Mapping forest types and tree species in temperate forests with satellite data from Landsat, Sentinel-2, and MODIS

By

Konrad Ireneusz Turlej

A dissertation completed in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Forestry)

at the

UNIVERSITY OF WISCONSIN – MADISON

2018

Date of final oral examination: 12/14/18

This dissertation is approved by the following committee:

Volker C. Radeloff, Department of Forest and Wildlife Ecology

Eric Kruger, Department of Forest and Wildlife Ecology

Annemarie Schneider, NELSON Institute

Mutlu Ozdogan, NELSON Institute / Department of Forest and Wildlife Ecology

Philip Townsend, Department of Forest and Wildlife Ecology

Abstract

Forests cover about 30% of Earth's land surface, affect the environment from local to global scales, and provide important services to humankind. Both ecological processes and ecosystem services depend greatly on which tree species are present. Unfortunately, mapping tree species across large areas is difficult. However, there are new opportunities for mapping tree species thanks to new satellites, more frequent satellite images, and new algorithms. The goal for my dissertation was to map main groups of tree species aggregated to forest types across large areas with satellite imagery. I analyzed data from three satellite programs: Landsat, Sentinel-2, and MODIS. I achieved three main outcomes. First, my results show a positive effect of the number of observations on classification accuracy. Increase in the number of image acquisitions improved the accuracy of my maps especially when the cloud-free scenes were not available. Second, I improved the methodology of generating dense time-series of vegetation indexes, which are useful for both monitoring forest phenology and forest type classifications. I did so by advancing the use of the STARFM algorithm, a tool for fusing Landsat and MODIS imagery, to conditions of low data availability, where the standard application of STARFM is suboptimal. Finally, I evaluated the usefulness of the imagery from constellation of three satellites: Landsat-8, Sentinel-2A and 2B. Forest type classifications based on imagery from all three satellites were as good as classifications based on Landsat imagery from three-year period. My results advance remote sensing technology, and are relevant for environmental sciences, forest management, and conservation. Managers can utilize forest type maps to assess the commercial value of the forests, develop climate change adaptation strategies, and manage wildlife habitat.

Acknowledgments

I would like to say thank you to all of those who have helped and supported me during my time as a graduate student at University of Wisconsin-Madison. I have learned a lot over last four and half year of pursuing my PhD degree. This would not be possible without wonderful people with whom I had a pleasure to work, meet, and interact.

I am especially grateful to my advisor Volker C. Radeloff for always providing me with useful ideas, constructive feedback, and amazing mentoring. Thank you for offering me this great opportunity of doing research at SILVIS Lab and for providing me with my first opportunities for teaching in an American university. Our numerous conversations opened my eyes on multiple topics including importance of scientific publishing, and encouraged me to continue my career in research.

I would like to express my deep appreciation to my committee members: Eric Kruger, Mutlu Ozdogan, Annemarie Schneider, and Philip Townsend for inspiration and providing me with critical comments on my ideas and manuscripts at various stages of development of my PhD thesis. I would like to thank Mutlu Ozdogan, for active collaboration on my chapters, for willingness to discuss approaches to image processing, and for mentoring both in terms of teaching and giving public presentations. Annemarie Schneider, for mentoring and providing me with opportunities for teaching, first as a grader, then teaching assistant, and finally as a lead class instructor. This experiences allowed me to grow both as a teacher and as a scientist, and supported me with funding necessary to finish my dissertation. I thank Philip Townsend for changing the way I see forests. Taking classes and talking with Phillip Townsend was an eye-opening experience for which I am deeply grateful. Last but not least, I would like to thank Eric Kruger, for stressing the necessity of looking at the forests not only from satellites but also from

the ground level. Thanks to this, I have become more critical while working with remote sensing data which helps me to interpret my results more accurately.

I would like to thank to other people that contributed to this work in various ways. To Murray Clayton for serving in my committee until my preliminary examination, and providing me with guidance on statistical approaches I used in this dissertation. Andrew Stoltman, for providing me with reference data for forest types in Wisconsin. Feng Gao, for providing me with useful comments on applying STARFM algorithm. Jeff Masek and Junchang Ju for providing me with access to Harmonized Landsat-8 and Sentinel-2 (HLS) product in version 1.4 before it was released. This allowed me to accomplish my work in timely manner. Mark Rickenbach, the Chair of Department of Forest and Wildlife Ecology, for providing me with funding for spring semester 2018. David Helmers from SILVIS Lab for being helpful and sharing his knowledge on GIS and programing in Python and managing computing servers.

Last but not least, I want to recognize my wonderful parents Bozena and Boguslaw Turlej for their love, support, and believing in me both in good and difficult times. And my mentor Katarzyna Dabrowska-Zielinska for inspiring me to become a scientist, encouraging me to pursue a PhD degree, and providing me with workspace during my visits in Warsaw, Poland.

Unfortunately, I cannot name everybody that had a direct or indirect contribution to my work and wellbeing during my time in the Madison. I will keep warm memories of friends and colleagues from SILVIS Lab, classmates, and all kind people I had a chance to meet.

Table of Contents

List of Figures

List of Tables

List of Appendices

Introduction

Problem statement

Forests cover about 30% of the Earth's land surface providing various services to natural systems and humankind largely depending on their tree species composition. In order to accurately assess the value of the forests, for example in terms of the timber resources, biofuel potential, influence on biodiversity and climate, or habitat quality for wildlife, detailed information on location of individual tree species is necessary. Unfortunately, such data is difficult to obtain for large areas. The standard on-the-ground methods for generating tree species maps are cost- and time-prohibitive, and the remote sensing approaches have been far from operational applications due to technical limitations of satellite systems. Fortunately, recent developments in satellite remote sensing, mainly new satellites, open access to unprecedented amounts of image data, and improvements in analysis techniques provide new opportunities for mapping trees from space and assessing their condition. Therefore, my overarching goal in this dissertation is to evaluate how these developments can improve the mapping main groups of tree species aggregated to forest types across large areas. To achieve this goal, I divided my research in to three chapters where I asked the following questions corresponding to mapping forest tree species: 1) how can we use the large amounts of partially cloudy and incomplete Landsat imagery? 2) can we improve the monitoring of single year phenology by fusing the data from MODIS and Landsat? and 3) can very frequent and harmonized data from Virtual Constellation of Landsat-8 and Sentinel-2 satellites improve the mapping?

Background

Spatially explicit maps of forest types corresponding to tree species location can improve many applications including environmental sciences, forest management, and conservation efforts. The

need for such maps is particularly pressing for areas that have a large number of tree species, for example the Great Lakes Region in Midwest USA, and for temperate forests that will be affected the most by climate change (Iverson and Prasad 1998). The lack of species-specific, up-to-date forest data hinders numerous natural resources management efforts including deer management (Rooney, 2001), the development of climate change adaptation strategies (Hanewinkel et al., 2012), and bioenergy initiatives (Jenkins et al., 2003).

Tree species vary in how they influence numerous ecological processes. For example, they differ in aboveground net primary production (Gower et al., 1993) and thus in the rates of carbon sequestration (Bergh et al., 2003). The chemical structure of their leaves and needles affects nutrition cycling in soils (Finzi, Van Breelmen, and Canham 1998; Hobbie et al., 2007), water chemistry (Lovett, Weathers, and Arthur, 2002), and fire regimes (Bergeron et al., 2004). Their composition shapes the Earth's environment on a global scale i.e. by affecting the rates of climate change (Bonan, 2008), as well as on a local scale impacting biodiversity (Barbier, Gosselin, and Balandier, 2008), and influencing the quality of water in streams and lakes (Lovett, Weathers, and Arthur, 2002).

In terms of forest management, tree species drive the monetary value of the forests because the dimensions and the quality of timber that can be produced from given species differ. They vary in terms of costs of harvesting, regeneration and regrowth time which affects the profitability and biofuel potential (Sims et al., 2001). Last but not least trees differ greatly in susceptibility to disturbance (He, Mladenoff, and Gustafson, 2002), and elasticity to changing climatic conditions (Hanewinkel et al., 2012).

Finally, in terms of conservation, trees create habitats for specific wildlife (Lee and Rotenberry, 2005; Wood et al., 2012; Kumar et al., 2018) determining the protection zones for unique species (Poole et al., 1996). Species composition can also influence the food availability in forests. For example, mixed forest stands have prolonged green-up period that sustain insects that feed on young leaves and provide food to insectivorous birds (Burger et al., 2012).

Mapping tree species is difficult. One reason is that standard on-the-ground methods for generating maps of tree species are time and cost prohibitive for large areas. Second, the remote sensing methods are still not mature. Remote sensing can identify the tree species accurately via wide range of techniques and provide the information at spatial scales ranging from individual trees to forest stands (Fassnacht et al., 2016), but it is not a routine forest application due numerous limitations of satellite systems: shortage of data, unfavorable spatial resolution, and insufficient revisit time. For example, trees can be recognized with great accuracy based on species-specific reflectance characteristics recorded by hyperspectral imagery (Fassnacht et al., 2014; Fagan et al., 2015), a trees' structural properties are captured by LiDAR (Jeronimo et al., 2018), but the access to such data is still too limited for operational use. Therefore, even though forests have been widely monitored with satellite imagery, the scope of that monitoring is largely limited to the detection of forest change (Hansen et al., 2013), forest disturbances, such as forest fires, wind-throws and harvests (Kennedy et al., 2010; Zhu et al., 2012), and the mapping of the broad forest categories (Griffiths et al., 2014).

Fortunately, two recent developments in optical remote sensing offer new opportunities to map tree species across large areas with moderate spatial resolution imagery, mainly: 1) much better data availability, and 2) new algorithms to process large numbers of satellite images.

First, data from Landsat satellite program is now openly available from USGS making it cost-effective to use both cloud-free and partially cloudy acquisitions (Wulder et al., 2012). Furthermore, the amount of Landsat-like data has increased in recent years thanks to successful launch of two Sentinel-2 satellites (A and B) as part of the European Copernicus program. Together, with NASA/USGS Landsat-8 satellite they create a Virtual Constellation of satellites providing surface reflectance data at global scale in unprecedented temporal resolution of 2-3 days (Claverie et al., 2018).

Second, the increasing amounts of data has driven the development of new algorithms that can take advantage of large image repositories (Wulder et al., 2016). New algorithms include: 1) the methodology for analyzing incomplete data sets that could tackle the problem of cloud cover (Schneider, 2012; Zhu and Woodcock, 2014), 2) techniques for fusing data of multiple sensors in order to combine their assets (Gao et al., 2006; Baumann et al., 2017), and 3) approaches to harmonize data from synergetic sensors (Claverie et al., 2018).

Landsat-like data has played a key role in efforts to making mapping of forest tree species across large areas a routine remote sensing application. Monitoring forest resources that has been one of strong focus of the Landsat program since its beginnings in 1972 including monitoring forest phenology (Fisher and Mustard, 2007), mapping forest cover (Wulder et al., 2003), forest disturbance (Kennedy et al., 2010; Zhu et al., 2012; Griffiths et al., 2014), and mapping forest tree species (Wolter et al., 1995; Wolter et al., 2008; Dymond, Mladenoff, & Radeloff, 2002).

The greatest advantages of Landsat-like imagery are: 1) moderate spatial resolution, 2) global coverage, and 3) fairly frequent repeat cycle. The Landsat-8 and Sentinel-2 systems are characterized by moderate spatial resolution of 10-30 meters which is sufficient for capturing

information from fairly small groups of trees and therefore for extracting information about the dominating tree species. The data provided by the systems consists visible (blue, green, and red) and infra-red (infra-red and short wave infra-red) reflectance bands crucial for characterization of species-specific reflectance characteristics (Iverson et al., 1989). The imagery is collected at global scale revisiting each area every 16 days in case of a single Landsat satellite (8 days for constellation of two Landsat satellites) and every 10 days for Sentinel-2 (5-days for Sentinel-2 A and B together). This allows for tracking changes in trees phenology that are crucial for capturing the inter-species differences in reflectance characteristics caused by synchronization of green-up, fall coloring, and senescence (Isaacson et al., 2012; Melaas et al., 2013; Baumann et al., 2017).

Development of the methods for mapping forest tree species with Landsat data

Over the years the Landsat-based approaches to mapping tree species have developed greatly, driven by the improvements in sensors installed onboard the satellites and data availability. The pioneering activities started in the eighties focused mostly on temperate forests of the US, especially in the Midwest region. The initial assessments of data from Landsat Thematic Mapper (TM) showed the usefulness of its bands, especially bands 3, 4, and 5, for separation of forest classes (Horler and Ahern, 1986). However, the first attempts to classify TM data with supervised approaches were of moderate accuracy. The classification of nine forest classes in area of northern Wisconsin based on single Landsat acquisitions reached overall accuracy of 69% (Hopkins et al., 1988). It is also worth noticing the promising work done on measuring the separation between the tree species of boreal forest in Minnesota using TM data simulated from aerial photography (Shen, Badhwar, and Carnes, 1985).

Later developments in image processing allowed for more complex classification activities, e.g. comparisons between sensors, or using additional environmental information. The TM sensor showed superiority over MSS data for mapping trees (Moore and Bauer, 1990). The usage of single TM imagery in combination with environmental information on soils, and topography increased the classification accuracy from 65-72% to 82-94% for 13 forest classes in northern Wisconsin (Bolstad and Lillesand, 1992).

The next improvement in mapping tree species with Landsat data was due to application of multi-seasonal acquisitions and a stepwise hierarchical approach. The base assumption of this methodology is capturing the inter-species spectral differences occurring on imagery from key periods where trees are in varying stages phenology. The first application of such approach on four MSS, and one TM images (from spring, summer, autumn and winter, enriched by use of Normalized Difference Vegetation Index NDVI), allowed for differentiation of 13 forest types in northern Lake States region with overall accuracy of 83% (Wolter et al., 1995). In this case land cover classes were delineated in stepwise process using band thresholding, and maximum likelihood classification. The stepwise, hierarchical approach was also used for mapping wetland tree species of southeastern USA reaching overall accuracy of 92% including non-forest classes (Townsend et al., 2001), and for forest land cover in outwash and loess plains of Northern Wisconsin, where the accuracies ranged from 33% up to 87% for different tree species (Dymond, Mladenoff, and Radeloff, 2002). In both cases the spectral data was enriched with NDVI, Tasseled Cap layers, and band differences.

In the nineties, the way of applying the multi-seasonal Landsat data for mapping tree species developed to use of the multi-band data stacks, taking advantage from multi-temporal signatures depicting tree phenology over the year. Usage of 18-layered image containing data from three

images capturing spring, summer, and fall phenology made it possible to distinguish 33 forest classes in northwestern Connecticut with overall accuracy of 79% (Mickelson et al., 1998). Here a minimum-distance-to-means classifier was applied, allowing for work with non-normal distribution of forest areas. A similar approach was used during the first edition of the WISCLAND project (1994-1998), fostering the process of mapping tree species to statewide scale using dual date image stacks acquired on spring and fall (Reese et al., 2002). Forests were classified according to Anderson levels I-III with accuracies of 95% for deciduous and 93% for coniferous forests at level I, and in range from 70% to 84% at lower levels, respectively. It is also worth noting that the combination of TM bands with Tasseled Cap layers, and band ratios classified in hybrid setting of ISODATA and maximal likelihood classifiers allowed Bauer et al. (1993) to reach accuracy ranging from 68% to 80% for 6 forest and 5 non-forest classes. The final conclusion from this classification activity was that Landsat data could be useful input for operational forest monitoring in Minnesota in 4-year intervals.

