Mapping understory plant communities in deciduous forests from Sentinel-2 time series

Xiucheng Yang [a,*], Shi Qiu [a], Zhe Zhu [a], Chadwick Rittenhouse [a], Dustin Riordan [a,b], Mari Cullerton [a]

[a] *Department of Natural Resources and the Environment, University of Connecticut, Storrs, CT 06269, USA*
[b] *Connecticut Department of Energy and Environmental Protection, Marlborough, CT 06447, USA*

ARTICLE INFO

ABSTRACT

Understory plant communities are an integral component of deciduous forests, playing a vital role in the overall health of the ecosystem. However, remote sensing of understory plant communities is challenging due to the obstruction by the forest canopy. In this study, we proposed an automated dense Sentinel-2 time series-based approach for understory plant communities and created maps of four understory classes (i.e., native shrubs of *greenbrier* and *mountain laurel*, and invasive shrubs of *barberry* and the assemblage of *mixed invasive*) at 10 m resolution in Connecticut's deciduous forests in 2020. A harmonic time series model and three years of Sentinel-2 time series from 2019 to 2021 were used to classify understory species based on their unique, intra-annual phenology characteristics. The time series model coefficients captured the subtle phenology differences and created synthetic cloud-free images within a short temporal window in the spring prior to canopy leaf-on (hereafter called "observation window"). During the observation window, Sentinel-2 data penetrated the deciduous overstory canopy and observed the unique trajectories of different understory species due to their phenology differences. We also calculated spatial texture features (i.e., mean, second moment, and contrast from gray level co-occurrence matrix) based on the synthetic images created within the observation window to capture the different conditions of leaf growth and distinct spatial patterns within deciduous forests. By using the spectral, temporal, and spatial features as input variables from dense Sentinel-2 data, auxiliary data (i.e., LiDAR and soil drainage layer), a random forest classifier, and a new strategy to iteratively select representative sample (namely ISRS), understory species maps were created with an overall accuracy of approximately 93%, and the user's and producer's accuracies varied from 39% to 99% for the three mapped understory species and one assemblage of species. The proposed method created an accurate binary map of understory presence with an overall accuracy of 95%, a producer's accuracy of 84%, and a user's accuracy of 68%. Additionally, we separated the invasive (i.e. *barberry* and *mixed invasive* of *multi-flora rose*, *oriental bittersweet*, *honeysuckle*, *winged euonymus*, and *autumn olive*) and native (*greenbrier* and *mountain laurel*) species with an overall accuracy of 94%. We estimated that the invasive species cover an area of 649.33 ± 140.59 km², which occupied a large proportion (~53%) of the shrub understory in Connecticut's deciduous forests.

1. Introduction

Understory, a vegetative layer of shrubs and young or small trees between the forest canopy and the forest floor, is crucial in the functioning of Earth's forest ecosystems. It contributes to plant species richness (Schafer et al., 2014), provides habitat and forage for many wildlife species (McDermid et al., 2009; Pisek et al., 2012), influences a variety of forest processes such as nutrient cycling and tree regeneration (Ploughe and Dukes, 2019), reflects sustaining ecosystem productivity and stability (Sanz et al., 2020; Sumnall et al., 2021), provides long-term stability in forest carbon cycling (Hubau et al., 2019), and responds to environmental changes (Scolastri et al., 2017). Given these crucial ecosystem services, it is essential to quantify the presence, spatial distribution, and species composition of understory vegetation in forested environments (Araujo et al., 2020; Lee et al., 2022; Utz and Fetsko, 2020; Van Doninck et al., 2020; Żołnierz et al., 2016).

Understory surveys often require labor-intensive fieldwork that is sometimes logistically unfeasible at large spatial extents (Tuanmu et al., 2010). Remote sensing technology has been a useful alternative tool in recent years. Understory mapping has been drawing increasing attention but still faces extreme challenges in the remote sensing community (See the summary of the previous studies in Table S1) because the overstory canopy may obscure the visibility of understory in remote sensing data sources. Light detection and ranging (LiDAR) can characterize the three-dimensional structure of forests (Scolastri et al., 2017) and is the most widely used dataset to map the presence of understory based on height information. However, this method cannot differentiate between understory species (Asner and Vitousek, 2005; Campbell et al., 2018; Crespo-Peremarch et al., 2018; Li et al., 2021b; Singh et al., 2015; Sumnall et al., 2021).

Passive optical remote sensing data, including very high resolution (VHR) images from satellites and drones (Araujo et al., 2020; Li et al., 2020; Meng et al., 2018; Oreti et al., 2020; Shouse et al., 2012; Welch et al., 2002), medium spatial resolution satellite images (e.g., Landsat and Sentinel-2)(Chastain and Townsend, 2007; Dai et al., 2020; Singh et al., 2018; Singh and Gray, 2020; Wang et al., 2009), and coarse spatial resolution satellite images (e.g., MODIS) (Tuanmu et al., 2010), also have demonstrated potential for understory mapping. The reflectance of forest stands in a multispectral optical image has two major vegetation components: overstory canopy and understory (Pisek et al., 2016; Rautiainen et al., 2009). The relative contributions of these two layers change with the growing season (Pisek et al., 2015; Rautiainen and Heiskanen, 2013). Differences in leaf phenology are key to analyzing the understory from optical images (Rautiainen et al., 2011; Shouse et al., 2012; Tuanmu et al., 2010; Wilfong et al., 2009). Previous studies have explored the contribution of understory on leaf area index estimation (Chen and Cihlar, 1996; Ganguly et al., 2012; Nikopensius et al., 2015) and estimated their spatial distribution in the sparse boreal evergreen forests (Kobayashi et al., 2016; Pisek et al., 2016; Rautiainen et al., 2011), where the presence of understory vegetation significantly influenced the reflectance spectrum of forests.

In deciduous forests, the signal of understory vegetation in satellite images becomes faint because of overstory canopy closure (Duncanson et al., 2014; Gibson et al., 2020). Some studies used multiple satellites (e.g., Landsat and Sentinel-2) to capture the seasonal variation and spectral difference between evergreen understory and deciduous canopy. Some evergreen understory vegetation, such as bamboo (Wang et al., 2019), *L. sinense* plant (Singh et al., 2018), and other evergreen shrubs (e.g.,

Pistacia lentiscus, Cistus ladanifer, and *Retama sphaerocarpa* (Fragoso-Campón et al., 2020)), were mapped in the previous studies. Other studies considered the most distinct spectral difference between deciduous forests with and without understory from optical satellite images acquired in spring green-up period (Rautiainen et al., 2011). The levels of light, water, and nutrients are greatest in the early spring compared to other seasons (Neufeld and Young, 2014), which allows understory vegetation to produce new leaves several weeks before deciduous trees in the same habitat (Fridley, 2012; McLachlan and Bazely, 2001). That is, understory vegetation becomes green while the deciduous tree canopy above is still in leaf-off condition (Hicks and Taylor, 2015). Optical remote sensing images can observe the greening signal of understory vegetation in this specific time window (hereafter called "observation window"). For example, Singh and Gray (2020) concluded that the use of selected single-date Landsat images acquired in early spring (March 29, 2011) can produce higher overall accuracy of understory mapping than the use of multitemporal images. The observation window (Fig. 1), which varies based on the ecoregions and target vegetation species, provides unique opportunity to observe distinct phenology characteristics between different understory plant communities and deciduous forest canopy.

The identification of specific species, along with the location of the understory plant communities, is important, particularly with respect to vegetation composition and spread of invasive species. The characterization of species and their distribution benefits the understanding of links among global environmental change, invasive species spread over natives (Wilde et al., 2015), and the extinction of local understory species (Landuyt et al., 2018). Specifically, the spread of invasive understory species threatens biodiversity and ecosystem functioning (O'Loughlin et al., 2019) by modifying and changing soil conditions, vegetation structure, and understory microclimate (Link et al., 2018; Maynard-Bean and Kaye, 2019). To our knowledge, understory species identification is still not well addressed in the remote sensing community and only a few studies are available (Table S1), most of which focused on mapping evergreen understory species in the deciduous forests. For example, Welch et al. (2002) used aerial images acquired in the March/April 1997–1998 to manually interpret the evergreen understory species. Chastain and Townsend (2007) classified two evergreen understory species (i.e., *rosebay rhododendron* and *mountain laurel*) from Landsat Enhanced Thematic Mapper Plus (ETM+) images. Tuanmu et al. (2010) extracted understory bamboo distributions and separated them into two understory bamboo species (i.e., *arrow* and *umbrella*)
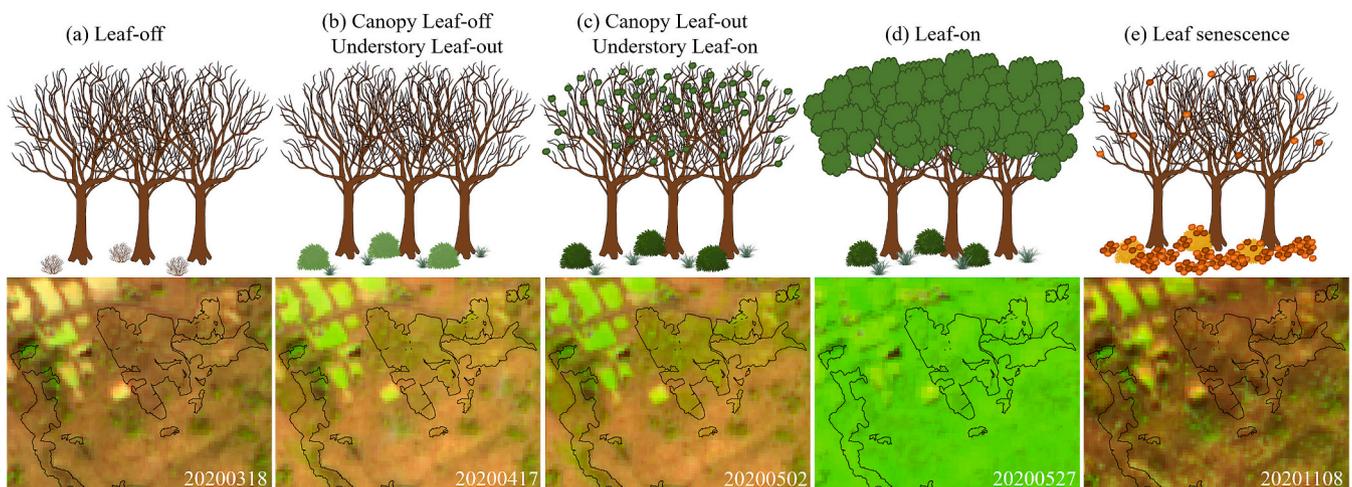
Fig. 1. An example of the distinct phenology between deciduous understory and forest canopy. The "observation window" from Sentinel-2 images is between (b) leaf-out of understory and (c) leaf-out of the canopy. The Sentinel-2 images in the second row depict each vegetation phenology phase at a location in the University of Connecticut Forest (-72.25, 41.83). The spatial distribution of the deciduous understory of *barberry* is delineated in black polygons and were established from fieldwork. The Sentinel-2 image chips are composed of bands 11 (SWIR), 8 (NIR), and 4 (red) with the same stretch, and thus they are directly comparable. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

using a single elevation variable based on the prior knowledge that they grow at different elevations.

Mapping deciduous understory in deciduous forests is much more challenging than mapping evergreen understory based on remote sensing data, as both the deciduous understory and overstory canopies could share similar phenology. Therefore, the images acquired in the narrow observation window become increasingly important to capture the early leaf-out phenology of the deciduous understory before the closure of the deciduous overstory. For example, Shouse et al. (2012) used an aerial image in early spring (late March and early April) to identify the canopy – *bush honeysuckle*, of which the spring leaf development is typically two to three weeks earlier than that of native shrubs and canopy. Thus, collecting distinct growing dates and selecting optimal satellite images is important for mapping these target understory plant communities. However, it is very likely that during the narrow observation window, there are no clear (free of cloud, cloud shadow, and snow) images available, making large-scale and operational mapping of understory species difficult, particularly for deciduous understory in deciduous forests.

The dense time series analysis by using Landsat and Sentinel-2 created a new opportunity to reconstruct the temporal trajectory for each pixel (Yang B. et al., 2022; Zhang et al., 2021a; Zhu, 2017), which does not require completely clear images. The temporal trajectory depicts the vegetation dynamics and its phenology patterns (e.g., spring leaf-out, summer leaf-on, autumn leaf-senescence and winter leaf-off) for each pixel throughout a year. Moreover, instead of using the original surface reflectance values for a certain date, the time series coefficients (Brown et al., 2020; Bullock et al., 2020; Cohen et al., 2020; Wang et al., 2020; Zhang et al., 2022) and other derived phenological metrics (Hermosilla et al., 2022; Huang et al., 2017; Pasquarella et al., 2018; Sun et al., 2021a) can be used as the input variables to conduct the land cover classification and identify different vegetation species. Some studies used phenology characteristics derived from time series analysis to map the understory, demonstrating that full-year phenology is much more effective to extract the understory compared to simply using multitemporal images (Becker et al., 2013; Tuanmu et al., 2010). However, the potential of dense time series analysis on understory species identification was not well explored.

Sentinel-2 satellite images provide a high level of detail, with a spatial resolution of 10 m and frequent revisit interval of five days. Some previous work demonstrated the advantages of using dense time series of Sentinel-2 to map salt marsh plant species (Sun et al., 2021a), temperate forest tree species (Hemmerling et al., 2021), floodplain grassland plant communities (Rapinel et al., 2019), and crop types (Belgiu and Csillik, 2018). The increasing spatial and temporal resolution of Sentinel-2 creates new possibilities for generating spatial distribution maps of understory plant communities. Ahl et al. (2006) suggested that satellite data should have a temporal resolution sufficient to capture leaf expansion (<1 week) for phenological studies on understory. Lee et al. (2022) recently also reported that the North American light window duration (i.e., observation window) between deciduous forest canopy and understory averages 11.7 ± 4.1 days. Understory plant communities are characterized by a wide variety of species and the high spatial resolution of the satellite images benefits the mapping and estimation of different understory classes (Su et al., 2022; Xi et al., 2022). For example, Shouse et al. (2012) investigated the efficacy of detecting understory shrubs by assessing the simulated spatial resolution of 10 m, 15 m and 30 m, and ultimately concluded that the Landsat resolution (30 m) is inadequate for mapping these shrubs.

Therefore, this study endeavors to investigate whether Sentinel-2 time series analysis can be used to accurately classify various types of understory plant communities in deciduous forests. We specifically proposed an automated approach to address the following challenges: (1) is the Sentinel-2 time series model capable of detecting the distinct phenology between the understory (particularly deciduous understory species) and deciduous forest canopy; (2) how effectively can the various understory species be separated and also labeled as native or invasive species?

2. Study area and materials

2.1. Study area

The study area covers the entire state of Connecticut (hereafter CT), with a total land area of 11,624 km^2. Deciduous forest covers approximately 40% of our study area (Fig. 2), and there are large areas of shrub understories beneath the deciduous forests, such as the native evergreen *mountain laurel*, native woody vining *greenbrier* and invasive deciduous shrubs composed of *barberry* and *mixed invasive* (such as *barberry, multiflora rose, honeysuckle, oriental bittersweet, winged euonymus,* and *autumn olive* in varying proportions). In CT, *mountain laurel* is a widely distributed and healthy native shrub which can create persistent understories (Brose, 2017). Unfortunately, the generation of the extremely dense thickets can cause a variety of forest management issues, such as hindrance to forest renewal and damage to overstory trees (Brose, 2016, 2017; Royo and Carson, 2006). *Greenbrier* is a native woody vine in coastal regions that can grow up to 3-6 m when invading woodlands. In open areas, it spreads over shrubs and herbaceous plants, blocking out sunlight and killing them in the process. This leads to a decrease in plant species diversity in coastal grasslands and heathlands (Ohman, 2006). The rapid growth rate of invasive species poses an urgent threat to forests, as they can outcompete native plants and quickly become the dominant species in understory ecosystems (Link et al., 2018; Utz and Fetsko, 2020). In CT, the rapid expansion of invasive understory species, particularly *barberry*, has resulted in their presence in nearly every county (EDDMapS, 2023; Kartesz, 2015). This has caused a decrease in the abundance of desirable native species, such as wildflowers and grasses (Linske et al., 2018; Ward et al., 2009; Williams and Ward, 2010). As such, it is essential to keep track of the spread of invasive species and the presence of native species in CT.

In this study, we aimed to detect and identify these four understory plant communities, hereafter called the "target understory" (Fig. 3 (a-d)). Notably, in addition to the bare forest floor beneath the deciduous forest canopy (second row in Fig. 3 (e-i)), a variety of other understory types, such as deciduous and coniferous tree saplings, non-target native shrub (e.g., *sweet pepperbush, highbush blueberry, lowbush blueberry, and northern spicebush*) and herbaceous ground covers (e.g., *skunk cabbage, moss,* and *fern*) may also exist. We combined them into a single category for this study – "others" in deciduous forests.

2.2. Reference data

Our study aims to map the understory species beneath deciduous forests. First, to define the potential regions of deciduous forests and limit the spatial distribution of understory, we collected the training sample of deciduous forests and non-deciduous forests based on manual interpretation from the VHR image and National Land Cover Database (NLCD) map (Dewitz and Survey, 2021). Second, because it is hard to reliably interpret understory vegetation directly from remote sensing images, we collected the reference data for understory mapping based on field visits. We collected high-quality training data for target understory plant communities and "others" in deciduous forests to identify understory species. Finally, to verify the accuracy of the predicted understory species classification map (Section 4.2), we used stratified random sampling (Olofsson et al., 2014) to create validation sample and interpreted their class through in-situ visits and on-screen work.

