other land cover types, makes the phenological detection difficult to be bridged to field phenological observations, as also suggested by previous studies (Donnelly et al., 2022; Elmore et al., 2016). The 250m MODIS data, on the other hand, capture phenological development of only a portion of the forest community, in which species composition and relative basal area may vary.

The hybrid deep learning model, with its demonstrated capability of fusing satellite imagery of varying spatial and temporal resolutions, is employed in the devised framework to generate temporally dense fusion data at 30-m spatial resolution. By integrating SRCNN for spatial relationship modeling and LSTM for temporally evolving feature learning, the hybrid deep learning model has been found advantageous in capturing rapid phenological changes among the imagery (Yang et al., 2021). The fusion data enable the reconstruction of high-quality phenological development trajectories of forest community in spring. Together with forest phenology modeling and phenological metric extraction, a range of satellite phenological metrics (i.e., threshold-, derivative-, curvature-, and Gu-based metrics) are extracted for the study years, which greatly expands limited phenological retrievals in previous studies (mostly field budburst/leafing out events) and is critical for satellite-field phenological bridging (Berra and Gaulton, 2021; Kang et al., 2003; Schwartz and Hanes, 2010). Compared to curve feature-based pheno-metric extraction methods, the threshold-based method achieves higher accuracy in retrieving most of target field events, partly due to the comprehensive and extensive thresholds (5%-95% thresholds with an interval of 5%) considered in this study for characterizing variation in spring canopy greenness. The target field events B1, E1, E3, F1 and F3 are best bridged to 10%, 30%, 45%, 60%, and 85% thresholds of NDVI time series, respectively, with high accuracy (MAE and bias less than 3 days). The M field event is more accurately bridged to the curvature-based Maturity metric with MAE and bias around 2 days. The inter-annual variation of these field events is well captured by the corresponding satellite events. With the wide use of curve feature-based metrics, these metrics are further bridged to corresponding best matching field events with relatively high accuracy achieved (MAE and bias around 2-3 days). Specifically, the curvature-based Greenup and Gu-based Upturn metrics approximate the field events B1 and B2, respectively. The derivative-based SOS corresponds to the field event E3. The Gu-based Saturation and curvature-based Maturity metrics connect with the field events F3 and M, respectively. The satellite detection of early field events (e.g., B1) may be subject to the influence of understory plants, soil background, and snow. From 2016 to 2020, understory herbs of Trelease Woods emerged on average around DOY 98 (standard deviation: 4.7 days), which is earlier than or around the budswell phase of forest community phenological development (Augspurger and Zaya, 2020). The subtle phenological changes of budswell also are more difficult to capture by satellite imagery, compared to other phenological phases. Due to the challenge in bridging satellite and field phenological measures, previous studies have been focused mainly on detection of the budburst/leafing out phase at the landscape level, which has been found to connect with the 20-30% threshold of vegetation index (e.g., NDVI) time series (Berra and Gaulton, 2021; Bórnez et al., 2020; Kowalski et al., 2020). With the devised satellite-field phenological bridging framework, our study not only corroborates the limited bridging relationship of previous studies, but also expands substantially the scope of satellite detection to all the critical spring phenological events at the forest community level. Additionally, the bridging relationships of all the critical events are successfully extended to seven large trees, with each occupying the majority of a pixel, indicating the strong potential of the devised framework in conducting phenological detection of individual trees at the crown level, given adequate spatial and temporal resolutions of satellite imagery (Zhao et al., 2023).