The opening the Landsat archives in 2008 started a new era for monitoring land cover allowing for analysis of dense data stacks and time series of Landsat data with machine learning and statistical tools. One example of such analysis in land cover classification is the use of Support Vector Machines (SVM) to a stack of 7 Landsat acquisitions for mapping pine, oak and mixed-mesophytic forests in area of varying topography located in southeastern Ohio (Zhu and Liu, 2014). This work pointed out the problem of overfitting the SVM with high dimensional data and need for feature selection. In the study, 15 features (from total set of 48) allowed the discrimination of 3 forest classes with an overall accuracy of 90%. A recent application of dense image stacks of Landsat data was performed for Wisconsin for the WISCLAND-2 project (Wiscland 2, 2016). Here, the data from Landsat-5, -7, and -8 were combined and classified with

a boosted technique applied to decision trees, providing classification accuracies from 70% to 90% depending on tree species. The authors used a hierarchical approach to classify trees at 4 levels, starting from forest cover at level 1, and ending at detailed tree species on level 4. Statistical tools have also been used for mapping distributions and abundance of tree species. For example, Partial Least Squares (PLS) regression has proven to be an effective tool for modeling forest structure within heterogeneous forests (Wolter et al., 2008). Finally, the analysis of averaged multi-year time series of Landsat observations made it possible to quantify both harmonic and phenological metrics in a set of spectral-temporal features allowing accurate mapping across multiple footprints (Pasquarella et al., 2018).

Despite the great potential of data from the Sentinel-2 satellites, the imagery has not yet been heavily tested for mapping tree species. Only one study focusing on temperate forests of southeastern Germany reported somewhat low levels of accuracy (65%) when a single late-summer image was classified (Immitzer, Vuolo and Atzberger, 2016). The most probable reason for this result was a limited amount of data and sub-optimal acquisition time. However, the accuracy expressed in kappa coefficient of 0.95 was achieved with 17 multi-seasonal images from spatially similar Formosat-2 satellite (spatial resolution of 8 meters) for 8 broadleaves and 5 coniferous tree species (Sheeren et al., 2016). This result shows the potential of Sentinel-2 data for mapping tree species composition.

Goal and objectives

My goal for this dissertation was to evaluate how currently available satellite data from Landsat, Sentinel-2, and MODIS can improve the applicability of mapping main groups of tree species aggregated to forest types across large areas with moderate spatial resolution of 30 meters. I focused on three particular objectives. In my first chapter, I looked at the relationship between

mapping accuracy and the availability of cloud-free and partially incomplete Landsat acquisitions due to clouds or the Scanning Line Corrector malfunction. In my second chapter, I focus on deriving a dense time series of forest phenology by blending Landsat and MODIS data to support the mapping process. Finally, in my third chapter, I evaluate the usefulness of the harmonized surface reflectance data from Virtual Constellation of Landsat-8 and Sentinel-2.

Chapter summaries

In the following pages I will provide a short synopsis for each chapter, and their contribution to previous studies (**Table 1**).

Chapter 1 summary

Mapping temperate tree species with Landsat-like data relies on the availability of cloud-free observations capturing inter-species differences in phenology. Unfortunately, the access to high quality Landsat observations is limited due to cloud cover and SLC gaps. Moreover, the relation between the mapping performance to the number of image acquisition, their seasonality, and levels of missing data has not yet been evaluated for tree species mapping.

My overarching goal for this chapter was to map tree species aggregated to forest types with Landsat imagery, and to maximize mapping accuracy by including imagery with gaps due to both cloud cover and SLC gaps. My specific objectives were to assess the effects of a) the number of acquisitions, b) the amount of missing data, and c) the seasons for with the imagery was available, on classification accuracy.

I achieved an overall accuracy of 88% for my final map of tree types for northern Wisconsin containing 23 classes. Depending on the type of sampling my maps provided a good combination

of species-specific accuracies (70-97%) for the most abundant forest types, or reasonable accuracies (70-85%) for all forest types regardless of their abundance.

My results show a positive effect of the number of observations on classification accuracy. This is emphasized the most for cloudy imagery, e.g. the high number of acquisitions allowed me to keep the level of accuracy close to 80% even when using data missing >=30% of pixels. In other words, a stack of imagery containing sufficiently high number of partially clouded acquisitions provided enough data for: a) covering the whole study area with at least single clear observation, and b) capturing the differences in reflectance occurring due to species-specific phenology. In terms of seasonality, the fall acquisitions added the most to the mapping performance providing the best accuracy than any combination of other seasons and just slightly better accuracy when combined with imagery from other seasons.

Interpreting my results, I drew guidelines for mapping forest type with Landsat data. First, that acceptable classification accuracy can be achieved only with handful of cloud-free acquisitions. Especially, if they cover spring and fall season. However, I recommend to use all available Landsat images from at least 3-year period if cloud-free acquisitions are not available.

Chapter 2 summary

Recent scientific literature report improvements in mapping temperate tree species with Landsat data coming from combination of surface reflectance and phenology information (Pasquarella et al., 2018). However, deriving the information on phenology from Landsat is difficult due to low frequency of available clear observations. This applies especially to areas affected by climate change and inter-annual variability in tree-phenology for which the single year phenology is required. Often, clear observations are too sparse for statistically valid modelling of phenology

or any remedy processing, e.g. gap filling (Baumann et al., 2017). Moreover, the lack of entirely cloud-free acquisitions over the year excludes also the classical applications of spatial-temporal blending of Landsat with spatially coarse but frequent MODIS data that could provide the sufficiently dense time series of satellite observations (Gao et al., 2006).

My goal for this chapter was to design an approach for generating a cloud-free time series of Landsat surface reflectance covering a period of single growing season, that could be used even the data availability is very low. My specific objectives were to: a) improve the applicability of spatial-temporal blending of Landsat and MODIS data to large areas, b) generate the cloud-free time series of Landsat surface reflectance based on compositing of the outcomes of numerous spatial-temporal models blending cloudy Landsat and MODIS, and c) assess the optimal approach to composite the outcomes of fusing Landsat and MODIS.

I have improved the applicability of spatial-temporal blending of the data from Landsat and MODIS, by removing the need for at least one cloud-free Landsat acquisition per year. My approach provided a quality time series of cloud-free surface reflectance filling the data gaps caused by clouds and SLC gaps and retaining all original Landsat observations.

My results suggest that simple averaging the predictions of Landsat surface reflectance from multiple fusing models based on numerous cloudy acquisitions is sufficient for generating cloud-free data sets. My predictions were fairly uniform among multiple models with standard deviations in surface reflectance varying from 2 up to 7 percentage points between my models. However, the cloud-free time series did not improve classifications of forest type.

Chapter 3 summary

Combination of imagery from Landsat and Sentinel-2 provides unprecedented temporal frequency of 2-3 days providing opportunity to increase the accuracy of mapping forest type. According to results in my first chapter, the high number of partially cloudy Landsat observations allow us to map forest type at levels of accuracy comparable to results obtained with almost cloud free acquisitions. Therefore, increased data availability resulting from combining the imagery from both systems should improve the mapping performance especially when only partially cloudy imagery is available.

My goal for this chapter was to evaluate how harmonized data from Landsat-8 and Sentinel-2 (HLS) affects the mapping tree species aggregated to temperate forest types in northern Wisconsin. My specific objectives were to: a) assess how single year of data from Landsat and Sentinel-2 improves the mapping accuracy comparing to imagery from individual systems, b) evaluate the usefulness of 8-day, and 16-day image composites for mapping tree species, and c) check whether HLS data is more resistant to the effects of cloud cover comparing to the less frequent individual systems.

My results did not show great improvements resulting from usage of HLS data compared to the imagery from either Landsat-8 or Sentinel-2. The differences in mapping accuracy between the data subsets was <5%. Similarly, compositing of the HLS imagery to either 8-day or 16-day periods did not improve the classification accuracy, but it accelerated the mapping process. Finally, all three data subsets performed similarly when all available acquisitions contained >25% of pixels clouded resulting in about 10 percentage points drop in mapping accuracy.

My findings from this chapter provide first insights to usage of harmonized data from the full Virtual Constellation of Landsat-8 and Sentinel-2 for mapping forest type. However, even though I found only minor mapping improvements resulted from increased temporal frequency of the data I conclude that the data can be useful for more accurate monitoring single year spatial-temporal and phenological patterns of woody vegetation.

Significance

Working on remote sensing methodology for mapping tree species aggregated to forest types, I have contributed to technology, science, forest management, and conservation.

In terms of technology, I provide useful insights how to map forest type with Landsat-type data, and make important arguments in discussion of future of satellite monitoring. Remote sensing is changing, providing us with both numerous advantages and challenges coming with new technology (Wulder et al., 2012; Wulder et al., 2016; Young et al., 2017). Thus it is important to validate new data, new ways how to process it, and to point out both problems and promises. The blessing of access to large archives of Landsat and Sentinel-2 repositories provide us with tremendous amounts of useful information, but also comes with the curse of dealing with its quality (Ju and Roy, 2008). By evaluating how statistical learning tools deals with data gaps caused by clouds and SLC gaps, I was able to draw guidelines for using such incomplete datasets and assess the necessity of filling these missing values. My solutions are directly applicable to results of current NASA's efforts focused on generating harmonized surface reflectance (HLS) product from Landsat-8 and Sentinel-2 satellites (Claverie et al., 2018). Finally, my conclusions on the relation between mapping accuracy and amount of satellite information, and its source, is important for discussion of future of Earth Observation missions. I conclude that for mapping forest type the access to just few good quality images is sufficient, but when cloud-free data is

not available, then the advantages of data from numerous synergetic satellites are notable. Considering the costs of building, launching, and maintaining satellite missions, such conclusions should be of great value.

More broadly, my results on mapping forest type are useful for all scientific areas that require spatially explicit information on location of main groups of tree species. Forests influence natural systems largely depending on their tree species composition. Trees vary in influence to numerous ecological processes and their composition shapes the Earth's environment at both global and local scales (Lovett, Weathers, and Arthur, 2002; Bergh et al., 2003; Barbier, Gosselin, and Balandier, 2008). More spatially detailed information on location of forest types can help to model these processes with greater accuracy. Because on-the-ground mapping of forest type is time and cost prohibitive, my results present an alternative for repeated mapping activities that could provide data to evaluate results of research focused on modeling climate change at wide range of scales. Considering estimated species migrations such information could be very useful argument in discussion on how to protect the environment.

Finally, by generating tools for image processing and maps of forest type distribution, I have contributed to forest management and conservation. Trees are important economically (Hanewinkel et al., 2012), and for human wellbeing (Bell and Thompson, 2013). Since tree species differ in their timber value and silvicultural methods, it is necessary to monitor their availability. History shows that usage of forests can be very rapid and lead to fast exhaustion of timber resources (Curtis, 1959; Radeloff et al., 1999; Rhemtulla et al., 2007; Rhemtulla et al., 2009). Therefore, development of time efficient ways to monitor forest resources is necessary for sustainable management, and to prevent illegal logging. Forests are also habitats for unique wildlife that is affected by tree species composition. For example, numerous species of birds

breed in areas of particular trees (Holmes and Robinson, 1981). Knowledge on current forest composition and monitoring its change can provide arguments for protecting certain areas. This is important considering the cultural value of forests for future generations.

References

Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Walsh, T.A., Walters, D.K., 1993. Satellite Inventory. Area.

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

Bell, S., Thompson, C.W., 2013. Human Engagement with Forest Environments: Implications for Physical and Mental Health and Wellbeing, in: Challenges and Opportunities for the World's Forests in the 21st Century. Springer International Publishing, pp. 71–92. https://doi.org/https://doi.org/10.1007/978-94-007-7076-8_5

Bergeron, Y., Gauthier, S., Flannigan, M.D., Kafka, V., 2004. Fire Regimes at the Transition between Mixedwood and Coniferous Boreal Forest in Northwestern Quebec Author (s): Yves Bergeron , Sylvie Gauthier , Mike Flannigan and Victor Kafka Published by : Wiley Stable URL : http://www.jstor.org/stable/3450365 REFE. Ecology 85, 1916–1932.

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.

Burger, C., Belskii, E., Eeva, T., Laaksonen, T., Mägi, M., Mänd, R., Qvarnström, A., Slagsvold, T., Veen, T., Visser, M.E., Wiebe, K.L., Wiley, C., Wright, J., Both, C., 2012. Climate change, breeding date and nestling diet: How temperature differentially affects seasonal changes in pied flycatcher diet depending on habitat variation. J. Anim. Ecol. 81, 926–936. https://doi.org/10.1111/j.1365-2656.2012.01968.x

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

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

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

Fagan, M.E., DeFries, R.S., Sesnie, S.E., Arroyo-Mora, J.P., Soto, C., Singh, A., Townsend, P.A., Chazdon, R.L., 2015. Mapping species composition of forests and tree plantations in northeastern Costa Rica with an integration of hyperspectral and multitemporal Landsat imagery. Remote Sens. 7, 5660–5696. https://doi.org/10.3390/rs70505660

Fassnacht, F.E., Lati, H., Sterenczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. Remote Sens. Environ. 186, 64–87.

Fassnacht, F.E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Kochb, 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, 2547–2561.

Fisher, J.I., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sens. Environ. 109, 261–273. https://doi.org/10.1016/j.rse.2007.01.004

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44, 2207–2218. https://doi.org/10.1109/TGRS.2006.872081

Gower, S.T., Reich, P.B., Son, Y., 1993. Canopy dynamics and aboveground production of five tree species with different leaf longevities. Tree Physiol. 12, 327–345.

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V.C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., Hostert, P., 2014. Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. Remote Sens. Environ. 151, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Hanewinkel, M., Cullmann, D.A., Schelhaas, M.-J., Nabuurs, G.-J., Zimmermann, N.E., 2012. Climate change may cause severe loss in the economic value of European forest land. Nat. Clim. Chang. 3, 203–207. https://doi.org/10.1038/nclimate1687

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. Science 850, 2011–2014. https://doi.org/10.1126/science.1244693

Holmes, R.T., Robinson, S.K., 1981. Tree species preferences of foraging insectivorous birds in a northern hardwoods forest. Oecologia 48, 31–35. https://doi.org/10.1007/BF00346985

Hopkins, P.F., Maclean, A.L., Lillesand, T.M., 1988. Assessment of Thematic Mapper Imagery for Forestry Applications under Lake States Conditions. Photogramm. Eng. Remote Sensing 54, 61–68.