2.2.1. Training data for deciduous forest mapping

To generate the deciduous forest map at a spatial resolution of 10 m, we collected the training data of deciduous forests and non-deciduous forests from manual interpretation of the VHR image and NLCD maps (See details in *Supplementary S1.1*). We interpreted 220 deciduous forest
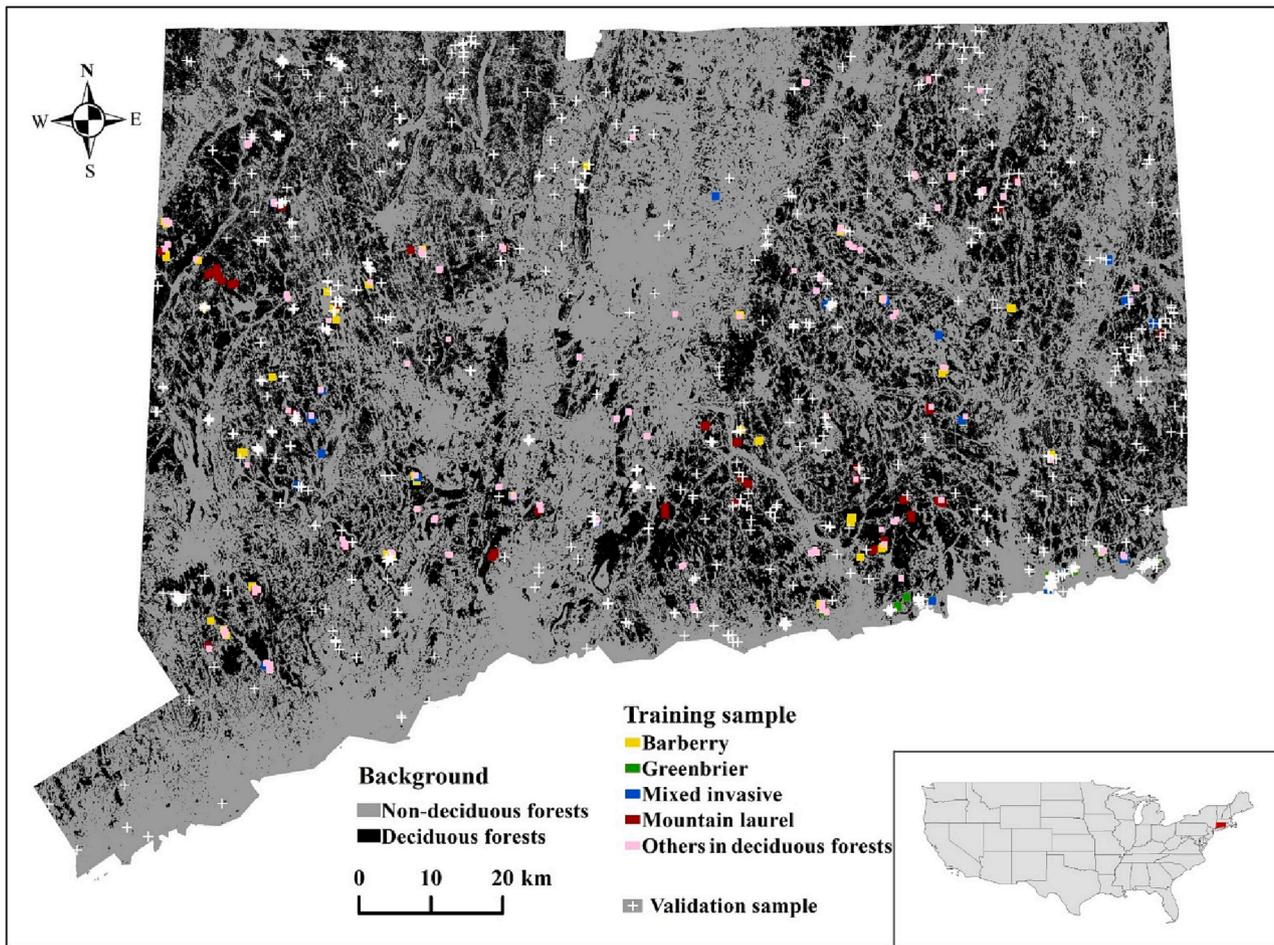
Fig. 2. Connecticut (CT) State overlaid with deciduous forests (black colour) map at 10 m resolution (generated in Section 3.2). The training data for four kinds of target understory (including *barberry*, *greenbrier*, *mixed invasive*, and *mountain laurel*) and "others" in deciduous forests were collected in November 2020 and April 2021. The white crosses are the independent validation sample that were collected in June 2021 based on the stratified random sampling strategy.

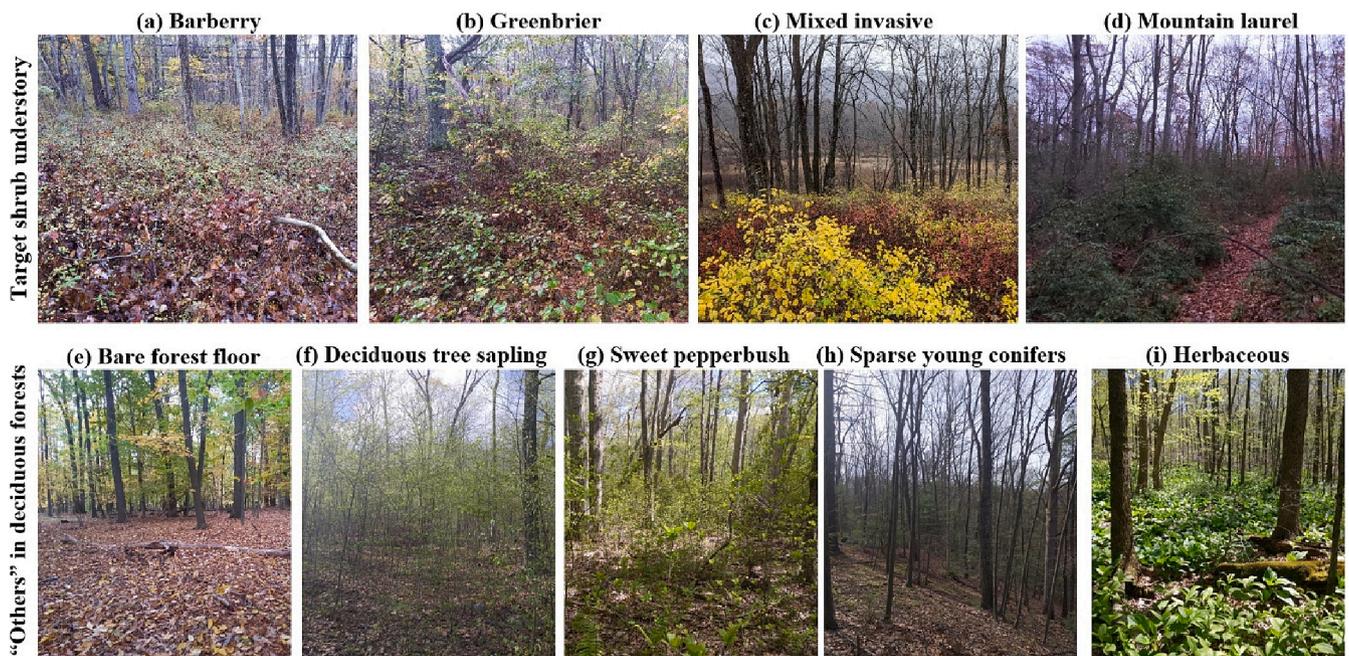**Fig. 3.** Photos of target shrub understory (*barberry, greenbrier, mixed invasive, and mountain laurel*) and "others" in deciduous forests (with different ground conditions beneath the forest canopy) taken in November 2020 and April 2021 in deciduous forests of Connecticut (CT).

polygons, 200 coniferous forest polygons and 198 mixed forest polygons based on human interpretation on the reference of VHR NAIP image. We also collected training data for non-forest classes (i.e., water, developed, barren, shrubland, herbaceous, planted/cultivated, wetlands) based on random sampling in the NLCD 2019 map.

Table 1
Definition of target understory and "others" in deciduous forests, as well as their in-field training data collected in this study.

Category		Number of pixels (Number of Polygonal objects)	Description
Target understory plant communities	*Mountain laurel*	14,199 (75)	*Mountain laurel* is an evergreen shrub species, native to CT that holds its leaves year-round and grows 3-9 m tall.
	Barberry	4830 (87)	*Barberry* is a deciduous invasive shrub that leaves out 2-3 weeks before deciduous trees and grows up to 2.5 m. *Barberry* forms a monoculture that negatively impacts seed germination of native plant species and provides microclimate conditions amenable to black-legged ticks.
	Greenbrier	987 (65)	*Greenbrier* is a native, deciduous to semi-evergreen woody vining species and can grow up to 6 m by climbing objects and vegetation. It has a green stem year-round and provides forage for wildlife.
	Mixed invasive	2406 (69)	*Mixed invasive*, used in this study to describe a plant community including *barberry, multi-flora rose, oriental bittersweet, honeysuckle, winged euonymus,* or *autumn olive* in differing proportions. Greenbrier sometimes co-occurs with invasive species and is included here. These species can be indicative of past land use and can negatively impact seed germination while also providing cover for wildlife species
"Others" in deciduous forests	Non-target shrub	1751 (45)	Deciduous forests without target understory. Shown in the second row of Fig. 3 (e-i), different kinds of conditions may exist beneath the canopies, including the bare land, deciduous tree saplings, non-target shrub understory (mainly consisting of native understory species such as *sweet pepperbush, highbush blueberry, lowbush blueberry, and northern spicebush*), sparse young conifers, and herbaceous (such as *skunk cabbage, moss,* and *fern*)
	Sparse coniferous	1288 (13)	
	Tree sapling	1973 (37)	
	Herbaceous	1194 (38)	
	Clear bare land	7500 (158)	

2.2.2. In-field training data collection for understory mapping

We collected training data (Table 1) in deciduous forests with and without target understory over multiple field visits carried out in November 2020 and April 2021. We collected training data for four target understory plant communities in high density regions, where the target understory communities comprised at least 75% cover of a Sentinel-2 pixel. The *mixed invasive* understory type included a mix of the deciduous species of *barberry, multi-flora rose, honeysuckle, oriental bittersweet, winged euonymus,* and *autumn olive* with varying proportions, where the proportion of *barberry* was <50%. A small proportion of *greenbrier* could be included in mixed invasive because it frequently co-occurred with *invasive species*. Additionally, we collected data for "others" in deciduous forests without the target understory plant communities, which includes different deciduous forests ground conditions such as bare land, deciduous tree saplings, non-target shrub understory (mainly consisting of native species such as *sweet pepperbush, highbush blueberry, lowbush blueberry,* and *northern spicebush*), sparse young co-nifers, and herbaceous vegetation (such as skunk cabbage, moss, and fern). Training data (polygons) were collected in the field using the Collector for ArcGIS on a smartphone paired to a GPS receiver or the smartphone's internal GPS receiver. Training data polygons were collected while the GPS location was reported with a root mean square error (RSME) of <5 m. In total, we collected 587 polygonal objects comprising 36,128 Sentinel-2 pixels at 10 m resolution.

2.2.3. Validation sample collection

We followed the "good practice" recommendations (Olofsson et al., 2014) and used the stratified random sampling design to generate the location of validation sample for accuracy assessment of the predicted understory maps. The strata were determined by the understory species map (four understory classes and "others" which included "others" in deciduous forests, non-deciduous forests, and non-forest classes). The number of collected validation pixels was 105 (*barberry*), 69 (*mixed invasive*), 82 (*greenbrier*), 124 (*mountain laurel*), and 295 ("others"), respectively. Finally, the accuracy assessment and area estimation of the understory vegetation map was conducted using 675 validation sample, of which 426 were collected from field visits at locations that were publicly accessible in June 2021 (mainly for deciduous forest with and without understory) and 249 were derived from aerial photo in-terpretations (mainly for non-deciduous forest sample). In the validation process, we defined a pixel as understory existence if the density of the target species is larger than 10% cover in ground. If multiple target understory communities were present in the validation plot, we defined the pixel as *mixed invasive* only if both *barberry* and *greenbrier* comprised <50% of the proportion. We defined the pixel as *barberry* or *greenbrier* if one of them comprised >50% of the region even though some other target understory plant communities existed in the plot.

2.3. Sentinel-2 data

The Sentinel-2 mission is currently composed of a two-satellite sys-tem, Sentinel-2A and Sentinel-2B, which were launched on 23 June 2015 and 07 March 2017 respectively, allowing for a high revisit fre-quency of five days at a high spatial resolution of up to 10 m (Drusch et al., 2012; Yang et al., 2018). In this study, we used six broad bands (i. e., visible (VIS) and near-infrared (NIR) bands at 10 m and two short-wave infrared (SWIR) bands at 10 m) and four narrow bands (three red edge bands and narrow NIR at 20 m) (See Table S2). The two SWIR, three eed edge and narrow NIR bands with a spatial resolution of 20 m were resized to 10 m by using cubic interpolation to maintain the same spatial resolution as the VIS and NIR bands.

Five popular vegetation indices (See Table S2), including normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), soil adjusted vegetation index (SAVI), normalized burn ratio (NBR), and red edge normalized difference vegetation index (RENDVI) were included as additional variables to highlight the differences among

different understory species in addition to Sentinel-2 surface reflectance bands.

We downloaded all the available Sentinel-2 Level-1C product, Top-Of-Atmosphere (TOA) reflectance, with cloud cover <100% between 2019 and 2021 in four Sentinel-2 tiles (T18TXL, T18TXM, T18TYL and T18TYM) from USGS Earth Explorer (https://earthexplorer.usgs.gov). Among the four tiles that cover CT, there are three Sentinel-2 orbits (i.e., R011, R054 and R111), but we only used a total of 745 Sentinel-2 images (cloud cover <100%) acquired from the single orbit (R011). This single orbit processing was used to remove the partial coverage images acquired by adjacent orbit paths (R054 and R111), which can lead to noticeable artifacts between orbit overlap and nonoverlap areas because of unbalanced observation densities (see details in *Supplementary S6*). For each Sentinel-2 image, the TOA data were converted to surface reflectance by using the version 2.8 Sen2Cor atmospheric correction module, and cloud, cloud shadow, and snow were masked by using Fmask 4.6 algorithm (Qiu et al., 2019).

Although Sentinel-2 has a repeat time of approximately five days, the temporal density of the clear observations is reduced substantially owing to cloud and cloud shadow. In this study, we assumed that the understory and deciduous forests in CT remain unchanged between 2019 and 2021 (with 2020 being the epoch year) and merged the whole Sentinel-2 clear observations in these three years to generate the three-fold dense Sentinel-2 observations in a single calendar year.

2.4. Auxiliary data

2.4.1. Soil drainage layer

A soil drainage layer was combined in the pre-definition of the potential layer of the understory distribution map in *Section 3.2*. Eight levels of drainage class of soils are defined by the USDA Natural Resources Conservation Service Soil Survey Geographic (SSURGO) database (https://websoilsurvey.sc.egov.usda.gov/; last access on 01/24/2023). The very poorly drained soil class is mostly wetland soils that have free water present at or very near the ground surface during much of the growing season. In these conditions, wetland plant communities including herbaceous species, such as *skunk cabbage* and *moss*, and shrub species, such as *highbush blueberry* and *sweet pepperbush*, would be the popular understory plants beneath the canopy (Kiviat, 2023; Paolucci and Stolt, 2018; Poindexter and Thompson, 2009). These types of vegetation could be misclassified as *barberry* and *mixed invasive* due to a similar early-spring leaf out phenology. *Barberry* and the other invasive plants included in the *mixed invasive* plant community, such as *multiflora rose*, *honeysuckle*, *bittersweet*, and *euonymus*, are not suited to
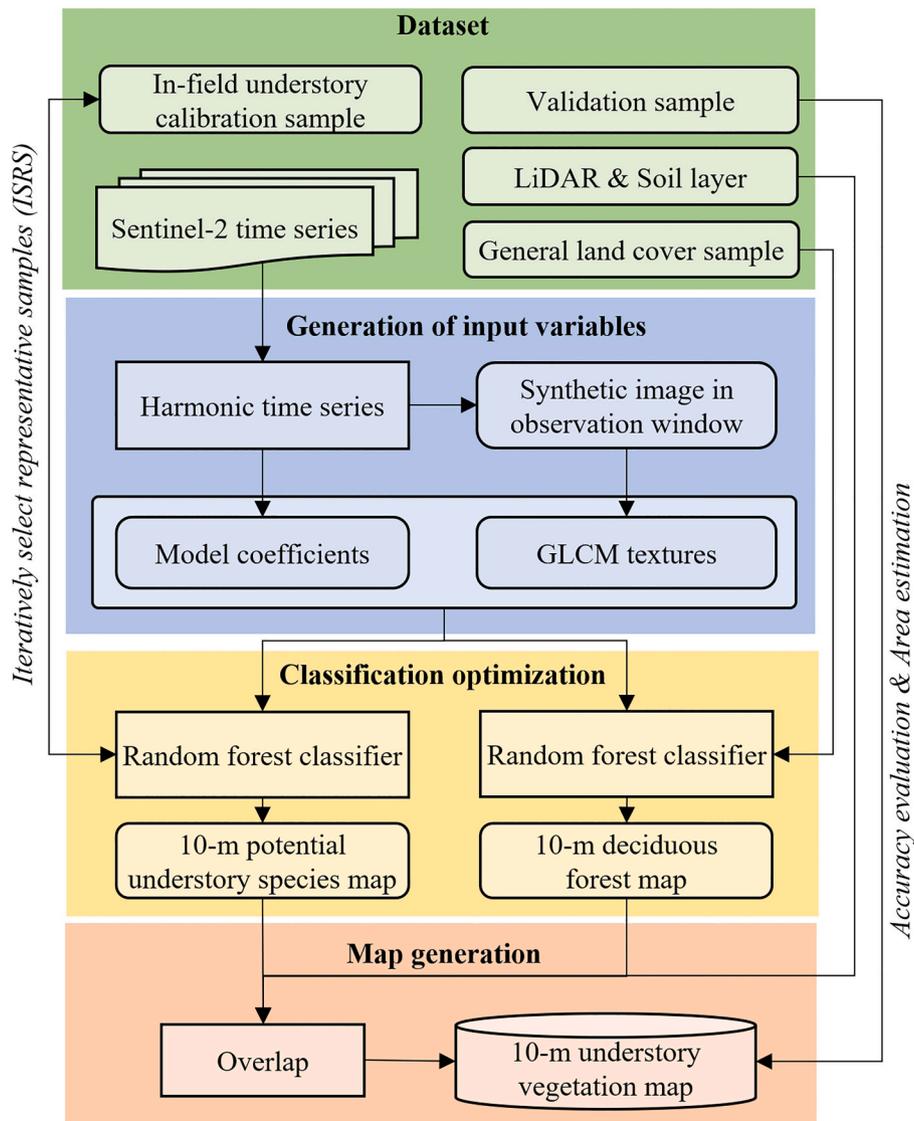

Fig. 4. Workflow of the understory species mapping approach with iteratively select representative sample (ISRS).

establish and persist in standing water or in very poorly drained soils. Thus, we exclude the very poorly drained soils that could not support the growth of the target understory plant communities.

