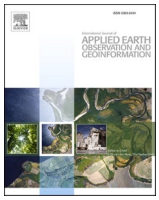
International Journal of Applied Earth Observations and Geoinformation 107 (2022) 102689

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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



Mapping forest types over large areas with Landsat imagery partially affected by clouds and SLC gaps

Contributors: Konrad Turlej *, Mutlu Ozdogan, Volker C. Radeloff

Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Dr, Madison, WI 53706, USA

ARTICLE INFO

Keywords:
Forest type
Landsat
Clouds
SLC gaps
Remote sensing
Classification
Forest inventory map
Wisconsin
USA

ABSTRACT

The ecosystem services that forests provide depend on tree species composition. Therefore, it is important to map not only forest extent and its dynamics, but also composition. Open access to Landsat has resulted in considerable improvements in remote sensing methods for mapping tree species, but most approaches fail to perform when there is a shortage of clear observations. Our main goal was to map forest composition with Landsat imagery in various data availability conditions, and to investigate how the missing data, either due to clouds or scan line problems affect classification accuracy. We tested a data driven approach that is based on multi-temporal analysis of the tree species' spectral characteristics making it applicable to regional-scale mapping even when the gap-free imagery is not available. Our study area consisted of one Landsat footprint (26/28) located in Northern Wisconsin, USA. We selected this area because of numerous tree species (23), heterogenic composition of forests where the majority of stands are mixed, and availability of high-quality reference data. We quantified how classification accuracy at the species level was affected by a) the amount of missing data due to cloud cover and Scanning Line Corrector (SLC) gaps, b) the number of acquisitions, and c) the seasonal availability of images. We applied a decision tree classifier, capable of handling missing data to both single- and a three-year Landsat-7 and Landsat-8 observations. We classified the dominant tree species in each pixel and grouped results to forest stands to match our reference data. Our results show four major findings. First, producer's and user's accuracies range from 46.2% to 96.2% and from 59.9% to 93.7%, respectively for the most abundant forest types in the study area (all types covering greater than 2% of the forest area). Second, all tree species were mapped with overall accuracy above 70% even in when we restricted our data set to images having gaps larger than 30% of the study area. Third, the classification accuracy improved with more acquisitions, especially when images were available for the fall, spring, and summer. Finally, producer's accuracies for pure-stands were higher than those for mixed stands by 10 to 30 percentage points. We conclude that inclusion of Landsat imagery with missing data allows to map forest types with accuracies that previously could be achieved only for those rare years for which several gap-free images were available. The approach presented here is directly applicable to Landsat-like observations and derived products such as seasonal composites and temporal statistics that miss 30% or more of the data for any single date to develop forest composition maps that are important for both forest management and ecology.

1. Introduction

Forests cover about one third of the Earth's land surface and provide numerous services, but these services depend on their tree species composition. For example, tree species vary in their responses to the environmental processes such as climate change (Bergh et al., 2003), nutrition cycling in soils (Finzi et al., 1998; Hobbie et al., 2007), and in their effects on the chemistry of soils and stream waters (Lovett et al., 2002). At global scales, tree species modulate the rates of climate change (Bonan, 2008). At local scales, they provide habitats for wildlife (Lee and Rotenberry, 2005; Wood et al., 2012) and affect biodiversity (Barbier et al., 2008). Knowledge about the location of certain tree species is important for their conservation and to model distributions of related wildlife species (Loiselle et al., 2003). Last but not least, tree species differ greatly in both their timber value and susceptibility of forests to disturbance causing a range of economic effects depending on the future shifts in their distributions (Hanewinkel et al., 2012).

Field-based mapping of tree species for large areas is generally cost-

* Corresponding author.

https://doi.org/10.1016/j.jag.2022.102689
Received 11 June 2021; Received in revised form 22 November 2021; Accepted 12 January 2022
Available online 29 January 2022

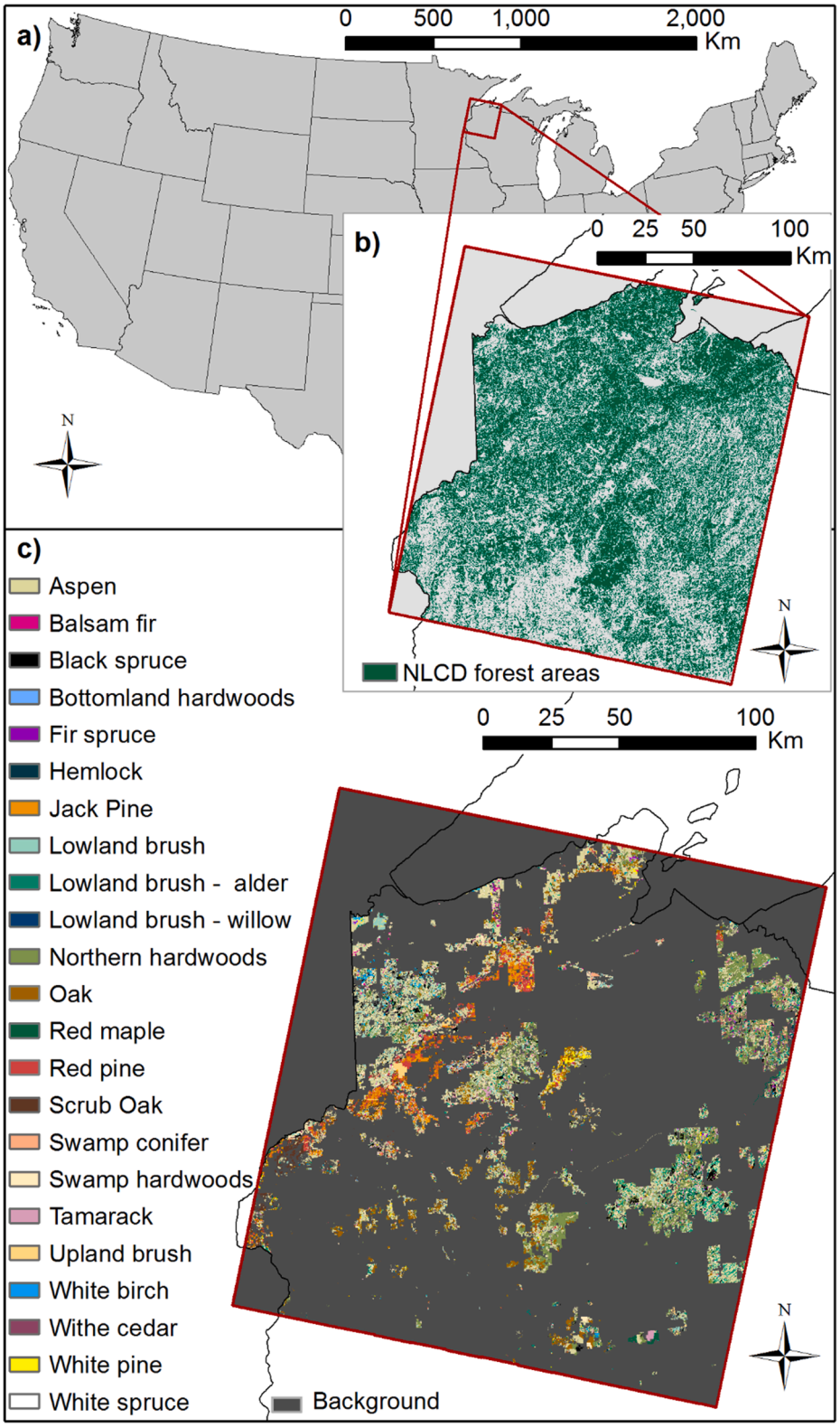


Fig. 1. Study area: a) location of the footprint path: 26, row: 28, b) forest pixels indicated by NLCD 2011, c) the forest types in RECON polygons.

prohibitive. However, recent technological developments in remote sensing make it possible to develop forest composition maps using a variety of data types (Fassnacht et al., 2014), including hyperspectral (Modzelewska et al., 2020), and airborne laser scanning data (Lindberg et al., 2021; Shi et al., 2018), even at level of individual trees (Shi et al., 2021). However, for large area mapping applications these data types are not as widely available as multispectral imagery. Tree species can be successfully separated by analyzing the differences in the surface reflectance absorption characteristics captured with multispectral information collected in the time of the year crucial to phenology. The mapping accuracy increases when analyzing multi-temporal acquisitions that capture differences in tree phenology between seasons. For

Table 1
The list of the acquisitions used in this study.

	Landsat 7 ETM+					Landsat 8 OLI				
2014										
Winter	7-Jan	23-Jan	8-Feb	24-Feb		15-Jan	31-Jan	16-Feb		
Spring	12-Mar	28-Mar	31-May			4-Mar	20-Mar	5-Apr	21-Apr	
Summer	16-Jun	2-Jul	18-Jul	3-Aug		24-Jun				
Fall	4-Sep	20-Sep	6-Oct	22-Oct		14-Oct	15-Nov			
2015										
Winter	10-Jan	27-Feb				3-Feb	19-Feb			
Spring	31-Mar	16-Apr	2-May			7-Mar	23-Mar	10-May		
Summer	19-Jun	5-Jul	21-Jul	6-Aug	22-Aug	27-Jun	13-Jul	29-Jul	14-Aug	30-Aug
Fall	7-Sep	23-Sep	9-Oct	25-Oct	10-Nov	15-Sep	1-Oct	17-Oct	18-Nov	
2016										
Winter	14-Dec					5-Jan	22-Dec			
Spring	1-Mar	2-Apr	4-May	20-May		25-Mar	10-Apr			
Summer	21-Jun	7-Jul	8-Aug	24-Aug		15-Jul	31-Jul	16-Aug		
Fall	9-Sep	12-Nov	28-Nov			1-Sep	3-Oct	19-Oct	4-Nov	20-Nov

example, Immitzer et al. (2016) reported overall accuracy of 66.2% for temperate forests in Germany as a result of classification of a single Sentinel-2 image while multi-temporal observations from Sentinel-2 resulted in accuracies ranging from 75.6% to 88.2% (Hościło and Lewandowska, 2019; Persson et al., 2018). Similarly, the use of Landsat data covering multiple stages of forest phenology resulted in mapping accuracy ranging from 83% to 96% depending on the number of tree species (Townsend and Walsh, 2001; Zhu and Liu, 2014). In rare cases, high classification accuracy can be achieved with only a few images from spring, summer, and fall (Mickelson et al., 1998). However, a larger number of acquisitions is an advantage as long as the additional imagery provides novel information on differences in tree phenology (Axelsson et al., 2021; Wolter et al., 1995; Zhu and Liu, 2014). Lastly, certain approaches are based on capturing phenology patterns from multi-year data (Pasquarella et al., 2018), and deriving spectral-temporal statistics (Grabska et al., 2020; Hemmerling et al., 2021; Schindler et al., 2021).