Horler, D.N.H., Ahern, F.J., 1986. Forestry information content of Thematic Mapper data. Int. J. Remote Sens. 7, 405–428. https://doi.org/https://doi.org/10.1080/01431168608954695

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

Isaacson, B.N., Serbin, S.P., Townsend, P.A., 2012. Detection of relative differences in phenology of forest species using Landsat and MODIS. Landsc. Ecol. 27, 529–543. https://doi.org/10.1007/s10980-012-9703-x

Iverson, L.R., Graham, R.L., Cook, E.A., 1989. Applications of satellite remote sensing to forested ecosystems. Landsc. Ecol. 3, 131–143. https://doi.org/10.1007/BF00131175

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., Birdsey, R.A., 2003. National-scale biomass estimators for United States tree species. For. Sci. 49, 12–35. https://doi.org/10.2737/NE-GTR-310

Jeronimo, S.M.A., Kane, V.R., Churchill, D.J., McGaughey, R.J., Franklin, J.F., 2018. Applying LiDAR individual tree detection to management of structurally diverse forest landscapes. J. For. 116, 336–346. https://doi.org/10.1093/jofore/fvy023

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, 1196–1211. https://doi.org/10.1016/j.rse.2007.08.011

Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Kumar, P., Sajjad, H., Jain, P., 2018. Assessment of Forest Species Diversity in Sariska Tiger Reserve, Rajasthan, India, in: Applications and Challenges of Geospatial Technology. Springer, Cham, pp. 257–273. https://doi.org/https://doi.org/10.1007/978-3-319-99882-4_15

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

Melaas, E.K., Friedl, M.A., Zhu, Z., 2013. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. Remote Sens. Environ. 132, 176–185. https://doi.org/10.1016/j.rse.2013.01.011

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.

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

Poole, K.G., Wakelyn, L.A., Nicklen, P.N., 1996. Habitat selection by lynx in the Northwest Territories. Can. J. Zool. 74, 845–850. https://doi.org/10.1139/z96-098

Radeloff, V.C., Mladenoff, D.J., He, H.S., Boyce, M.S., 1999. Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. Can. J. For. Res. 29, 1649–1659. https://doi.org/10.1139/x99-089

Reese, H.M., Lillesand, T.M., Nagel, D.E., Stewart, J.S., Goldmann, R.A., Simmons, T.E., Chipman, J.W., Tessar, P.A., 2002. Statewide land cover derived from multiseasonal Landsat TM data: A retrospective of the WISCLAND project. Remote Sens. Environ. 82, 224–237. https://doi.org/10.1016/S0034-4257(02)00039-1

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, 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, 57–75. https://doi.org/10.1007/s10980-007-9117-3

Rooney, T.P., 2001. Deer impacts on forest ecosystems: a North American perspective. Forestry 74. https://doi.org/10.1093/forestry/74.3.201

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

Shen, S.S., Badhwar, G.D., Carnes, J.G., 1985. Separability of boral fores species in the Lake Jennette Area, Minnesota. Photogramm. Eng. Remote Sensing 51, 1775–1783.

Sims, R.E.H., Maiava, T.G., Bullock, B.T., 2001. Short rotation coppice tree species selection for woody biomass production in New Zealand. Biomass and Bioenergy 20, 329–335. https://doi.org/10.1016/S0961-9534(00)00093-3

Townsend, P. a, Walsh, S.J., Ecology, S.P., Dec, N., Townsend, A., Stephen, J., J, S., 2001. Remote Sensing of Forested Wetlands: Application of Multitemporal and Multispectral

Satellite Imagery to Determine Plant Community Composition and Structure in Southeastern Stable URL : http://www.jstor.org/stable/20051167 and sensing of forested wetland. Plant Ecol. 157, 129–149. https://doi.org/10.1023/A:1013999513172

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.

Wolter, P.T., Townsend, P.A., Sturtevant, B.R., Kingdon, C.C., 2008. Remote sensing of the distribution and abundance of host species for spruce budworm in Northern Minnesota and Ontario. Remote Sens. Environ. 112, 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., Dechka, J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudoin, A., Franklin, S.E., 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. For. Chron. 79, 1075–1083. https://doi.org/10.5558/tfc791075-6

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

Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R., Evangelista, P.H., 2017. A survival guide to Landsat preprocessing. Ecology 98, 920–932. https://doi.org/10.1002/ecy.1730

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., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

Zhu, Z., Woodcock, C.E., Olofsson, P., 2012. 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

List of tables

Table 1. The overview of the remote sensing studies focused on mapping tree species communities including the findings from this dissertation.

Study	Location	Number of forest classes	Data	Methods	Accuracy / Conclusions
Horler and Ahern, 1986	US	1	Landsat TM	The assessment of TM data for forest monitoring	Usefulness of TM bands: 3, 4, and 5
Shen, Badhwar, and Carnes, 1985	Minnesota, US	1	Aerial imagery	Using simulated TM data from aerial imagery	Promising possibility of using real TM data for mapping forests
Hopkins, Maclean, and Lillesand, 1988	Wisconsin, US	9	Single Landsat TM image	Classification of multiple forest classes	0.69
Moore and Bauer, 1990	Northern Wisconsin, US	3	Landsat MSS, Landsat TM	Comparison of imagery from Landsat sensors	Superiority of TM data over MSS
Bolstad and Lillesand, 1992	Northern Wisconsin, US	13	Landsat TM	Addition of environmental information (soils, topography)	82-94%
Bauer et al., 1993	Minnesota, US	6	Landsat TM	multi-layer image stacks classified with hybrid method combining ISODATA and Maximal Likelihood classifiers	68% - 80% for forest and non-forest classes
Wolter et al., 1995	Northern Lake State Region, US	13	Landsat MSS, Landsat TM	Multi-temporal imagery; hierarchical stepwise classification	83% / problems with access cloud free to data
Mickelson, Civco, and Silander, 1998	North-western Connecticut, US	33	Landsat TM	18-layer image stack classified with minimal distance classifier	Overall accuracy 78,9%
Townsend et al., 2001	South-eastern US	21	Landsat TM	hierarchical stepwise classification; addition of vegetation indices, and Tasseled Cap layers	92.1% including non-forest classes
Dymond, Mladenoff, and Radeloff, 2002	Wisconsin, US	12	Landsat TM	hierarchical stepwise classification; addition of vegetation indices, and Tasseled Cap layers	Accuracy for forest classes: 33-84 %
Reese et al., 2002	Wisconsin, US	Multiple at 4 levels	Landsat TM	Large scale mapping activity covering the whole state of Wisconsin	Accuracy for tree species: 70-84%

Study	Location	Number of forest classes	Data	Novelty	Accuracy / Conclusions
Zhu and Liu, 2014	South-eastern Ohio, US	3	Landsat ETM+	Using dense Landsat image stack with SVM classifier; feature selection	Overall accuracy including non-forest classes 96.27%
Gao et al., 2015	North China	1	Landsat OLI	Addition of multiscale texture measures	92.70%
Wiscland 2 Land, 2016	Wisconsin, US	Multiple at 4 levels	Landsat TM, ETM+, OLI	Boosted Decision Trees applied to large stacks of multi-temporal image acquisitions and vegetation indices	Accuracy for tree species: 70-90%
Immitzer, Vuolo, and Atzberger, 2016	South-eastern Germany	7	Sentinel-2A MSI	First published use of Sentinel-2A data for classifying tree species	Overall accuracy 65% due to un-favored timing of image acquisition
Pasquarrella et al., 2018	Massachusetts, US	8	Landsat MSS, Landsat TM, Landsat ETM+, Landsat OLI	Use of spectral and temporal features derived from all available observations from 1985 to 2015	83.39 ±2.31%
Chapter 1	Wisconsin, US	23	Landsat-7, Landsat-8	Use of partially clouded imagery; multi-temporal image stacks	Overall accuracy of 88.1%; no need for cloud free data
Chapter 2	Wisconsin, US	23	Landsat-7, Landsat-8, MODIS 43A4	Actual Landsat observations filled with STARFM predictions	The outcomes of SATARFM do not improve classification accuracy
Chapter 3	Wisconsin, US	23	Harmonized Landsat-8 and Sentinel-2	Use of 8-day and 16-day intra-annual image composites	16-day intra annual composites improve the efficiency of mapping tree types without decrease in classification accuracy (83.5%)

Chapter 1. Mapping forest types over large areas with Landsat imagery partially affected by clouds and SLC gaps

Contributors: Konrad Turlej[1], Mutlu Ozdogan[1], Volker C. Radeloff[1]

[1] Department of Forest and Wildlife Ecology, University of Wisconsin-Madison

Abstract

Forests provide numerous services, but which services they provide depends greatly on their tree species composition. That makes it important to map not only forest extent and its dynamics, something that remote sensing excels in, but also to map tree species. Our main goal was to map tree species aggregated to forest types with Landsat imagery, and to test when adding imagery with missing data, either due to clouds or scan line problems, improves classification accuracy. Our study area covered one Landsat footprint (26/28) in Northern Wisconsin, USA, with predominantly temperate forests. We selected this area because of spatially heterogeneous forests and availability of high-quality reference data. We quantified how species-level classification accuracy was affected by a) the amount of missing data due to cloud cover and SLC gaps, b) the number of acquisitions, and c) the seasons for which images were available. We classified data for both a single year and a three-year period of Landsat 7 and 8 images with a decision tree classifier and mapped the dominant tree species in each pixel and in each forest stand of our reference data. We obtained three major results. First, we achieved producer's accuracies ranging from 72.4% to 97% and user's accuracies ranging from 79.8% to 93.4% for the most abundant forest types in the study area (>2% of the forest area). Second, classification accuracy improved with more acquisitions, especially when images were available for spring, summer, and fall, even when we included only imagery with more than 30% pixels missing pixels. Finally, classifications for pure stands were 10 to 30 percentage points better than those for mixed stands.

We conclude that inclusion of Landsat imagery with missing data allows to map forest types with accuracies that were previously possible only for those rare years for which several cloud-free images were available. Our resulting maps with high accuracy provide important data for both forest management and ecology.

Keywords

Forest type, Landsat, clouds, SLC gaps, remote sensing, classification, forest inventory map, Wisconsin, USA

Introduction

Forests cover about one third of the Earth's land surface and provide numerous services, but which services they provide exactly depends on tree species composition. Tree species vary, for example, in their rates of carbon sequestration (Bergh et al., 2003), nutrition cycling in soils (Finzi, Van Breelmen, and Canham 1998; Hobbie et al., 2007), and their effects on soil water chemistry (Lovett et al., 2002). Tree species composition shapes the Earth's environment at a global scale, e.g., by affecting the rates of climate change (Bonan, 2008), and at a local scale, e.g., by providing habitats for wildlife (Wood et al., 2012; Lee & Rotenberry, 2005), and affecting biodiversity (Barbier, Gosselin, and Balandier, 2008), and runoff water quality (Lovett, Weathers, and Arthur, 2002). Last but not least, tree species differ greatly in their timber value and susceptibility to disturbance (Hanewinkel et al., 2012). Therefore, tree species maps are necessary for science, where they constitute important input data for modeling of various environmental variables, and for forest management and the sustainable use of forest resources (Portoghesi, 2006). Finally, 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).

On-the-ground mapping of tree species is cost-prohibitive for large areas. However, recent technological developments in remote sensing may make it possible to map tree species efficiently. Tree species can be successfully separated by measuring the reflectance and its annual changes due to phenology. The best results are obtained with hyperspectral data, which provide species-specific reflectance and absorption characteristics (Fassnacht et al., 2014), but hyperspectral imagery is not widely available. For both hyperspectral and multispectral data, mapping accuracy increases when using multi-temporal acquisitions that capture differences in tree phenology are used (Elatawneh et al., 2013). For example, the use of a single Sentinel-2 image resulted in an overall accuracy of 66.2% for temperate forests in Germany (Immitzer et al., 2016), while multi-temporal data from Formosat-2 satellite imagery resulted in reported accuracies ranging from 0.90 to 0.96 for the kappa coefficient (Sheeren et al., 2016). Similarly, the use of Landsat data covering multiple stages of forest phenology results in mapping accuracy ranging from 83% to 96% depending on the number of tree species (Townsend et al., 2001; Zhu and Liu, 2014). In rare cases, high classification accuracy can be achieved with only a handful images from spring, summer, and fall (Mickelson, Civco, and Silander, 1998). However, a large number of acquisitions is an advantage as long as the additional imagery introduces more information on differences in tree phenology (Wolter et al., 1995; Zhu and Liu, 2014).

When the goal is to map tree species for large areas, Landsat and Sentinel-2 imagery is the only viable choice because of its spatial and temporal resolution, coverage of spectral ranges essential for monitoring tree phenology, radiometric stability, and deep image archives. First, the spatial resolution of 10-30 meters of these images allows for collecting information over extensive areas, yet captures the reflectance from a sufficiently small group of trees so that the dominant species for each pixel can be assessed (Wolter et al., 1995). Second, the 5-8-day acquisition

frequency of two Landsat satellites, plus additional Sentinel-2 images for recent years, make it possible to capture rapid changes in canopy development during the growing season, which is what makes inter-species differences most apparent (Mickelson, Civco, and Silander, 1998). Finally, the radiometric quality of all Landsat bands is stable over the entire historical archive (Wulder et al., 2012) making it possible to classify changes in tree species over time (Pasquarella et al., 2018).

The challenge is that in order to map tree species accurately using Landsat data, multi-temporal imagery is necessary, but most images have gaps due to clouds, their shadows, and scan line corrector (SLC) gaps on Landsat 7. These gaps preclude the use of many mapping approaches, including those that employ thresholds for the imagery (Wolter et al., 1995; Wolter et al., 2008; Dymond, Mladenoff, & Radeloff, 2002), statistical classifiers (Mickelson, Civco, and Silander, 1998), and random forest classifiers (Zhu and Liu, 2014) that require completely cloud-free imagery. 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, but the probability of two cloud-free observations within 16 days from Landsat satellite being available is low for conterminous United States (Ju and Roy, 2008).

One possible solution to the shortage of cloud-free Landsat data is to analyze image with gaps via machine learning algorithms that can handle missing data. Indeed, Landsat data with gaps can be useful when mapping general land cover categories (Schneider, 2012), as well as tree species (*Wiscland 2 Land Cover User Guide*, 2016), and so is partially cloudy imagery from the Formosat-2 satellites (Sheeren et al., 2016). However, the question is up to which level of missing pixels imagery is still useful to map tree species, and how both the number of acquisitions and the seasons for which imagery is available affects classification accuracy.

The main goal of our work was to map tree species aggregated to forest types with Landsat imagery, and to maximize mapping accuracy by including imagery with gaps due to both cloud cover and SLC gaps. 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.

Methods

Study area

Our study area was in northern Wisconsin and consisted of the forested areas of one Landsat footprint (path 26, row 28) (**Figure 1. 1a)**. The area is divided by three ecoregions: Lake Superior Lowland, Northern Highland, and western part of Central Plain (Martin, 1965). 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.

Tree species composition in the local forests stands is mixed. 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*), 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., 2007; Rhemtulla et al., 2009). The local species composition prior to European settlement was dominated by hemlock and northern hardwood species in till plains and mesic moraines, and pines on sandy outwash soils (Curtis, 1959). Pines were accompanied by less abundant aspen also present in composition with oaks in the areas of the forest-savanna (Schulte et al., 2002). The changes resulted from deforestation, i.e., land conversion from forest to croplands followed by rapid logging between 19th and 20th centuries, and from disturbance, i.e., insect outbreaks and return of severe fire events in middle of 20th century. The regrowth of the forests since the 1930s did not follow the species distribution characteristic prior to European settlement.

Data

We focus here on Landsat data only excluding the imagery from Sentinel-2 satellite because of differences in resolution, and band wavelength ranges that can affect our results on the data availability. We analyzed Landsat-7 and Landsat-8 imagery from 1/1/2014 to 12/31/2016. In total, 82 acquisitions, out of 133, provided at least some cloud free pixels for our 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. 1**). The Landsat-7 data were processed via the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Department of the Interior U.S. Geological Survey, 2012; Department of the Interior U.S. Geological Survey, 2017a) and the Landsat-8 data via the Landsat Surface Reflectance Code (LaSRC) (Department of the Interior U.S. Geological Survey, 2017b). We removed pixels that were: a) flagged as clouds, cloud shadows, or water based on the quality band, b) that exceeded the valid range of surface reflectance, and c) all non-forest pixels according to the 2011 National Land Cover Database (NLCD, **Figure 1. 1c,** Homer et al., 2015).