2.4.2. LiDAR

We generated a canopy height model by subtracting the bare earth Digital Elevation Models (DEM) from the height of LiDAR cloud points. The LiDAR data were downloaded from Connecticut Environmental Conditions Online (CT ECO), which were collected by the Leica ALS70 sensor onboard airplane between March 11, 2016 and April 16, 2016 (http://www.cteco.uconn.edu/data/lidar/info_lidar.htm; last access on 01/24/2023). CT ECO also provides the bare earth DEM data with 1-m resolution based on LiDAR data. The canopy height information is incorporated to define the regions of deciduous forests.

3. Methodology

The methods consist of five distinct steps described in the following sections (Fig. 4): estimating the harmonic time series model and generation of the spectral, temporal and spatial variables in Section 3.1, definition of the potential distribution regions of the understory plant communities in deciduous forests in Section 3.2, building of an optimized random forest classifier by Iteratively Selecting Representative Sample (namely ISRS) in Section 3.3, removal of the isolated pixels in the post-processing step of Section 3.4, and accuracy assessment and area estimation of the spatial distribution of the understory plant communities in CT in Section 3.5.

3.1. Input variables

Unlike most of the understory mapping approaches that used satellite images acquired at selected dates to highlight the reflectance difference between the understory and overstory, our algorithm used all available Sentinel-2 observations to build harmonic time series models. The input variables include the spectral and temporal features derived from the harmonic time series model and spatial information measured by the texture feature of synthetic images in the observation window. The spectral bands, including blue, green, red and NIR broadbands, two SWIR broadbands and four red edge narrowbands, and five vegetation indices (i.e., NDVI, EVI, SAVI, NBR, and RENDVI) were considered. For each spectral band or index, a harmonic model is created based on the time series of Sentinel-2 data, and its coefficients can represent the spectral and temporal features. Based on the harmonic model, cloud-free synthetic Sentinel-2 images in the observation window can be predicted

Table 2
A total of 360 potential input variables were extracted from the spectral, temporal, and spatial data dimensions for ten surface reflectance bands and five vegetation indices of Sentinel-2 data. Please refer to Eq. (1) for specific variable names.

Dimension	Source	Variables	Number	Descriptions
Spectral/ Temporal	Time series model of all observations	a_0	15	Overall surface reflectance
		$a_{k,i}$ and $b_{k,i}$	240	Intra-annual phenology
		RMSE	15	Residuals omitted by that model fitting
Spatial	Synthetic images in observation window, 100th and 120th DOYs	Mean	30	Local mean of the gray levels in the kernel
		Contrast	30	Variability or lack of similarity in the kernel
		Second moment	30	Degree of pixel pair repetition in the kernel

and used to calculate texture metrics (Table 2).

3.1.1. Spectral and temporal features derived from time series model

To extract the spectral and temporal features, a harmonic time series model (Eq. 1) with Fourier terms was estimated based on all clear Sentinel-2 observations for each pixel and each spectral band or index (Fig. 5). The harmonic model has often been used to process time series of satellite data (Verbesselt et al., 2012; Zhu et al., 2015, 2020) for capturing the intra-annual variability of vegetation (Zhang et al., 2021a). The harmonic time series model was controlled by the number of Fourier terms (i.e., the number of sine and cosine pairs), and the short-term dynamics can be handled by increasing Fourier terms (de Livera et al., 2011; Hyndman and Athanasopoulos, 2018a). In this study, we used a harmonic time series model with eight pairs of Fourier terms (Eq. 1) to capture the phenology of the understory, especially within the observation window, and different understory species presented distinct signals. The complex harmonic time series models with a large number of Fourier terms would incur overfitting problems. We estimated the time series models using the Least Absolute Shrinkage and Selection Operator (LASSO) regression method (Tibshirani, 2011). The LASSO regression method minimizes the residual sum of squared errors with a bound on the sum of the absolute values of the coefficients. In this way, some of the model coefficients will be exactly zero and overfitting will be greatly constrained (Yang X. et al., 2022; Zhu et al., 2020).

For each spectral band and index, there were 18 potential variables, with 17 of them from the time series model coefficients to indicate the spectral and temporal characteristics and one of them from Root Mean Square Error (RMSE). The first variable (a_0) represents the overall value of the time series and indicates the mean value of the entire year. The eight pairs of harmonic variables provide temporal information about intra-annual seasonal patterns, such as the annual (a_1 and b_1), bi-annual (a_2 and b_2), tri-annual (a_3 and b_3) and bi-monthly (a_6 and b_6) changes. The additional variable (RMSE) was calculated during model regression and measures how well the time series model fits the data.

$$\widehat{\rho}_{i,x} = a_{0,i} + \sum_{k=1}^{n} \left\{ a_{k,i} cos\left(\frac{2k\pi}{T}x\right) + b_{k,i} sin\left(\frac{2k\pi}{T}x\right) \right\} \quad (1)$$

where,

x is day of year (DOY);

T is the number of days per year ($T = 365.25$);

k is the temporal frequency of the harmonic component on an annual basis (k = 1, 2, …, n);

n is the number of pairs of Fourier terms ($n = 8$), and each pair of terms designates the number of complete cycles completed by a wave over the calendar year (e.g., the eighth pair of terms completes eight cycles) (Jakubauskas et al., 2001);

$a_{0,i}$ is the coefficient for overall value; ($a_{k,i}$, $b_{k,i}$) are coefficients for intra-annual seasonality;

i indicates the variables, which can be the Sentinel-2 spectral bands or indices;

$\widehat{\rho}_{i,x}$ is predicted reflectance or index value for the ith variable at x DOY date from model prediction.

Fig. 5 illustrates the harmonic time series models for different understory species and "others" in deciduous forests (See Fig. S4 with estimated coefficient values). Although the shapes of the curves are similar within the leaf-on season of deciduous forests, different characteristics of understory coverage show somewhat different shapes in the leaf-off season, especially during the observation window. Take *barberry* and "others" in deciduous forests (Fig. 5(a)) as examples, *barberry* becomes greener earlier than the "others" in deciduous forests in the observation window. By using the harmonic time series model, synthetic images can be generated for any DOYs (Zhu et al., 2015). Fig. 6 shows an example of deciduous forests that are partially covered with evergreen *mountain laurel* and deciduous *barberry*. The location with *mountain laurel* remains green over the entire year. *Barberry* has leaf-off
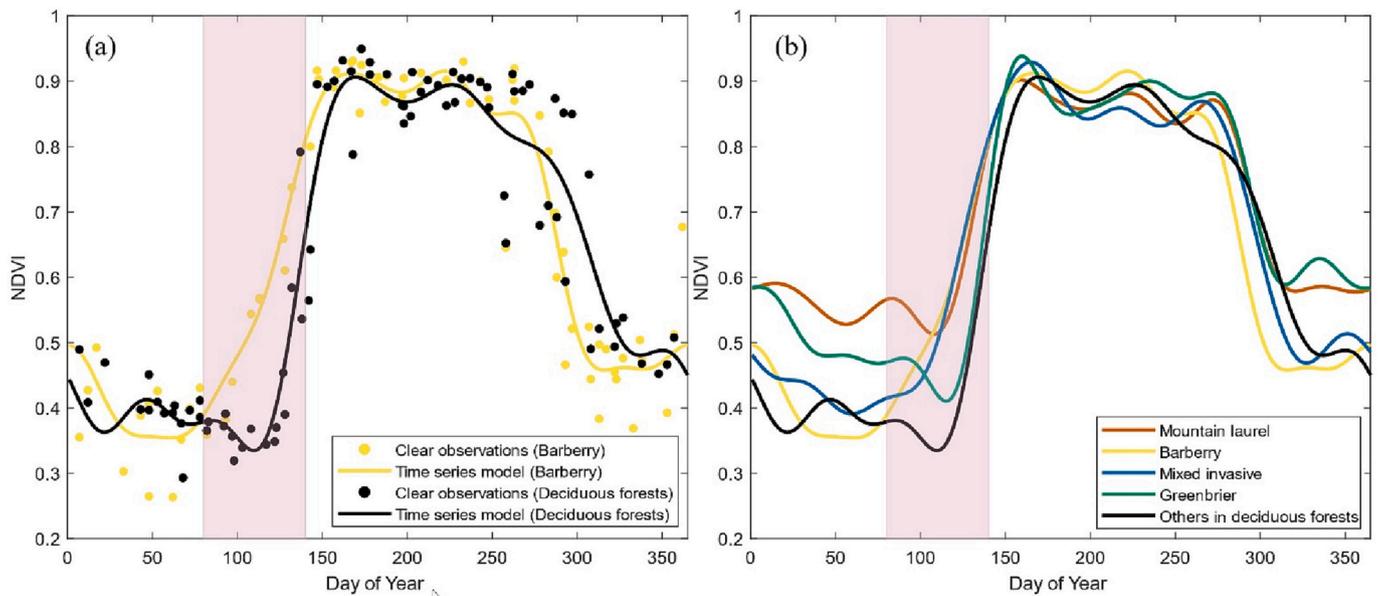
Fig. 5. Time series models are estimated based on Sentinel-2 clear observations collected between 2019 and 2021. (a) shows the clear observations and estimated harmonic models for deciduous forests with *barberry* and "others" in deciduous forest without target understory plants, respectively. (b) shows the different temporal trajectories for the target understory classes and "others" in deciduous forests pixels. The pink rectangular regions indicate the observation window in the spring prior to canopy leaf-on. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
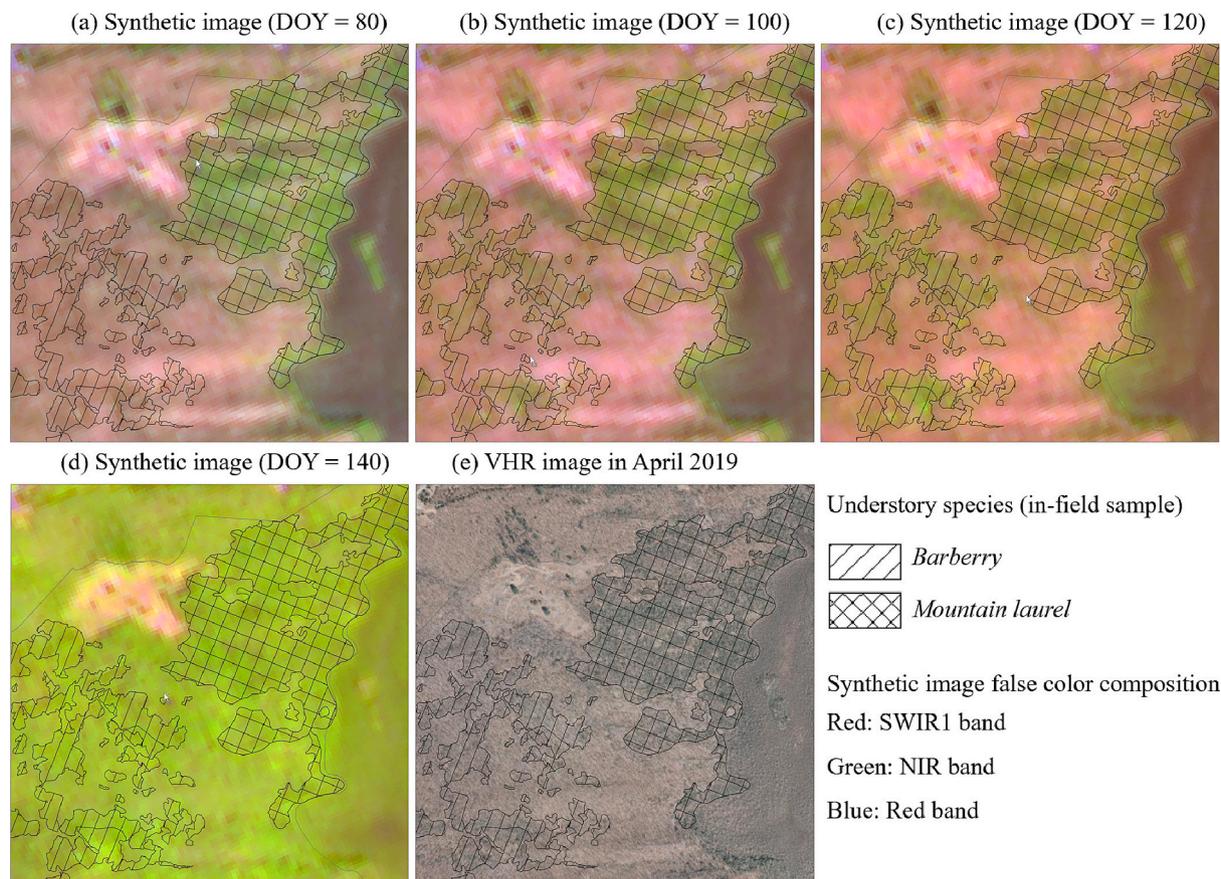
Fig. 6. Synthetic Sentinel-2 images ($\widehat{\rho}_{i,x}$ in Eq. 1) created at different DOYs in the observation window. (e) shows the very high resolution (VHR) image and in-field training data collected for evergreen *mountain laurel* in cross-hatch and deciduous *barberry* in diagonal stripes. The synthetic image chips (a-d) are composed of Sentinel-2 SWIR1, NIR, and red bands with the same stretch, and thus they are directly comparable. Healthy vegetation is bright green, and soils appear pink. In these synthetic images, the green pixels indicate the greening of the vegetation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

seasons similar to "others" in deciduous forests (Fig. 6 (a)) but produces leaves earlier than the overstory canopy (Fig. 6 (b-d)). The different phenology is the main basis for mapping the distribution of understory and identifying their species.

3.1.2. Spatial features derived from synthetic images

The time series coefficients captured the spectral and temporal information at pixel-level yet ignored the spatial information among pixels. Thus, we computed the widely used texture measures, derived from the gray level co-occurrence matrix (GLCM) (Dinstein et al., 1973), based on synthetic Sentinel-2 images (Fig. 6) created in the observation window. The synthetic Sentinel-2 images were created based on the surface reflectance estimated from the time series models ($\widehat{\rho}$ in Eq. 1) for each pixel and each spectral band (Zhu et al., 2015). We used these synthetic images to calculate the texture features because they are free of clouds, cloud shadow, and snow for any given DOY. The synthetic Sentinel-2 images within the observation window (e.g., DOYs of 100 and 120) were employed, because the understory started to produce new leaves while the overstory trees remained dormant in the study area.

The GLCM is created by calculating how often pairs of pixels with specific values in a specified spatial relationship occur in an image. GLCM texture has proved useful in many forest classification studies (Baron and Hill, 2020; Chrysafis et al., 2019; Ferreira et al., 2019). Understory patterns beneath the canopy can lead to different vegetation structures and distinctive textures. Statistical measures such as mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation can be computed from a GLCM (Clausi, 2002). To reduce the number of candidate variables, we selected three GLCM features with a small correlation, namely, mean, contrast, and second moment (Hall-Beyer, 2017; Murray et al., 2010). The mean feature refers to the local mean of the gray levels within the kernel. Contrast is a measure of variability and can be interpreted as a measure of lack of local similarity. The second moment, also known as uniformity or energy, measures the degree of pixel pair repetition in the kernel. These three variables are calculated with a window size of 9 × 9 pixels over four directions (0°, 45°, 90°, and 135°) (Zhu et al., 2012b). The average value of these four directions is then calculated to ensure the omnidirectional characteristics of selected texture metrics (Baron and Hill, 2020).

3.2. Creation of deciduous forest map

Our goal is to identify the distribution and species of the understory plant communities in deciduous forests. Therefore, we pre-defined a potential layer to limit the study area, which supports the growth of target understory vegetation in mature deciduous forests. This was necessary because other classes such as shrubland, herbaceous, and mixed forests could have a similar spectral signature as the target understory (Rittenhouse et al., 2022). We generated a new deciduous forest map (see *Supplementary S3*) at 10 m resolution by using a random forest classifier with variables derived from a harmonic time series model with three terms ($n = 3$ in Eq. 1). The classes include the first-level non-forest classes (i.e., water, developed, barren, shrubland, herbaceous, planted/cultivated, wetlands) and second-level subclasses of forests (deciduous, coniferous and mixed forests), following the NLCD 2019 map. The training data for the non-forest classes are randomly generated from the NLCD 2019 map, and subclasses of forests are manually interpreted as described in Section 2.2.1. Compared to the deciduous forest map in NLCD 2019, the updated map had a clear boundary between deciduous forests and non-deciduous forests owing to the higher spatial resolution of 10 m and reduced the omission errors. These improvements are important to ensure the coverage of the potential domain for understory classification. In addition, to reduce the commission error of shrubs and wetlands misclassified as the understory of the deciduous forests mapping (see *Supplementary S3.5*), we excluded these regions in the potential layer where the forest canopy height is not higher than 10 m based on

the LiDAR-based canopy height dataset (Bouvier et al., 2015; Reinmann and Hutyra, 2017; Senécal et al., 2018) and the very poorly drained soils cannot support the growth of the target understory plant communities (Kiviat, 2023; Paolucci and Stolt, 2018; Poindexter and Thompson, 2009).

3.3. Random forest classifier with representative sample

The random forest classifier was used to perform understory classification because of its relatively high accuracy and computational efficiency (Breiman, 2001). The random forest classifier is an ensemble supervised learning method, in which many classification trees are constructed, and a majority voting rule is applied to determine the final classification type. We used the default number of trees (500) to build the random forest classifier.

Training sample selection is usually one of the most critical elements for supervised classification (Heydari and Mountrakis, 2018; Li et al., 2021a; Zhu et al., 2016). When training a random forest classifier, it is important to select representative training sample that accurately reflect the target classes (Belgiu and Drăgu, 2016; Ishida, 2019; Li et al., 2017). For example, it is necessary to have training data that is representative of the different species compositions and cover densities of the *mixed invasives* to achieve optimal results with a random forest model. This is because the model will learn to recognize patterns from the data it is trained on; thus, the more representative the training data is, the more accurately the model will be able to generalize and predict unseen data.