When the goal is to map tree species for large areas, Landsat and Sentinel-2 imagery are the only viable choices because of their spatial and temporal resolution, spectral ranges essential for monitoring tree phenology, radiometric stability, and open access to image archives. First, the 10–30 m ground sampling distance of the sensors collecting the imagery globally allows to map tree species for large areas because the images capture differences in reflectance from a sufficiently small number of trees that the dominant species for each pixel can be assessed (Wolter et al., 1995). Second, the five to eight day acquisition frequency of two Landsat satellites, plus additional Sentinel-2 images since 2015, make it possible to capture rapid changes in canopy development during the growing season that reflect phenological differences among species (Mickelson et al., 1998). Third, the radiometric quality of Landsat bands is largely stable over the entire historical archive (Wulder et al., 2012) making it possible to use multi-year observations from various sensors on different Landsat missions (Pasquarella et al., 2018). Finally, the data is openly available which reduces the costs of mapping activity.

The challenge is that existing approaches to mapping tree species with Landsat data can be limited by the availability of gap-free acquisitions or low numbers of annual per-pixel observations, which is typical for large portions of individual scenes (Wulder et al., 2016). For example, having two images from spring is highly desirable, because bud burst is often rapid and differences in its timing among tree species can be subtle (Lechowicz, 1984), but the probability of two cloud-free observations from a single Landsat satellite within a 16-day window is low for the conterminous United States (Ju and Roy, 2008). Most images have gaps due to clouds, their shadows, and Scan Line Corrector (SLC) failure issue on Landsat 7. These gaps preclude the use of many mapping approaches that require input data to be completely gap-free, including those that employ image thresholds (Dymond et al., 2002; Wolter et al., 2008, 1995), statistical classifiers (Mickelson et al., 1998), as well as non-parametrical approaches such as Random Forests and Support Vector Machines (Grabska et al., 2019; Hemmerling et al., 2021; Hościło and Lewandowska, 2019; Zhu and Liu, 2014). Even those approaches relaying on band compositing (Grabska et al., 2020) or modelling multi-year per pixel observations (Pasquarella et al., 2018) to provide gap-free data for larger extents can suffer if there is not enough observations for: 1) the compositing process or 2) the modelling of multi-year phenology patterns. Moreover, the use of multi-year data covering long time periods can introduce problems resulting from the variability in the phenology timing caused by interannual shifts in climate, land cover changes, or climate change (Garcia and Townsend, 2016).

One way to address the shortage of gap-free Landsat observations is to employ a machine learning algorithm that can handle missing data. Indeed, Landsat data with gaps have been successfully used when mapping general land cover categories (Schneider, 2012), as well as tree species (Wiscland 2 Land Cover User Guide, 2016). The same is also true for partially cloudy imagery from Formosat-2 satellites (Sheeren et al., 2016). What is less known however is how the magnitude of missing data, both in terms of the number of acquisitions and the seasons for which imagery is available affects tree species classification accuracy. To this end, our main goals were: 1) to improve mapping of tree species aggregated to forest types with Landsat imagery when data availability is limited, and 2) to evaluate the usefulness of a machine learning classifier when classifying imagery is affected by gaps due to both cloud cover and SLC failure. Our specific objectives were to assess the effects of a) the amount of missing data, b) the number of image acquisitions, and c) the seasonality of the imagery on classification accuracy.

2. Materials and methods

2.1. Study area

The study area was in northern Wisconsin and consisted of the forested areas of a single Landsat footprint (path 26, row 28) (Fig. 1a). The area is divided into three ecoregions: Lake Superior Lowland, Northern Highland, and western part of Central Plain (Martin, 1965). Forests cover just over 1.6 million hectares (roughly 30% percent of the footprint) (Fig. 1b). The local climate is temperate continental and influenced by three air masses passing over the area: the cold and dry arctic, warm and moist subtropical, and very dry continental. The flat topography, mostly a rolling plain shaped by glaciers during the Pleistocene, provides no natural obstacles for air masses causing characteristic zonal distribution of vegetation (Curtis, 1959). In addition to air masses, climate is influenced by Lake Superior in the north causing local temperature gradients.

The composition of forest tree species in our study area is mixed. Most abundant tree species include aspen spp. (*Populous*), oak spp. (*Quercus*), pine spp. (*Pinus*), spruce spp. (*Picea*), maple spp. (*Acer*), and miscellaneous hardwoods, accompanied by tamarack (*Larix larcina*),

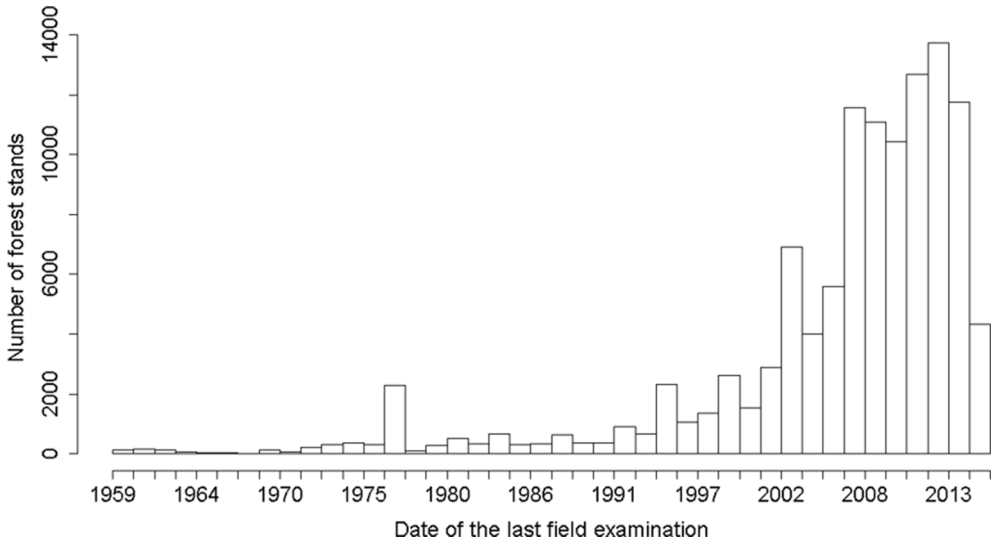


Fig. 2. Dates of last field examination for forest stands in the reference data set.

eastern hemlock (*Tsuga canadensis*), and white cedar (*Thuja occidentalis*). Contemporary tree species distributions have been shaped mainly by land use, and differs substantially from the period prior to onset of European settlers in the mid-19th century (Curtis, 1959; Radeloff et al., 1999; Rhemtulla et al., 2009, 2007).

2.2. Data

We focus here on Landsat data only and analyzed observations from Landsat-7 and Landsat-8 acquired from 1/1/2014 to 12/31/2016. In total, 82 acquisitions, out of 133, provided at least some clear pixels over the study area. Specifically, we analyzed the USGS Collection-1 Surface Reflectance product; bands 2–7 of Landsat-8 OLI, and bands 1–5 and 7 of Landsat-7 ETM+ (Table 1). The Landsat-7 data were processed via the Landsat Ecosystem Disturbance Adaptive Processing System (Schmidt et al., 2013; Department of the Interior U.S. Geological Survey, 2017) and the Landsat-8 data via the Landsat Surface Reflectance Code (Department of the Interior U.S. Geological Survey, 2020). We removed pixels that were: a) flagged as clouds, cloud shadows, or water based on the quality band, b) that exceeded the valid range from 0 to 10,000 scaled surface reflectance, and c) all non-forest pixels according to the 2011 National Land Cover Database (NLCD, (Homer et al., 2015), Fig. 1b).

We used the Reconnaissance forest inventory data (RECON) of the Wisconsin Forest Inventory & Reporting System (WisFIRS) provided by the Wisconsin Department of Natural Resources and County Forests (State of Wisconsin Department of Natural Resources, 2013) as ground reference data (Fig. 1c). This dataset provides stand-level information on forest types (Appendix B): primary tree type (≥50% of the basal area), secondary tree type, understory type, tree height, tree density, total basal area, year of stand establishment, and year of the last field examination. We separated pure stands, where primary, secondary, and understory vegetation type were all of the same tree species, from mixed stands, which included all remaining stands. The dataset does not provide detailed information on the composition of tree species in the mixed stands. The majority of the RECON forest stands were examined in the field after 2000, <16 years before the satellite observations used here (Fig. 2).

2.3. Classification

Our primary goal was to obtain an accurate pixel-level classification for 23 forest types and tree species considering a range of data availability conditions for the forested pixels indicated by the NLCD. To achieve this goal, we conducted a number of classifications to test how three aspects of data availability affect the classification accuracy: 1) the influence of missing data due to clouds and SLC gaps, 2) the influence of the number of image acquisitions, and 3) the usefulness of imagery from different seasons and their combinations. We repeated each test 10 times for the 2016-only imagery versus 2014–2016 imagery, as well as for pure stands only versus pure-plus-mixed stands. For each of the 10 iterations, we randomly divided the forest stands from the RECON into training and testing subsets, each time drawing 25% of the pure stands of each forest type into training and leaving remaining 75% of pure stands plus all mixed stands for testing. For each repetition, we built a classification model using pixels belonging to the training stands excluding the edge pixels, and applied the model to all pixels from the test stands. We determined edge pixels by applying a 30-m buffer inside the polygons.

We used the C5.0 Decision Trees and Rule-Based Models, implemented in the R statistical software, as our tool for classification (Kuhn et al., 2021). We selected this classifier mainly due to its ability to handle missing data, which is a strength over the widely-used C4.5 algorithm (Farhangfar et al., 2008; Quinlan, 1993). In similarly, other classifiers commonly used in remote sensing applications, such as SVM (Cortes and Vapnik, 1995) or Random Forest (Breiman, 2001), require the data to be complete and all gaps filled. We built each model using the C5.0 default settings (Kuhn et al., 2021), and ran 100 boosting trials, the maximum in C5.0, because higher numbers improved the classification accuracy in initial tests (results not shown). The actual number of trials varied between the models depending on the amount of data and the level of data gaps in the training data set. The classifier's training procedure stops once the predictive power of a model is either highly effective or if it is generally ineffective (Kuhn and Johnson, 2016). The boosting procedure relies on fitting models sequentially and adjusting the case weights based on the sample's prediction accuracy. The output predicted class is assigned based on the highest average confidence value from all boosting trials.