Our ground reference data were the Reconnaissance forest inventory data of the Wisconsin Forest Inventory & Reporting System (WisFIRS) of the Wisconsin Department of Natural Resources and County Forests (State of Wisconsin Department of Natural Resources, 2013, **Figure 1b**). The dataset provides stand-level information on forest types (**Appendix 1. 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 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 mixed stands. The majority of the Reconnaissance forest stands were examined in the field after 2000, less than 16 years before the satellite data we used for this study were acquired (**Figure 1. 2**).

Classification

Our main goal was to obtain an accurate pixel-level classification of forest type for all forests according to the National Land Cover Database (NLCD). In order to do so, we conducted numerous classifications to test: 1) the influence of missing data due to cloud cover and SLC scanning gaps on classification accuracy, 2) the influence of number of image acquisitions on classification accuracy, and 3) the usefulness of imagery from different seasons and their combinations. We performed cross validation repeating each tests 10 times for the only-2016 imagery versus 2014-2016 imagery, as well as for pure stands only versus pure-plus-mixed stands. For each of the 10 repetitions, we randomly divided the forest stands from 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 all pixels belonging to the training stands, and

applied the model to the pixels from the test stands. Finally, we used the model providing the best accuracy to map the forest types within forests according to NLCD database.

We used the C5.0 Decision Trees and Rule-Based Models, implemented in R statistical software, as our tool for classification because it can handle incomplete datasets (Farhangfar et al., 2008), such as partially clouded imagery. We built each model requiring 100 boosting iterations, the maximum in C5.0, because higher numbers improved the classification accuracy. However, the actual number of iterations varied between the models depending on the amount of data and the level of missing pixels in the training data set.

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

To evaluate the influence of missing data on the classification accuracy, we ran a series of classifications with changing thresholds for the minimum and maximum percentage of pixels missing values. For each threshold, we recorded: 1) the overall accuracy, 2) the number of image acquisitions meeting the threshold, 3) the percent of the pixels in the study area with at least one cloud-free observation, 4) the percent ratio between the number of cloud-free observations and the number of the pixels in the study area (e.g., 100% means that on average there was one cloud free 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, containing >=90% and <=100% values missing. Next, we broaden the range of missing values to add images providing values until we included all acquisitions, from 0% to 100% pixel values missing. Finally, we narrowed the range by removing the most incomplete imagery form the analysis. We finished with the imagery with the data missing least (0% to 10% of pixel values). For this test we used only imagery from 2016.

Test 2: Number of image acquisitions

To test the effects of higher number of images on classification accuracy, we extended the image acquisition time to three years (2014 to 2016), and repeated the tests with different levels of missing data (see above). 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 percentages of pixels missing values, 2) the percent of the pixels in the study area for which the data provided at least one cloud-free observation, 3) the percent ratio between the number of cloud-free observations and the number of the pixels in the study area, and 4) the overall percentage of missing data within the image stack.

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 tested two data sets containing all images from 2016 only versus 2014-2016, and compared 1) the overall accuracy, 2) the number of image acquisitions meeting current percentages of pixels missing values, 3) the percent ratio between the number of cloud-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 cloud-free observations available, and 5) the overall percentage of missing data within the image stacks.

Accuracy assessment, classification probability

We recorded the area-weighted overall accuracy of classification 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 our classifications to the forest stands,

labeling each stand based on its most common class. We excluded edge pixels within 30 meter of the stand border, calculated the overall classification accuracy, user's and producer's accuracy (Congalton, 1991), and derived graphs and error matrices to present the minimal, mean, and maximal value of overall accuracy for each data set related to cloud cover and seasonal distribution of the imagery for both all stands and pure stands only. Finally, to estimate class-level accuracies at the forest-stand level, we calculated the area-weighted 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 for pure stands, mixed stands and for all stands together.

For our final map, we calculated class-specific inverted probability for the entire study area, which we defined at pixel level as 100% minus the percent probability of the pixel classification. We compared its values for individual forest types in box and whisker plots. Finally, we presented a map depicting the spatial distribution of the inverted classification probability.

Additionally, we analyzed how the time since the last field examination of our reference stands affected the results of our classification. In order to do that we performed logistic regression whether our classification at the stand level was correct or not (values of 1 and 0) and the number of years from last field visit for pure forest stands.

Results

We were able to classify forest types with high accuracy. Our classification accuracy at the stand level reach 86.6% for pure stands, and 74.6% for all stands when analyzing all available data from 2014-2016, and 84.2% for pure stands, and 73.6% for all stands based on data for 2016 alone (**Figure 1. 3**).

Test 1 - Missing data due to cloud cover and SLC scanning gaps

Several clear results emerged when we tested the value of including imagery with missing data in the classifications. First of all, when cloud-free imagery was available, i.e., included in our classifications, the addition of imagery with missing data improved classification accuracies only marginally (**Figure 1. 3d**). This was even the case when analyzing data for 2016 only, for which only 2 cloud-free images were available (**Figure 1. 3b**). Furthermore, once cloud-free images were included as input, there was only a minor improvement in classification accuracy when using data from 2014-2016 versus 2016 only.

However, when no cloud-free images were included (i.e., when simulating a situation where no cloud-free imagery is available) then having data for three years was clearly advantageous. With three years of data, classification accuracy remained high even when the best images had >=20% data gaps (**Figure 1. 3c**), and the eight best images were removed from the image stack (**Figure 1. 3a**). However, classification accuracies for imagery from 2016 alone dropped precipitously as soon as the cloud-free imagery was removed (**Figure 1. 3c**).

When comparing classifications for pure stands (**Figure 1. 3c-d**) with those for all stands (**Figure 1. 3e-f**), accuracies for pure stands alone were generally 10 percentage points higher than those for all stands, but the general patterns when including imagery with missing data were very similar.

Test 2 – number of acquisitions

Analyzing data for three years instead of only one improved image availability from 24 to 84 acquisitions, with some cloud-free pixels (**Figure 1. 3a-b**). However, the increase the number of

acquisitions varied depending on the cloud cover range and shrank to 12 and 4 for 0-30% and 0-10% cloudy imagery respectively.

Data from three years provided at least one cloud-free observation for each pixel even with imagery that had 60-100% missing values. In contrast for data from one year, imagery with 20-100% missing data was necessary. Second, the maximal number of observations per pixel was 59 considering all acquisitions from three-year period (**Figure 1. 1c**). The average number of cloud-free observations per pixel grew rapidly from <1 for the most clouded imagery (10 acquisitions missing >= 90% and < 100% pixel values) to 35 for data set containing all available imagery (0-100%) from years 2014-2016 and 9 for 2016 only. Finally, the overall level of missing data decreased gradually when we added less clouded imagery (**Figure 1. 3a**), and when we removed clouded imagery (**Figure 1. 3b**). In both cases the overall level of missing data was similar for both data set from 2014-2016 and 2016 only.

As mentioned above, the increase in classification accuracy due to additional data acquisitions was marginal when cloud-free imagery was included (**Figure 1. 3d**). Even with just four images with 0-30% missing data for 2016 alone, classification accuracy was 85.2%, compared to 86.5% for the 82 acquisitions for 2014-2016, and 86.4% for the 53 acquisitions for the data with 0-70% missing values, which was the highest mean accuracy. However, without the cloud-free imagery, differences between 2016-only versus 2014-16 data were stark, especially in the range of 10-100 to 40-100% missing values, where the three-year data sets with many more observations increased classification accuracy up to level of 76% (**Figure 1. 3c**). Results for pure and mixed stands were again similar in trends but lower in absolute values (**Figure 1. 3e-f**).

Test 3 – seasonal data distribution

The season during which the images were recorded affected classification accuracy greatly (**Figure 1. 4a-b**). Fall imagery was by far the most valuable for mapping forest types, providing by itself a level of accuracy just slightly lower than most of the combination of fall imagery with data from other seasons. For pure stands, we achieved the highest accuracy using combinations of all seasons when using the imagery from the three-year period 2014-2016 (86.1%) and from 2016 only (85.2%). In other words, as long as a fall image was available, then data from a single year performed almost as well as data from three years.

When we simulated situations where no fall image was available, then the difference in the classifications for 2016-only versus 2014-16 data was pronounced (**Figure 1. 4a**), presumably because of the large number of image acquisitions for the three-year period (**Figure 1. 4c**). When we tested the value of imagery for different seasons for pure versus all stands, the latter had again lower accuracy by 10-15 percentage points (**Figure 1. 4b**). The most useful imagery was again from fall, and as long as fall imagery was included, the 2014-2016 dataset resulted in only marginally better classifications (74.4%) than the 2016-only data (72.8%).

Accuracy assessment

The overall accuracy at the stand level weighted by the number of classes' pixel counts of our final classification within the boundaries of reference RECON stands was 86.6% for pure stands (88.1% when the least abundant classes are aggregated for purpose of presentation), 63.3% for mixed stands, and 74% for all stands. User's accuracy for individual forest types generally decreased with decreasing species acreage, but producer's accuracy showed no visible pattern. 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 almost 30 percentage points for most of the classes (**Figure 1. 5a**).

For pure stands of forest types that covered >=2% of the area (**Table 1. 2**), we obtained generally high user's accuracy: oak (93.4%), brush (93.2%), red pine (92.5%), northern hardwoods (91.3%), scrub oak (87.5%), swamp hardwoods (85.5%), jack pine (84.8%), aspen (83.9%), and black spruce (79.8). With exception of jack pine (79.4%), scrub oak (72.4%), swamp hardwoods (69.2%), and oak (46.9%), we achieved also high producer's accuracy for these classes: northern hardwoods (97%), aspen (95,7%), black spruce (91.7%), red pine (88.8%), and brush (87%). For the classes covering <2% of the area, we achieved low range accuracies (**Appendix 1. A**), missing them 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 check the overall accuracy of the map, because we did not have independent ground truth data to do so, we confirmed that the prediction probabilities were moderately high (**Figure 1. 1e**). For most of the mapped area, the inverted probability was lower than 60%, and for a large portion <10%. It varied among classes, but more abundant tree species were generally classified with greater probability (**Figure 1. 5b**). However, for most classes the maximal range of inverted probability ranged from 0 to 90%.

The relation between the time of last field examination of our reference forest stands and our classification varied among forest type. For example, we found no significant relation between correct classification of oak stands and the time when they were last examined (**Table 1. 3**). Whereas, for northern hardwoods such relation existed showing significant negative effect on correct stand classification (**Table 1. 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 (**Figure 1. 6**).

Discussion

We were able to map forest types accurately by analyzing Landsat imagery with missing values concomitantly with cloud-free images, and by applying the C5.0 decision trees algorithm, which can handle data with gaps. Our approach is thus an advancement over previous approaches such as image thresholding (Wolter et al., 1995s; Wolter et al., 2008; Dymond, Mladenoff, & Radeloff, 2002), traditional classification algorithms (e.g. Minimum Distance To Means Classifier, Mickelson, Civco, and Silander, 1998), or other machine learning algorithms (e.g. Random Forest Classifier, Zhu and Liu, 2014). Furthermore, our approach does not require gap filling, making it more efficient especially for large-area mapping. In general, we found that 1) forest types in temperate forests can be classified with accuracies in the range of 72.4%-97%, 2) that imagery from multiple years is crucial when cloud-free imagery is lacking, and 3) that data from fall is by far the most important for high classification accuracies.

When cloud-free imagery is lacking, we recommend to map forest types based on Landsat data from multiple years to increase the number of acquisitions, and to ensure that each pixel has cloud-free observations from different stages of phenology. However, when multiple cloud-free images are available for a single year, then that will probably suffice, as long as one image is from fall. Similar findings were reported for data from WorldView-2 where 7 out of 20 cloud free images provided already the maximal classification accuracy (Elatawneh et al., 2013). For tests with Landsat-like data, the number of images that provided satisfactory results ranged from three (Mickelson, Civco, and Silander, 1998), four (Dymond et al., 2002), seven (Zhu and Liu, 2014), 8 (Townsend et al., 2001), to ten clear images (Wolter et al., 1995). Pasquarella et al.

(2018) obtained similar results to ours using pairs of cloud-free imagery from multiple years, and up to 5% better accuracy using spectral and temporal indices derived from all available observations from 1985 to 2015. However, the highest improvement in that study was due to the addition of ancillary data, such as topography and wetland probability, which we did not test 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, the imagery from 2016 only, and for the three-year period 2014-2016, fall imagery 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. Our findings generally follow those from prior research identifying spring and fall imagery as the most useful for discriminating forest types (Wolter et al., 1995; Mickelson, Civco, and Silander, 1998), but summer imagery being useful if its quality is high (Elatawneh et al., 2013).

The high classification accuracy for pure forest stands provides strong support for of our approach 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 remotely sensed imagery. In total we used 2,786,133 pixels to validate the classification within reference forest stands. Validation datasets in prior studies were generally smaller, i.e., 370 (Bolstad and Lillesand, 1992), 1,211 (Wolter et al., 1995), 322 (Mickelson et al., 1998), 529 (Townsend et al., 2001), 528 (Dymond et al., 2002), 95 (Zhu and Liu, 2014), but Pasquarella et al., (2018) had 161,880 samples. Since we did not have access to detailed information tree species composition within each stand, we calculated classification accuracy based on the dominant tree species in each stand (Bolstad & Lillesand, 1992; Wolter et al., 1995; Dymond et

al., 2002; Zhu & Liu, 2014), as is typical for forest inventory maps. An alternative would be to use fuzzy logic to depict the actual species composition (Mickelson, Civco, and Silander, 1998; Townsend et al., 2001), but we did not pursue this. It is worth to mention that the RECON data set is highly unbalanced in terms of the class abundance, reflecting species distributions on the ground, which leads to better recognition of the more abundant forest types. In tests we were able to improve the accuracy for less abundant forest types by up to 13.3% by leveling the number of training samples per class (results not shown). 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 important implications for science, forest industry, and conservation planning, and provide a promising methodology for large scale, detailed, and accurate mapping of temperate forest types. First, our approach makes it possible to map forest types for large areas because it does not require cloud free images (Ju and Roy, 2008). Second, our maps are at 30-m resolution, thereby providing detailed input for modeling of various environmental variables (Richardson et al., 2012). Third, our maps could be used to update stand-level forest inventories and hence support forest management (Portoghesi, 2006) and conservation efforts.

We conclude that even Landsat imagery with missing observations due to cloud cover and SLC scanning gaps can be used for the operational mapping of the temperate forest types. We found that just a handful of fairly complete and cloud-free acquisitions combined with imagery with missing values, especially from fall, can result in accurate forest types maps, which represent a major advancement over prior approaches that required cloud-free imagery.

Acknowledgments

The study was supported by the USDA McIntire-Stennis Program. The authors thank Eric Kruger, Annemarie Schneider, Philip Townsend from University of Wisconsin-Madison for ideas and constructive feedback at various stages of work on this manuscript; and S.P. Engle from The CALS Statistical Consulting Group at the University of Wisconsin-Madison for valuable discussions and feedback.