In this study, we proposed an automated way to iteratively select representative sample (called ISRS) to build a robust random forest classification model (Fig. 7). We collected a large amount of high-quality training data in fieldwork and took them as the training sample pool (Fig. 2 and Table 1). ISRS seeks to ensure that the training data contains representative sample of each understory species, promoting diversity and capturing the full range of patterns and characteristics while avoiding domination by any particular conditions. At first, we randomly selected an equal number of training sample per class to train an initial random forest classifier. We then applied the classifier to generate predicted classes for the training sample pool and found the misclassified sample (hereafter referred to as "disagreement sample"). These disagreement sample likely originate from patterns ignored in the currently selected sample. It is thus essential to add more "representative sample" from the disagreement sample that are typically located where the classified labels diverge from the training data labels (Zhou and Li, 2010). We updated the training data with an iterative replacement procedure until the improvement plateaus. Every iteration replaced a small proportion (5%) of the previously selected training data with the disagreement sample (an equal number of disagreement sample for each class). This small proportion was adopted to ensure a gradual updating process and avoid excessively different maps by replacing a large proportion of the sample. The inclusion of the disagreement sample gradually corrects the issues in the previous run (as the disagreement pixels may differ in each run). After several iterations, further improvements by ISRS become limited and the selected set of training data should have accurately identified the whole patterns in the large training sample pool.

3.4. Removal of the isolated pixels

The proposed time series analysis approach is based on individual pixels and will have salt-and-pepper noise in the classification map (Pasquarella et al., 2018; Yang et al., 2020). Based on a sensitivity analysis, we used a two-step eight-connected sieving filter to eliminate these isolated pixels in the understory classification map (Senf and Seidl, 2021; Ye et al., 2021). The first sieving filter was applied to the binary understory presence map (i.e., presence or absence of any target understory plant communities). We removed the small understory objects smaller than 10 Sentinel-2 pixels (minimum mapping unit (MMU) of
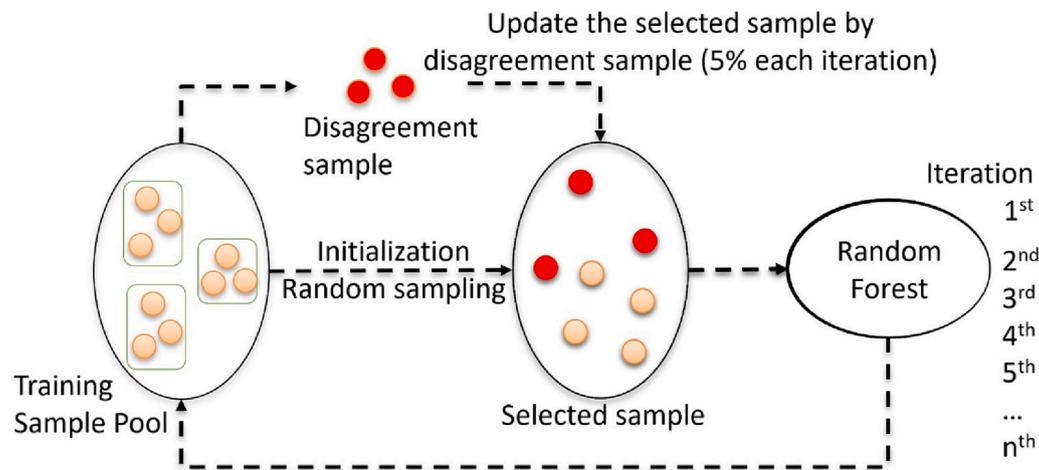
Fig. 7. Illustration of the process for automated selection of representative training sample by iteratively updating the selected sample with disagreement sample.

1000 m^2). Then, taking this noise-removed map as the mask, we employed the second sieving filter on the multi-class understory species map (i.e., four target understory classes) to modify small objects smaller than five pixels. We extracted the boundary pixel across the small object, and then counted the most frequent classes among the boundary pixels and defined the small object to this class. The two-step process was designed to maintain the patches larger than the MMU that consisted of different understory species, in which the area of the single understory class was smaller than the MMU.

3.5. Accuracy assessment and area estimation

The understory species map derived from satellite images contained prediction errors and biases. Thus, the estimation of areas was better accomplished by applying an unbiased estimator to adjust area for map error and providing confidence intervals for the error-adjusted area estimates (Bullock et al., 2020; Olofsson et al., 2013). A confusion matrix, based on the stratified validation sample in Section 2.2.3, can be calculated including the user's accuracy, producer's accuracy, and overall accuracy. Then, we followed the "good practice" recommendations presented in (Olofsson et al., 2013) to calculate a post-stratified estimator to translate the matrix into terms of unbiased accuracies and area proportions. The uncertainty of the area and accuracy estimator was quantified by a 95% confidence interval (CI). Our aim was to map the spatial distribution of four kinds of understory plant communities. Our final maps combine the "others" in deciduous forest and regions outside of the potential layer into the background class of "others". That is, we conducted the accuracy assessment and area estimation with five classes, including the four target understory plant communities and "others".

4. Results

4.1. Random Forest classifier calibration

Within the potential layer, we built a random forest classifier by using the proposed ISRS solution. Five classes (Fig. 3 (e-i) and Table 1), including the four target understory classes and "others" in deciduous forests were considered and an equal number of training sample were used to train the random forest classification model. That is, each target understory plant community and "others" in deciduous forests accounted for one-fifth of the total training data. The stratum of "others" in deciduous forests consisted of five major subclasses (i.e., deciduous sapling, sparse coniferous sapling, non-target native shrub understory, herbaceous ground covers and bare floor). We used an equal number of training data for each subclass to avoid certain ground conditions

dominating the variables of the category of "others" in deciduous forests.

The different number of representative training data and input variables can lead to random forest classification models with different accuracies. We used the in-field training data to select different numbers of representative training sample and input variables for creating the optimal random forest classification model by an object-level repeated validation. Each time, a total of 80% of the ground reference data were randomly selected to train the classifier, and the remaining 20% were used to evaluate map accuracy, and this process was repeated 50 times to estimate the overall performance. To mitigate the prevalence of spatial autocorrelation within the calibration process, a training data splitting design at the polygonal object level was used to ensure spatial independence between the training and test sets during the calibration process (Karasiak et al., 2022; Roberts et al., 2017). That is, the polygonal sample were initially randomly divided into two sets 80% of the polygons for training and the remaining 20% of the polygons for testing). Then a specified number of the training sample were randomly chosen from the training polygons, and the whole sample in the test polygons was used to calculate the accuracy.

To build an optimal random forest classification model based on the training data and input variables, we calibrated the classification algorithm by optimizing the overall accuracy and minimum accuracy. The minimum accuracy, the minimum value of the user's and producer's accuracies across all classes, was used to measure the maximum possible per-class commission and omission error (Zhu et al., 2016).

4.1.1. Influence of sample size and iterative times

We calibrated the representative training sample size and iterative times to build the random forest classifier with the proposed ISRS solution. Fig. 8 shows the overall accuracies and minimum accuracies in each iteration. Both the overall accuracy and minimum accuracy increased substantially at the early iteration stage (<5 iterations), regardless of the total number of training sample used. Both the overall and minimum accuracies plateaued after approximately ten iterations and when >8000 training sample were used. Therefore, we selected a total of 8000 representative sample with ten iterations, which improved the overall accuracy by 1% and minimum accuracy by 10%, compared to the use of all training data (~30,000 sample). That is, a more concise random forest classification model was generated with a small number of more representative training sample and achieved better results than using all the available training data.

4.1.2. Influence of input variables

We calculated the time series coefficients and texture metrics of Sentinel-2 bands and spectral indices (Table 2) and compared whether
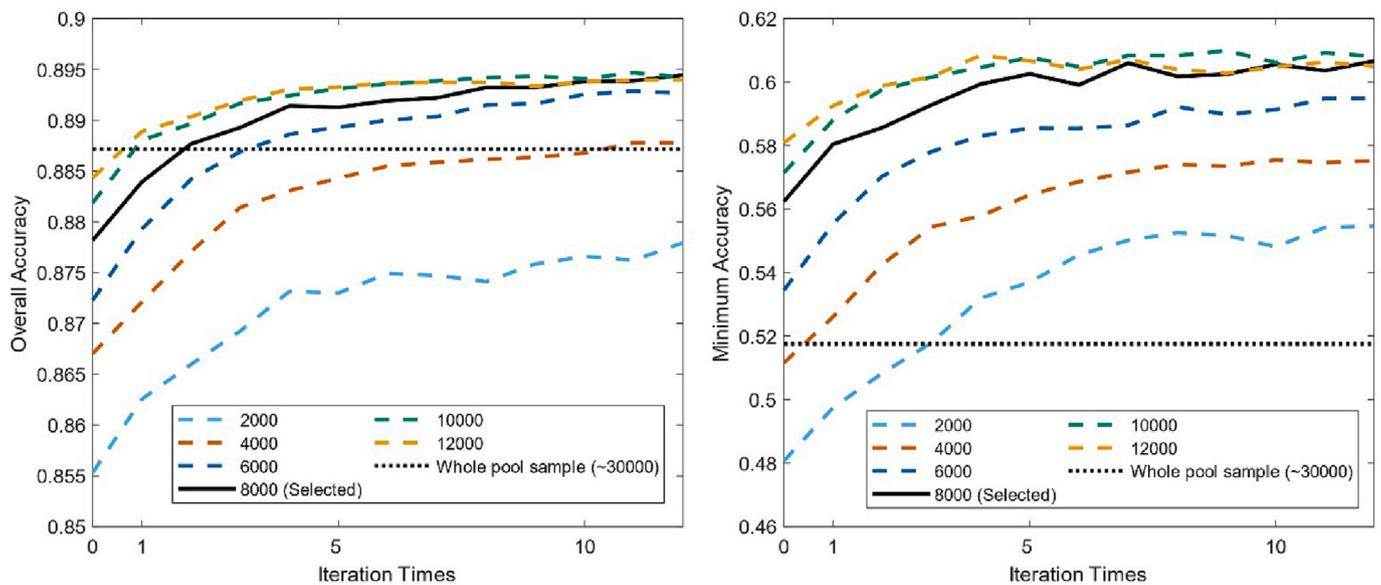
Fig. 8. Sensitivity analyses of selecting representative training sample by overall accuracy and minimum accuracy. The black dotted line indicates the accuracy when using all the reference training data (~30,000 sample). A total of 8000 representative training sample and ten iterations were determined as the optimal strategy for this iterative sample selection process.

some of them may be redundant and have limited impacts on the classification results (Pal, 2006; Zhu et al., 2012a, 2012b). To select useful input variables for classification, we designed several scenarios by considering with/without indices, with/without textures and with/without narrowband (Table 3) and compared their accuracies by using 8000 representative training sample (Fig. 9). We used synthetic Sentinel-2 images on DOYs of 100 and 120 (see details in *Supplementary S7*) to calculate the GLCM metrics. Two DOYs in the observation window are to add information on different phases of spring greening for deciduous understory species.

The use of indices (Scenario 3) improved both the overall accuracy and minimum accuracy dramatically. The three scenarios excluding indices (Scenario 1, 2, 5) achieved low accuracies compared to the others using indices. The texture information improved the accuracy further, such as the comparison of Scenario 1 versus 5, Scenario 3 versus 6 and Scenario 4 versus 7. The combined use of indices and their texture metrics (Scenario 6) achieved the best overall accuracy even without using the spectral bands, but the inclusion of broadbands (Scenario 7) improved the minimum accuracy by 4%. Thus, our model employed broadband, spectral indices and their texture information as the input variables (a total of 264 variables) to classify the understory species, but excluded the narrowband owing to the saturated accuracy (Scenario 8).

We employed 264 input variables (Scenario 7 in Table 3) derived from spectral, temporal and spatial domains to build the random forest classifier and classify the understory species. Random forest classifier can effectively measure the variable importance by considering not only

the influence of each predictor variable separately but also the multivariate interactions with other predictor variables (Archer and Kimes, 2008; Breiman, 2001; Chan and Paelinckx, 2008; Gislason et al., 2006; Zhu et al., 2016). Variable importance shows how much the prediction accuracy would drop if a given metric was excluded (Belgiu and Csillik, 2018). According to the importance estimation of the random forest classification model, the top 64 variables were derived from a variety of spectral bands/indices and different temporal and spatial features (See *Supplementary S5*), of which 49 variables were time series model coefficients and RMSE from model fits, and 15 variables were GLCM metrics from synthetic images.

A large proportion of the variables (51 variables) were from spectral indices. The use of five spectral indices as input variables can achieve a 5% increase of overall accuracy and 8% increase of minimum accuracy compared to the use of surface reflectance bands (Fig. 9). RENDVI was the most important variable, which provided five of the top 10 variables (a_0, a_1, b_2, and GLCM mean values in both 100th and 120th DOYs). One reason lies in the exclusion of the red edge bands and increased independence of RENDVI. For broadband surface reflectance bands, infrared bands (13 variables) provided more useful information than VIS bands (not included in the top 64 variables). The inclusion of the spatial texture features would lead to the increase of overall accuracy by 1% and minimum accuracy by 5% (Fig. 9).

4.1.3. Influence of harmonic terms
It is important to select the appropriate number of Fourier terms to

Table 3
Classification scenarios for selecting different groups of input variables. Texture features were calculated based on the synthetic Sentinel-2 image created on 100th and 120th DOY. * indicates the optimal scenario.

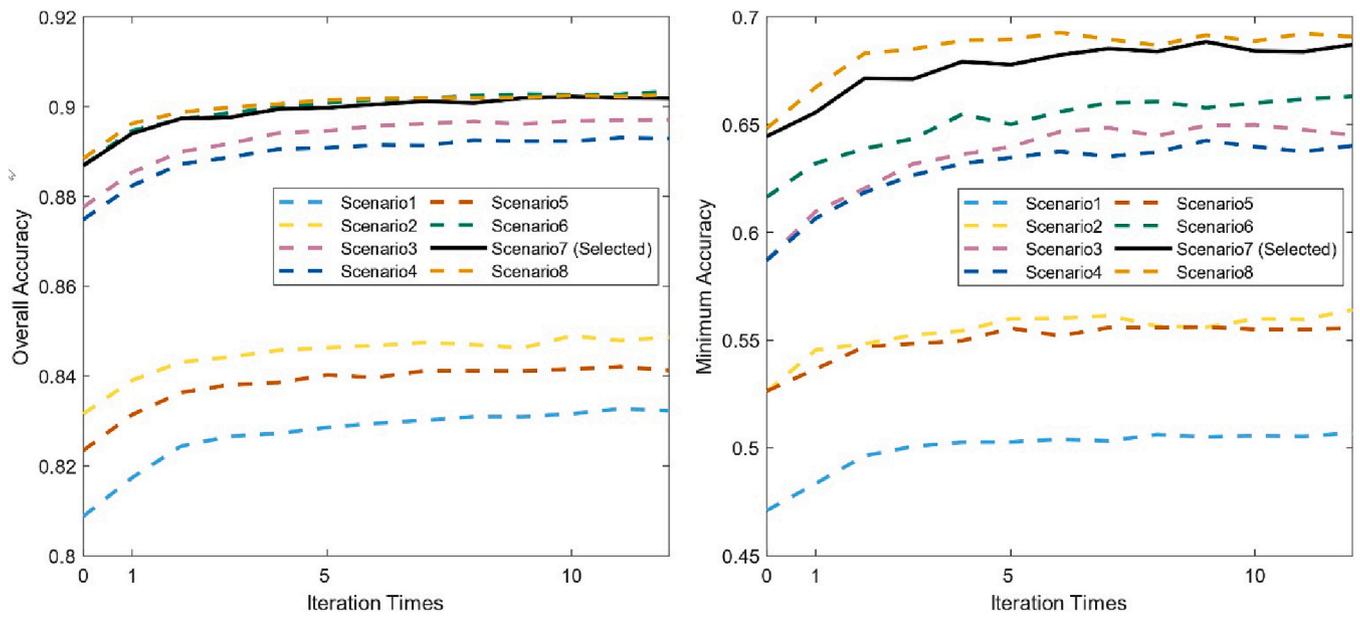| Scenarios	Broadband	Narrowband	Indices	Textures	Number of input variables
Scenario-1	**Yes**	No	No	No	108
Scenario-2	**Yes**	**Yes**	No	No	180
Scenario-3	No	No	**Yes**	No	90
Scenario-4	**Yes**	No	**Yes**	No	198
Scenario-5	**Yes**	No	No	**Yes**	144
Scenario-6	No	No	**Yes**	**Yes**	120
Scenario-7*	**Yes**	No	**Yes**	**Yes**	264
Scenario-8	**Yes**	**Yes**	**Yes**	**Yes**	360