The ground phenological observations that are representative of the variability in phenology events of the forest community are essential for evaluating the accuracy of satellite-based phenological retrieval. However, the community-level forest phenological observation remains a great challenge due to the heterogeneity in phenological development of tree species as well as species composition within a community (Fisher et al., 2006; Polgar and Primack, 2011). The different phenological response of individual trees to environmental conditions (e.g., temperature, photoperiod, and soil moisture) can lead to large intra-specific and inter-specific phenological variation of the forest community, which necessitates the systematic and consistent field observations of all the key phenological events for a large number of trees of diverse species with relative basal area surveyed (Donnelly et al., 2017; Vitasse et al., 2009). Such phenological information is key to scaling up for measuring the community-level forest phenology, but has seldom been collected in previous studies. With 10 spring phenological events of 148 haphazardly selected individual canopy trees of 15 common species observed via a consistent monitoring protocol, we innovatively generate three community-level field phenological indices: DOY-FCP, Event-FCP, and Canopy Leafing-Light index. The strong agreement between DOY-FCP and Event-FCP indicates the effectiveness of our field phenological surveying approach in capturing community-wide phenological variation and characterizing corresponding forest phenological development trajectory. Both field- and satellite- phenological observations of the forest community throughout canopy development are modeled using the logistic function with a strong fit, which facilitates the comparison of phenological curves generated in-situ with those from satellite imagery. Compared to EVI2, NDVI time series shows better alignment with the field FCP, due partly to the early phenological events (e.g., early bud/leaf development events) considered in the field phenological monitoring protocol and the reduced sensitivity of EVI2 to these early events, with EVI2 being designed to suppress the soil background signal (Gao et al., 2020; Huete et al., 2002). Under-canopy light levels have been found to be related to visual phenological observations of canopy trees in previous studies (Schwartz et al., 2013). Our study further extends the canopy light transmission analysis from the canopy tree level to the community level and generates the Canopy Leafing-Light index. The temporal patterns of the Canopy Leafing-Light Index are generally comparable to those of the two FCP indices, yet with the delayed start in early spring. It takes into account both the state of the entire forest's canopy and canopy closure-induced shading in the understory. This canopy index may further help with the study of understory phenology in response to canopy closure and light interception.

Despite the significant value of the devised framework in bridging all critical field and satellite spring phenological events at the forest community level, the study's framework also has limitations. One limitation stems from the spatial extent of field phenological observations. Although systematic tree phenological observations that are representative of community-wide forest phenological variation are essential to bridge satellite-field phenological measures at the community level, the field phenological observations of our study site may not capture the diverse phenological development patterns of forests of varying species compositions and richness across habitats and microclimates. Field phenological efforts similar to our study over extended regions of diverse forest characteristics are important in the future to further evaluate the community-level satellite-field phenological bridging relationship. Such efforts may also benefit from the incorporation of phenological observations via citizen science/crowdsourcing initiatives (e.g., US NPN) and Phenocam near-surface cameras across diverse ecosystems and regions (Peng et al., 2017; Richardson et al., 2018), for which more rigorous designs of tree phenological monitoring representative of community-wide canopy development would be needed. Another limitation lies in the uncertainty associated with the detection of the earliest field phenological events (e.g., budswell). In our study site, the emergence of understory herbs and/or saplings in early spring may complicate the phenological signal collected via remote sensing. It would be beneficial in the future to evaluate the devised framework and assess the influence of understory on forest phenological detection under varying densities of understory herbs and/or saplings. Despite the focus on spring phenology, our devised framework could further be extended to autumn phenology of the forest community in future studies for building corresponding satellite-field phenological bridging relationships, which can enhance our understanding of forest phenological responses under the changing climate (Xie et al., 2018).

## 6. Conclusions

In this study, we develop a novel satellite-field phenological bridging framework to characterize a variety of key spring phenological events at the community level for a fragmented forest, using multi-scale remote sensing imagery. The framework encompasses deep learning-based spatiotemporal image fusion, satellite-based forest phenology modeling, satellite-based forest phenological metric extraction, and field-based forest community phenological characterization. By fusing the satellite imagery of varying resolution characteristics, the deep learning-based fusion model reconstructs temporally dense satellite time series at 30-m spatial resolution for tracking the community-level forest phenological dynamics in spring. Along with forest phenology modeling and pheno-metric extraction, a multitude of satellite phenological metrics (i.e., threshold-, curvature-, derivative-, and Gu-based metrics) are retrieved. With the phenology of a large number of trees of diverse species of a forest community being observed in-situ, we devise three summative field indices (i.e., DOY-FCP, Event-FCP, and Canopy Leafing-Light Index) to quantify community-level phenological states throughout canopy development in spring. Through bridging satellite and field measures, the framework successfully detects all critical spring forest phenological phases (i.e., bud swell, budburst/leafing out, leaf expansion, and leaf maturity) at the community level, yet the detection of the earliest phenological events may be subject to the influence of understory vegetation. Compared to curve feature-based metrics, the threshold-based metrics achieve higher accuracy in phenological detection. The multi-scale satellite-derived phenology analysis further indicates the importance of satellite imagery fusion in retrieving the phenology of fragmented forests in heterogeneous landscapes, particularly when forest green-up is under cloud cover and rainy conditions. This study innovatively scales up the field phenological observations from the species to the community level, and broadens the scope of phenological monitoring from limited events to all the critical ones of the forest community. The community-level bridging of satellite and field metrics for the key spring phenological events can be used to assess forest phenological responses to climate change and weather extremes. The diversity of monitored events can also help evaluate the rate of forest phenology progression and the duration of each phenological phase, facilitating the understanding of forest phase-specific phenological response to environmental changes, and improving models of carbon-water-energy cycling.