2.4. Test 1: Missing data due to cloud cover and SLC gaps

To evaluate the influence of data gaps on classification accuracy, we ran a series of classifications with imagery selected based on various thresholds for the minimum and maximum percentage of pixels affected by missing data (Fig. 3a-b). For example, the thresholds ≥ 50% and < 100% restricted the data set to 17 images for the year 2016 and 51

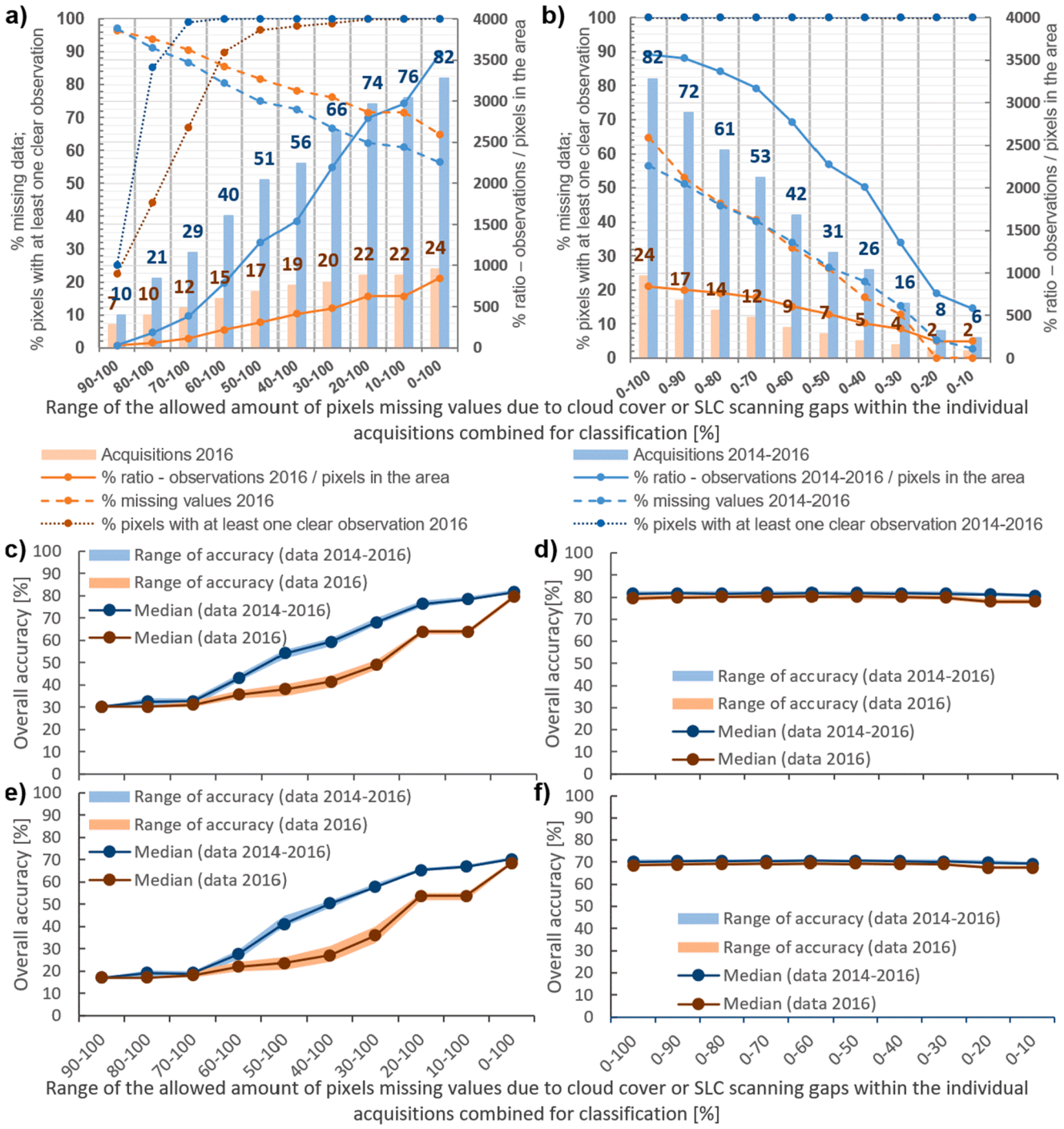


Fig. 3. The data availability and classification accuracy for tests of mapping forest types with image stacks containing imagery from 2016 and 2014–2016 characterized by data gaps of various extents: a) data availability when imagery containing fewer gaps is added, b) data availability when imagery with more gaps is removed, c) classification accuracy corresponding to data sets from chart a for pure forest stands only, d) classification accuracy corresponding to data sets from chart b for pure forest stands only, e) classification accuracy corresponding to data sets from chart a for both pure and mixed stands, f) classification accuracy corresponding to data sets from chart b for both pure and mixed stands.

images for the period 2014–2016. Spatially, each image in this subset was missing values for 50% or more of the pixels belonging to study area. The overall level of missing data for our study area in such image stacks was equal to 81% and 75% respectively.

For each set of thresholds, we recorded: 1) the overall accuracy, 2) the number of images meeting the thresholds, 3) the percentage of pixels in the study area with at least one clear observation, 4) the percent ratio between the number of available observations and the number of the pixels in the study area (e.g., 100% means that on average there was one clear value for each pixel of the study area, but some pixels may have had multiple observations, and others none), and 5) the overall percentage of missing data within the image stack. In total, we tested 19 thresholds separated by 10% steps. We started with the most incomplete imagery, where the data gaps affect $\geq$ 90% and $<$ 100% of the pixels (Fig. 3a). Next, we decreased the threshold for minimal gap cover from 90% to 0% in order to broaden the range of thresholds which resulted in addition of images less affected by the gaps. Once the threshold reached 0% all acquisitions were included, containing from 0% to $<$ 100% pixels

with missing values in any date. Finally, we narrowed the range by decreasing the threshold for maximal gap cover from 100% to 10% that resulted in removal of the most incomplete imagery from the analysis (Fig. 3b). We finished with the imagery with the least amount of missing data (0% to < 10% of the pixels in any acquisition), which included images only from 2016.

2.5. Test 2: Number of image acquisitions

To test the effects of number of images on classification accuracy, we extended the image acquisition time to three years (2014 to 2016), and repeated the tests with varying levels of data gaps as described earlier. For each threshold, we assessed the improvement in classification accuracy when analyzing three years of data, and compared the differences in 1) the number of image acquisitions meeting current thresholds for data gaps, 2) the percent of pixels in the study area for which the data provided at least one clear observation, 3) the percent ratio between the number of clear observations and the number of the pixels in the study area, and 4) the overall percentage of missing data within the image stack.

2.6. Test 3: Seasonality

To determine the seasons most important for the forest types classifications, we performed 15 tests of all possible combinations of data from winter (December, January, February), spring (March, April, May), summer (June, July, August), and fall (September, October, November). We assigned the seasons to months roughly corresponding to major phenology processes (e.g., leaf-out, fall senescence) in our study area (Curtis, 1959), and made the time-windows wide to ensure that they captured these processes even in years with relatively early or relatively late phenology. We did so to ensure that, for example, leaf-out did not occur in some years in our spring, and in other years in our summer. We tested two data sets, one containing all images from 2016 and the other containing those from the period between 2014 and 2016. We recorded: 1) the overall accuracy, 2) the number of image acquisitions meeting each of the data gaps thresholds, 3) the percent ratio between the number of gap-free observations and the number of the pixels in the study area, 4) the percent ratio between the total number of the pixels and the number of gap-free observations available, and 5) the overall percentage of missing data within the image stacks.

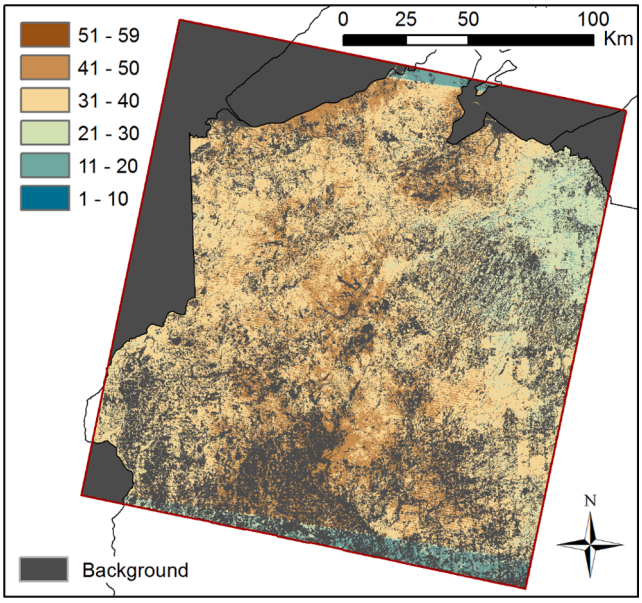


Fig. 4. Number of clear per pixel observations from 2014 to 2016.

2.7. Accuracy assessment, classification uncertainty

We calculated the overall accuracy of all classifications for all iterations of our tests and presented the detailed class-level accuracies for the classification generated with the model based on the best selection of image acquisitions. We aggregated pixel-level classification results to forest stands, labeling each stand based on the most frequently occurring category. However, to account for the area of each class we counted all of the pixels belonging to the test stands instead of operating with the stand counts for the accuracy assessments. We calculated: 1) the overall classification accuracy, 2) user's, and 3) producer's accuracy (Congalton, 1991). We derived graphs and error matrices to present the minimum, median, and maximum value of overall accuracy for each data set related to cloud cover and seasonal distribution of the imagery based on the full extents of forest stands. We presented the results for both: all stands and pure stands only. Finally, to estimate class-level accuracies at the forest-stand level, we calculated the accuracy for our best classification by using all 2,786,133 pixels for the 44,741 forest stands remaining for the accuracy assessment. We calculated accuracy separately for pure and mixed stands as well as for all stands together.

As addition to the standard accuracy measures for the final map, we calculated a measure of uncertainty across the study area. We defined it at the pixel level as 100% minus the probability of the C5.0 pixel's classification assignment. We compared its values for individual forest types in box and whisker plots and presented in form of a map depicting its spatial distribution.

As a robustness check, we analyzed how the time since the last field examination of our reference stands affected the results of our classification. To do so, we performed logistic regression of whether our classification at the stand level was correct or not (1 versus 0) as a function of the number of years from the last field visit for pure forest stands.