Fig. 9. Comparisons of different scenarios with/without indices, with/without textures, and with/without narrowband. Notably, the accuracy was evaluated based on 8000 representative training sample. See Table 5 for a description of each scenario.
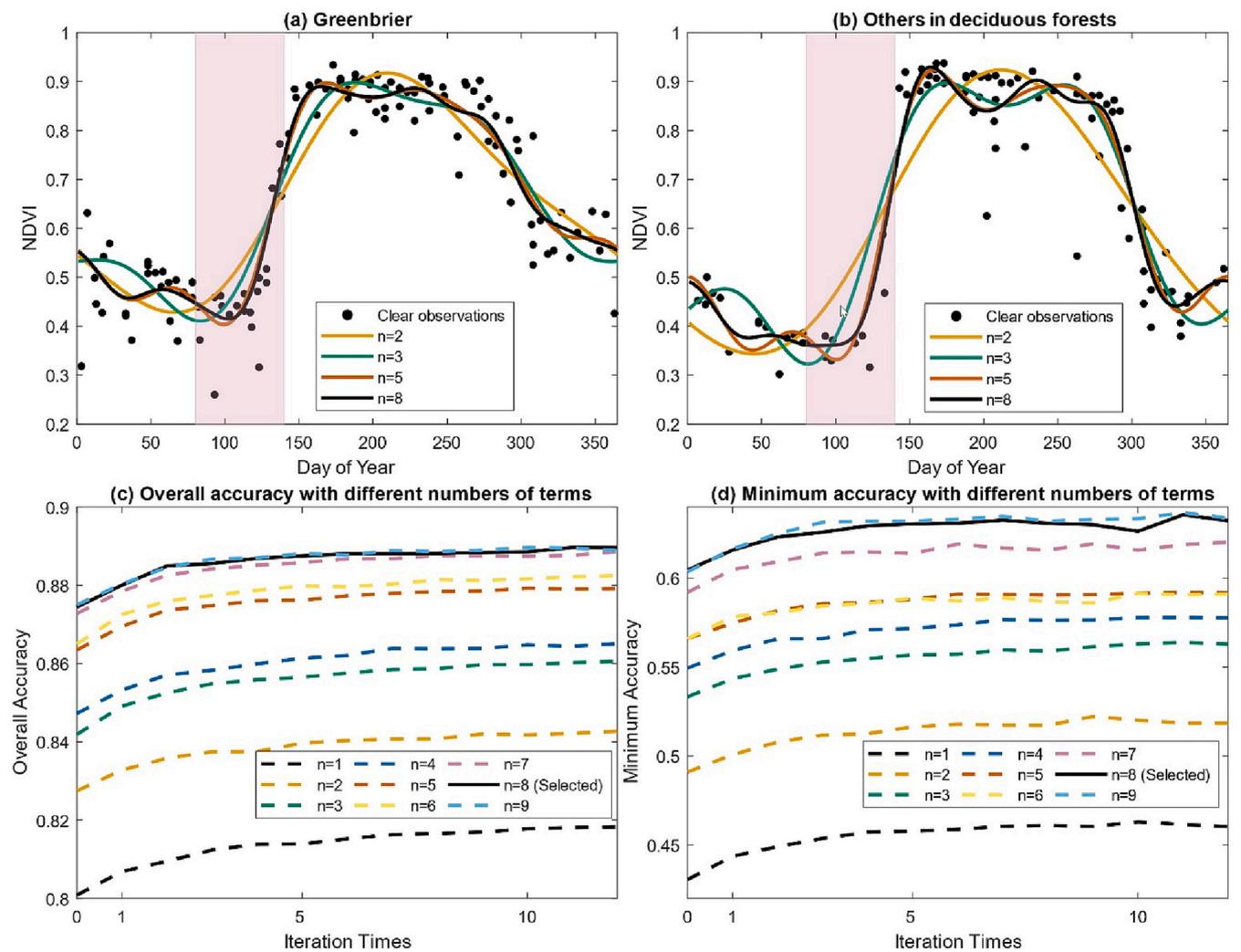

Fig. 10. Time series models (Eq. 1) and performances by using different numbers of Fourier terms (varying from 1 to 8). Notably, the accuracies were calculated for the Scenario-4 (Table 3) with broadbands and indices.

develop the harmonic model and capture the seasonality (Hyndman and Athanasopoulos, 2018b). We varied the number of Fourier terms (*n*) from 1 to 9. Fig. 10 shows the seasonal patterns and understory species classification performances. Notice that as *n* increases the Fourier terms capture the early greening of the understory (from third pair) and difference between the understory and "others" in deciduous forests (from fifth pair). In this study, we used eight terms because the accuracy did not improve with additional terms (i.e., ninth pair). The overall accuracy and minimum accuracy decreased by 1% and 5% when using five pairs of terms, and further reduced by 2% and 3% when using three pairs of terms.

4.1.4. Influence of observation density

The density of the clear observations within the observation window is important to estimate an accurate harmonic time series model that captures the subtle spectral-temporal differences among the different understory species (Zhang et al., 2021a). We combined three years of Sentinel-2 data acquired between 2019 and 2021 to yield three-fold the data density as using just one year of data, assuming that the understory and deciduous forests were stable and did not undergo changes during these three years. Fig. 11 depicted the harmonic time series models (Eq. 1) and performance with clear observations acquired within different years (collapse into a single year). In general, the denser the time series observations were, the better the results were owing to the increased availability of clear observations within the observation window. The denser time series observations by combining three years improved the

overall accuracy and minimum accuracy by approximately 2% and 6%, respectively. The improvement was trivial when observations were acquired from more than three years.

4.2. Classification maps

We used 264 variables in scenario 7 (Table 3) and selected 8000 representative sample after the 10th iterations to build the random forest classification model and applied the model to the entire area of CT to obtain the understory map. Fig. 12 shows the distribution of target understory plant communities in CT, United States. An interactive display of the understory species map and synthetic images in observation window is available at: https://gers.users.earthengine.app/view/understory.

The native *mountain laurel* is widely distributed across CT, particularly in the western part, while the *greenbrier* is more concentrated in the coastal east. The invasive species of *barberry* has spread throughout the state, especially in the east of the Connecticut River. *Mixed invasive* understory species can be found in eastern CT and central parts spanning the Connecticut River.

4.3. Quantitative evaluation

We assessed the accuracy and estimated the biases of the generated understory map for three themes, in which the first is the presence and absence of the understory (understory presence map in Fig. S8-1), the
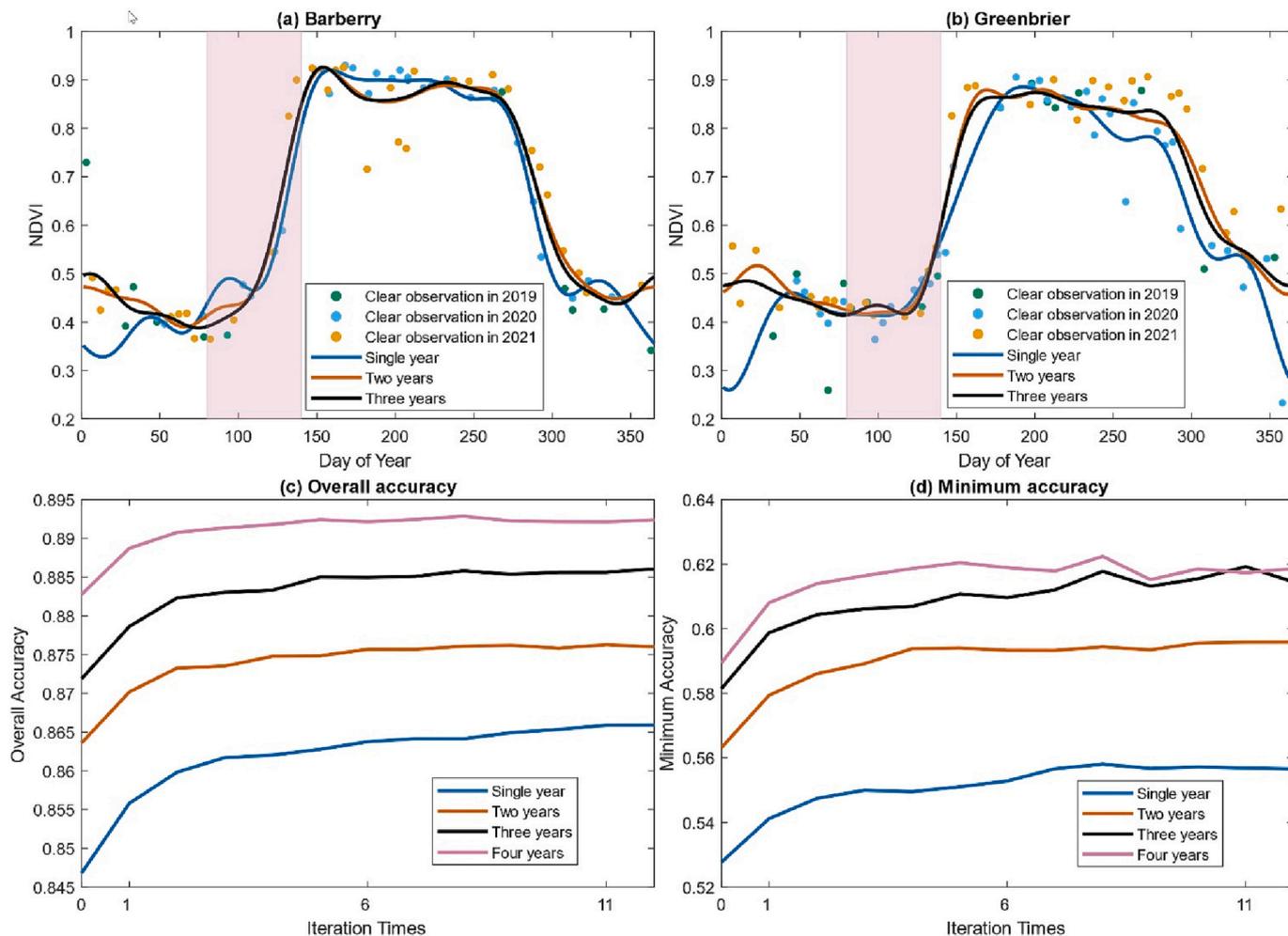

Fig. 11. Accuracy by using clear observations acquired in different years. Notably, the accuracies were calculated for the Scenario-4 (Table 3) with broadbands and indices.
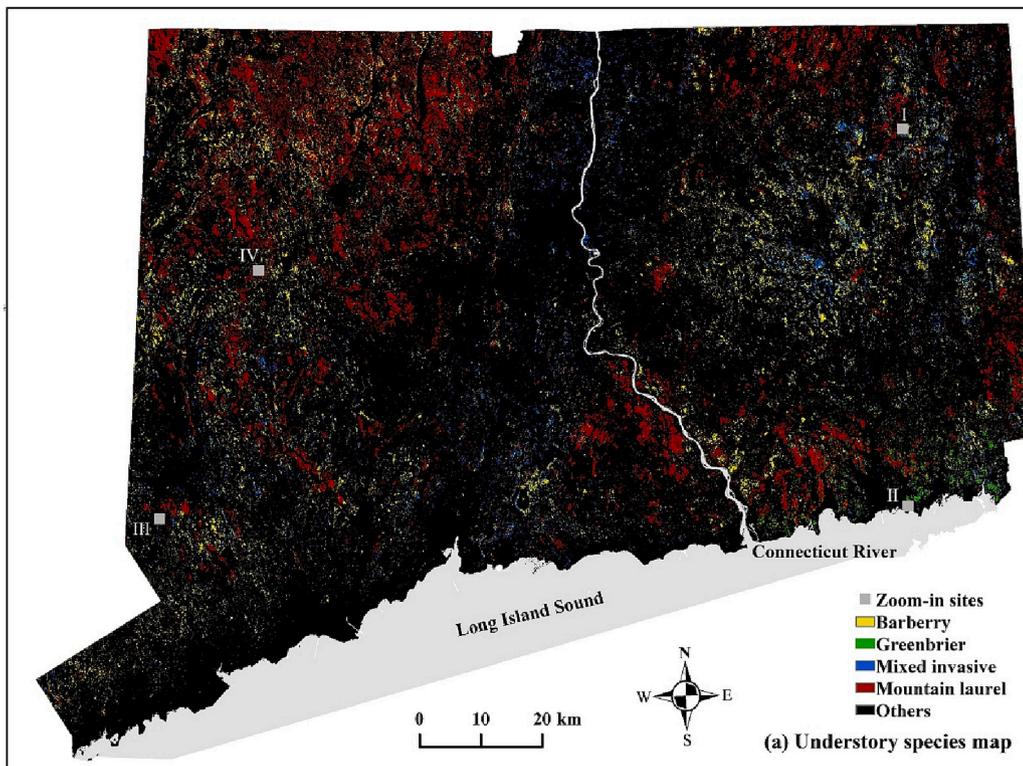
Fig. 12. (a) shows understory species distribution in CT and four zoom-in validation sites (I, II, III, and IV) are displayed in (b)-(e), which correspond to the species map, Sentinel-2 synthetic image on 80th DOY, Sentinel-2 synthetic image on 120th DOY and VHR images, respectively. Column (f) shows the in-situ photos (July 2021) of the validation pixel location (white cross in center) in (b)-(e). The four rows of (b-e) are examples of *barberry*, *greenbrier*, *mixed invasive* and *mountain laurel*, respectively. DOY: Day-Of-Year; VHR: Very High Resolution.
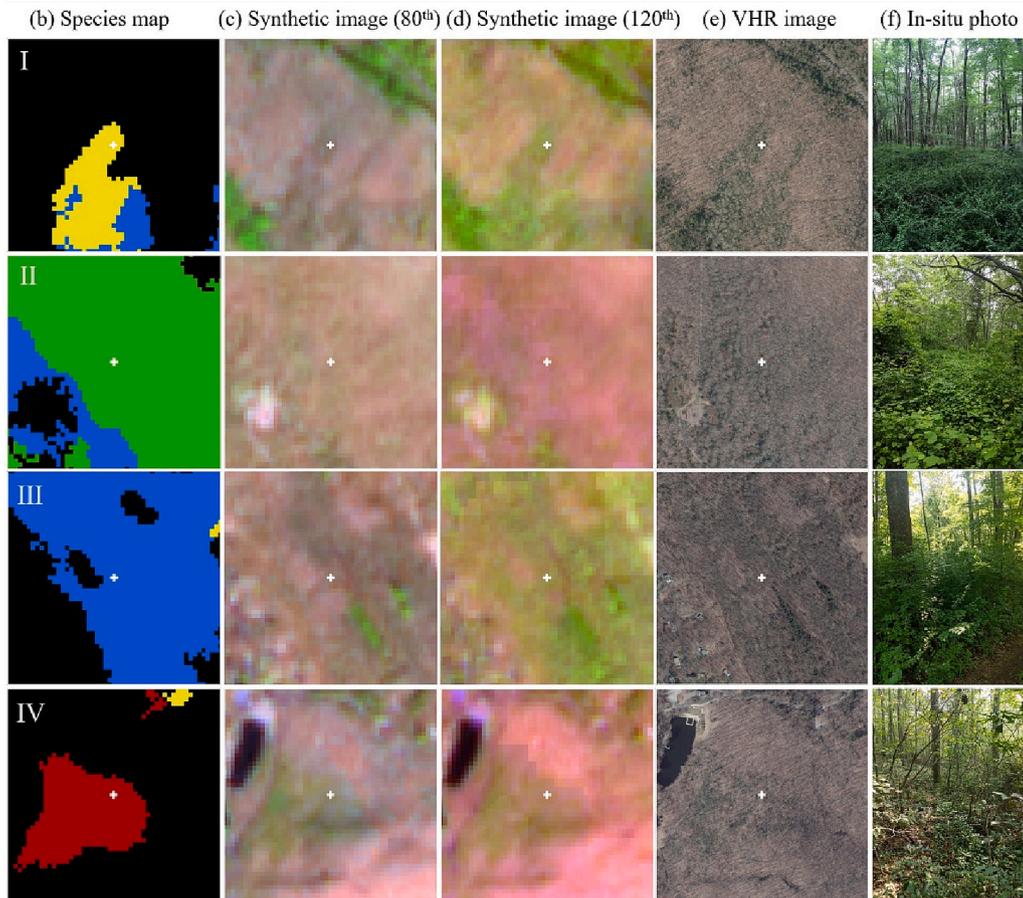
Table 4
Confusion matrices and estimates of accuracy, area, and their uncertainty (with 95% CI) for the understory present map. Columns represent reference labels and area estimation, and rows represent map strata.

Map class	Reference data	
	Presence of understory	Absence of understory
	Confusion matrix, sample counts	
Presence of understory	260	120
Absence of understory	5	290
	Confusion matrix, area proportions	
Presence of understory	0.0789	0.0364
Absence of understory	0.0150	0.8697
	Accuracy and area estimates	
Area proportion [0-1]	0.0939	0.9061
Area [km^2]	1222.96±183.95	11,800.04±183.95
Map bias [%]	+279.04	-279.04
User's acc. [%]	68.42±2.39	98.31±0.75
Producer's acc. [%]	84.03±5.99	95.98±0.29
Overall acc. [%]	94.86±0.72	

second is the measurement of the invasive species and native species (invasive/native understory map in Fig. S8-2), and the third is the identification of understory species (understory species map in Fig. 12).

4.3.1. Understory presence map

The results of the accuracy assessment and unbiased area estimation for the understory presence map (i.e., presence and absence of the target understory) are illustrated in Table 4. The user's and producer's accuracy (and 95% CI) for the presence of target understory is 68.42% (±2.39%) and 84.03% (±5.99%) respectively, while for the absence of target understory it is 98.31% (±0.75%) and 95.98% (±0.29%) respectively. The overall accuracy (and 95% CI) is 94.86% (±0.72%). The area of target understory is 1222.96±183.95 km^2 and occupied approximately 9.39% of CT's land area and one fourth of the deciduous forests.

4.3.2. Invasive/native understory map

The accuracy assessment and unbiased area estimation for the invasive/native understory map (i.e., invasive species of *barberry* and *mixed invasive*, native species of *mountain laurel* and *greenbrier,* and "others") are illustrated in Table 5. The overall accuracy of the invasive/native understory map is 94.70% ± 0.72%. The invasive species had both good user's and producer's accuracies (approximately 74% ~ 81%), and the native species had a moderate user's accuracy of 61.17% ± 3.40%, but good producer's accuracy of 83.62% ± 8.13%. The invasive species (649.33 ± 140.59 km^2) occupied more than half of the

shrub understory in deciduous forests of CT, compared to the native species (577.09 ± 121.26 km^2).

4.3.3. Understory species map

The accuracy assessment and unbiased area estimation for the understory species map (i.e., four target understory classes and "others") are in Table 6. The overall accuracy of the species map (93.12% ± 0.75%) is slightly lower than that of the binary understory presence map. The "others" (i.e., "others" in deciduous forests and outside of potential layer) had both high user's and producer's accuracies (>95%) in the understory species map. User's accuracy was moderate for all four target understory classes, ranging from approximately 50% to 77%. Producer's accuracy was high for *mountain laurel* (98.76% ± 1.10%) and moderate for *barberry* (73.68% ± 8.89%) and *mixed invasive* (51.31%± 8.25%). *Mountain laurel* is the native and most widespread of the target understory plant communities in CT and occupies 399.46 ± 70.25 km^2. *Barberry* and *mixed invasive* understory species are distributed at similar levels of approximately 300 km^2. *Greenbrier* is an important native understory species that is mostly distributed in coastal regions with an estimated area of approximately 123.94 ± 108.62 km^2.