## 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.

## Acknowledgements

This study is partly supported by the National Science Foundation (grant number 1951657) and the National Aeronautics and Space Administration (grant number 80NSSC22K1548). We also acknowledge the Blue Waters petascale supercomputer from the National Center for Supercomputing Applications to fuse the multi-scale satellite imagery with the help of Zijun Yang.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2024.03.018.

## References

Ahl, D.E., Gower, S.T., Burrows, S.N., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., 2006. Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. Remote Sensing of Environment 104, 88–95. https://doi.org/10.1016/j.rse.2006.05.003.

Atkinson, P.M., Jeganathan, C., Dash, J., Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote Sensing of Environment 123, 400–417. https://doi.org/10.1016/j.rse.2012.04.001.

Augspurger, C.K., Cheeseman, J.M., Salk, C.F., 2005. Light Gains and Physiological Capacity of Understorey Woody Plants during Phenological Avoidance of Canopy Shade. Functional Ecology 19, 537–546.

Augspurger, C.K., Zaya, D.N., 2020. Concordance of long-term shifts with climate warming varies among phenological events and herbaceous species. Ecological Monographs 90, e01421.

Badeck, F.-W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., Sitch, S., 2004. Responses of spring phenology to climate change. New Phytologist 162, 295–309. https://doi.org/10.1111/j.1469-8137.2004.01059.x.

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 Sensing of Environment 100, 321–334. https://doi.org/10.1016/j.rse.2005.10.021.

Berra, E.F., Gaulton, R., 2021. Remote sensing of temperate and boreal forest phenology: A review of progress, challenges and opportunities in the intercomparison of in-situ and satellite phenological metrics. Forest Ecology and Management 480, 118663. https://doi.org/10.1016/j.foreco.2020.118663.

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 Sensing of Environment 240, 111685. https://doi.org/10.1016/j.rse.2020.111685.

Bórnez, K., Descals, A., Verger, A., Peñuelas, J., 2020. Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. International Journal of Applied Earth Observation and Geoinformation 84, 101974. https://doi.org/10.1016/j.jag.2019.101974.

Both, C., Van Asch, M., Bijlsma, R.G., Van Den Burg, A.B., Visser, M.E., 2009. Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? Journal of Animal Ecology 78, 73–83. https://doi.org/10.1111/j.1365-2656.2008.01458.x.

Caparros-Santiago, J.A., Rodriguez-Galiano, V., Dash, J., 2021. Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review. ISPRS Journal of Photogrammetry and Remote Sensing 171, 330–347. https://doi.org/10.1016/j.isprsjprs.2020.11.019.

Chmielewski, F.-M., Rötzer, T., 2001. Response of tree phenology to climate change across Europe. Agricultural and Forest Meteorology 108, 101–112. https://doi.org/10.1016/S0168-1923(01)00233-7.

Chuine, I., 2010. Why does phenology drive species distribution? Philosophical Transactions of the Royal Society B: Biological Sciences 365, 3149–3160. https://doi.org/10.1098/rstb.2010.0142.

Chuine, I., Beaubien, E.G., 2001. Phenology is a major determinant of tree species range. Ecology Letters 4, 500–510. https://doi.org/10.1046/j.1461-0248.2001.00261.x.

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 Sensing of Environment 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., Schwartz, M.D., 2007. Shifting plant phenology in response to global change. Trends in Ecology & Evolution 22, 357–365. https://doi.org/10.1016/j.tree.2007.04.003.