References

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, 1–15. 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, 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 (80-.). 320, 1444–1449. https://doi.org/10.1126/science.1155121

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

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

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

Department of the Interior U.S. Geological Survey, 2017b. Product Guide: LANDSAT 8 SURFACE REFLECTANCE CODE (LASRC) PRODUCT 27. https://doi.org/10.1080/1073161X.1994.10467258

Department of the Interior U.S. Geological Survey, 2012. Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description. Open-file Rep. 2013-1057 1–20. https://doi.org/No. 2013-1057

Dymond, C.C., Mladenoff, D.J., Radeloff, V.C., 2002. Phenological differences in Tasseled Cap indices improve deciduous forest classification. Remote Sens. Environ. 80, 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. 5. RESA Work. 4/2013 1–10.

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

Fassnacht, F.E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Kochb, 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, 2547–2561.

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.

Hanewinkel, M., Cullmann, D.A., Schelhaas, M.-J., Nabuurs, G.-J., Zimmermann, N.E., 2012. Climate change may cause severe loss in the economic value of European forest land. Nat. Clim. Chang. 3, 203–207. https://doi.org/10.1038/nclimate1687

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, 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. Pe & Rs. Photogramm. Eng. Remote Sensing 81, 345–354.

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, 1196–1211. https://doi.org/10.1016/j.rse.2007.08.011

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

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, 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.

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

Portoghesi, L., 2006. European Forest Types. Categories and types for sustainable forest management reporting and policy, Forest@ - Rivista di Selvicoltura ed Ecologia Forestale. https://doi.org/10.3832/efor0425-003

Radeloff, V.C., Mladenoff, D.J., He, H.S., Boyce, M.S., 1999. Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. Can. J. For. Res. 29, 1649–1659. https://doi.org/10.1139/x99-089

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, 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, 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, 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

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

Schulte, L.A., Mladenoff, D.J., Nordheim, E. V, 2002. Quantitative classification of a historic northern Wisconsin (U.S.A.) landscape: mapping forests at regional scales. Can. J. For. Res. 32, 1616–1638. https://doi.org/10.1139/x02-082

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

State of Wisconsin Department of Natural Resources, 2013. Public Forest Lands Handbook.

Townsend, P. a, Walsh, S.J., Ecology, S.P., Dec, N., Townsend, a, Stephen, J., J, S., 2001. Remote Sensing of Forested Wetlands : Application of Multitemporal and Multispectral Satellite Imagery to Determine Plant Community Composition and Structure in Southeastern Stable URL : http://www.jstor.org/stable/20051167 and sensing of forested wetland. Plant Ecol. 157, 129–149. https://doi.org/10.1023/A:1013999513172

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.

Wolter, P.T., Townsend, P.A., Sturtevant, B.R., Kingdon, C.C., 2008. Remote sensing of the distribution and abundance of host species for spruce budworm in Northern Minnesota and Ontario. Remote Sens. Environ. 112, 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

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

List of tables

Table 1. 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

Table 1. 2. The error matrix for the classification at the stand level weighted by the number of pixel 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 cloud cover ranging from 0% up to 70%.

	ASPEN	BLACK SPRUCE	JACK PINE	NORTHERN HARDWOODS	OAK	RED PINE	SCRUB OAK	SWAMP HARDWOODS	BRUSH	OTHER	***Row sum***	***User's accuracy (%)***
ASPEN	**330311**	1538	2381	11945	9248	2381	4220	7319	12086	12224	393653	**83.9**
BLACK SPRUCE	357	**35567**	40	84	-	18	16	47	1469	6961	44559	**79.8**
JACK PINE	201	1	**34068**	9	4	5217	365	-	204	112	40181	**84.8**
NORTHERN HARDWOODS	6788	348	6	**440618**	9454	629	146	8202	2354	14123	482668	**91.3**
OAK	355	10	111	638	**17809**	34	8	9	74	23	19071	**93.4**
RED PINE	373	41	4094	26	11	**73640**	88	10	123	1184	79590	**92.5**
SCRUB OAK	501	1	561	-	76	194	**13535**	2	526	81	15477	**87.5**
SWAMP HARDWOODS	589	165	11	390	1365	7	31	**39860**	2138	2065	46621	**85.5**
BRUSH	1101	509	1642	191	10	778	286	1990	**130755**	3045	140307	**93.2**
OTHER	4494	623	5	252	-	5	3	203	527	**9958**	16070	**62.0**
Column sum	345070	38803	42919	454153	37977	82903	18698	57642	150256	49776	***Overall Accuracy***	**88.1**
Producer's accuracy (%)	**95.7**	**91.7**	**79.4**	**97.0**	**46.9**	**88.8**	**72.4**	**69.2**	**87.0**	**20.0**		

Table 1. 3. Logistic regression table between correctly classified stands and numbers of years from last field examination for pure oak forest stands

Coefficients:

Model	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 1. 4. Logistic regression table between correctly classified stands and numbers of years from last field examination for pure north hardwoods forest stands

Coefficients:

Model	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

List of appendices

Appendix 1. A. The full error matrix for the classification at the stand level weighted by the number of pixel 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 cloud cover ranging from 0% up to 70%.

	ASPEN	BALSAM FIR	BLACK SPRUCE	BOTTOMLAND HARDWOODS	FIR SPRUCE	HEMLOCK	JACK PINE	LOWLAND BRUSH	LOWLAND BRUSH ALDER
ASPEN	**330311**	446	1538	788	353	468	2381	5388	5181
BALSAM FIR	-	**-**	-	-	-	-	-	-	-
BLACK SPRUCE	357	108	**35567**	3	155	3	40	287	1180
BOTTOMLAND HARDWOODS	-	-	-	**-**	-	-	-	-	-
FIR SPRUCE	-	-	-	-	**-**	-	-	-	-
HEMLOCK	-	-	-	-	-	**-**	-	-	-
JACK PINE	201	-	1	-	13	-	**34068**	28	3
LOWLAND BRUSH	209	-	278	170	-	-	10	**52432**	5418
LOWLAND BRUSH ALDER	837	17	229	-	5	-	206	10464	**45103**
LOWLAND BRUSH WILLOW	-	-	-	-	-	-	-	-	-
NORTHERN HARDWOODS	6788	13	348	68	501	350	6	609	1446
OAK	355	-	10	-	-	-	111	36	18
RED MAPLE	-	-	-	-	-	-	-	-	-
RED PINE	373	-	41	-	159	-	4094	40	24
SCRUB OAK	501	-	1	-	-	-	561	28	22
SWAMP CONIFER	2	-	61	-	5	-	-	10	42
SWAMP HARDWOODS	589	7	165	71	97	33	11	910	1112
TAMARACK	4475	-	534	-	-	-	5	24	224
UPLAND BRUSH	55	-	2	-	-	-	1426	-	9
WHITE BIRCH	-	-	-	-	-	-	-	-	-
WHITE CEDAR	17	5	28	-	19	-	-	17	44
WHITE PINE	-	-	-	-	-	-	-	-	1
WHITE SPRUCE	-	-	-	-	-	-	-	-	-
C.sum	345070	596	38803	1100	1307	854	42919	70273	59827
Producer's accuracy (%)	**95.7**	**0.0**	**91.7**	**0.0**	**0.0**	**0.0**	**79.4**	**74.6**	**75.4**

	LOWLAND BRUSH WILLOW	NORTHERN HARDWOODS	OAK	RED MAPLE	RED PINE	SCRUB OAK	SWAMP CONIFER	SWAMP HARDWOODS	TAMARACK
ASPEN	19	11945	9248	1326	2381	4220	686	7319	1410
BALSAM FIR	-	-	-	-	-	-	-	-	-
BLACK SPRUCE	1	84	-	3	18	16	2231	47	3707
BOTTOMLAND HARDWOODS	-	-	-	-	-	-	-	-	-
FIR SPRUCE	-	-	-	-	-	-	-	-	-
HEMLOCK	-	-	-	-	-	-	-	-	-
JACK PINE	-	9	4	-	5217	365	-	-	8
LOWLAND BRUSH	742	55	9	-	35	214	113	746	658
LOWLAND BRUSH ALDER	-	133	1	227	21	1	464	1243	775
LOWLAND BRUSH WILLOW	**-**	-	-	-	-	-	-	-	-
NORTHERN HARDWOODS	1	**440618**	9454	8276	629	146	2748	8202	103
OAK	-	638	**17809**	-	34	8	1	9	6
RED MAPLE	-	56	-	**-**	-	1	-	1	-
RED PINE	-	26	11	-	**73640**	88	6	10	3
SCRUB OAK	-	-	76	-	194	**13535**	-	2	19
SWAMP CONIFER	-	65	-	-	1	-	**2389**	152	-
SWAMP HARDWOODS	-	390	1365	22	7	31	741	**39860**	117
TAMARACK	-	131	-	4	-	2	-	14	**887**
UPLAND BRUSH	-	3	-	-	722	71	-	1	-
WHITE BIRCH	-	-	-	-	-	-	-	4	-
WHITE CEDAR	-	-	-	-	-	-	400	32	-
WHITE PINE	-	-	-	-	4	-	-	-	-
WHITE SPRUCE	-	-	-	-	-	-	-	-	-
C.sum	763	454153	37977	9858	82903	18698	9779	57642	7693
Producer's accuracy (%)	**0.0**	**97.0**	**46.9**	**0.0**	**88.8**	**72.4**	**24.4**	**69.2**	**11.5**

	UPLAND BRUSH	WHITE BIRCH	WHITE CEDAR	WHITE PINE	WHITE SPRUCE	***R.sum***	***User's accuracy (%)***
ASPEN	1498	1908	845	3099	895	393653	**83.9**
BALSAM FIR	-	-	-	-	-	-	**-**
BLACK SPRUCE	1	10	564	3	174	44559	**79.8**
BOTTOMLAND HARDWOODS	-	-	-	-	-	-	**-**
FIR SPRUCE	-	-	-	-	-	-	**-**
HEMLOCK	-	-	-	-	-	-	**-**
JACK PINE	173	-	-	65	26	40181	**84.8**
LOWLAND BRUSH	567	11	44	1	2	61714	**85.0**
LOWLAND BRUSH ALDER	199	-	469	5	84	60483	**74.6**
LOWLAND BRUSH WILLOW	-	-	-	-	-	-	**-**
NORTHERN HARDWOODS	298	916	923	54	171	482668	**91.3**
OAK	20	12	-	4	-	19071	**93.4**
RED MAPLE	-	-	-	-	48	106	**0.0**
RED PINE	59	26	20	929	41	79590	**92.5**
SCRUB OAK	476	52	-	10	-	15477	**87.5**
SWAMP CONIFER	-	-	1838	-	1	4566	**52.3**
SWAMP HARDWOODS	116	-	962	5	10	46621	**85.5**
TAMARACK	165	-	-	-	45	6510	**13.6**
UPLAND BRUSH	**15821**	-	-	-	-	18110	**87.4**
WHITE BIRCH	-	**-**	-	-	-	4	**0.0**
WHITE CEDAR	-	-	**4011**	-	-	4573	**87.7**
WHITE PINE	-	-	-	**291**	-	296	**98.3**
WHITE SPRUCE	-	-	-	-	**15**	15	**100.0**
C.sum	19393	2935	9676	4466	1512	***Overall***	***86.6***
Producer's accuracy (%)	**81.6**	**0.0**	**41.5**	**6.5**	**1.0**	***Accuracy***	

Appendix 1. B. Tree species composition included in 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	silver maple (*Acer saccharinum*), green ash (*Fraxinus pennsylvanica*), swamp white oak (*Quercus bicolor*), American elm (*Ulmus americana*), river birch (*Betula nigra*), cottonwood (*Populus deltoides*)
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*), red maple (*Acer rubrum*), silver maple (*Acer saccharinum*), swamp white oak (*Quercus bicolor*), American elm (*Ulmus americana*)
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*)

List of figures

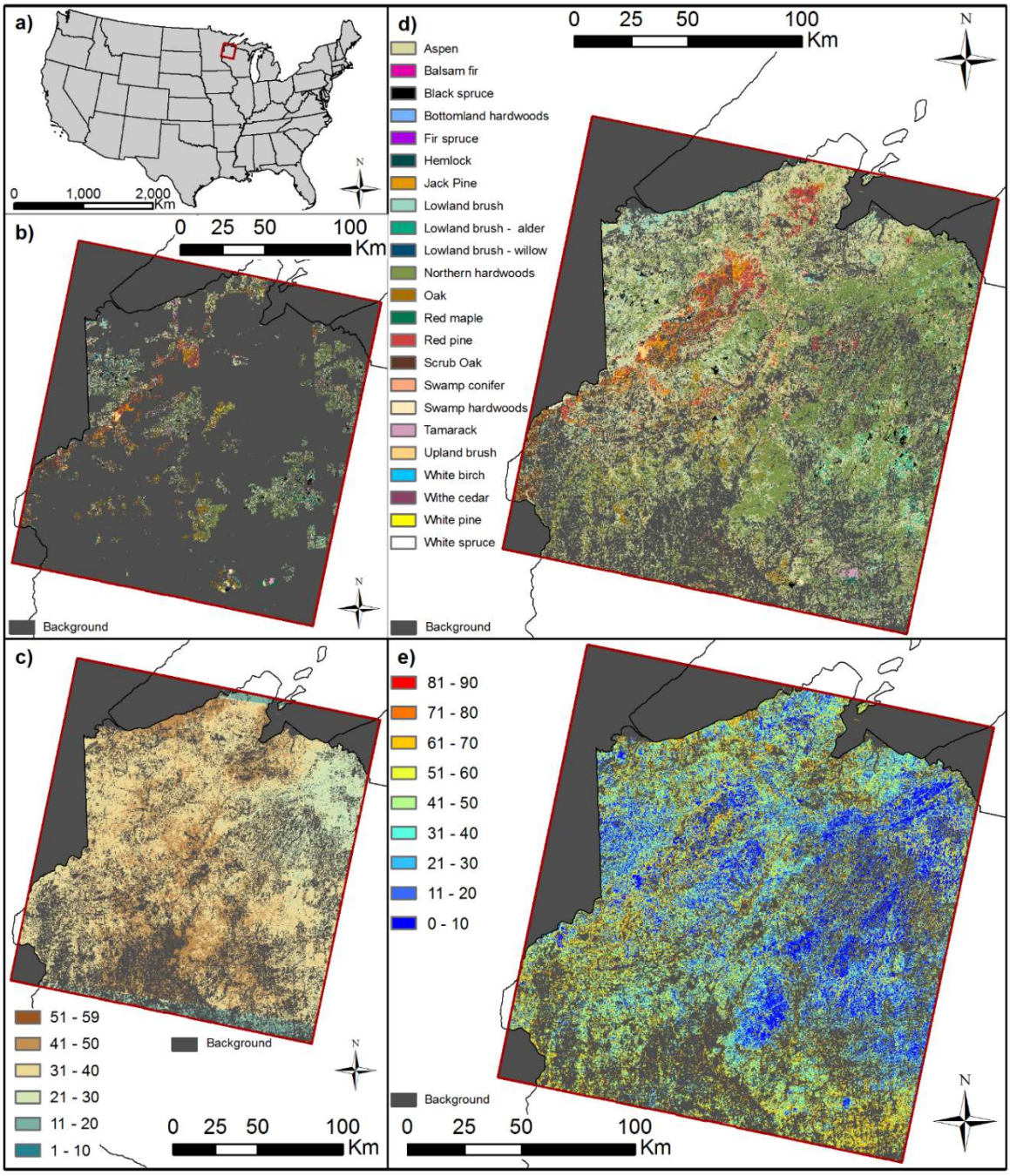

Figure 1. 1 Study area a) location of the footprint path: 26 row: 28, b) the forest types in RECON polygons, c) number of cloud-free observations per forest pixels in period 2014-2016, d) classification of tree species, e) inverted classification probability.