4.4. Error sources

The commission error of the understory maps was mainly from the confusion of the mixed forests, wetlands and disturbed deciduous forests (typically harvest). The misclassification of mixed forests as *mountain*

Table 5
Confusion matrices and estimates of accuracy, area, and their uncertainty (with 95% CI) for the invasive and native understory species mapping. Columns represent reference labels and area estimation and rows represent map strata.

Map class	Reference data		
	invasive understory	native understory	Others
	Confusion matrix, sample counts		
invasive understory	128	4	42
native understory	2	126	78
Others	3	2	290
	Confusion matrix, area proportions		
invasive understory	0.0403	0.0013	0.0132
native understory	0.0006	0.0371	0.0229
Others	0.0090	0.0060	0.8697
	Accuracy and area estimates		
Area proportion [0-1]	0.0499	0.0443	0.9058
Area [km^2]	649.33 ± 140.59	577.09 ± 121.26	11,796.58 ± 183.60
Map bias [%]	+63.67	+211.91	-275.58
User's acc. [%]	73.56 ± 3.35	61.17 ± 3.40	98.31 ± 0.75
Producer's acc. [%]	80.78 ± 8.45	83.62 ± 8.13	96.01 ± 0.29
Overall acc. [%]	94.70 ± 0.72		

Table 6

Confusion matrices and estimates of accuracy, area, and their uncertainty (with 95% CI) for the understory species map. Columns represent reference labels and area estimation and rows represent map strata.

	Reference data				
Map class	Barberry	Greenbrier	Mixed invasive	Mountain laurel	Others
	Confusion matrix, sample counts				
Barberry	58	0	23	1	23
Greenbrier	0	63	2	1	16
Mixed invasive	8	3	39	0	19
Mountain laurel	0	0	0	62	62
Others	1	2	2	0	290
	Confusion matrix, area proportions				
Barberry	0.0195	0	0.0077	0.0003	0.0077
Greenbrier	0	0.0028	0.0001	0	0.0007
Mixed invasive	0.0018	0.0007	0.0089	0	0.0044
Mountain laurel	0	0	0	0.0303	0.0303
Others	0.0030	0.0060	0.0060	0	0.8697
	Accuracy and area estimates				
Area proportion [0-1]	0.0261	0.0096	0.0234	0.0548	0.8862
Area [km^2]	316.48±89.61	123.94±108.62	296.26±116.63	399.46±70.25	11,886.86±188.64
Map bias [%]	+142.52	-75.94	-90.26	+389.54	-365.86
User's acc. [%]	55.24±4.88	76.83±4.69	56.52±6.01	50.00±4.51	98.31±0.75
Producer's acc. [%]	80.11±10.24	29.75±13.37	39.30±8.15	98.76±1.10	95.28±0.34
Overall acc. [%]	93.12±0.75				

laurel understory in deciduous forest was the major commission error in the understory map, because they were both evergreen vegetation and shared similar spectral characteristics. Wetlands with *skunk cabbage*, *fern*, *moss*, and/or *spicebush* were misclassified as understory sometimes, namely barberry or mixed invasive, because they share a similar early greening phenology (Neufeld and Young, 2014). In addition, we assumed that the deciduous forests and understory were stable during the study period (between 2019 and 2021), and if some forest disturbance (typically harvest) occurred after 2019, the time series may still capture the understory signal, but our validation process (in July 2021) would regard them as "others" (i.e., non-deciduous forests). If a more accurate map of the deciduous forests were available, the commission error could be reduced to <15%. That said, there were still some commission errors in the deciduous forests, mainly owing to the existence of non-target understory vegetation (such as *lowbush* and *blueberry*). Most of these misclassified validation sample also included sparse target understory plant communities (<10% cover). These mixed pixels could provide mixed signals in the time series model and led to misclassification.

The omission error of the understory presence map was approximately 15%, which indicated that the proposed approach can detect the understory beneath the canopy well. Compared to mapping understory presence, identification of the different understory species, especially the deciduous species, is a more difficult task. In the understory species map, the evergreen understory of *mountain laurel* was well extracted with a small omission error of <2%. The errors mainly came from the confusion of different kinds of deciduous understory (i.e., *barberry*, *greenbrier* and *mixed invasive*). If we extracted the distribution of invasive understory, we achieved both the user's and producer's accuracies of approximately 75–80%. However, the labeling of different types of the invasive understory was extremely difficult because they frequently grew nearby and were presented as mixed pixels in Sentinel-2 images. For example, the *mixed invasive* class could include a proportion of *barberry* (<50%) and a small proportion of native *greenbrier* (due to co-occurrence) and therefore was easily misclassified with *barberry* and *greenbrier*.

In the validation process, we randomly selected sample within publicly accessible properties that could be reached in a reasonable amount of time and distance from a road or parking area for logistical reasons. Consequently, some validation sample correspond to edge pixels between the understory and non-understory. According to the previous analyses on the selection of validation sample (Heydari and Mountrakis, 2018; Powell et al., 2004), the accuracy of the understory map could be higher if we were to collect a greater number of validation sample from interior portions of mapped target understory patches. For the understory species map, the user's accuracy and producer's accuracy of the *barberry* and *mixed invasive* would increase by approximately 5% if excluding these edge pixels with high uncertainty (See Table S8).

5. Discussion

5.1. Automated detection of understory plant communities

Despite numerous attempts to monitor forests through multispectral satellite imagery (Anderegg et al., 2022; Hansen et al., 2013; Senf and Seidl, 2021; Zhu et al., 2012a), the detection of understory plant communities within the forest ecosystem remains a difficult task, and the identification of different understory species is even more challenging. To this end, we developed an automated time series algorithm (open-source package at https://github.com/GERSL/understory_mapping) that took into account the distinct phenology between understory plant communities and forest canopy in the spring prior to canopy leaf-on (defined as the observation window in this study). The observation window varies depending on the ecoregion, and a random forest classifier can be built and applied accordingly at the ecoregional scale. In this study, the study area of CT is mainly located in the Northeastern Coastal Zone (https://www.epa.gov/eco-research/ecoregions), where the greening of the understory plant communities generally begins in early April, while the deciduous forest canopy greens up in late April and early May. Consequently, the calibrated random forest model can be directly applied to the Northeastern Coastal Zone in the United States (such as Rhode Island and eastern Massachusetts) given their similar spring phenology (Dannenberg et al., 2015; Reed, 2006; Seyednasrollah et al., 2018; White et al., 2009). Additionally, this approach is transferable to other regions by recalibrating the random forest classifier, provided that local training sample are available for the ecoregion. Similar to some previous methods (Pasquarella et al., 2018; Sun et al., 2021b; Tuanmu et al., 2010), our approach leveraged dense Sentinel-2 time series to capture the phenological differences between over and understory plants and is thus characterized by its flexibility and expansibility.

5.2. Spatial distribution of different understory species

We generated the distribution maps of understory plant communities at the state scale of CT, United States (Yang et al., 2023). We mapped the

understory plant communities in three levels, including understory presence map, invasive/native species map, and understory species map, and all three maps achieved high overall accuracies (>93%). Approximately one-quarter of the deciduous forest ecosystems in CT exhibited the targeted understory plant communities, with more than half of them comprising invasive species.

To the best of our knowledge, this was the first map of understory plant communities with multiple classes (notably deciduous understory species) at a regional scale. To date, there were few resources available to provide the spatial distribution of the understory plant communities, with the current estimation largely being conducted through modeling the probability of occurrence of understory based on random plots investigation in the field (Naqinezhad et al., 2022). In the United States, the U.S. Forest Service's Forest Inventory and Analysis program (FIA) collected the forest and understory vegetation data through field inventory and used random plots to measure the density and diversity of the understory (Gray et al., 2012; Hoover et al., 2022; Patterson and O'brien, 2011). However, these plots-based investigations did not provide the distribution of the understory plant communities seamlessly in space. The maps will assist in pinpointing the locations of understory plant communities and informing local invasive plant species management and intervention practices.

Moreover, we identified the species of the understory plant communities and thus provided the distribution of the invasive species. The previous work mainly focused on the mapping of the presence of understory plant communities (Dai et al., 2020), and some works recognized the different classes of evergreen understory vegetation beneath the deciduous forest canopy (Chastain and Townsend, 2007; Singh and Gray, 2020; Tuanmu et al., 2010; Welch et al., 2002). Our species map underscored that the dense Sentinel-2 time series can be used to capture the subtle yet distinct phenology difference between deciduous understory shrubs and deciduous forest canopy, and between different understory species. This implies that the proposed approach has the potential to be applied to detect the different vegetation classes with distinct phenology, such as the invasive plant classes, and the classification of different crop types.

5.3. Impact of density on time series analysis

The density of the time series observations was pivotal for capturing the spring phenology within the narrow observation window. Sentinel-2 has a revisit frequency of every five days and theoretically provides sufficient temporal resolution (Lee et al., 2022; Ahl et al., 2006) to detect understory plant communities. However, the impact of cloud cover reduced the density of the clear observations. The clear observations in a single year would be not sufficient to capture the distinct phenology between deciduous understory plant communities and deciduous forest canopy. Furthermore, the performance of time series analysis was drastically dependent on the observation density (Zhang et al., 2021a). To address this issue, the combination of observations acquired over multiple years had been used by some other work to construct a time series model with better performance (Baumann et al., 2017). The major drawback of such a combination was the assumption that the forests did not experience dramatic disturbances within three years.

Therefore, there was a delicate balance between capturing the understory signal better and the assumption of unchanged forests. In this study, we combined the observations in three years to generate the threefold denser observations based on the calibration of training data (Section 4.1.4). Compared to the time series constructed from the observation in a single year, the threefold observations improved the overall accuracy and minimum accuracy by approximately 2% and 6%, respectively. The assumption of a long-term stable period might lead to commission errors if the understory or deciduous forests underwent changes during this period. In our validation process, 10% of the misclassified sample were attributed to a forest harvest between 2019 and 2021. We obtained a balance between the assumption of a stable period

(three years in this study) and sufficient clear observations for capturing the understory phenology.

This limitation could be further alleviated by using NASA's Harmonized Landsat and Sentinel-2 (HLS) data (Babcock et al., 2021; Claverie et al., 2018; Shang and Zhu, 2019). The HLS data enabled repeated coverage every 2-3 days and provided higher denser observations to capture the spring phenology. However, HLS data has a spatial resolution of 30 m, which is not sufficient (up to 10 m) for accurate mapping understory (Shouse et al., 2012). Given the development of data fusion techniques (Shao et al., 2019; Zhang et al., 2021b; Zhu et al., 2022), the emergence of HLS data at 10 m with both high temporal and spatial resolution (Song and Zhu, 2022) offers new opportunity for mapping understory plant communities at the annual scale and even track the spatial change of the understory, such as the proliferation and extinction of understory plant communities.

5.4. Selection of representative sample

The paper proposed an iterative solution to select representative sample for random forest classification, particularly in the context of understory mapping. The selection of representative sample among a large amount of training data was important for random forest classification (Li et al., 2016, 2017; Liang et al., 2016; Liu et al., 2020; Wen et al., 2022). Understory plant communities are distributed with heterogeneous patterns (der Sluijs et al., 2016; Ozdemir, 2014; Tinya and Ódor, 2016), such as the varying densities over space and a mixed composition of the understory species (especially for the class of *mixed invasive*). Thus, the selected sample should reflect the phenological and spectral characteristics of abundant phenomena of the same understory plant classes (Li et al., 2020).

In this study, we selected 8000 representative sample and calibrated an optimal random forest classifier (Section 4.1.1). If we randomly selected 8000 sample, the overall accuracy and minimum accuracy of the calibrated random forest classifier decreased 1.5% and 5%, respectively. Moreover, the calibration results proved that the addition of representative sample did not further improve the accuracy of the model. This indicated that the sample were sufficiently representative to cover the typical conditions for each understory class. Consequently, the proposed ISRS approach can help to generate representative sample, particularly in classification tasks where varying conditions exist within the same class. It should be noted that the solution depends on the high-quality training data. If the training data contained a lot of noise and low-confidence sample, the iterative selection process would propagate errors. For example, ISRS was not adopted to map the deciduous forests in this study, because the sample for the background (such as developed land, open water, and soil) were randomly generated from the NLCD 2019 map. The uncertainty of these sample would lead to accumulated errors when the random forest model was iteratively updated (See *Supplementary S3.2*).

6. Conclusion

In conclusion, our proposed method, based on dense Sentinel-2 time series observations, effectively addressed the challenges of mapping understory plant communities from satellite images. A harmonic time series model was used to capture the distinct phenology in the narrow observation window between understory plant communities and forest canopy, and between different species of understory plants, allowing us to overcome the obstruction of the forest canopy to the understory layer. We demonstrated that the use of vegetation indices and GLCM texture features benefited the understory classification, which improved overall accuracy by 5% and the minimum accuracy by 12%. Moreover, the proposed ISRS approach improved the accuracy (e.g., by 10% of the minimum accuracy) and helped to build a concise random forest classification model with a small number of training sample. The final binary understory presence map, the invasive/native understory map and

the understory species map in CT, United States, achieved overall accuracies of approximately 94.86%, 94.70% and 93.12%, respectively. Notably, all the input variables of the random forest classifier were extracted from estimated harmonic time series model through a fully automated process. The approach is transferable to other ecoregions after the recalibration of the random forest classifier with the use of local training data. Therefore, Sentinel-2 time series is a feasible option for mapping understory plant communities and has the potential for operational understory mapping for other places and at a much larger scale.

CRediT authorship contribution statement

Xiucheng Yang: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Shi Qiu:** Conceptualization, Methodology, Software, Investigation, Data curation, Writing – review & editing. **Zhe Zhu:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Chadwick Rittenhouse:** Conceptualization, Validation, Resources, Writing – review & editing, Project administration, Funding acquisition. **Dustin Riordan:** Validation, Data curation, Resources, Investigation, Writing – review & editing. **Mari Cullerton:** Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Code for iteratively selecting representative sample is available at GitHub (https://github.com/GERSL/understory_mapping). Products of understory distribution in CT is available to download at https://data.mendeley.com/datasets/rschxhwgvw/3. Online interactive understory species map is available at https://gers.users.earthengine.app/view/understory.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2023.113601.

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 Sens. Environ. 104, 88–95. https://doi.org/10.1016/j.rse.2006.05.003.

Anderegg, W.R.L., Wu, C., Acil, N., Carvalhais, N., Pugh, T.A.M., Sadler, J.P., Seidl, R., 2022. A climate risk analysis of Earth's forests in the 21st century. Science 1979 (377), 1099–1103.

Araujo, R.F., Chambers, J.Q., Celes, C.H.S., Muller-Landau, H.C., dos Santos, A.P.F., Emmert, F., Ribeiro, G.H.P.M., Gimenez, B.O., Lima, A.J.N., Campos, M.A.A., Higuchi, N., 2020. Integrating high resolution drone imagery and forest inventory to distinguish canopy and understory trees and quantify their contributions to forest structure and dynamics. PLoS One 15, 1–16. https://doi.org/10.1371/journal.pone.0243079.

Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52, 2249–2260. https://doi.org/10.1016/j.csda.2007.08.015.

Asner, G.P., Vitousek, P.M., 2005. Remote analysis of biological invasion and biogeochemical change. Proc. Natl. Acad. Sci. U. S. A. 102, 4383–4386. https://doi.org/10.1073/pnas.0500823102.

Babcock, C., Finley, A.O., Looker, N., 2021. A bayesian model to estimate land surface phenology parameters with harmonized landsat 8 and Sentinel-2 images. Remote Sens. Environ. 261 https://doi.org/10.1016/j.rse.2021.112471.

Baron, J., Hill, D.J., 2020. Monitoring grassland invasion by spotted knapweed (Centaurea maculosa) with RPAS-acquired multispectral imagery. Remote Sens. Environ. 249, 112008 https://doi.org/10.1016/j.rse.2020.112008.

Baumann, M., Ozdogan, M., Richardson, A.D., Radeloff, V.C., 2017. Phenology from landsat when data is scarce: using MODIS and dynamic time-warping to combine multi-year landsat imagery to derive annual phenology curves. Int. J. Appl. Earth Obs. Geoinf. 54, 72–83. https://doi.org/10.1016/j.jag.2016.09.005.

Becker, R.H., Zmijewski, K.A., Crail, T., 2013. Seeing the forest for the invasives: mapping buckthorn in the oak openings. Biol. Invasions 15, 315–326. https://doi.org/10.1007/s10530-012-0288-8.

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sens. Environ. 204, 509–523. https://doi.org/10.1016/j.rse.2017.10.005.

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bouvier, M., Durrieu, S., Fournier, R.A., Renaud, J.P., 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sens. Environ. 156, 322–334. https://doi.org/10.1016/j.rse.2014.10.004.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brose, P.H., 2017. An evaluation of seven methods for controlling mountain laurel thickets in the mixed-oak forests of the central Appalachian Mountains, USA. For. Ecol. Manag. 401, 286–294. https://doi.org/10.1016/j.foreco.2017.06.041.

Brose, P.H., 2016. Origin, development, and impact of mountain laurel thickets on the mixed-oak forests of the central Appalachian Mountains, USA. For. Ecol. Manag. 374, 33–41. https://doi.org/10.1016/j.foreco.2016.04.040.

Brown, J.F., Tollerud, H.J., Barber, C.P., Zhou, Q., Dwyer, J.L., Vogelmann, J.E., Loveland, T.R., Woodcock, C.E., Stehman, S.V., Zhu, Z., Pengra, B.W., Smith, K., Horton, J.A., Xian, G., Auch, R.F., Sohl, T.L., Sayler, K.L., Gallant, A.L., Zelenak, D., Reker, R.R., Rover, J., 2020. Lessons learned implementing an operational continuous United States national land change monitoring capability: the land change monitoring, assessment, and projection (LCMAP) approach. Remote Sens. Environ. 238, 111356 https://doi.org/10.1016/j.rse.2019.111356.

Bullock, E.L., Woodcock, C.E., Souza, C., Olofsson, P., 2020. Satellite-based estimates reveal widespread forest degradation in the Amazon. Glob. Chang. Biol. 26, 2956–2969. https://doi.org/10.1111/gcb.15029.