Delbart, N., Le Toan, T., Kergoat, L., Fedotova, V., 2006. Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004). Remote Sensing of Environment 101, 52–62. https://doi.org/10.1016/j.rse.2005.11.012.

Diao, C., 2019a. Complex network-based time series remote sensing model in monitoring the fall foliage transition date for peak coloration. Remote Sensing of Environment 229, 179–192. https://doi.org/10.1016/j.rse.2019.05.003.

Diao, C., 2019b. Innovative pheno-network model in estimating crop phenological stages with satellite time series. ISPRS Journal of Photogrammetry and Remote Sensing 153, 96–109. https://doi.org/10.1016/j.isprsjprs.2019.04.012.

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

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

Didan, K., Barreto, A., 2016. NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology EVI2 Yearly Global 0.05Deg CMG. Doi: 10.5067/MEaSUREs/VIP/VIPPHEN_EVI2.004.

Donnelly, A., Yu, R., Caffarra, A., Hanes, J., Liang, L., Desai, A.R., Liu, L., Schwartz, M.D., 2017. Interspecific and interannual variation in the duration of spring phenophases in a northern mixed forest. Agricultural and Forest Meteorology 243, 55–67. https://doi.org/10.1016/j.agrformet.2017.05.007.

Donnelly, A., Yu, R., Jones, K., Belitz, M., Li, B., Duffy, K., Zhang, X., Wang, J., Seyednasrollah, B., Gerst, K.L., Li, D., Kaddoura, Y., Zhu, K., Morisette, J., Ramey, C., Smith, K., 2022. Exploring discrepancies between in situ phenology and remotely derived phenometrics at NEON sites. Ecosphere 13, e3912.

Dukes, J.S., Pontius, J., Orwig, D., Garnas, J.R., Rodgers, V.L., Brazee, N., Cooke, B., Theoharides, K.A., Stange, E.E., Harrington, R., Ehrenfeld, J., Gurevitch, J., Lerdau, M., Stinson, K., Wick, R., Ayres, M., 2009. Responses of insect pests, pathogens, and invasive plant species to climate change in the forests of northeastern North America: What can we predict?This article is one of a selection of papers from