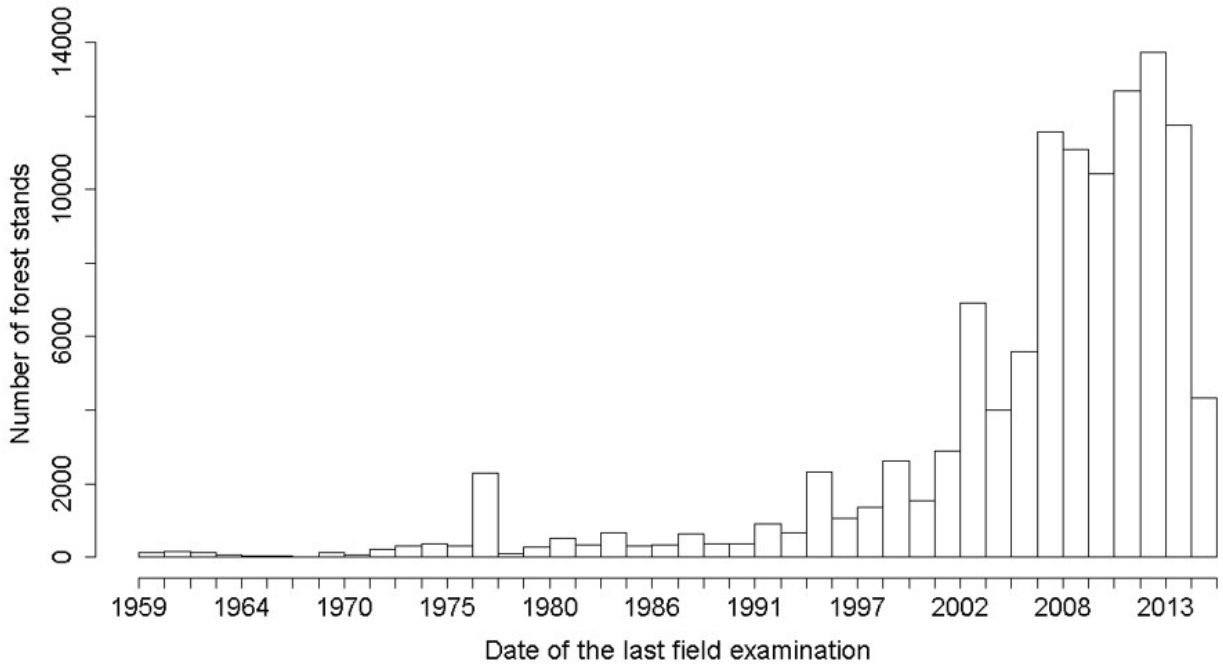

Figure 1. 2. Dates of last field examination for forest stands in the reference data set.

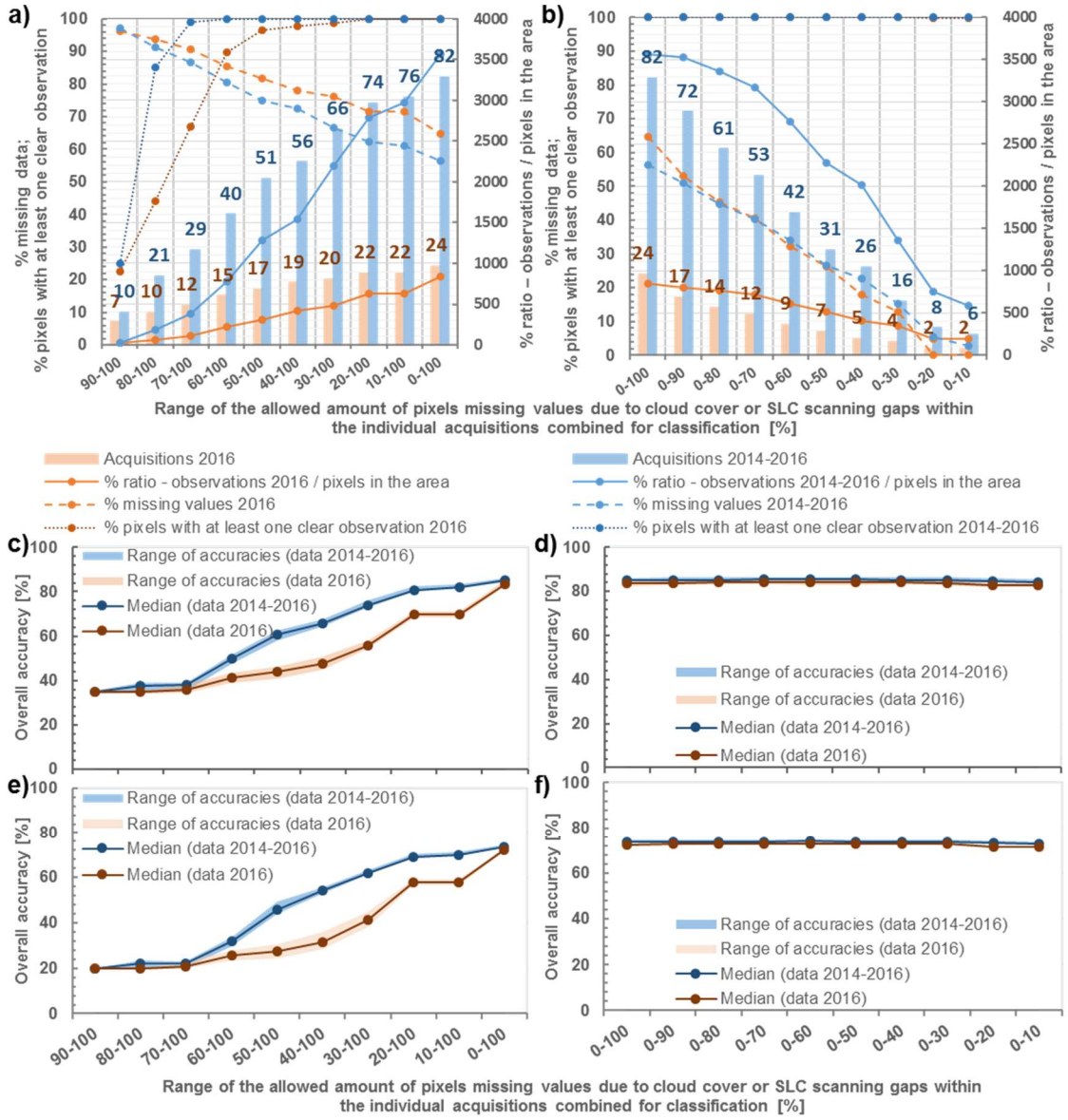

Figure 1. 3. The data availability and classification accuracy for tests of mapping forest types with image stacks containing imagery from 2016 and 2014-2016 characterized by various percentage of pixels with missing values: a) data availability along the addition of imagery containing less missing values, b) data availability along the removal of imagery with large percentage of missing values, c) classification accuracy along the addition of imagery containing less missing values – pure forest stands, d) classification accuracy along the removal of imagery with large percentage of missing values – pure forest stands, e) classification accuracy along the addition of imagery containing less missing values – pure and mixed stands, f) classification accuracy along the removal of imagery with large percentage of missing values – pure and mixed stands.

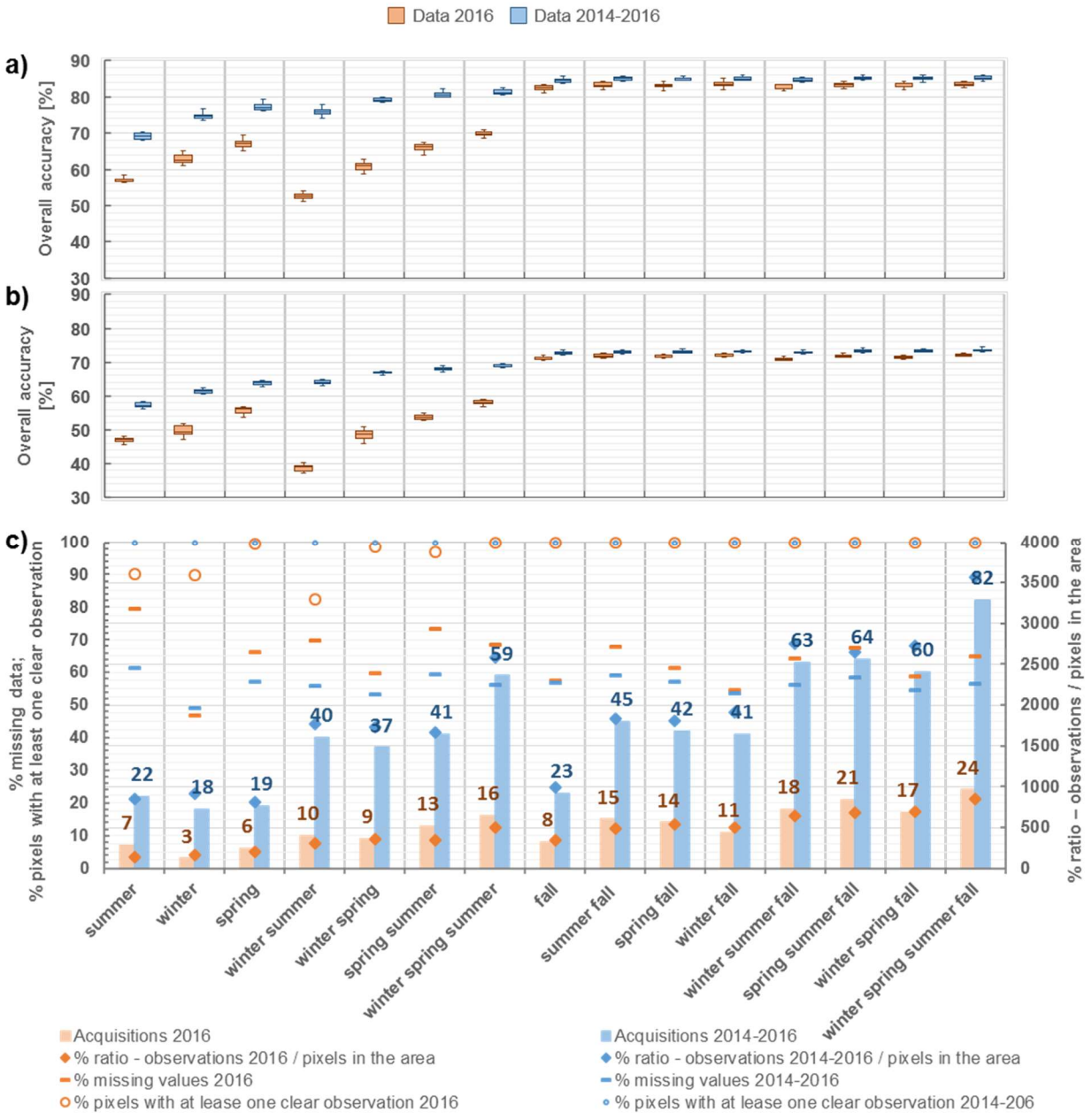

Figure 1. 4. The classification accuracy (a – pure forest stands, b – pure and mixed stands) and data availability (c) for tests of mapping forest types with image stacks containing imagery from various seasonal combinations from 2016 and 2014-2016.

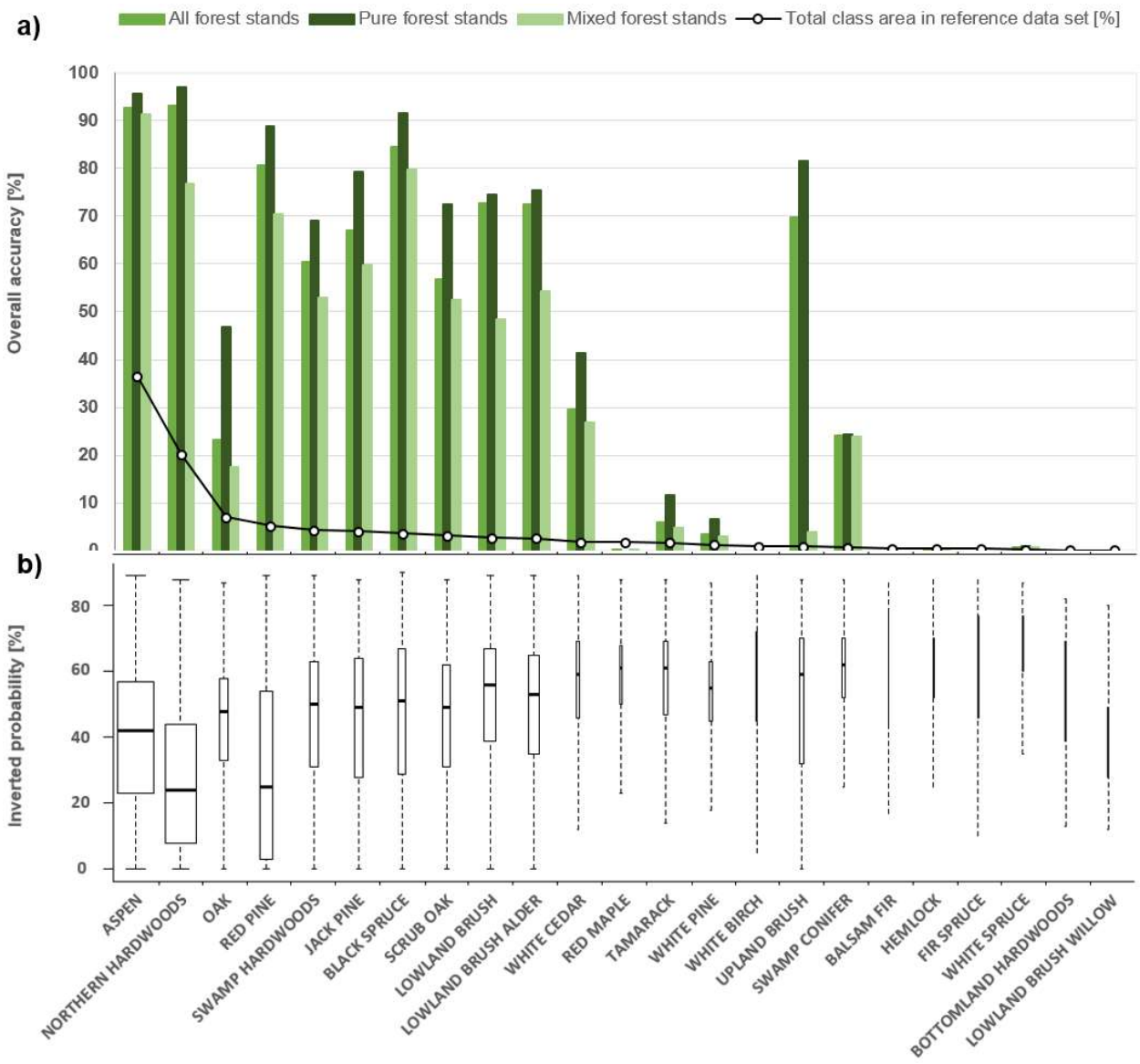

Figure 1. 5. The accuracy assessment: a) the agreement between classification and the RECON data set - producer's accuracy for forest types for pure, mixed and all stands; b) the inverted probability of pixel classification for tree types in the final map, the width of the boxes depicts the variability in the number of the pixels belonging to classes in the final map.