Campbell, M.J., Dennison, P.E., Hudak, A.T., Parham, L.M., Butler, B.W., 2018. Quantifying understory vegetation density using small-footprint airborne lidar. Remote Sens. Environ. 215, 330–342. https://doi.org/10.1016/j.rse.2018.06.023.

Chan, J.C.W., Paelinckx, D., 2008. Evaluation of random Forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens. Environ. 112, 2999–3011. https://doi.org/10.1016/j.rse.2008.02.011.

Chastain, R.A., Townsend, P.A., 2007. Use of landsat ETM and topographic data to characterize evergreen understory communities in appalachian deciduous forests. Photogramm. Eng. Remote Sens. 73, 563–575. https://doi.org/10.14358/pers.73.5.563.

Chen, J.M., Cihlar, J., 1996. Retrieving leaf area index of boreal conifer forests using landsat TM images. Remote Sens. Environ. 55, 153–162. https://doi.org/10.1016/0034-4257(95)00195-6.

Chrysafis, I., Mallinis, G., Tsakiri, M., Patias, P., 2019. Evaluation of single-date and multi-seasonal spatial and spectral information of Sentinel-2 imagery to assess growing stock volume of a Mediterranean forest. Int. J. Appl. Earth Obs. Geoinf. 77, 1–14. https://doi.org/10.1016/j.jag.2018.12.004.

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote. Sens. 28, 45–62. https://doi.org/10.5589/m02-004.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S.V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens. Environ. 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002.

Cohen, W.B., Healey, S.P., Yang, Z., Zhu, Z., Gorelick, N., 2020. Diversity of algorithm and spectral band inputs improves landsat monitoring of forest disturbance. Remote Sens. 12, 1–15. https://doi.org/10.3390/rs12101673.

Crespo-Peremarch, P., Tompalski, P., Coops, N.C., Ruiz, L.Á., 2018. Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data. Remote Sens. Environ. 217, 400–413. https://doi.org/10.1016/j.rse.2018.08.033.

Dai, J., Roberts, D.A., Stow, D.A., An, L., Hall, S.J., Yabiku, S.T., Kyriakidis, P.C., 2020. Mapping understory invasive plant species with field and remotely sensed data in Chitwan, Nepal. Remote Sens. Environ. 250 https://doi.org/10.1016/j.rse.2020.112037.

Dannenberg, M.P., Song, C., Hwang, T., Wise, E.K., 2015. Empirical evidence of El Niño-southern oscillation influence on land surface phenology and productivity in the

western United States. Remote Sens. Environ. 159, 167–180. https://doi.org/10.1016/j.rse.2014.11.026.

de Livera, A.M., Hyndman, R.J., Snyder, R.D., 2011. Forecasting time series with complex seasonal patterns using exponential smoothing. J. Am. Stat. Assoc. 106, 1513–1527. https://doi.org/10.1198/jasa.2011.tm09771.

der Sluijs, J., Hall, R.J., Peddle, D.R., 2016. Influence of field-based species composition and understory descriptions on spectral mixture analysis of tree species in the Northwest Territories, Canada. Can. J. Remote Sens. 42, 591–609. https://doi.org/10.1080/07038992.2016.1196581.

Dewitz, J., Survey, U.S.Geological, 2021. National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021). https://doi.org/10.5066/P9KZCM54. U.S. Geological Survey data release [WWW Document].

Dinstein, I., Shanmugam, K., Haralick, R.M., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. SMC-3, 610–621.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution Mission for GMES operational services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

Duncanson, L.I., Cook, B.D., Hurtt, G.C., Dubayah, R.O., 2014. An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. Remote Sens. Environ. 154, 378–386. https://doi.org/10.1016/j.rse.2013.07.044.

EDDMapS, 2023. Early Detection & Distribution Mapping System [WWW Document]. URL. The University of Georgia - Center for Invasive Species and Ecosystem Health (accessed 1.23.23). https://www.eddmaps.org/distribution/uscounty.cfm?sub=3010&map=distribution.

Ferreira, M.P., Wagner, F.H., Aragão, L.E.O.C., Shimabukuro, Y.E., de Souza Filho, C.R., 2019. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. ISPRS J. Photogramm. Remote Sens. 149, 119–131. https://doi.org/10.1016/j.isprsjprs.2019.01.019.

Fragoso-Campón, L., Quirós, E., Mora, J., Gutiérrez Gallego, J.A., Durán-Barroso, P., 2020. Overstory-understory land cover mapping at the watershed scale: accuracy enhancement by multitemporal remote sensing analysis and LiDAR. Environ. Sci. Pollut. Res. 27, 75–88. https://doi.org/10.1007/s11356-019-04520-8.

Fridley, J.D., 2012. Extended leaf phenology and the autumn niche in deciduous forest invasions. Nature 5–10. https://doi.org/10.1038/nature11056.

Ganguly, S., Nemani, R.R., Zhang, G., Hashimoto, H., Milesi, C., Michaelis, A., Wang, W., Votava, P., Samanta, A., Melton, F., Dungan, J.L., Vermote, E., Gao, F., Knyazikhin, Y., Myneni, R.B., 2012. Generating global leaf area index from landsat: algorithm formulation and demonstration. Remote Sens. Environ. 122, 185–202. https://doi.org/10.1016/j.rse.2011.10.032.

Gibson, R., Danaher, T., Hehir, W., Collins, L., 2020. A remote sensing approach to mapping fire severity in South-Eastern Australia using sentinel 2 and random forest. Remote Sens. Environ. 240, 111702 https://doi.org/10.1016/j.rse.2020.111702.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011.

Gray, A., Brandeis, T., Shaw, J., McWilliams, W., Miles, P., 2012. Forest inventory and analysis database of the United States of America (FIA). Biodivers. Ecol. 4, 225–231. https://doi.org/10.7809/b-e.00079.

Hall-Beyer, M., 2017. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. Int. J. Remote Sens. 38, 1312–1338. https://doi.org/10.1080/01431161.2016.1278314.

Hemmerling, J., Pflugmacher, D., Hostert, P., 2021. Mapping temperate forest tree species using dense Sentinel-2 time series. Remote Sens. Environ. 267 https://doi.org/10.1016/j.rse.2021.112743.

Hermosilla, T., Bastyr, A., Coops, N.C., White, J.C., Wulder, M.A., 2022. Mapping the presence and distribution of tree species in Canada's forested ecosystems. Remote Sens. Environ. 282 https://doi.org/10.1016/j.rse.2022.113276.

Heydari, S.S., Mountrakis, G., 2018. Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 landsat sites. Remote Sens. Environ. 204, 648–658. https://doi.org/10.1016/j.rse.2017.09.035.

Hicks, D.J., Taylor, M.S., 2015. Effects of Aesculus glabra canopy on understory community structure and environment in a temperate deciduous Forest. Castanea 80, 8–19. https://doi.org/10.2179/14-024R3.

Hoover, C.M., Bartig, J.L., Bogaczyk, B., Breeden, C., Iverson, L.R., Prout, L., Sheffield, R. M., 2022. Forest inventory and analysis data in action: Examples from eastern national forests. Trees For. People 7. https://doi.org/10.1016/j.tfp.2021.100178.

Huang, H., Chen, Y., Clinton, N., Wang, J., Wang, X., Liu, C., Gong, P., Yang, J., Bai, Y., Zheng, Y., Zhu, Z., 2017. Mapping major land cover dynamics in Beijing using all landsat images in Google earth engine. Remote Sens. Environ. 202, 166–176. https://doi.org/10.1016/j.rse.2017.02.021.

Hubau, W., De Mil, T., Van den Bulcke, J., Phillips, O.L., Angoboy Ilondea, B., Van Acker, J., Sullivan, M.J.P., Nsenga, L., Toirambe, B., Couralet, C., Banin, L.F., Begne, S.K., Baker, T.R., Bourland, N., Chezeaux, E., Clark, C.J., Collins, M., Comiskey, J.A., Cuni-Sanchez, A., Deklerck, V., Dierickx, S., Doucet, J.L., Ewango, C. E.N., Feldpausch, T.R., Gilpin, M., Gonmadje, C., Hall, J.S., Harris, D.J., Hardy, O.J., Kamdem, M.N.D., Kasongo Yakusu, E., Lopez-Gonzalez, G., Makana, J.R., Malhi, Y., Mbayu, F.M., Moore, S., Mukinzi, J., Pickavance, G., Poulsen, J.R., Reitsma, J., Rousseau, M., Sonké, B., Sunderland, T., Taedoumg, H., Talbot, J., Tshibamba Mukendi, J., Umunay, P.M., Vleminckx, J., White, L.J.T., Zemagho, L., Lewis, S.L., Beeckman, H., 2019. The persistence of carbon in the African forest understory. Nat. Plants 5, 133–140. https://doi.org/10.1038/s41477-018-0316-5.

Hyndman, R.J., Athanasopoulos, G., 2018. Forecasting: Principles and Practice, Principles of Optimal Design.

Hyndman, R.J., Athanasopoulos, G., 2018. Forecasting: Principles and Practice, Principles of Optimal Design.

Ishida, E.E.O., 2019. Machine learning and the future of supernova cosmology. Nat. Astron. https://doi.org/10.1038/s41550-019-0860-6.

Jakubauskas, M.E., Legates, D.R., Kastens, J.H., 2001. Harmonic analysis of time-series AVHRR NDVI data mark. Photogramm. Eng Remote Sens. 67, 461–470.

Karasiak, N., Dejoux, J.F., Monteil, C., Sheeren, D., 2022. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. Mach. Learn. 111, 2715–2740. https://doi.org/10.1007/s10994-021-05972-1.

Kartesz, J.T., 2015. The Biota of North America Program (BONAP) Version 1.0. [WWW Document]. URL (accessed 1.23.23). http://bonap.net/Napa/TaxonMaps/Genus/State/Berberis.

Kiviat, E., 2023. Wetland Assessment of the Proposed Mill Brook Greenway, Village and Town of New Paltz, Ulster County, New York Report to the Village of New Paltz and the Town of New Paltz.

Kobayashi, H., Yunus, A.P., Nagai, S., Sugiura, K., Kim, Y., Van Dam, B., Nagano, H., Zona, D., Harazono, Y., Bret-Harte, M.S., Ichii, K., Ikawa, H., Iwata, H., Oechel, W.C., Ueyama, M., Suzuki, R., 2016. Latitudinal gradient of spruce forest understory and tundra phenology in Alaska as observed from satellite and ground-based data. Remote Sens. Environ. 177, 160–170. https://doi.org/10.1016/j.rse.2016.02.020.

Landuyt, D., Perring, M.P., Seidl, R., Taubert, F., Verbeeck, H., Verheyen, K., 2018. Modelling understorey dynamics in temperate forests under global change–Challenges and perspectives. Perspect. Plant Ecol. Evol. Syst. 31, 44–54. https://doi.org/10.1016/j.ppees.2018.01.002.

Lee, B.R., Miller, T.K., Rosche, C., Yang, Y., Heberling, J.M., Kuebbing, S.E., Primack, R. B., 2022. Wildflower phenological escape differs by continent and spring temperature. Nat. Commun. 13, 7157. https://doi.org/10.1038/s41467-022-34936-9.

Li, C., Gong, P., Wang, J., Yuan, C., Hu, T., Wang, Q., Yu, L., Clinton, N., Li, M., Guo, J., Feng, D., Huang, C., Zhan, Z., Wang, X., Xu, B., Nie, Y., Hackman, K., 2016. An all-season sample database for improving land-cover mapping of Africa with two classification schemes. Int. J. Remote Sens. 37, 4623–4647. https://doi.org/10.1080/01431161.2016.1213923.

Li, C., Gong, P., Wang, J., Zhu, Z., Biging, G.S., Yuan, C., Hu, T., Zhang, H., Wang, Q., Li, X., Liu, X., Xu, Y., Guo, J., Liu, C., Hackman, K.O., Zhang, M., Cheng, Y., Yu, L., Yang, J., Huang, H., Clinton, N., 2017. The first all-season sample set for mapping global land cover with Landsat-8 data. Sci. Bull. (Beijing) 62, 508–515. https://doi.org/10.1016/j.scib.2017.03.011.

Li, C., Xian, G., Zhou, Q., Pengra, B.W., 2021. A novel automatic phenology learning (APL) method of training sample selection using multiple datasets for time-series land cover mapping. Remote Sens. Environ. 266, 112670 https://doi.org/10.1016/j.rse.2021.112670.

Li, L., Chen, J., Mu, X., Li, W., Yan, G., Xie, D., Zhang, W., 2020. Quantifying understory and overstory vegetation cover using UAV-based RGB imagery in forest plantation. Remote Sens. 12, 1–18. https://doi.org/10.3390/rs12020298.

Li, Q., Wong, F.K.K., Fung, T., 2021. Mapping multi-layered mangroves from multispectral, hyperspectral, and LiDAR data. Remote Sens. Environ. 258, 112403 https://doi.org/10.1016/j.rse.2021.112403.

Liang, J., Hawbaker, T.J., Zhu, Z., Li, X., Gong, P., 2016. Forest disturbance interactions and successional pathways in the southern Rocky Mountains. For. Ecol. Manag. 375, 35–45. https://doi.org/10.1016/j.foreco.2016.05.010.

Link, A.F., Turnblacer, T., Snyder, C.K., Daugherty, S.E., Utz, R.M., 2018. Low recruitment of native trees in a deciduous Forest associated with Japanese barberry (Berberis thunbergii) invasion. Invasive Plant Sci. Manag. 11, 20–26. https://doi.org/10.1017/inp.2018.1.

Linske, M.A., Williams, S.C., Ward, J.S., Stafford, K.C., 2018. Indirect effects of japanese barberry infestations on white-footed mice exposure to borrelia burgdorferi. Environ. Entomol. 47, 795–802. https://doi.org/10.1093/ee/nvy079.

Liu, X., Zhu, A.X., Yang, L., Pei, T., Liu, J., Zeng, C., Wang, D., 2020. A graded proportion method of training sample selection for updating conventional soil maps. Geoderma 357. https://doi.org/10.1016/j.geoderma.2019.113939.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century Forest cover change. Science 1979 (342), 850–853. https://doi.org/10.1126/science.1244693.

Maynard-Bean, E., Kaye, M., 2019. Invasive shrub removal benefits native plants in an eastern deciduous forest of North America. Invasive Plant Sci. Manag. 12, 3–10. https://doi.org/10.1017/inp.2018.35.

McDermid, G.J., Hall, R.J., Sanchez-Azofeifa, G.A., Franklin, S.E., Stenhouse, G.B., Kobliuk, T., LeDrew, E.F., 2009. Remote sensing and forest inventory for wildlife habitat assessment. For. Ecol. Manag. 257, 2262–2269. https://doi.org/10.1016/j.foreco.2009.03.005.

McLachlan, S.M., Bazely, D.R., 2001. Recovery patterns of understory herbs and their use as indicators of deciduous forest regeneration. Conserv. Biol. 15, 98–110. https://doi.org/10.1046/j.1523-1739.2001.98145.x.

Meng, R., Wu, J., Zhao, F., Cook, B.D., Hanavan, R.P., Serbin, S.P., 2018. Measuring short-term post-fire forest recovery across a burn severity gradient in a mixed pine-oak forest using multi-sensor remote sensing techniques. Remote Sens. Environ. 210, 282–296. https://doi.org/10.1016/j.rse.2018.03.019.

Murray, H., Lucieer, A., Williams, R., 2010. Texture-based classification of sub-Antarctic vegetation communities on heard island. Int. J. Appl. Earth Obs. Geoinf. 12, 138–149. https://doi.org/10.1016/j.jag.2010.01.006.

Naqinezhad, A., de Lombaerde, E., Gholizadeh, H., Wasof, S., Perring, M.P., Meeussen, C., de Frenne, P., Verheyen, K., 2022. The combined effects of climate and canopy cover changes on understorey plants of the hyrcanian forest biodiversity hotspot in northern Iran. Glob. Chang. Biol. 28, 1103–1118. https://doi.org/10.1111/gcb.15946.

Neufeld, H.S., Young, D.R., 2014. Ecophysiology of the Herbaceous Layer in Temperate Deciduous Forests. In: The Herbaceous Layer in Forests of Eastern North America (Edited by Frank s. Gilliam), pp. 35–95.

Nikopensius, M., Pisek, J., Raabe, K., 2015. Spectral reflectance patterns and seasonal dynamics of common understory types in three mature hemi-boreal forests. Int. J. Appl. Earth Obs. Geoinf. 43, 84–91. https://doi.org/10.1016/j.jag.2014.11.012.

Ohman, M., 2006. Characteristics of fuel beds invaded by smilax rotundifolia. University of Massachusetts Amherst.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.

Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sens. Environ. 129, 122–131. https://doi.org/10.1016/j.rse.2012.10.031.

O'Loughlin, L.S., Gooden, B., Foster, C.N., MacGregor, C.I., Catford, J.A., Lindenmayer, D.B., 2019. Invasive shrub re-establishment following management has contrasting effects on biodiversity. Sci. Rep. 9, 1–11. https://doi.org/10.1038/s41598-019-40654-y.

Oreti, L., Barbati, A., Marini, F., Giuliarelli, D., 2020. Very high-resolution true color leaf-off imagery for mapping Taxus baccata L. And Ilex aquifolium L. Understory population. Biodivers. Conserv. 29, 2605–2622. https://doi.org/10.1007/s10531-020-01991-x.

Ozdemir, I., 2014. Linear transformation to minimize the effects of variability in understory to estimate percent tree canopy cover using rapideye data. GIsci Remote Sens. 51, 288–300. https://doi.org/10.1080/15481603.2014.912876.

Pal, M., 2006. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. Int. J. Remote Sens. 27, 2877–2894.

Paolucci, A.J., Stolt, M.H., 2018. Assessing dynamic soil properties in southern New England forests within an ecological site framework. Soil Sci. Soc. Am. J. 82, 1191–1202. https://doi.org/10.2136/sssaj2018.01.0047.

Pasquarella, V.J., Holden, C.E., Woodcock, C.E., 2018. Improved mapping of forest type using spectral-temporal landsat features. Remote Sens. Environ. 210, 193–207. https://doi.org/10.1016/j.rse.2018.02.064.

Patterson, P.L., O'brien, R.A., 2011. Understory Vegetation Data Quality Assessment for the Interior West Forest Inventory and Analysis Program.