3. Results

We were able to classify forest types with high accuracy even when all the input data had large gaps. Our classification accuracy at the stand level reached 82.8% for pure stands and 71.4% for all stands when analyzing all available data from 2014 to 2016. The accuracies are slightly lower (81.4% for pure stands, and 70.1% for all stands) when using data for 2016 alone (Fig. 3c-d).

3.1. Test 1 - missing data due to cloud cover and SLC gaps

Several results emerged when we tested the added value of including imagery with gaps in the classifications. First, when gap-free images were available, the addition of imagery with gaps improved classification accuracies only marginally (Fig. 3d). This was particularly important when analyzing data for 2016 only, for which only two relatively gap-free images were available (Fig. 3b). Furthermore, once images with the least amount of missing data were included as input, there was only a minor improvement in classification accuracy when using data from 2014 to 2016 versus 2016 only.

When gap-free images were excluded (i.e., when simulating a situation where no gap-free imagery is available), having data for three years was advantageous. With three years of observations, classification accuracy remained high even when the best images had gaps $\geq$ 30% of the study area (Fig. 3c), i.e. when the available eight best images were removed from the image stack (Fig. 3a). However, classification accuracies for observations from 2016 only dropped precipitously as soon as the imagery with the smallest gaps were removed (Fig. 3c).

The patterns for the overall level of missing data were similar for both datasets, that is the data for a single year, and for three years, and was not clearly related with the patterns of the overall accuracy of classification. The overall level of missing data in the image stacks decreased gradually when we added imagery with fewer gaps from 97% to 64% for 2016 and from 98% to 58% for 2014–2016 (Fig. 3a). When

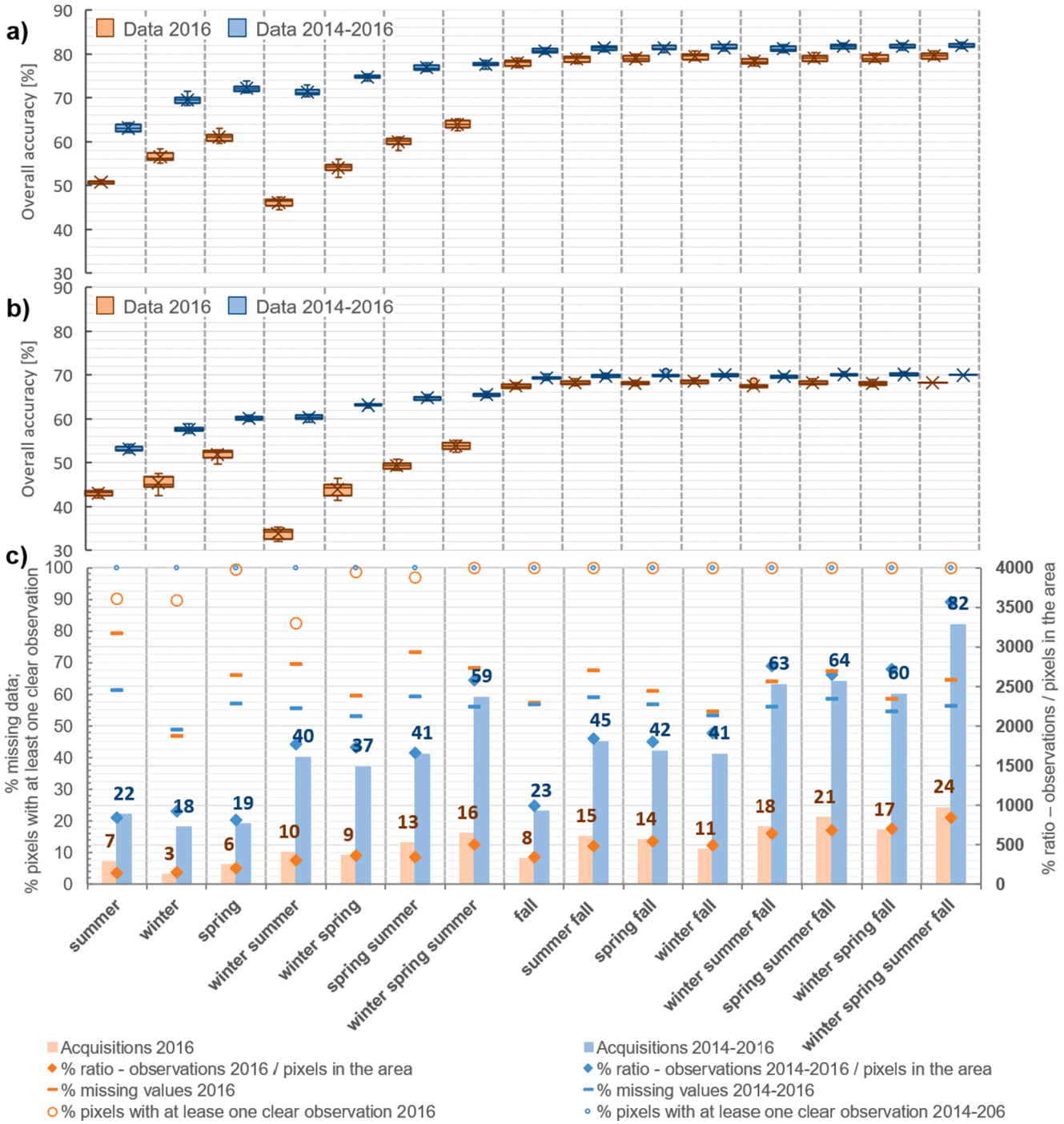


Fig. 5. The classification accuracy (a – pure forest stands, b – pure and mixed stands) and data availability (c) for tests of mapping forest types with image stacks containing imagery from various seasonal combinations from 2016 and 2014–2016.

doing so, there was a gradual increase in classification accuracy (Fig. 3c). The decrease in the level of missing data continued when we removed imagery with gaps from the stacks from 64% to 0% and from 58% to 2% for 2016 and 2014–2016 respectively (Fig. 3b). However, this did not result in any improvements in the classification accuracy (Fig. 3d).

When comparing classifications for pure stands (Fig. 3c-d) with those for all stands (Fig. 3e-f), accuracies for pure stands alone were generally 10 percentage points higher than those for all stands, but the overall patterns when including imagery with gaps were very similar.

3.2. Test 2 – Number of acquisitions

Analyzing data for three years instead of only one increased image availability from 24 to 84 acquisitions with some gap-free pixels (Fig. 3a-b). However, the increase in the number of acquisitions varied depending on the gap cover range and shrank to 12 and 4 for 0–30% and 0–10% thresholds, respectively.

Data from three years provided at least one clear observation for each pixel even with imagery where the minimal gap cover was 60% of the area. In contrast, for data from a single year, the minimal gap cover of 20% was necessary to provide data for the entire study area. Second, the maximum number of observations per pixel was 59 considering all

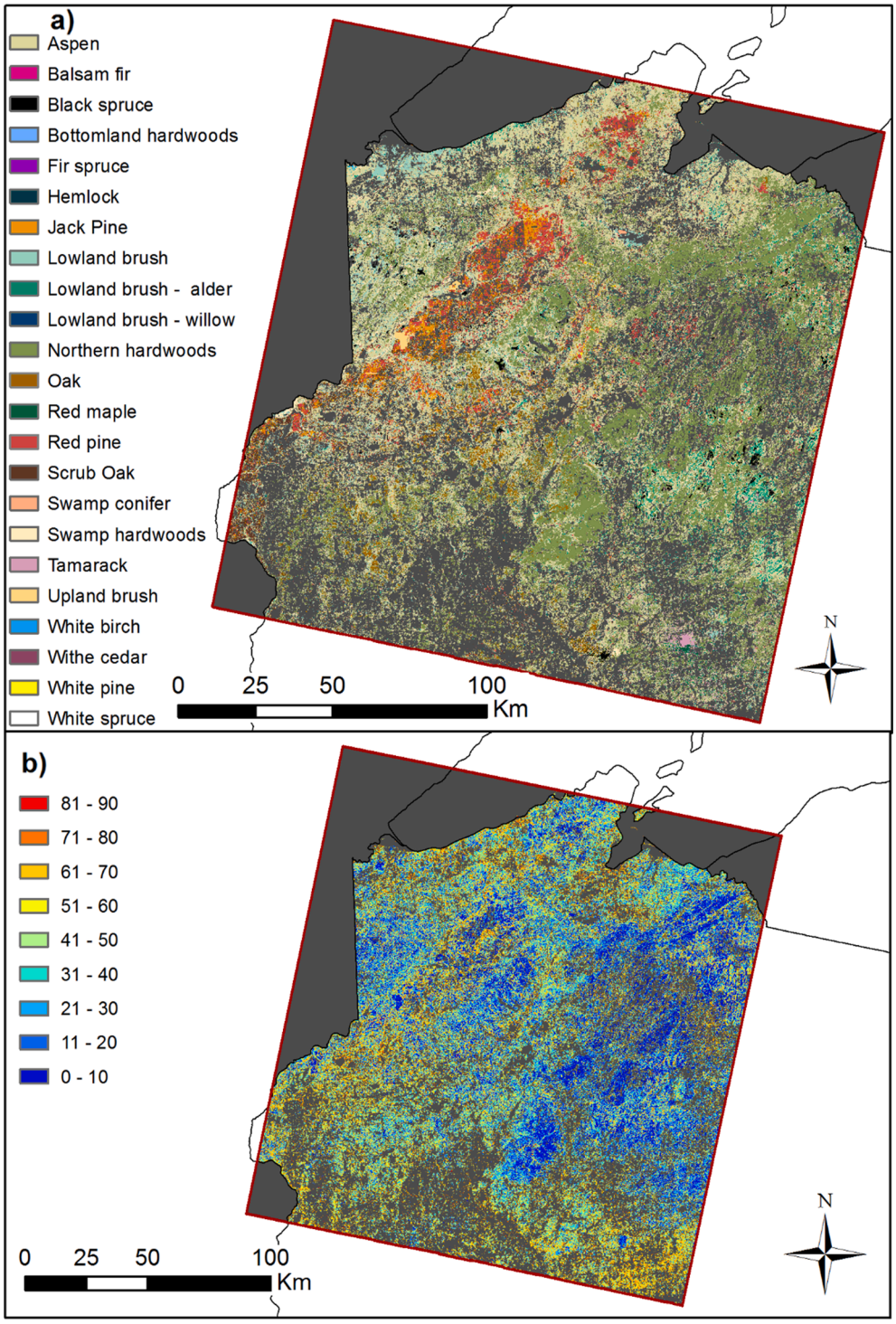


Fig. 6. Tree species distribution: a) the final map, and b) associated classification uncertainty (100% minus C5.0 classification assignment probability).

acquisitions from three-year period (Fig. 4). The average number of clear observations per pixel grew rapidly from < 1 for the most incomplete imagery (10 acquisitions missing $\geq 90\%$ and $< 100\%$ pixel's values) to 35 for all available imagery (0–100%) from years 2014–2016 and to 9 for 2016 only.