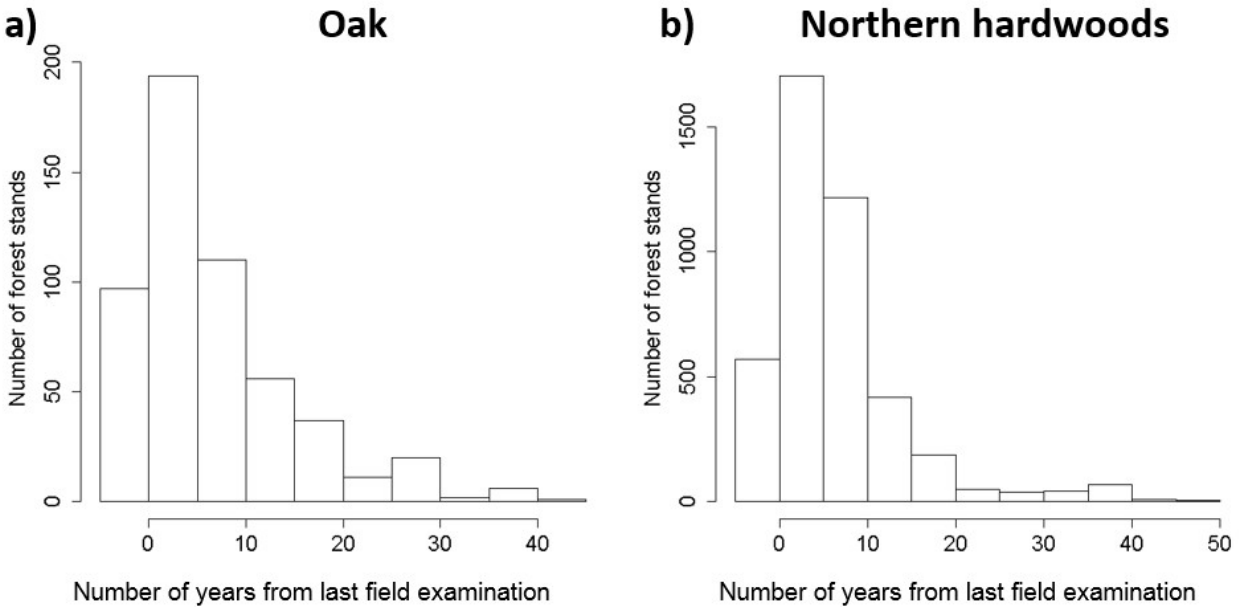

Figure 1. 6. Number of years from last field examination for pure reference forest stands of a) oak and b) northern hardwoods forest types.

Chapter 2 – Fusing MODIS and Landsat data for analyzing phenology and mapping forest types

Contributors: Konrad Turlej[1], Feng Gao[2], Mutlu Ozdogan[1], Volker C. Radeloff[1]

[1] SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison

[2] USDA-ARS Hydrology and Remote Sensing Laboratory

Abstract

Forest phenology greatly affects various ecological processes occurring both at local and global scales including carbon sequestration and therefore climate change. However, tracking annual phenology cycles is difficult at medium resolutions (i.e., around 30 m). Our goal was to generate an 8-day time series of cloud-free surface reflectance product that can capture phenology for a single year at 30 m resolution. In order to achieve our goal, we combined satellite observations from different types of sensors because Landsat data alone was too infrequent. We composited the available cloud-free Landsat observations with synthetic data generated by fusing Landsat imagery with MODIS imagery using the STARFM algorithm. In contrast to the standard application of STARFM, which pairs only one or two cloud-free Landsat acquisitions, we applied STARFM to all available Landsat images, even those that were incomplete due to cloud cover and gaps resulted from the Scanning Line Corrector (SLC) malfunction, and composited the resulting, partially-complete synthetic images with four different compositing algorithms. We created high-quality data on surface reflectance, which allowed us to generate a smooth time series of Enhanced Vegetation Index (EVI) that reflected forest phenology. STARFM predictions for surface reflectance were fairly consistent during the growing season (standard deviation of surface reflectance < 0.05) but varied considerably in winter (up to 0.15). However, the resulting

dense time series of actual and synthetic Landsat images did not improve forest type classifications compared with those based only on actual Landsat images. Our application of STARFM has two main advantages over its standard implementation: 1) it retains all available Landsat acquisitions, and 2) it can generate pixel-level, single-year phenology curves for large areas with low data availability and no cloud-free acquisitions, where both standard applications of STARFM and per-pixel modelling fail.

Keywords

Forest type, phenology, Landsat, MODIS, STARFM, image fusion, cloud cover, Wisconsin, USA

Introduction

Information on annual forest green leaf phenology is important for modeling various ecological processes including carbon sequestration and climate change. As an adaptation to annual seasonal cycle (Lechowicz, 1984), the spatial and temporal patterns of forest phenology are controlled by climate, day-length, species type and age, as well as substrate (Schaber & Badeck, 2003; Wolfe et al., 2005). The growing season length, in turn, is a key determinant of the Carbon Uptake Period (CUP), i.e., the period when trees are a net carbon sink, and thus can approximate the annual carbon exchange (Churkina et al., 2005). At a local scale, detailed data on tree phenology, and its spatial variability is useful information for modelling food availability for wildlife in the forest ecosystems, for example migrating insectivorous birds (Burger et al., 2012). This relates directly to the quality of forest habitats because the abundance of certain insects supporting insectivorous birds correlates with the variability in spatial pattern of phenology (Jones et al., 2003). Last but not least, key phenology dates vary among tree species and, when

combined with information on trees' reflectance characteristics, can improve forest type classifications based on remotely sensed imagery (Pasquarella et al., 2018).

Remote sensing is a source of information on forest phenology at global, regional, and local scales (Fisher and Mustard, 2007; Helman, 2018). Because phenology advances rapidly, especially during green-up, most phenology studies based on remote sensing employ satellite systems that provide frequent observations, such as MODIS on Terra/Aqua, AVHRR on NOAA's satellites, and Sentinel-3, which provide 250-1000-m data. Based on dense time-series of these satellite data, often converted into a vegetation index it is then possible to measure key phenology dates such as the 1) start of the season, 2) maturity, 3) total seasonal greenness, 4) senescence, 5) end of the season, and 6) dormancy (Elmore et al., 2012).

However, characterizing single year forest phenology over large areas at moderate resolutions, i.e., around 30-m pixel size, is still difficult because phenology information at moderate spatial resolution is necessary for numerous ecological applications, and generally correlates better with on-the-ground measurements (Fisher et al., 2006). The problem is that there are rarely enough satellite observations for a single year at 30-m resolution because of the lower temporal resolution of satellites such as Landsat, and frequent cloud cover. To some extent, the availability of two Landsat and Sentinel-2 satellites has alleviated this problem and increased the data availability in recent years (Claverie and Masek, 2016), but this does not allow for assessment of changes in phenology for the pre-Sentinel-2 era, and even in the current record, problems with clouds and gaps resulted by malfunction of Scanning Line Corrector remain for the most cloudy areas (Ju and Roy, 2008).

One option to overcome the lack of sufficient observations for a single year is to combine acquisitions from several years to assess an averaged multi-year phenology. Such a cumulative

multi-year satellite dataset is typically organized according to the day of the year when the imagery was acquired, considerably increasing the number of per-pixel observations (Elmore et al., 2012; Melaas et al., 2016). A refinement of this approach is to calibrate the date of the Landsat data based on MODIS-derived phenology (Isaacson et al., 2012; Baumann et al., 2017), to account, for example, for an early spring in one year. However, any assessment of phenology based on multiple years of satellite data requires no changes in land cover among the years of the satellite observations, which is especially problematic when analyzing agricultural areas where crops differ from year to year.

An alternative to the analysis of multi-year data is to generate additional, synthetic Landsat-style observations. Indeed, when one or two cloud-free Landsat acquisitions are available then single-year phenology can be obtained at 30-m resolution by fusing imagery from infrequent sensors, such as Landsat, with imagery from coarse-resolution by temporally frequent sensors, such as MODIS. There are several algorithms that can combine cloud-free Landsat and MODIS imagery as long as there is at least one (e.g. STARFM - Spatial and Temporal Adaptive Reflectance Fusion Model (Gao et al., 2006)) or two (e.g. STAARCH - Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (Hilker et al., 2009); ESTARFM -Enhanced STARFM (Zhu et al., 2010)) pairs of Landsat-MODIS images. If such pairs of imagery are available, then it is possible to generate synthetic Landscape observations for any date and area for which cloud-free MODIS data is available. However, the application of STARFM or similar algorithms has two shortcomings. First of all, they require at least one cloud-free Landsat image, which is not always available. Second, once such an image is available, the observations from other Landsat images, either cloud-free or partially clouded, are not retained. Assuming that the original satellite observations are better than synthetic predictions (Senf et al., 2015), and given

that there are typically many images available that provide partial coverage, that second limitation seems particularly unfortunate.

Therefore, our goal was to generate a single-year, 30-m resolution time-series of EVI that captures single-year phenology even when no cloud-free acquisitions are available and that retain all available Landsat observations. We developed a new approach to fuse Landsat and MODIS data based on the STARFM algorithm. We had two specific objectives: 1) to composite a time series of surface reflectance from both Landsat and MODIS data, and 2) to test if that time series improves forest types classifications.

Methods

Study area

Our study area was one Landsat footprint (path 26, row 28) located in northern Wisconsin, USA (**Figure 2. 1**). We selected it due to its high spatial heterogeneity, and considerable variability of climatic conditions, which create a challenging environment to assess forest phenology.

Geographically, our study area includes parts of three distinct ecoregions: the Central Plain in the east, the Northern Highlands in the west, and the Lake Superior Lowland in the north (Martin, 1965). The topography of the area is flat, consisting mostly of rolling plains stemming from Pleistocene and its last glacial expansion, and creating no natural obstacles for air masses that cause the major zonal distribution of the vegetation (Curtis, 1959). Local climatology is characterized by strong seasonality, considerable inter-annual variability, and climatic changes. There is considerable inter-annual variability in phenology during spring and fall. Winter conditions depend largely on climatological teleconnections, especially El Niño (warm years) and La Niña (cold years) (Trenberth et al., 1998; Mcphaden et al., 2014). In the northern part of

the region seasonality is also influenced by Lake Superior, which creates local temperature gradients due to lake effects. In addition, the study area has been affected by climate change over the last 30 years including regional warming, cooler springs, considerable fall warming, and an average increase in the length of the growing season by +0.27 days/year (Garcia and Townsend, 2016).

The region is characterized by spatially heterogeneous forests. The most abundant trees species are aspen spp. (*Populus*), oak spp. (*Quercus*), pine spp. (*Pinus*), spruce spp. (*Picea*), maple spp. (*Acer*), accompanied by tamarack (*Larix larcina*), eastern hemlock (*Tsuga canadensis*), and white cedar (*Thuja occidentalis*). Current tree species distributions differ substantially from those prior to European settlement in the mid-19th century. Previously, pines dominated forests on sandy outwash soils (Curtis, 1959), sometimes accompanied by aspen and oaks in forest-savannas (Schulte et al., 2002), whereas the till plains and mesic moraines were mostly dominated by hemlock and northern hardwood species (Curtis, 1959; Radeloff et al., 1999; Rhemtulla et al., 2007; Rhemtulla et al., 2009). Current tree species distributions resulted from intensive logging and land conversion to croplands accompanied by insect outbreaks (McCullough et al., 1998) and severe fires (Schulete and Mladenoff, 2005). Forest have regrown in many areas but with novel species composition (Goring et al., 2015; Goring et al., 2016).

Image data

We combined 1) the USGS Collection-1 surface reflectance data from Landsat-7, and -8, and 2) 16-day MODIS BRDF adjusted surface reflectance data (MCD43A4). For each data set, we obtained all available acquisitions for 2016.

Landsat

The USGS Collection-1 surface reflectance is processed via the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) for Landsat-7 (Department of the Interior U.S. Geological Survey, 2012, Department of the Interior U.S. Geological Survey, 2017a) and the Landsat Surface Reflectance Code (LaSRC) for Landsat-8 imagery (Department of the Interior U.S. Geological Survey, 2017b) both of which apply radiometric, atmospheric, and geometric image corrections. We also used the information on clouds and cloud-shadows generated via the F-mask algorithm, which is embedded in the product's quality band.

MODIS

We analyzed 16-day MODIS reflectance (MCD43A4), which is a Nadir Bidirectional Reflectance Distribution Function (BRDF) adjusted product containing 16-day image composites for MODIS bands 1-4 and 6-7 that correspond to the Landsat spectral bands. The data are provided every 8 days at 500-m resolution. In total, we had 46 MODIS composites for 2016. We preprocessed the data to STARFM requirements, which included the reprojection from native sinusoidal to Landsat's UTM, resampling to 30-meters spatial resolution, and cropping to the Landsat footprint boundaries.

Reference data on tree species composition

For our classifications, we used the Reconnaissance forest inventory data (RECON) as training and validation data, which is available via the Wisconsin Forest Inventory & Reporting System (WisFIRS) of the Wisconsin Department of Natural Resources. The data set contains forest inventory information for forest stands, and a wide range of forest characteristics including the primary tree type defined as tree species representing ≥ 50% of the stand's basal area, the secondary tree type, understory type, the year of stand establishment, year of last field

examination, and information on trees' height, density, and basal area. Based on that information we divided stands into pure and mixed stands. We defined pure stands as those for which the primary tree species, secondary species, and understory type was the same, or for which only the primary tree species was specified. We defined all other stands as mixed.

Predicting surface reflectance via STARFM

We generated the STARFM predictions for each of the 46 dates for which the MODIS MCD43A4 product provides reflectance data. In total, we made 18 STARFM predictions for each of the 46 dates, where each prediction was based on a different, unique Landsat-MODIS pair of acquisitions. We made predictions based on all available Landsat acquisitions containing >= 20% of cloud-free pixels, and made STARFM predictions for all pixels that were cloud-free in both the 18 Landsat and MODIS image pairs, and cloud-free in the 46 MODIS images for which the predictions were performed.

Compositing Landsat and STARFM data

In order to generate a dense time series of 30-m reflectance data for each pixel, we composited the original Landsat data and STARFM predictions (**Figure 2. 2**). We evaluated 2 temporal schemes and 2 compositing schemes. Our temporal schemes were: 1) the Landsat acquisition scheme where we filled gaps in all of the 44 original Landsat acquisitions from 2016 with composited STARFM predictions (i.e., including Landsat dates which were entirely cloudy), and 2) the MODIS MCD43A4 dates where we composited the Landsat data and STARFM predictions for each of the 46 dates. In both temporal schemes we retained all available Landsat observations and used STARFM predictions only for the areas where Landsat data were missing. The advantage of the first approach is that it matches the original Landsat dates exactly. The advantage of the second approach lies in harmonizing data to the global MODIS 8-day scheme,

which makes it possible to generate large-area phenology datasets with consistent dates, since the Landsat dates differ among neighboring paths.