NE Forests 2100: A Synthesis of Climate Change Impacts on Forests of the Northeastern US and Eastern Canada. Can. J. For. Res. 39, 231–248. https://doi.org/10.1139/X08-171.
Elmore, A.J., Stylinski, C.D., Pradhan, K., 2016. Synergistic Use of Citizen Science and Remote Sensing for Continental-Scale Measurements of Forest Tree Phenology. Remote Sensing 8. https://doi.org/10.3390/rs8060502.
Filippa, G., Cremonese, E., Migliavacca, M., Galvagno, M., Forkel, M., Wingate, L., Tomelleri, E., Morra di Cella, U., Richardson, A.D., 2016. Phenopix: A R package for image-based vegetation phenology. Agricultural and Forest Meteorology 220, 141–150. https://doi.org/10.1016/j.agrformet.2016.01.006.
Fisher, J.I., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sensing of Environment 109, 261–273. https://doi.org/10.1016/j.rse.2007.01.004.
Fisher, J.I., Mustard, J.F., Vadeboncoeur, M.A., 2006. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. Remote Sensing of Environment 100, 265–279. https://doi.org/10.1016/j.rse.2005.10.022.
Fitchett, J.M., Grab, S.W., Thompson, D.I., 2015. Plant phenology and climate change: Progress in methodological approaches and application. Progress in Physical Geography: Earth and Environment 39, 460–482. https://doi.org/10.1177/0309133315578940.
Galvão, L.S., dos Santos, J.R., Roberts, D.A., Breunig, F.M., Toomey, M., de Moura, Y.M., 2011. On intra-annual EVI variability in the dry season of tropical forest: A case study with MODIS and hyperspectral data. Remote Sensing of Environment 115, 2350–2359. https://doi.org/10.1016/j.rse.2011.04.035.
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 Sensing of Environment 114, 1805–1816. https://doi.org/10.1016/j.rse.2010.04.005.
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 Sensing of Environment 242, 111752. https://doi.org/10.1016/j.rse.2020.111752.
Gu, L., Post, W.M., Baldocchi, D.D., Black, T.A., Suyker, A.E., Verma, S.B., Vesala, T., Wofsy, S.C., 2009. Characterizing the Seasonal Dynamics of Plant Community Photosynthesis Across a Range of Vegetation Types. In: Noormets, A. (Ed.), Phenology of Ecosystem Processes: Applications in Global Change Research. Springer, New York, New York, NY, pp. 35–58. https://doi.org/10.1007/978-1-4419-0026-5_2.
Gunderson, C.A., Edwards, N.T., Walker, A.V., O'Hara, K.H., Campion, C.M., Hanson, P. J., 2012. Forest phenology and a warmer climate – growing season extension in relation to climatic provenance. Global Change Biology 18, 2008–2025. https://doi.org/10.1111/j.1365-2486.2011.02632.x.
Henebry, G.M., De Beurs, K.M., 2013. Remote sensing of land surface phenology: A prospectus, in: Phenology: An Integrative Environmental Science. Springer, pp. 385–411.
Hird, J.N., McDermid, G.J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sensing of Environment 113, 248–258. https://doi.org/10.1016/j.rse.2008.09.003.
Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., de Grandcourt, A., Burban, B., Flechard, C., Granier, A., Gross, P., Heinesch, B., Longdoz, B., Moureaux, C., Ourcival, J.-M., Rambal, S., Saint André, L., Soudani, K., 2013. Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. Remote Sensing of Environment 132, 145–158. https://doi.org/10.1016/j.rse.2013.01.010.
Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.
Jonsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE transactions on Geoscience and Remote Sensing 40, 1824–1832.
Kang, S., Running, S.W., Lim, J.-H., Zhao, M., Park, C.-R., Loehman, R., 2003. A regional phenology model for detecting onset of greenness in temperate mixed forests, Korea: an application of MODIS leaf area index. Remote Sensing of Environment 86, 232–242. https://doi.org/10.1016/S0034-4257(03)00103-2.
Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., Munger, J. W., O'Keefe, J., Schmid, H.P., Wing, I.S., Yang, B., Richardson, A.D., 2014. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nature Climate Change 4, 598–604. https://doi.org/10.1038/nclimate2253.
Klosterman, S.T., Hufkens, K., Gray, J.M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M.A., Richardson, A.D., 2014. Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11, 4305–4320. https://doi.org/10.5194/bg-11-4305-2014.
Kowalski, K., Senf, C., Hostert, P., Pflugmacher, D., 2020. Characterizing spring phenology of temperate broadleaf forests using Landsat and Sentinel-2 time series. International Journal of Applied Earth Observation and Geoinformation 92, 102172. https://doi.org/10.1016/j.jag.2020.102172.
Kramer, K., Leinonen, I., Loustau, D., 2000. The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and Mediterranean forests ecosystems: an overview. International Journal of Biometeorology 44, 67–75. https://doi.org/10.1007/s004840000066.
Liang, L., Schwartz, M.D., Fei, S., 2011. Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. Remote Sensing of Environment 115, 143–157. https://doi.org/10.1016/j.rse.2010.08.013.