The increase in classification accuracy associated with additional data acquisitions was marginal when the least gap-affected imagery was included (Fig. 3d). Even with just four images with 0–30% missing data for 2016 alone, classification accuracy already reached 80.8%, compared to 82.2% for the 82 acquisitions for three-year set, and 81.9% for the 53 acquisitions for the data with 0–60% missing values, which was the highest median accuracy in our tests. However, without the gap-free imagery, differences between the single and three year sets were stark, especially in the range of 10–100% to 40–100% of missing values, where the three-year data sets with many more observations increased classification accuracy up to 79.4% (Fig. 3c). Results for pure and mixed stands were similar in trends but lower in absolute values (Fig. 3e-f).

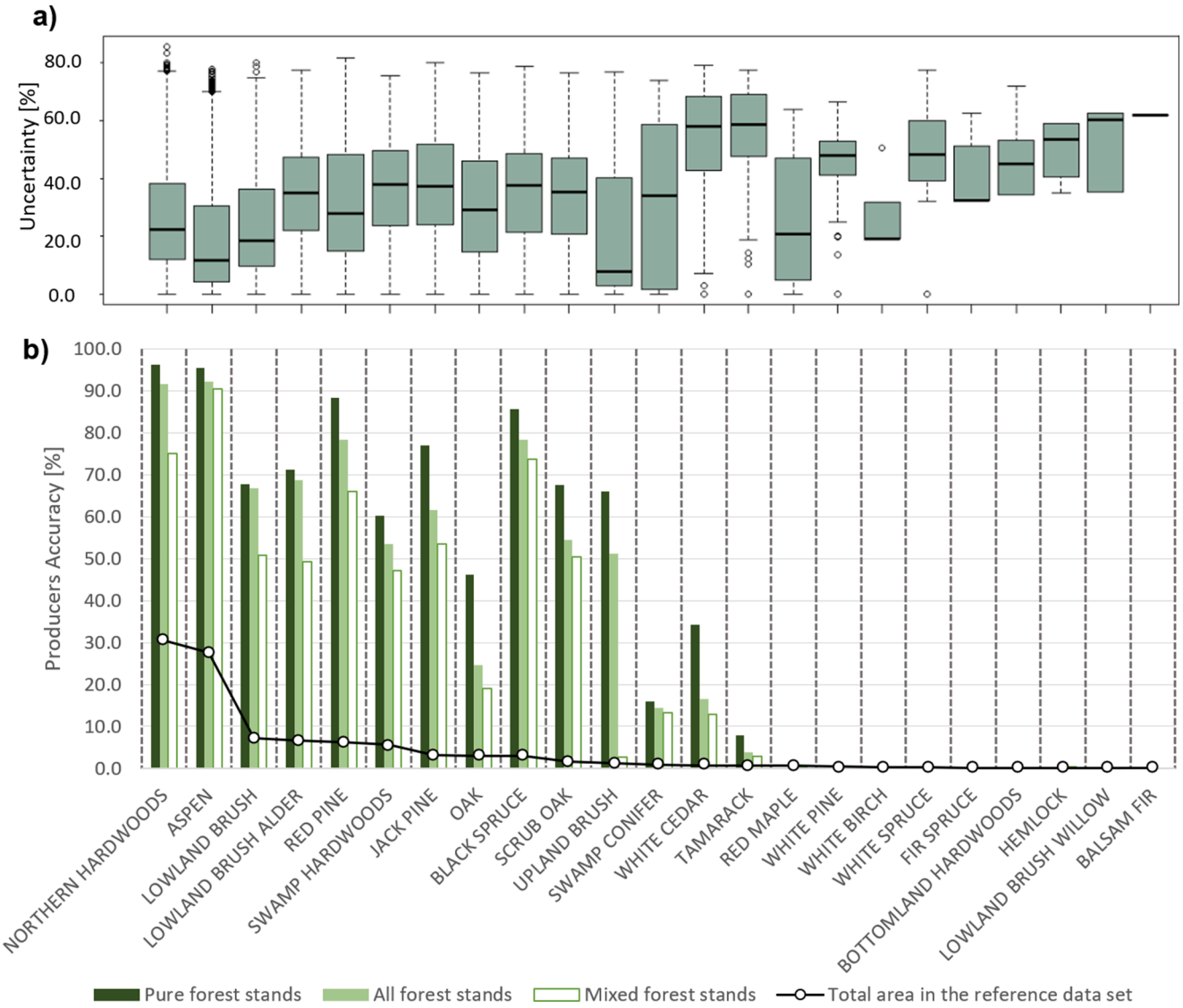


Fig. 7. The accuracy assessment: a) the uncertainty for tree types within the reference data set, the order of the boxes depicts the number of the pixels belonging to each class; b) the agreement between classification and the RECON data set - producer's accuracy for forest types for pure, mixed and all stands.

3.3. Test 3 – Seasonal data distribution

The season during which the images were recorded affected classification accuracy (Fig. 5a-b). Fall imagery was, by far, the most valuable for mapping forest types, providing by itself a level of accuracy slightly lower than combination with data from other seasons. For pure stands, we achieved the highest accuracy using combinations of all seasons when using the imagery from both the three-year period (82.6%) and the single year (80.6%). In other words, as long as a fall image was available, data from a single year performed almost as well as observations from three years.

When we simulated situations where no fall images were available, then the difference in the classifications between the single- and three-year sets was pronounced (Fig. 5a-b). This difference was due to the larger number of image acquisitions for the three-year period and its lower levels of missing data (Fig. 5c). When we tested the value of imagery for different seasons for pure versus all stands, the latter had again a lower accuracy. We observed the largest difference in the median overall accuracy (ca. 25 percentage points) when using a combination of imagery from winter and summer (Fig. 5a). The most useful imagery was again from the fall, and as long as fall imagery was included, the 2014–2016 dataset resulted in only marginally better classifications (71.1%) than the 2016-only data (69.4%).

3.4. Final map - accuracy assessment

The overall accuracy at the stand level for the final map (Fig. 6a) was 82.8% for pure stands (84.5% when rare forest types are aggregated for the purpose of presentation), 60.7% for mixed stands, and 70.5% for all stands. User's accuracy for individual forest types generally decreased with decreasing species acreage, but producer's accuracy showed no visible patterns. However, for most forest types the producer's accuracy was noticeably higher for pure stands than for mixed stands, with differences ranging from 10 to 30 percentage points for most of the types (Fig. 7b).

For pure stands of forest types that covered ≥ 2% of the area (Table 2), we obtained generally high user's accuracy: oak (93.7%), lowland brush alder (76.1%), red pine (91.2%), northern hardwoods (87.8%), scrub oak (85.5%), swamp hardwoods (85.4%), jack pine (86.0%), aspen (77.8%), and black spruce (75.0). With exception of jack pine (76.9%), scrub oak (67.6%), swamp hardwoods (60.3%), and oak (46.2%), we also achieved high producer's accuracy for these classes: northern hardwoods (96.2%), Aspen (95.5%), red pine (88.3%), black spruce (85.6%), jack pine (85.6%), lowland brush alder (76.9%). For the

Table 2
The error matrix for the classification at the stand level weighted by the number of pixels belonging to the pure forest stands in the reference RECON validation set; the classification was generated with Landsat imagery acquired in years 2014–2016 characterized by data gaps covering from 0% up to 70% of the study area.

	ASPEN	BLACK SPRUCE	JACK PINE	NORTHERN HARDWOODS	OAK	RED PINE	SCRUB OAK	SWAMP HARDWOODS	BRUSH	OTHER	*R.sum*	*User's accuracy*
ASPEN	**621,820**	4984	6071	19,568	19,840	7454	10,724	20,632	54,642	33,291	799,026	**77.8**
BLACK SPRUCE	504	**58,666**	39	4334	–	28	–	35	3438	11,170	78,214	**75.0**
JACK PINE	367	–	**57,699**	–	17	7665	980	–	149	199	67,076	**86.0**
NORTHERN HARDWOODS	17,846	1019	–	**695,185**	16,631	761	267	23,650	8062	28,080	791,501	**87.8**
OAK	672	–	158	1188	**33,044**	–	49	48	64	24	35,247	**93.7**
RED PINE	707	119	8311	51	–	**131,946**	222	–	349	2920	144,625	**91.2**
SCRUB OAK	1052	–	1524	–	287	832	**27,484**	124	734	123	32,160	**85.5**
SWAMP HARDWOODS	851	376	–	1404	1665	–	19	**79,539**	4388	4859	93,101	**85.4**
BRUSH	2582	2350	1197	275	–	810	931	7188	**281,056**	6917	**303,306**	**92.7**
OTHER	4975	1055	–	581	–	–	–	664	1615	**13,301**	22,191	**59.9**
C.sum	651,376	68,569	74,999	722,586	71,484	149,496	40,676	131,880	354,497	100,884	***Overall***	**84.5**
Producer's accuracy	**95.5**	**85.6**	**76.9**	**96.2**	**46.2**	**88.3**	**67.6**	**60.3**	**79.3**	**13.2**	***Accuracy***	

Table 3
Logistic regression table between correctly classified stands and numbers of years from last field examination for pure oak forest stands.

Model	Coefficients: Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−0.087	0.116	−0.748	0.454
years from last examination	−0.009	0.011	−0.826	0.409
Null deviance: 737.28 on 533 degrees of freedom				
Residual deviance: 736.59 on 532 degrees of freedom				
AIC: 740.59				

Table 4
Logistic regression table between correctly classified stands and numbers of years from last field examination for pure north hardwoods forest stands.

Model	Coefficients: Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.852	0.164	17.419	< 2e-16
years from last examination	−0.044	0.012	−3.604	0.0003
Null deviance: 532.77 on 999 degrees of freedom				
Residual deviance: 521.65 on 998 degrees of freedom				
AIC: 525.65				

classes covering < 2% of the area, accuracies were generally lower (Appendix A), missing the rare tree species entirely or having producer's accuracy at maximum level of 24.4%. However, we mapped most of these classes with high user's accuracy of ≥ 80%.