Pisek, J., Chen, J.M., Kobayashi, H., Rautiainen, M., Schaepman, M.E., Karnieli, A., Sprinstin, M., Ryu, Y., Nikopensius, M., Raabe, K., 2016. Retrieval of seasonal dynamics of forest understory reflectance from semiarid to boreal forests using MODIS BRDF data. J. Geophys. Res. Biogeosci. 121, 855–863. https://doi.org/10.1002/2016JG003322.

Pisek, J., Rautiainen, M., Heiskanen, J., Mõttus, M., 2012. Retrieval of seasonal dynamics of forest understory reflectance in a northern european boreal forest from MODIS BRDF data. Remote Sens. Environ. 117, 464–468. https://doi.org/10.1016/j.rse.2011.09.012.

Pisek, J., Rautiainen, M., Nikopensius, M., Raabe, K., 2015. Estimation of seasonal dynamics of understory NDVI in northern forests using MODIS BRDF data: semi-empirical versus physically-based approach. Remote Sens. Environ. 163, 42–47. https://doi.org/10.1016/j.rse.2015.03.003.

Ploughe, L.W., Dukes, J.S., 2019. Understory plant composition and nitrogen transformations resistant to changes in seasonal precipitation. Ecosphere 10. https://doi.org/10.1002/ecs2.2747.

Poindexter, D.B., Thompson, R.L., 2009. Vascular flora and plant habitats of Wallace woods, a hemlock-northern hardwoods palustrine forest, Crawford County, Pennsylvania. Rhodora 111, 231–260. https://doi.org/10.3119/08-10.1.

Powell, R.L., Matzke, N., De Souza, C., Clark, M., Numata, I., Hess, L.L., Roberts, D.A., Clark, M., Numata, I., Hess, L.L., Roberts, D.A., 2004. Sources of error in accuracy assessment of thematic land-cover maps in the brazilian Amazon. Remote Sens. Environ. 90, 221–234. https://doi.org/10.1016/j.rse.2003.12.007.

Qiu, S., Zhu, Z., He, B., 2019. Fmask 4.0: improved cloud and cloud shadow detection in landsats 4–8 and Sentinel-2 imagery. Remote Sens. Environ. 231, 111205 https://doi.org/10.1016/j.rse.2019.05.024.

Rapinel, S., Mony, C., Lecoq, L., Clément, B., Thomas, A., Hubert-Moy, L., 2019. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. Remote Sens. Environ. 223, 115–129. https://doi.org/10.1016/j.rse.2019.01.018.

Rautiainen, M., Heiskanen, J., 2013. Seasonal contribution of understory vegetation to the reflectance of a boreal landscape at different spatial scales. IEEE Geosci. Remote Sens. Lett. 10, 923–927. https://doi.org/10.1109/LGRS.2013.2247560.

Rautiainen, M., Mõttus, M., Heiskanen, J., Akujärvi, A., Majasalmi, T., Stenberg, P., 2011. Seasonal reflectance dynamics of common understory types in a northern european boreal forest. Remote Sens. Environ. 115, 3020–3028. https://doi.org/10.1016/j.rse.2011.06.005.

Rautiainen, M., Nilson, T., Lükk, T., 2009. Seasonal reflectance trends of hemiboreal birch forests. Remote Sens. Environ. 113, 805–815. https://doi.org/10.1016/j.rse.2008.12.009.

Reed, B.C., 2006. Trend analysis of time-series phenology of North America derived from satellite data. GIsci Remote Sens. 43, 24–38. https://doi.org/10.2747/1548-1603.43.1.24.

Reinmann, A.B., Hutyra, L.R., 2017. Edge effects enhance carbon uptake and its vulnerability to climate change in temperate broadleaf forests. Proc. Natl. Acad. Sci. U. S. A. 114, 107–112. https://doi.org/10.1073/pnas.1612369114.

Rittenhouse, C.D., Berlin, E.H., Mikle, N., Qiu, S., Riordan, D., Zhu, Z., 2022. An object-based approach to map young Forest and shrubland vegetation based on multi-source remote sensing data. Remote Sens. 14, 1–20. https://doi.org/10.3390/rs14051091.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913–929. https://doi.org/10.1111/ecog.02881.

Royo, A.A., Carson, W.P., 2006. On the formation of dense understory layers in forests worldwide: consequences and implications for forest dynamics, biodiversity, and succession. Can. J. For. Res. 36, 1345–1362.

Sanz, B., Malinen, J., Heiskanen, J., Tokola, T., 2020. Need for pre-harvest clearing of understory vegetation determined by airborne laser scanning. Forests 11, 1–14. https://doi.org/10.3390/f11030294.

Schafer, A., Man, R., Chen, H.Y.H., Lu, P., 2014. Effects of post-windthrow management interventions on understory plant communities in aspen-dominated boreal forests. For. Ecol. Manag. 323, 39–46. https://doi.org/10.1016/j.foreco.2014.03.030.

Scolastri, A., Bricca, A., Cancellieri, L., Cutini, M., 2017. Understory functional response to different management strategies in Mediterranean beech forests (central apennines, Italy). For. Ecol. Manag. 400, 665–676. https://doi.org/10.1016/j.foreco.2017.06.049.

Senécal, J.F., Doyon, F., Messier, C., 2018. Disentangling the causes of canopy height increase in managed and unmanaged temperate deciduous forests using multi-temporal airborne laser scanning. Remote Sens. Environ. 217, 233–243. https://doi.org/10.1016/j.rse.2018.08.023.

Senf, C., Seidl, R., 2021. Mapping the forest disturbance regimes of Europe. Nat. Sustain. 4, 63–70. https://doi.org/10.1038/s41893-020-00609-y.

Seyednasrollah, B., Swenson, J.J., Domec, J.C., Clark, J.S., 2018. Leaf phenology paradox: why warming matters most where it is already warm. Remote Sens. Environ. 209, 446–455. https://doi.org/10.1016/j.rse.2018.02.059.

Shang, R., Zhu, Z., 2019. Harmonizing landsat 8 and Sentinel-2: a time-series-based reflectance adjustment approach. Remote Sens. Environ. 235, 111439 https://doi.org/10.1016/j.rse.2019.111439.

Shao, Z., Cai, J., Fu, P., Hu, L., Liu, T., 2019. Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product. Remote Sens. Environ. 235, 111425 https://doi.org/10.1016/j.rse.2019.111425.

Shouse, M., Liang, L., Fei, S., 2012. Identification of understory invasive exotic plants with remote sensing: in urban forests. Int. J. Appl. Earth Obs. Geoinf. 21, 525–534. https://doi.org/10.1016/j.jag.2012.07.010.

Singh, K.K., Chen, Y.H., Smart, L., Gray, J., Meentemeyer, R.K., 2018. Intra-annual phenology for detecting understory plant invasion in urban forests. ISPRS J. Photogramm. Remote Sens. 142, 151–161. https://doi.org/10.1016/j.isprsjprs.2018.05.023.

Singh, K.K., Davis, A.J., Meentemeyer, R.K., 2015. Detecting understory plant invasion in urban forests using LiDAR. Int. J. Appl. Earth Obs. Geoinf. 38, 267–279. https://doi.org/10.1016/j.jag.2015.01.012.

Singh, K.K., Gray, J., 2020. Mapping understory invasive plants in urban forests with spectral and temporal unmixing of landsat imagery. Photogramm. Eng. Remote Sens. 86, 509–518. https://doi.org/10.14358/pers.86.8.509.

Song, K., Zhu, Z., 2022. Improved subtle change detection using Landsat and Sentinel-2 data fusion: A study of spongy moth outbreaks in New England forests. In: AGU Fall Meeting Abstracts, 2022 pp. B43B-04.

Su, X., Zheng, G., Chen, H.Y.H., 2022. Understory diversity are driven by resource availability rather than resource heterogeneity in subtropical forests. For. Ecol. Manag. 503 https://doi.org/10.1016/j.foreco.2021.119781.

Sumnall, M.J., Trlica, A., Carter, D.R., Cook, R.L., Schulte, M.L., Campoe, O.C., Rubilar, R.A., Wynne, R.H., Thomas, V.A., 2021. Estimating the overstory and understory vertical extents and their leaf area index in intensively managed loblolly pine (Pinus taeda L.) plantations using airborne laser scanning. Remote Sens. Environ. 254 https://doi.org/10.1016/j.rse.2020.112250.

Sun, C., Li, J., Liu, Yongxue, Liu, Yongchao, Liu, R., 2021a. Plant species classification in salt marshes using phenological parameters derived from Sentinel-2 pixel-differential time-series. Remote Sens. Environ. 256 https://doi.org/10.1016/j.rse.2021.112320.

Sun, C., Li, J., Liu, Yongxue, Liu, Yongchao, Liu, R., 2021b. Plant species classification in salt marshes using phenological parameters derived from Sentinel-2 pixel-differential time-series. Remote Sens. Environ. 256 https://doi.org/10.1016/j.rse.2021.112320.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 273–282. https://doi.org/10.1111/j.1467-9868.2011.00771.x.

Tinya, F., Ódor, P., 2016. Congruence of the spatial pattern of light and understory vegetation in an old-growth, temperate mixed forest. For. Ecol. Manag. 381, 84–92. https://doi.org/10.1016/j.foreco.2016.09.027.

Tuanmu, M.N., Viña, A., Bearer, S., Xu, W., Ouyang, Z., Zhang, H., Liu, J., 2010. Mapping understory vegetation using phenological characteristics derived from remotely sensed data. Remote Sens. Environ. 114, 1833–1844. https://doi.org/10.1016/j.rse.2010.03.008.

Utz, R.M., Fetsko, M.N., 2020. Exploratory survey of salamanders in pennsylvanian forests with dense understories of Berberis thunbergii (Japanese Barberry), an invasive shrub. Northeast Nat. (Steuben) 27, 299–306. https://doi.org/10.1656/045.027.0211.

Van Doninck, J., Jones, M.M., Zuquim, G., Ruokolainen, K., Moulatlet, G.M., Sirén, A., Cárdenas, G., Lehtonen, S., Tuomisto, H., 2020. Multispectral canopy reflectance improves spatial distribution models of Amazonian understory species. Ecography 43, 128–137. https://doi.org/10.1111/ecog.04729.

Verbesselt, J., Zeileis, A., Herold, M., 2012. Near real-time disturbance detection using satellite image time series. Remote Sens. Environ. 123, 98–108. https://doi.org/10.1016/j.rse.2012.02.022.

Wang, H., Zhang, M., Nan, H., 2019. Abiotic and biotic drivers of species diversity in understory layers of cold temperate coniferous forests in North China. J. For. Res. (Harbin) 30, 2213–2225. https://doi.org/10.1007/s11676-018-0795-2.

Wang, J.A., Sulla-Menashe, D., Woodcock, C.E., Sonnentag, O., Keeling, R.F., Friedl, M. A., 2020. Extensive land cover change across Arctic-boreal northwestern North America from disturbance and climate forcing. Glob. Chang. Biol. 26, 807–822. https://doi.org/10.1111/gcb.14804.

Wang, T., Skidmore, A.K., Toxopeus, A.G., Liu, X., 2009. Understory bamboo discrimination using a winter image. Photogramm. Eng. Remote Sens. 75, 37–47. https://doi.org/10.14358/PERS.75.1.37.

Ward, J.S., Worthley, T.E., Williams, S.C., 2009. Controlling japanese barberry (Berberis thunbergii DC) in southern New England, USA. For. Ecol. Manag. 257, 561–566. https://doi.org/10.1016/j.foreco.2008.09.032.

Welch, R., Madden, M., Jordan, T., 2002. Photogrammetric and GIS techniques for the development of vegetation databases of mountainous areas: Great Smoky Mountains National Park. ISPRS J. Photogramm. Remote Sens. 57, 53–68. https://doi.org/10.1016/S0924-2716(02)00118-1.

Wen, Y., Li, X., Mu, H., Zhong, L., Chen, H., Zeng, Y., Miao, S., Su, W., Gong, P., Li, B., Huang, J., 2022. Mapping corn dynamics using limited but representative samples with adaptive strategies. ISPRS J. Photogramm. Remote Sens. 190, 252–266. https://doi.org/10.1016/j.isprsjprs.2022.06.012.

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.D., Brown, J.F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Glob. Chang. Biol. 15, 2335–2359. https://doi.org/10.1111/j.1365-2486.2009.01910.x.

Wilde, H.D., Gandhi, K.J.K., Colson, G., 2015. State of the science and challenges of breeding landscape plants with ecological function. Hortic Res. 2 https://doi.org/10.1038/hortres.2014.69.

Wilfong, B.N., Gorchov, D.L., Henry, M.C., 2009. Detecting an invasive shrub in deciduous Forest understories using remote sensing. Weed Sci. 57, 512–520. https://doi.org/10.1614/ws-09-012.1.

Williams, S.C., Ward, J.S., 2010. Effects of japanese barberry (Ranunculales: Berberidaceae) removal and resulting microclimatic changes on ixodes scapularis (Acari: Ixodidae) abundances in Connecticut, USA. Environ. Entomol. 39, 1911–1921. https://doi.org/10.1603/EN10131.

Xi, Y., Tian, Q., Zhang, W., Zhang, Z., Tong, X., Brandt, M., Fensholt, R., 2022. Quantifying understory vegetation density using multi-temporal Sentinel-2 and GEDI LiDAR data. GIsci Remote Sens. 59, 2068–2083. https://doi.org/10.1080/15481603.2022.2148338.

Yang, B., Qin, L., Liu, J., Liu, X., 2022. UTRNet: an unsupervised time-distance-guided convolutional recurrent network for change detection in irregularly collected images. IEEE Trans. Geosci. Remote Sens. 60 https://doi.org/10.1109/TGRS.2022.3174009.

Yang, X., Qin, Q., Grussenmeyer, P., Koehl, M., 2018. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. Remote Sens. Environ. 219, 259–270. https://doi.org/10.1016/j.rse.2018.09.016.

Yang, X., Qin, Q., Yésou, H., Ledauphin, T., Koehl, M., Grussenmeyer, P., Zhu, Z., 2020. Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data. Remote Sens. Environ. 244, 111803 https://doi.org/10.1016/j.rse.2020.111803.

Yang, X., Qiu, S., Zhu, Z, Rittenhouse, C., Riordan, D, Cullerton, M., 2023. Understory species map in Connecticut US. Mendeley Data V3,. https://doi.org/10.17632/rschxhwgvw.3.

Yang, X., Zhe, Qiu, S., Kroeger, K.D., Zhu, Zhiliang, Covington, S., 2022. Detection and characterization of coastal tidal wetland change in the northeastern US using landsat time series. Remote Sens. Environ. 276, 113047 https://doi.org/10.1016/j.rse.2022.113047.

Ye, S., Rogan, J., Zhu, Z., Hawbaker, T.J., Hart, S.J., Andrus, R.A., Meddens, A.J.H., Hicke, J.A., Eastman, J.R., Kulakowski, D., 2021. Detecting subtle change from dense landsat time series: case studies of mountain pine beetle and spruce beetle disturbance. Remote Sens. Environ. 263 https://doi.org/10.1016/j.rse.2021.112560.

Zhang, J., Shang, R., Rittenhouse, C., Witharana, C., Zhu, Z., 2021. Evaluating the impacts of models, data density and irregularity on reconstructing and forecasting dense landsat time series. Sci. Remote Sens. 4, 100023 https://doi.org/10.1016/j.srs.2021.100023.

Zhang, Y., Ling, F., Wang, X., Foody, G.M., Boyd, D.S., Li, X., Du, Y., Atkinson, P.M., 2021. Tracking small-scale tropical forest disturbances: fusing the landsat and Sentinel-2 data record. Remote Sens. Environ. 261, 112470 https://doi.org/10.1016/j.rse.2021.112470.

Zhang, Y., Woodcock, C.E., Chen, S., Wang, J.A., Sulla-Menashe, D., Zuo, Z., Olofsson, P., Wang, Y., Friedl, M.A., 2022. Mapping causal agents of disturbance in boreal and arctic ecosystems of North America using time series of landsat data. Remote Sens. Environ. 272, 112935 https://doi.org/10.1016/j.rse.2022.112935.

Zhou, Z.H., Li, M., 2010. Semi-supervised learning by disagreement. Knowl. Inf. Syst. 24, 415–439. https://doi.org/10.1007/s10115-009-0209-z.

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 Sens. Environ. 274 https://doi.org/10.1016/j.rse.2022.113002.

Zhu, Z., 2017. Change detection using landsat time series: a review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogramm. Remote Sens. 130, 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013.

Zhu, Z., Gallant, A.L., Woodcock, C.E., Pengra, B., Olofsson, P., Loveland, T.R., Jin, S., Dahal, D., Yang, L., Auch, R.F., 2016. Optimizing selection of training and auxiliary data for operational land cover classification for the LCMAP initiative. ISPRS J. Photogramm. Remote Sens. 122, 206–221. https://doi.org/10.1016/j.isprsjprs.2016.11.004.

Zhu, Z., Woodcock, C.E., Holden, C., Yang, Z., 2015. Generating synthetic landsat images based on all available landsat data: predicting landsat surface reflectance at any given time. Remote Sens. Environ. 162, 67–83. https://doi.org/10.1016/j.rse.2015.02.009.

Zhu, Z., Woodcock, C.E., Olofsson, P., 2012a. Continuous monitoring of forest disturbance using all available landsat imagery. Remote Sens. Environ. 122, 75–91. https://doi.org/10.1016/j.rse.2011.10.030.

Zhu, Z., Woodcock, C.E., Rogan, J., Kellndorfer, J., 2012b. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using landsat and SAR data. Remote Sens. Environ. 117, 72–82. https://doi.org/10.1016/j.rse.2011.07.020.

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A.H., Cohen, W.B., Qiu, S., Zhou, C., 2020. Continuous monitoring of land disturbance based on landsat time series. Remote Sens. Environ. 238, 111116 https://doi.org/10.1016/j.rse.2019.03.009.

Zołnierz, L., Weber, J., Gilewska, M., Straczyńska, S., Pruchniewicz, D., 2016. The spontaneous development of understory vegetation on reclaimed and afforested post-mine excavation filled with fly ash. Catena (Amst.) 136, 84–90. https://doi.org/10.1016/j.catena.2015.07.013.