Liang, L., Schwartz, M.D., Wang, Z., Gao, F., Schaaf, C.B., Tan, B., Morisette, J.T., Zhang, X., 2014. A Cross Comparison of Spatiotemporally Enhanced Springtime Phenological Measurements From Satellites and Ground in a Northern U.S. Mixed Forest. IEEE Transactions on Geoscience and Remote Sensing 52, 7513–7526. https://doi.org/10.1109/TGRS.2014.2313558.
Linderholm, H.W., 2006. Growing season changes in the last century. Agricultural and Forest Meteorology 137, 1–14. https://doi.org/10.1016/j.agrformet.2006.03.006.
Liu, L., Liang, L., Schwartz, M.D., Donnelly, A., Wang, Z., Schaaf, C.B., Liu, L., 2015. Evaluating the potential of MODIS satellite data to track temporal dynamics of autumn phenology in a temperate mixed forest. Remote Sensing of Environment 160, 156–165. https://doi.org/10.1016/j.rse.2015.01.011.
Menzel, A., Fabian, P., 1999. Growing season extended in Europe. Nature 397, 659. https://doi.org/10.1038/17709.
Moon, M., Richardson, A.D., Friedl, M.A., 2021. Multiscale assessment of land surface phenology from harmonized Landsat 8 and Sentinel-2, PlanetScope, and PhenoCam imagery. Remote Sensing of Environment 266, 112716. https://doi.org/10.1016/j.rse.2021.112716.
Parry, M.L., 2007. Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC. Cambridge University Press.
Peng, D., Wu, C., Li, C., Zhang, X., Liu, Z., Ye, H., Luo, S., Liu, X., Hu, Y., Fang, B., 2017. Spring green-up phenology products derived from MODIS NDVI and EVI: Intercomparison, interpretation and validation using National Phenology Network and AmeriFlux observations. Ecological Indicators 77, 323–336. https://doi.org/10.1016/j.ecolind.2017.02.024.
Peñuelas, J., Rutishauser, T., Filella, I., 2009. Phenology Feedbacks on Climate Change. Science 324, 887–888. https://doi.org/10.1126/science.1173004.
Piao, S., Friedlingstein, P., Ciais, P., Viovy, N., Demarty, J., 2007. Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. Global Biogeochemical Cycles 21. https://doi.org/10.1029/2006GB002888.
Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: Current progresses and challenges. Global Change Biology 25, 1922–1940. https://doi.org/10.1111/gcb.14619.
Polgar, C.A., Primack, R.B., 2011. Leaf-out phenology of temperate woody plants: from trees to ecosystems. New Phytologist 191, 926–941. https://doi.org/10.1111/j.1469-8137.2011.03803.x.
Post, E., Pedersen, C., Wilmers, C.C., Forchhammer, M.C., 2008. Warming, plant phenology and the spatial dimension of trophic mismatch for large herbivores. Proceedings of the Royal Society B: Biological Sciences 275, 2005–2013. https://doi.org/10.1098/rspb.2008.0463.
Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P., Ollinger, S.V., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications 19, 1417–1428. https://doi.org/10.1890/08-2022.1.
Richardson, A.D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., Hollinger, D.Y., Kutsch, W.L., Longdoz, B., Luyssaert, S., Migliavacca, M., Montagnani, L., William Munger, J., Moors, E., Piao, S., Rebmann, C., Reichstein, M., Saigusa, N., Tomelleri, E., Vargas, R., Varlagin, A., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 3227–3246. https://doi.org/10.1098/rstb.2010.0102.
Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agricultural and Forest Meteorology 169, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012.
Richardson, A.D., Hufkens, K., Milliman, T., Aubrecht, D.M., Chen, M., Gray, J.M., Johnston, M.R., Keenan, T.F., Klosterman, S.T., Kosmala, M., Melaas, E.K., Friedl, M. A., Frolking, S., 2018. Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. Scientific Data 5, 180028. https://doi.org/10.1038/sdata.2018.28.
Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., Ohno, H., 2005. A crop phenology detection method using time-series MODIS data. Remote Sensing of Environment 96, 366–374. https://doi.org/10.1016/j.rse.2005.03.008.
Schwartz, M.D., 2003. Phenology: an integrative environmental science. Springer.
Schwartz, M.D., Crawford, T.M., 2001. DETECTING ENERGY-BALANCE MODIFICATIONS AT THE ONSET OF SPRING. Physical Geography 22, 394–409. https://doi.org/10.1080/02723646.2001.10642751.
Schwartz, M.D., Hanes, J.M., 2010. Intercomparing multiple measures of the onset of spring in eastern North America. International Journal of Climatology 30, 1614–1626. https://doi.org/10.1002/joc.2008.
Schwartz, M.D., Reed, B.C., White, M.A., 2002. Assessing satellite-derived start-of-season measures in the conterminous USA. International Journal of Climatology 22, 1793–1805. https://doi.org/10.1002/joc.819.
Schwartz, M.D., Hanes, J.M., Liang, L., 2013. Comparing carbon flux and high-resolution spring phenological measurements in a northern mixed forest. Agricultural and Forest Meteorology 169, 136–147. https://doi.org/10.1016/j.agrformet.2012.10.014.
Tan, B., Morisette, J.T., Wolfe, R.E., Gao, F., Ederer, G.A., Nightingale, J., Pedelty, J.A., 2010. An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4, 361–371.