While we did not have independent ground truth data to check the overall accuracy of the map, we confirmed that the prediction probabilities were moderately high (Fig. 6b). For most of the forest stands, the uncertainty was lower than 60%, and for a large portion < 10%. Prediction probabilities varied among classes, and more abundant tree species were generally classified with greater probability (Fig. 7a). For most classes the maximum level of uncertainty ranged from 0% to 80% and median was lower than 40%.

The relation between the time of last field examination of the reference forest stands and classification results varied among forest type. For example, we found no significant relationship between correct classification of oak stands and the time when they were last examined (Table 3). In contrast, for northern hardwoods the time of field examination had significant negative effect on correct stand classification (Table 4). In both oak and northern hardwoods cases, the majority of the stands were visited within 15 years from the date of acquisition of our satellite imagery (Fig. 8).

4. Discussion

We were able to map forest types accurately with Landsat data even when observations had substantial gaps by applying the C5.0 decision trees algorithm, which can handle missing data. Our use of the C5.0 algorithm for tree species mapping represents an improvement over previously used approaches such as image thresholding (Dymond et al., 2002; Wolter et al., 2008, 1995), traditional classification algorithms (e.g. Minimum Distance To Means Classifier (Mickelson et al., 1998)), and machine learning algorithms requiring the data to be gap-free such as Random Forests (Zhu and Liu, 2014). Furthermore, removing the need for gap filling makes our approach for extracting information on forest types more efficient for large areas, especially in locations where a low number of observations can degrade the performance of gap filling algorithms (Zhu et al., 2015). In general, we found that: 1) forest types in temperate forests can be classified with accuracy up to 84.5%, 2) imagery from multiple years is crucial when gap-free imagery is lacking, and 3) data from fall is by far the most important for high classification accuracies.

When gap-free observations are not available, we recommend mapping forest types based on data from multiple years to increase the

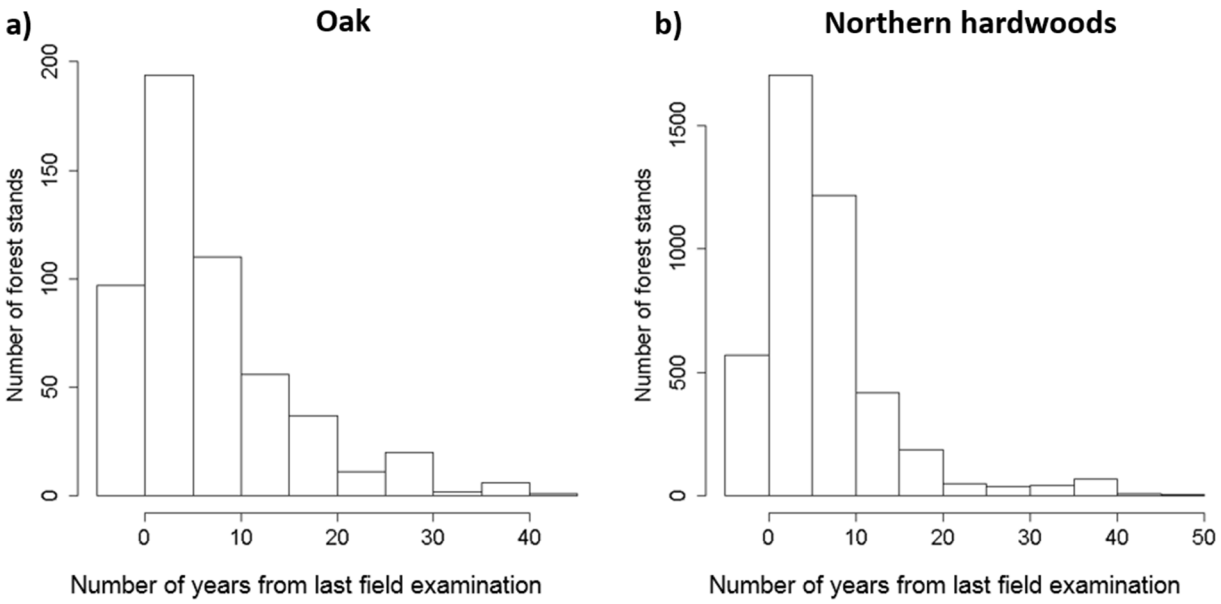


Fig. 8. Number of years from last field examination for pure reference forest stands of a) oak and b) northern hardwoods forest types.

number of acquisitions, and to ensure that each pixel has clear observations from different stages of phenology. However, when multiple gap-free images are available for a single year, the results are expected to be satisfactory, as long as at least one image is from the fall. Similar findings were reported for data from WorldView-2 where usage of only 7 out of 20 gap free images resulted in the maximal classification accuracy (Elatawneh et al., 2013). Similarly, Pasquarella et al. (2018) obtained up to 5% better accuracy using spectral and temporal indices derived from all available observations from 1985 to 2015, based on pairs of gap-free imagery from multiple years. However, the largest improvement in that study was due to the addition of ancillary data, such as topography and wetland probability, which were not tested here.

The data from all seasons used together provided the best classification accuracy, but the fall imagery was by far the most important. For both, fall imagery from single- and three-year periods by itself resulted in a higher accuracy than imagery from spring, summer, and winter combined. In general, the season with more acquisitions and lower percentage of missing values improved classification accuracy the most. Presumably, the larger number of image acquisitions and lower levels of missing data allowed to capture the differences in the timing of tree species' phenology occurring within seasons. Our findings are generally in line with previous research identifying spring and fall imagery as the most useful for discriminating forest types (Mickelson et al., 1998; Wolter et al., 1995), although the summer imagery has been found to be useful if it is of high quality (Elatawneh et al., 2013).

The high classification accuracy for pure forest stands provides strong support for using the C5.0 algorithm (Kuhn et al., 2021) with incomplete Landsat data, especially considering the large extent of our reference data set. However, the RECON data set provides information at the stand level, which does not perfectly match the scale of the remotely sensed observations. Depending on the random subset of the reference data set in our trials, we used somewhere from 2,784,346 to 2,814,616 pixels to validate classification results within the reference forest stands. Validation datasets in prior studies were generally smaller, e.g., 370 (Bolstad and Lillesand, 1992), 1,211 (Wolter et al., 1995), 322 (Mickelson et al., 1998), 529 (Townsend and Walsh, 2001), 528 (Dymond et al., 2002), 95 (Zhu and Liu, 2014), 2181 (Karasiak et al., 2017), 1587 (Hościło and Lewandowska, 2019), 1427 (Grabska et al., 2020), but Pasquarella et al., (2018) had 161,880 samples. Because we did not have access to detailed information about tree species composition within each stand, we calculated classification accuracy based on the dominant tree species in each stand (Bolstad and Lillesand, 1992; Dymond et al., 2002; Wolter et al., 1995; Zhu and Liu, 2014), as is typical for forest inventory maps. An alternative would be to use fuzzy logic to depict the actual species composition (Townsend and Walsh, 2001; Mickelson et al., 1998), but we did not pursue this. It is worth noting that the RECON data set is highly unbalanced in terms of classes, reflecting species distributions on the ground, which leads to better recognition of more abundant forest types in classification. This tendency of classifiers to bias towards the dominant classes is common (Kuhn and Johnson, 2016) and has been reported in other studies involving tree species mapping (Hemmerling et al., 2021). In tests we were able to improve the accuracy for less abundant forest types by leveling the number of training samples per class by sampling even number of training cases with replacement for each class (results not shown), We recommend doing so when the goal is to map rare forest types accurately. Lastly, our results may be also affected by the date of the RECON data. Depending on the forest type, stands that had not been recently examined in the field were less likely to be correctly classified.

Our results have implications for remote sensing, forest industry, and conservation planning. We tested a promising methodology for large scale, detailed, and accurate mapping of temperate forest types. That methodology makes it possible to produce reliable forest composition maps because it does not require completely gap free images, which are rare (Ju and Roy, 2008; Wulder et al., 2016) nor does it depend on gap filling. In general, this approach should also be applicable to other products derived from Landsat-like imagery e.g., data stacks containing seasonal composites or temporal statistics that lack information in 30% or more of the pixels in any individual image but provide coverage of the data to the entire mapped area. Second, our maps are at 30-m resolution, thereby providing a detailed input for modeling of environmental variables that differ among tree species (Richardson et al., 2012). Third, our maps could be used to update stand-level forest inventories and hence support forest management and conservation efforts (Loiselle et al., 2003).

We conclude that Landsat imagery with missing observations because of clouds and SLC-Off gaps can be used for operational mapping of the temperate forest types. We show that just a handful of fairly complete and cloud-free acquisitions, especially from fall, combined with imagery with various levels of missing values, can result in accurate forest types maps. This is an important improvement over prior approaches that either require gap-free data or a large number of annual

per pixel observations for gap filling.

Funding

The study was supported by the USDA McIntire-Stennis Program grant (WI03072).

CRediT authorship contribution statement

Contributors: Konrad Turlej: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft. **Mutlu Ozdogan:** Funding acquisition, Writing – review & editing. **Volker C. Radeloff:** Conceptualization, Methodology, Supervision, Writing – review & editing, Funding acquisition.

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.

Acknowledgments

The authors thank: E. Kruger, A. Schneider, and P. Townsend from University of Wisconsin-Madison for ideas and constructive feedback at various stages of work on this manuscript; the editor and reviewers for great feedback on the manuscript; and S.P. Engle from The CALS Statistical Consulting Group at the University of Wisconsin-Madison for valuable discussions and feedback.

Appendix A. The full error matrix for the classification at the stand level weighted by the number of pixels belonging to the pure forest stands in the reference RECON validation set; the classification was generated with Landsat imagery acquired in years 2014–2016 characterized by data gaps covering from 0% up to 70% of the study area.