Tang, J., Körner, C., Muraoka, H., Piao, S., Shen, M., Thackeray, S.J., Yang, X., 2016. Emerging opportunities and challenges in phenology: a review. Ecosphere 7, e01436.

Tian, J., Zhu, X., Shen, M., Chen, J., Cao, R., Qiu, Y., Xu, Y.N., 2024. Effectiveness of spatiotemporal data fusion in fine-scale land surface phenology monitoring: A simulation study. Journal of Remote Sensing. https://doi.org/10.34133/remotesensing.0118.

Vitasse, Y., Delzon, S., Dufrêne, E., Pontailler, J.-Y., Louvet, J.-M., Kremer, A., Michalet, R., 2009. Leaf phenology sensitivity to temperature in European trees: Do within-species populations exhibit similar responses? Agricultural and Forest Meteorology 149, 735–744. https://doi.org/10.1016/j.agrformet.2008.10.019.

White, K., Pontius, J., Schaberg, P., 2014. Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. Remote Sensing of Environment 148, 97–107. https://doi.org/10.1016/j.rse.2014.03.017.

White, M.A., Thornton, P.E., Running, S.W., 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles 11, 217–234. https://doi.org/10.1029/97GB00330.

White, M.A., Running, S.W., Thornton, P.E., 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. International Journal of Biometeorology 42, 139–145. https://doi.org/10.1007/s004840050097.

White, M.A., Hoffman, F., Hargrove, W.W., Nemani, R.R., 2005. A global framework for monitoring phenological responses to climate change. Geophysical Research Letters 32. https://doi.org/10.1029/2004GL021961.

White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'keefe, J., Zhang, G., Nemani, R.R., van Leeuwen, W.J., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Global Change Biology 15, 2335–2359.

Xiao, X., Hagen, S., Zhang, Q., Keller, M., Moore, B., 2006. Detecting leaf phenology of seasonally moist tropical forests in South America with multi-temporal MODIS images. Remote Sensing of Environment 103, 465–473. https://doi.org/10.1016/j.rse.2006.04.013.

Xie, Y., Wang, X., Silander, J.A., 2015. Deciduous forest responses to temperature, precipitation, and drought imply complex climate change impacts. Proceedings of the National Academy of Sciences 112, 13585–13590. Doi: 10.1073/pnas.1509991112.

Xie, Y., Wang, X., Wilson, A.M., Silander, J.A., 2018. Predicting autumn phenology: How deciduous tree species respond to weather stressors. Agricultural and Forest Meteorology 250–251, 127–137. https://doi.org/10.1016/j.agrformet.2017.12.259.

Yang, Z., Diao, C., Li, B., 2021. A Robust Hybrid Deep Learning Model for Spatiotemporal Image Fusion. Remote Sensing 13. https://doi.org/10.3390/rs13245005.

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 Sensing of Environment 237, 111511. https://doi.org/10.1016/j.rse.2019.111511.

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 Sensing of Environment 84, 471–475. https://doi.org/10.1016/S0034-4257(02)00135-9.

Zhang, X., Friedl, M.A., Schaaf, C.B., 2009. Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. International Journal of Remote Sensing 30, 2061–2074. https://doi.org/10.1080/01431160802549237.

Zhang, X., Liu, L., Yan, D., 2017. Comparisons of global land surface seasonality and phenology derived from AVHRR, MODIS, and VIIRS data. Journal of Geophysical Research: Biogeosciences 122, 1506–1525. https://doi.org/10.1002/2017JG003811.

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

Zhao, Y., Diao, C., Augspurger, C.K., Yang, Z., 2023. Monitoring spring leaf phenology of individual trees in a temperate forest fragment with multi-scale satellite time series. Remote Sensing of Environment 297, 113790. https://doi.org/10.1016/j.rse.2023.113790.

Zhu, X., Zhan, W., Zhou, J., Chen, X., Liang, Z., Xu, S., Chen, J., 2022. A novel framework to assess all-round performances of spatiotemporal fusion models. Remote Sensing of Environment 274, 113002. https://doi.org/10.1016/j.rse.2022.113002.