	ASPEN	BALSAM FIR	BLACK SPRUCE	BOTTOMLAND HARDWOODS	FIR SPRUCE	HEMLOCK	JACK PINE	LOWLAND BRUSH	LOWLAND BRUSH ALDER
ASPEN	**621,820**	1231	4984	2236	1600	1151	6071	27,559	22,400
BALSAM FIR	–	**–**	–	–	–	–	–	–	–
BLACK SPRUCE	504	219	**58,666**	–	426	8	39	576	2862
BOTTOMLAND HARDWOODS	–	–	–	**–**	–	–	–	–	–
FIR SPRUCE	–	–	–	–	**–**	–	–	–	–
HEMLOCK	–	–	–	–	–	**–**	–	–	–
JACK PINE	367	45	–	–	–	–	**57,699**	31	–
LOWLAND BRUSH	976	57	827	71	–	–	9	**114,332**	12,449
LOWLAND BRUSH ALDER	1606	47	1523	–	164	–	289	21,535	**110,943**
LOWLAND BRUSH WILLOW	–	–	–	–	–	–	–	–	–
NORTHERN HARDWOODS	17,846	106	1019	65	1265	1317	–	1717	4429
OAK	672	–	–	–	–	–	158	51	–
RED MAPLE	–	–	–	–	–	–	–	–	–
RED PINE	707	–	119	–	245	–	8311	39	70
SCRUB OAK	1052	–	–	–	–	–	1524	–	–
SWAMP CONIFER	15	–	72	–	25	–	–	98	894
SWAMP HARDWOODS	851	–	376	127	151	16	–	2597	1481
TAMARACK	4943	–	897	–	–	–	–	73	156
UPLAND BRUSH	–	–	–	–	–	–	899	–	–
WHITE BIRCH	–	–	–	–	–	–	–	–	–
WHITE CEDAR	–	–	86	–	46	–	–	64	101
WHITE PINE	17	–	–	–	–	–	–	–	–
WHITE SPRUCE	–	–	–	–	–	–	–	–	–
C.sum	651,376	1705	68,569	2499	3922	2492	74,999	168,672	155,785
Producer's accuracy	**95.5**	**–**	**85.6**	**–**	**–**	**–**	**76.9**	**67.8**	**71.2**

	LOWLAND BRUSH WILLOW	NORTHERN HARDWOODS	OAK	RED MAPLE	RED PINE	SCRUB OAK	SWAMP CONIFER	SWAMP HARDWOODS	TAMARACK
ASPEN	225	19,568	19,840	3329	7454	10,724	2632	20,632	5292
BALSAM FIR	–	–	–	–	–	–	–	–	–
BLACK SPRUCE	–	4334	–	–	28	–	3298	35	5839
BOTTOMLAND HARDWOODS	–	–	–	–	–	–	–	–	–
FIR SPRUCE	–	–	–	–	–	–	–	–	–
HEMLOCK	–	–	–	–	–	–	–	–	–
JACK PINE	37	–	17	–	7665	980	–	–	–
LOWLAND BRUSH	1479	102	–	–	97	922	328	2747	1658
LOWLAND BRUSH ALDER	–	173	–	137	32	–	1490	4441	1857
LOWLAND BRUSH WILLOW	**–**	–	–	–	–	–	–	–	–
NORTHERN HARDWOODS	–	**695,185**	16,631	12,120	761	267	6632	23,650	549
OAK	–	1188	**33,044**	–	–	49	–	48	14
RED MAPLE	–	200	–	**–**	–	–	–	–	–

(*continued on next page*)

(*continued*)

	LOWLAND BRUSH WILLOW	NORTHERN HARDWOODS	OAK	RED MAPLE	RED PINE	SCRUB OAK	SWAMP CONIFER	SWAMP HARDWOODS	TAMARACK
RED PINE	–	51	–	–	**131,946**	222	–	–	38
SCRUB OAK	–	–	287	–	832	**27,484**	–	124	–
SWAMP CONIFER	–	151	–	–	–	–	**3146**	664	13
SWAMP HARDWOODS	–	1404	1665	–	–	19	1713	**79,539**	318
TAMARACK	–	230	–	–	–	–	–	–	**1344**
UPLAND BRUSH	–	–	–	–	681	9	–	–	–
WHITE BIRCH	–	–	–	–	–	–	–	–	–
WHITE CEDAR	–	–	–	–	–	–	568	–	–
WHITE PINE	–	–	–	–	–	–	–	–	–
WHITE SPRUCE	–	–	–	–	–	–	–	–	–
C.sum	1741	722,586	71,484	15,586	149,496	40,676	19,807	131,880	16,922
Producer's accuracy	**–**	**96.2**	**46.2**	**–**	**88.3**	**67.6**	**15.9**	**60.3**	**7.9**

	UPLAND BRUSH	WHITE BIRCH	WHITE CEDAR	WHITE PINE	WHITE SPRUCE	***R.sum***	***User's accuracy***
ASPEN	4458	4285	2208	6523	2804	799,026	**77.8**
BALSAM FIR	–	–	–	–	–	0	**–**
BLACK SPRUCE	–	13	773	16	578	78,214	**75.0**
BOTTOMLAND HARDWOODS	–	–	–	–	–	0	**–**
FIR SPRUCE	–	–	–	–	–	0	**–**
HEMLOCK	–	–	–	–	–	0	**–**
JACK PINE	81	–	–	92	62	67,076	**86.0**
LOWLAND BRUSH	1143	51	54	13	19	137,334	**83.3**
LOWLAND BRUSH ALDER	486	17	794	–	160	145,694	**76.1**
LOWLAND BRUSH WILLOW	–	–	–	–	–	0	**–**
NORTHERN HARDWOODS	1916	1643	3871	131	381	791,501	**87.8**
OAK	13	–	–	10	–	35,247	**93.7**
RED MAPLE	–	–	–	–	–	200	**0.0**
RED PINE	240	–	–	2386	251	144,625	**91.2**
SCRUB OAK	734	123	–	–	–	32,160	**85.5**
SWAMP CONIFER	–	–	1606	–	–	6684	**47.1**
SWAMP HARDWOODS	310	–	2534	–	–	93,101	**85.4**
TAMARACK	229	–	–	–	185	8057	**16.7**
UPLAND BRUSH	**18,689**	–	–	–	–	20,278	**92.2**
WHITE BIRCH	–	**–**	–	–	–	0	**–**
WHITE CEDAR	–	–	**6143**	–	–	7008	**87.7**
WHITE PINE	–	–	–	**225**	–	242	**93.0**
WHITE SPRUCE	–	–	–	–	**–**	0	**–**
C.sum	28,299	6132	17,983	9396	4440	***Overall Accuracy***	***82.8***
Producer's accuracy	**66.0**	**–**	**34.2**	**–**	**–**		

Appendix B. Tree species composition included in the forest types mapped in this study

Forest type	Tree species
ASPEN	aspen (*Populus* spp.)
BALSAM FIR	balsam fir (*Abies balsamea*)
BLACK SPRUCE	black spruce (*Picea mariana*)
BOTTOMLAND HARDWOODS	green ash (*Fraxinus pennsylvanica*), silver maple (*Acer saccharinum*), swamp white oak (*Quercus bicolor*), river birch (*Betula nigra*), cottonwood (*Populus deltoides*), American elm (*Ulmus americana*)
FIR SPRUCE	fir (*Abies* spp.), spruce (*Picea* spp.)
HEMLOCK	hemlock (*Tsuga canadensis*)
JACK PINE	jack pine (*Pinus banksiana*)
LOWLAND BRUSH	alder (*Alnus* spp.), willow (*Salix* spp.), bog birch (*Betula pumila*)
LOWLAND BRUSH ALDER	alder (*Alnus* spp.)
LOWLAND BRUSH WILLOW	willow (*Salix* spp.)
NORTHERN HARDWOODS	sugar maple (*Acer saccharum*), beech (*Fagus grandifolia*), basswood (*Tilia americana*), white ash (*Fraxinus americana*), yellow birch (*Betula alleghaniensis*)
OAK	oak (*Quercus* spp.)
RED MAPLE	red maple (*Acer rubrum*)
RED PINE	red pine (*Pinus resinosa*)
SCRUB OAK	black oak (*Quercus velutina*), white oak (*Quercus alba*), northern pin oak (*Quercus ellipsoidalis*), bur oak (*Quercus macrocarpa*)
SWAMP CONIFER	white cedar (*Thuja occidentalis*), black spruce (*Picea mariana*), tamarack (*Larix laricina*), balsam fir (*Abies balsamea*)
SWAMP HARDWOODS	black ash (*Fraxinus nigra*), green ash (*Fraxinus pennsylvanica*), American elm (*Ulmus americana*), red maple (*Acer rubrum*), silver maple (*Acer saccharinum*), swamp white oak (*Quercus bicolor*)
TAMARACK	tamarack (*Larix laricina*)
UPLAND BRUSH	hazel (*Corylus* spp.), dogwood (*Cornus* spp.), juneberry (*Amelanchier* spp.), sumac (*Rhus* spp.), ninebark (*Physocarpus* spp.), prickly ash (*Aralia spinosa*)
WHITE BIRCH	white birch (*Betula papyrifera*)
WHITE CEDAR	white cedar (*Thuja occidentalis*)
WHITE PINE	white pine (*Pinus strobus*)
WHITE SPRUCE	white spruce (*Picea glauca*)

References

Axelsson, A., Lindberg, E., Reese, H., Olsson, H., 2021. Tree species classification using Sentinel-2 imagery and Bayesian inference. Int. J. Appl. Earth Obs. Geoinf. 100, 102318. https://doi.org/10.1016/j.jag.2021.102318.

Barbier, S., Gosselin, F., Balandier, P., 2008. Influence of tree species on understory vegetation diversity and mechanisms involved-A critical review for temperate and boreal forests. For. Ecol. Manage. 254 (1), 1–15. https://doi.org/10.1016/j.foreco.2007.09.038.

Bergh, J., Freeman, M., Sigurdsson, B., Kellomäki, S., Laitinen, K., Niinistö, S., Peltola, H., Linder, S., 2003. Modelling the short-term effects of climate change on the productivity of selected tree species in Nordic countries. For. Ecol. Manage. 183 (1-3), 327–340. https://doi.org/10.1016/S0378-1127(03)00117-8.

Bolstad, P., Lillesand, T., 1992. Improved classification of forest vegetation in Northern Wisconsin through a rule-based combination of soils, terrains, and Landsat Thematic Mapper data. For. Sci. 38, 5–20.

Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science 320 (5882), 1444–1449. https://doi.org/10.1126/science.1155121.

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

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37 (1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. https://doi.org/10.1007/BF00994018.

Curtis, J.T., 1959. The Vegetation of Wisconsin. The University of Wisconsin Press, Madison, Wisconsin.

Department of the Interior U.S. Geological Survey, 2017. Product Guide: Landsat 4-7 Surface Reflectance (LEDAPS) Product.

Department of the Interior U.S. Geological Survey, 2020. Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) Product Guide Version 3.0. USGS.

Dymond, C.C., Mladenoff, D.J., Radeloff, V.C., 2002. Phenological differences in Tasseled Cap indices improve deciduous forest classification. Remote Sens. Environ. 80 (3), 460–472. https://doi.org/10.1016/S0034-4257(01)00324-8.

Elatawneh, A., Rappl, A., Rehush, N., Schneider, T., Knoke, T., 2013. Forest tree species communities identification using multi phenological stages RapidEye data : case study in the forest of Freising. Proceedings of the 5th RESA Workshop, Neustrelitz, Germany. 2013 23–28.

Farhangfar, A., Kurgan, L., Dy, J., 2008. Impact of imputation of missing values on classification error for discrete data. Pattern Recognit. 41 (12), 3692–3705. https://doi.org/10.1016/j.patcog.2008.05.019.

Fassnacht, F.E., Neumann, C., Forster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Koch, B., 2014. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three Central European test sites. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6), 2547–2561. https://doi.org/10.1109/JSTARS.2014.2329390.

Finzi, A.C., Van Breelmen, N., Canham, C.D., 1998. Canopy Tree – Soil Interactions Within Temperate Forests : Species Effects on Soil Carbon and Nitrogen 8, 440–446.

Garcia, M., Townsend, P.A., 2016. Recent climatological trends and potential influences on forest phenology around western Lake Superior, USA. J. Geophys. Res. Atmos. 121 (22), 13,364–13,391. https://doi.org/10.1002/2016JD025190.

Grabska, E., Frantz, D., Ostapowicz, K., 2020. Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. Remote Sens. Environ. 251, 112103. https://doi.org/10.1016/j.rse.2020.112103.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest stand species mapping using the sentinel-2 time series. Remote Sens. 11, 1–24. https://doi.org/10.3390/rs11101197.

Hanewinkel, M., Cullmann, D.A., Schelhaas, M.-J., Nabuurs, G.-J., Zimmermann, N.E., 2013. Climate change may cause severe loss in the economic value of European forest land. Nat. Clim. Chang. 3 (3), 203–207. https://doi.org/10.1038/nclimate1687.

Hemmerling, J., Pflugmacher, D., Hostert, P., 2021. Mapping temperate forest tree species using dense Sentinel-2 time series. Remote Sens. Environ. 267, 112743. https://doi.org/10.1016/j.rse.2021.112743.

Hobbie, S.E., Ogdahl, M., Chorover, J., Chadwick, O.A., Oleksyn, J., Zytkowiak, R., Reich, P.B., 2007. Tree species effects on soil organic matter dynamics: The role of soil cation composition. Ecosystems 10 (6), 999–1018. https://doi.org/10.1007/s10021-007-9073-4.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a Decade of Land Cover Change Information. Photogramm. Eng. Remote Sensing 81, 345–354. https://doi.org/10.1016/S0099-1112(15)30100-2.

Hościło, A., Lewandowska, A., 2019. Mapping forest type and tree species on a regional scale using multi-temporal sentinel-2 data. Remote Sens. 11 (8), 929. https://doi.org/10.3390/rs11080929.

Immitzer, M., Vuolo, F., Atzberger, C., 2016. First experience with sentinel-2 data for crop and tree species classifications in central europe. Remote Sens. 8, 166. https://doi.org/10.3390/rs8030166.

Ju, J., Roy, D.P., 2008. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. Remote Sens. Environ. 112 (3), 1196–1211. https://doi.org/10.1016/j.rse.2007.08.011.

Karasiak, N., Sheeren, D., Fauvel, M., Willm, J., Dejoux, J.F., Monteil, C., 2017. Mapping tree species of forests in southwest France using Sentinel-2 image time series. 2017 9th Int. Work. Anal. Multitemporal Remote Sens. Images, MultiTemp 2017 4–7. 10.1109/Multi-Temp.2017.8035215.

Kuhn, M., Weston, S., Culp, M., Coulter, N., Quinlan R., 2021. Package 'C50'.

Kuhn, M., Johnson, K., 2016. Applied Predictive Modeling, 5th printi. ed. Springer.

Lechowicz, M.J., 1984. Why do Temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities. Am. Nat. 124 (6), 821–842. https://doi.org/10.1086/284319.

Lee, P.Y., Rotenberry, J.T., 2005. Relationships between bird species and tree species assemblages in forested habitats of eastern North America. J. Biogeogr. 32, 1139–1150. https://doi.org/10.1111/j.1365-2699.2005.01254.x.

Lindberg, E., Holmgren, J., Olsson, H., 2021. Classification of tree species classes in a hemi-boreal forest from multispectral airborne laser scanning data using a mini raster cell method. Int. J. Appl. Earth Obs. Geoinf. 100, 102334. https://doi.org/10.1016/j.jag.2021.102334.

Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G., Williams, P.H., 2003. Avoiding pitfalls of using species distribution models in conservation planning. Conserv. Biol. 17, 1591–1600. https://doi.org/10.1111/j.1523-1739.2003.00233.x.

Lovett, G.M., Weathers, K.C., Arthur, M.A., 2002. Control of nitrogen loss from forested watersheds by soil carbon: Nitrogen ratio and tree species composition. Ecosystems 5 (7), 712–718. https://doi.org/10.1007/s10021-002-0153-1.

Martin, L., 1965. The Physical Geography of Wisconsin, 1st ed. The University of Wisconsin Press, Madison, Wisconsin.

Mickelson, J.G., Civco, D.L., Silander, J. a, 1998. Delineating Forest Canopy Species in the Northeastern United States Using Multi-Temporal TM Imagery. Photogramm. Eng. Remote Sens. 64, 891–904.

Modzelewska, A., Fassnacht, F.E., Stereńczak, K., 2020. Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 84, 101960. https://doi.org/10.1016/j.jag.2019.101960.

Pasquarella, V., Holden, C.E., Woodcock, C.E., 2018. Improved mapping of forest types using spectral-temporal Landsat features. Forests 210, 193–207. https://doi.org/10.1016/j.rse.2018.02.064.

Persson, M., Lindberg, E., Reese, H., 2018. Tree species classification with multi-temporal Sentinel-2 data. Remote Sens. 10, 1–17. https://doi.org/10.3390/rs10111794.

Quinlan, R.J., 1993. C4. 5: Programming For Machine Learning. Morgan Kauffmann.

Radeloff, V.C., Mladenoff, D.J., Boyce, M.S., 1999. Detecting jack pine budworm defoliation using spectral mixture analysis. Remote Sens. Environ. 69 (2), 156–169. https://doi.org/10.1016/S0034-4257(99)00008-5.

Rhemtulla, J.M., Mladenoff, D.J., Clayton, M.K., 2009. Legacies of historical land use on regional forest composition and structure in Wisconsin, USA (mid-1800s–1930s–2000s). Ecol. Appl. 19 (4), 1061–1078. https://doi.org/10.1890/08-1453.1.

Rhemtulla, J.M., Mladenoff, D.J., Clayton, M.K., 2007. Regional land-cover conversion in the U.S. upper Midwest: Magnitude of change and limited recovery (1850–1935-1993). Landsc. Ecol. 22 (S1), 57–75. https://doi.org/10.1007/s10980-007-9117-3.

Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K.J., Desai, A.R., Dietze, M.C., Dragoni, D., Garrity, S.R., Gough, C.M., Grant, R., Hollinger, D.Y., Margolis, H.A., McCaughey, H., Migliavacca, M., Monson, R.K., Munger, J.W., Poulter, B., Raczka, B.M., Ricciuto, D. M., Sahoo, A.K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., Xue, Y., 2012. Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American Carbon Program Site Synthesis. Glob. Chang. Biol. 18 (2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x.

Schindler, J., Dymond, J., Wiser, S.K., Shepherd, J.D., 2021. Method for national mapping spatial extent of southern beeches using temporal sepctral signatures. Int. J. Appl. Earth Obs. Geoinf. submitted, 102408 https://doi.org/10.1016/j.jag.2021.102408.

Schmidt, G.L., Jenkerson, C.B., Masek, J., Vermote, E., Gao, F., 2013. Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description. US Geol. Surv. Open-File Rep. 2013–1057.

Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sens. Environ. 124, 689–704. https://doi.org/10.1016/j.rse.2012.06.006.

Sheeren, D., Fauvel, M., Josipović, V., Lopes, M., Planque, C., Willm, J., Dejoux, J.-F., 2016. Tree species classification in temperate forests using formosat-2 satellite image time series. Remote Sens. 8, 734. https://doi.org/10.3390/rs8090734.

Shi, Y., Skidmore, A.K., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X., Heurich, M., 2018. Tree species classification using plant functional traits from LiDAR and hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 73, 207–219. https://doi.org/10.1016/j.jag.2018.06.018.

Shi, Y., Wang, T., Skidmore, A.K., Holzwarth, S., Heiden, U., Heurich, M., 2021. Mapping individual silver fir trees using hyperspectral and LiDAR data in a Central European mixed forest. Int. J. Appl. Earth Obs. Geoinf. 98, 102311. https://doi.org/10.1016/j.jag.2021.102311.

State of Wisconsin Department of Natural Resources, 2013. Public Forest Lands Handbook.
Wiscland 2 Land Cover User Guide, 2016. Madison, Wisconsin.
Wolter, P.T., Mladenoff, D.J., Host, G.E., Crow, T.R., 1995. Improved Forest Classification in the Northern Lake States Using Multi-Temporal Landsat Imagery. Photogramm. Eng. Remote Sens.
Townsend, P.A., Walsh, S.J., 2001. Remote sensing of forested wetlands: application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA. Plant Ecol 157, 129–149. https://doi.org/10.1023/A:1013999513172.
Wolter, P., Townsend, P., Sturtevant, B., Kingdon, C., 2008. Remote sensing of the distribution and abundance of host species for spruce budworm in Northern Minnesota and Ontario. Remote Sens. Environ. 112 (10), 3971–3982. https://doi.org/10.1016/j.rse.2008.07.005.
Wood, E.M., Pidgeon, A.M., Liu, F., Mladenoff, D.J., 2012. Birds see the trees inside the forest: The potential impacts of changes in forest composition on songbirds during spring migration. For. Ecol. Manage. 280, 176–186. https://doi.org/10.1016/j.foreco.2012.05.041.
Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10. https://doi.org/10.1016/j.rse.2012.01.010.
Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. Remote Sens. Environ. 185, 271–283. https://doi.org/10.1016/j.rse.2015.11.032.
Zhu, X., Liu, D., 2014. Accurate mapping of forest types using dense seasonal landsat time-series. ISPRS J. Photogramm. Remote Sens. 96, 1–11. https://doi.org/10.1016/j.isprsjprs.2014.06.012.
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.