Our image compositing schemes were: 1) selection of the predictions generated with the Landsat-MODIS input pair acquired closest to the prediction date, and 2) the mean value from all predictions available for a given period. The advantage of the first scheme is that STARFM predictions based on the closest Landsat-MODIS pair should be most accurate (Wang et al., 2014). The advantage of the second approach is that the mean accounts for variability among predictions generated with STARFM models based on different Landsat-MODIS image pairs.

Evaluation of the quality of STARFM predictions and compositing process

We examined the variability of STARFM predictions by quantifying how much the predictions based on different Landsat-MODIS image pairs diverged. To do this we analyzed the distributions (standard deviation) of the predictions generated by our models. We calculated these standard deviations at the pixel level for each band and each date.

Classification and accuracy assessment

Our second objective was to see if our time series improve classifications of forest type compared to classifications using only the original Landsat data. For this part of our analysis we examined only areas identified as forests in the National Land Cover Database (NLCD 2011). As input for the classifications, we used our composited data sets of Landsat/STARFM surface reflectance. We tested the data generated with all four compositing approaches. We performed cross-validation classifying each composited data set ten times, training our models with randomly selected subsets of training and validation areas. Each time we drew 25% of the pure

forest stands of each tree species for training, and left the remaining 75% of the pure stands plus all mixed stands for validation.

We used the Trees and Rule-Based Models from the C5.0 package, implemented in R statistical software, to classify our imagery (Quinlan, 1993; Friedl and Brodley, 1997; DeFries and Chan, 2000). The algorithm has gradient boosting capabilities that improve learning and it can handle incomplete datasets (Farhangfar et al., 2008). For each of our models we conducted 100 boosting iterations, the highest possible setting in C5.0, because it improved the classification accuracy.

We calculated the area-weighted overall accuracies for all trials of our classifications, and present class-level accuracies for the model that achieved the best overall accuracy. We calculated user's and producer's accuracy (Congalton, 1991; Foody, 2002) separately for three types of validation data sets: 1) pure forest stands only, 2) mixed forest stands only, and 3) both pure and mixed forest stands together. Each time, we excluded edge pixels within 30 m of the stand border. We aggregated our classifications to forest stands and weighted the class area with the number of pixels belonging to each stand. In total, we analyzed 2,786,133 pixels belonging to 44,741 forest stands.

Results

Compositing Landsat and STARFM data

Our compositing of actual Landsat and STARFM-predicted Landsat-style data resulted in a nearly complete time series of surface reflectance. By filling the data gaps in the original Landsat acquisitions from 2016 with STARFM predictions, we increased the number of images providing at least some information for parts of the image from 24 to 44, i.e., to all dates for which Landsat acquisitions were made (**Figure 2. 3**). Furthermore, after compositing, we had 28

dates for which there was surface reflectance information for >90% of the footprint. Gaps in the MODIS imagery, especially in winter images, were the cause for missing predictions in the remaining dates.

Variability between synthetic STARFM time series

Two results emerged from our analysis of the variability of the STARFM predictions based on the 18 Landsat-MODIS image pairs (**Figure 2. 4**). First, variability among predictions was generally low, especially during spring, summer, and fall (SD of surface reflectance <0.05), but higher during winter (SD of surface reflectance up to 0.15). Second, variability among predictions was most pronounced for the visible spectrum: i.e., for bands 2-4 (Blue, Green, Red) SD varied from 0.02 to 0.15, but it was less prominent for the near infrared band (5), and virtually non-existent for the short-wave infra-red bands (6 and 7).

Phenology

The dense stack of composited Landsat/STARFM surface reflectance data resulted in a smooth time series of phenology information. There were three clear patterns when comparing the phenological curves resulting from our four compositing schemes. First, in general, our time-series were fairly similar. Second, the type of compositing scheme, i.e., based on the nearest date versus the mean value from all predictions, caused differences, especially at the start of the season and during senescence (**Figure 2. 5**). In most cases the curves generated based on the nearest date matched the original Landsat acquisitions better. In contrast, the curves generated in the "mean" mode smoothed the phenology curves, and tended to result in rapid changes in the EVI close to the original Landsat observations. This was especially striking when time series were based on a larger number of actual Landsat observations. Third, differences between the two temporal schemes (Landsat vs. MODIS) were generally smaller than those

based on closest date versus mean prediction. In spring, the MODIS temporal scheme showed earlier green-up in general, and in fall, later senescence on the order of approximately 4 days, equal to the 4-day shift between the temporal schemes of the Landsat and MODIS datasets. However, this trend occurred only for data predicted based on the compositing to the nearest date. Moreover, an artifact is that a single Landsat observation is sometimes used for two subsequent MODIS periods.

Our dataset allowed us to capture species-specific phenologies and environmental gradients. The phenology visible in averaged EVI time series follow the general order in which tree species shoot and shade leaves (**Figure 2. 6**). Birch has the earliest green-up followed by maple and oak species. We also noticed a considerable delay in phenology for stands located in the north (**Figure 2. 7**), as well in for the stands located near the shore of Lake Superior (**Figure 2. 8**).

Classification

The results of our classifications provided us with overall accuracy >80% for pure forest stands of 23 forest types (**Figure 2. 9**). None of the four compositing schemes ultimately resulted in a classification accuracy exceeding that of the original Landsat images by themselves, which was 88.1% **(Chapter 1)**. We achieved the best results with the data stack of original gap filled Landsat acquisitions corresponding to the growing season (max overall accuracy of 83.5%) followed by the data set composited to the MODIS temporal scheme based on the closest date (max overall accuracy of 83.3%). The results for mixed forest stands followed the same trends compared to pure stands, but had substantially lower accuracy (around 60% for mixed forest stands, and 70% for all stands together).

Discussion

We successfully generated 30-m resolution time-series of surface reflectance based on satellite data for a single year based on the combination of Landsat and MODIS data. Compared to prior approaches, our time series have the advantage that they retain all available Landsat observations. The standard application of STARFM (Gao et al., 2006) and most of its improvements (Zhu et al., 2010; Wang et al., 2014), or modifications (Hilker et al., 2009; Weng et al., 2014) rely only on one or two cloud-free Landsat acquisitions and discard the other potential acquisition during the year, which can provide valuable data even when these images have only partial coverage.

Our approach can successfully perform for the areas that are often cloudy, because it does not require a fully cloud-free Landsat acquisition. We found that the STARFM predictions based on different Landsat-MODIS input image pairs were generally very similar for the growing season, but not for winter. This suggests that it is valid to composite the predictions coming from various STARFM models in order to cover larger areas for which the separate models cannot generate predictions. As long as all the partially-cloudy Landsat imagery in total cover the entire footprint, and there is a cloud-free MODIS observation available for each pixel for the date of the corresponding Landsat observation, it is possible to generate surface reflectance.

Thanks to our composited data sets, we generated a dense time series of single year forest phenology. Ultimately, we increased the number of dates for which there was information on surface reflectance available for > 90% of the image from 2 to 28 for 2016. By using data for a single year, we avoided the problems related to inter-annual variability in

forest phenology, that are common for temperate forests. This is an improvement over methods relying on multi-year data (Fisher et al., 2006; Elmore et al., 2012; Melaas et al., 2013).

Our results suggest that the mode of compositing Landsat data with the predictions based on STARFM, i.e., either using the prediction based on the image pair from the nearest date or the mean of the predictions of all image pairs, affects the resulting time series more than the compositing temporal scheme (Landsat vs. MODIS dates). The phenology curves based on the nearest dates were more in line with the actual Landsat observations confirming that predictions from models based on input Landsat-MODIS imagery closer to the prediction dates are advantageous.

In regards to the two temporal composition schemes, i.e., whether to predict data based on the Landsat acquisition date or the MODIS acquisition date, we found only minor differences between the two. The advantage of the Landsat dates is that this scheme retains the exact data of the original Landsat surface reflectance data. However, it makes it more difficult to mosaic data for neighboring paths due to varying acquisition dates. The advantage of harmonizing information to the global MODIS 8-day scheme is that it allows us to generate wall-to-wall information for large areas. However, the MODIS temporal scheme tends to smooth the phenology curves more than the Landsat acquisition scheme.

Unexpectedly, the much denser time series resulting from the compositing of Landsat and MODIS data did not improve mapping accuracy compared to classifying only the original Landsat imagery (**Chapter 1**). The blending of Landsat and MODIS data do not increase the level of inter-species heterogeneity beyond the level already present in the original acquisitions. This generally follows the findings from prior research on spatial and temporal fusing where original Landsat data typically resulted in better classification accuracy than their synthetic

alternative, but allowed to achieve acceptable results in situations of when actual acquisitions were not available (Senf et al., 2015).

Our results are relevant for environmental science, forest management, and conservation because they provide spatially-detailed information on vegetation phenology for large areas. Our approach can provide single-year information on surface reflectance and vegetation phenology even when actual Landsat observations are sparse and no cloud-free imagery exists.

Acknowledgments

The study was supported by the USDA McIntire-Stennis Program. The authors thank Eric Kruger, Annemarie Schneider, and Philip Townsend from University of Wisconsin-Madison for ideas and constructive feedback at various stages of work on this manuscript.

References

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

Burger, C., Belskii, E., Eeva, T., Laaksonen, T., Mägi, M., Mänd, R., Qvarnström, A., Slagsvold, T., Veen, T., Visser, M.E., Wiebe, K.L., Wiley, C., Wright, J., Both, C., 2012. Climate change, breeding date and nestling diet: How temperature differentially affects seasonal changes in pied flycatcher diet depending on habitat variation. J. Anim. Ecol. 81, 926–936. https://doi.org/10.1111/j.1365-2656.2012.01968.x

Churkina, G., Schimel, D., Braswell, B.H., Xiao, X., 2005. Spatial analysis of growing season length control over net ecosystem exchange. Glob. Chang. … 11, 1777–1787.

https://doi.org/10.1111/j.1365-2486.2005.01012.x

Claverie, M., Masek, J.G., 2016. Harmonized Landsat-8 Sentinel-2 (HLS) Product User â€™ s Guide Web site : Authors : 2, 1–16.

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

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

DeFries, R.S., Chan, J.C.W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515. https://doi.org/10.1016/S0034-4257(00)00142-5

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, 2012. Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description. Open-file Rep. 2013-1057 1–20. https://doi.org/No. 2013-1057

Elmore, A.J., Guinn, S.M., Minsley, B.J., Richardson, A.D., 2012. Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests. Glob. Chang. Biol. 18, 656–674. https://doi.org/10.1111/j.1365-2486.2011.02521.x

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

Fisher, J.I., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sens. Environ. 109, 261–273. https://doi.org/10.1016/j.rse.2007.01.004

Fisher, J.I., Mustard, J.F., Vadeboncoeur, M.A., 2006. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. Remote Sens. Environ. 100, 265–279. https://doi.org/10.1016/j.rse.2005.10.022

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

Friedl, M. a. M.A., Brodley, C.E.C.E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sens. Environ. 61, 399–409. https://doi.org/10.1016/S0034-4257(97)00049-7

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44, 2207–2218. https://doi.org/10.1109/TGRS.2006.872081

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, 13,364–13,391. https://doi.org/10.1002/2015JC011421.Received

Goring, S., Mladenoff, D.J., Cogbill, C. V, Record, S., Paciorek, C.J., Jackson, S.T., Dietze, M.C., Dawson, A., Matthes, J., McLachlan, J.S., Williams, J.W., 2015. Changes in Forest Composition, Stem Density, and Biomass from the Settlement Era (1800s) to Present in the Upper Midwestern United States. bioRxiv 1–34. https://doi.org/10.1101/026575

Goring, S.J., Mladenoff, D.J., Cogbill, C. V., Record, S., Paciorek, C.J., Jackson, S.T., Dietze, M.C., Dawson, A., Matthes, J.H., McLachlan, J.S., Williams, J.W., 2016. Novel and lost forests in the upper Midwestern United States, from new estimates of settlement-era composition, stem density, and biomass. PLoS One 11, 1–34. https://doi.org/10.1371/journal.pone.0151935

Helman, D., 2018. Land surface phenology: What do we really "see" from space? Sci. Total

Environ. 618, 665–673. https://doi.org/10.1016/j.scitotenv.2017.07.237

Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sens. Environ. 113, 1613–1627. https://doi.org/10.1016/j.rse.2009.03.007

Isaacson, B.N., Serbin, S.P., Townsend, P.A., 2012. Detection of relative differences in phenology of forest species using Landsat and MODIS. Landsc. Ecol. 27, 529–543. https://doi.org/10.1007/s10980-012-9703-x

Jones, J., Doran, P.J., Holmes, R.T., 2003. Climate and food synchronise regional forest bird abundances. Ecology 84, 3024–3032. https://doi.org/10.1890/02-0639

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

McCullough, D.G., Werner, R.A., Neumann, D., 1998. Fire and Insects in Northern and Boreal Forest Ecosystems of North America. Annu. Rev. Entomol. 43, 107–127. https://doi.org/10.1146/annurev.ento.43.1.107

Mcphaden, M.J., Zebiak, S.E., Glantz, M.H., Mcphaden, M., 2014. ENSO as an Concept Integrating in Earth Science 314, 1740–1745.

Melaas, E.K., Friedl, M.A., Zhu, Z., 2013. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. Remote Sens. Environ. 132, 176–185. https://doi.org/10.1016/j.rse.2013.01.011

Melaas, E.K., Sulla-Menashe, D., Gray, J.M., Black, T.A., Morin, T.H., Richardson, A.D., Friedl, M.A., 2016. Multisite analysis of land surface phenology in North American temperate and boreal deciduous forests from Landsat. Remote Sens. Environ. 186, 452–464. https://doi.org/10.1016/j.rse.2016.09.014

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

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

Radeloff, V.C., Mladenoff, D.J., He, H.S., Boyce, M.S., 1999. Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. Can. J. For. Res. 29, 1649–1659. https://doi.org/10.1139/x99-089

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, 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, 57–75. https://doi.org/10.1007/s10980-007-9117-3

Schulete, L.A., Mladenoff, D.J., 2005. Severe Wind and Fire Regimes in Northern Forests : Historical Variability At the Regional Scale. Ecology 86, 431–445.

Schulte, L.A., Mladenoff, D.J., Nordheim, E. V, 2002. Quantitative classification of a historic northern Wisconsin (U.S.A.) landscape: mapping forests at regional scales. Can. J. For. Res. 32, 1616–1638. https://doi.org/10.1139/x02-082

Senf, C., Leit??o, P.J., Pflugmacher, D., van der Linden, S., Hostert, P., 2015. Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery. Remote Sens. Environ. 156, 527–536. https://doi.org/10.1016/j.rse.2014.10.018

Trenberth, K.E., Branstator, G.W., Karoly, D., Kumar, A., Lau, N.-C., Ropelewski, C., 1998. Progress during TOGA in understanding and modeling global teleconnections associated

with tropical sea surface temperatures. J. Geophys. Res. Ocean. 103, 14291–14324. https://doi.org/10.1029/97JC01444

Wang, P., Gao, F., Masek, J.G., 2014. Operational data fusion framework for building frequent landsat-like imagery. IEEE Trans. Geosci. Remote Sens. 52, 7353–7365. https://doi.org/10.1109/TGRS.2014.2311445

Weng, Q., Fu, P., Gao, F., 2014. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. Remote Sens. Environ. 145, 55–67. https://doi.org/10.1016/j.rse.2014.02.003

Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sens. Environ. 114, 2610–2623. https://doi.org/10.1016/j.rse.2010.05.032

List of figures

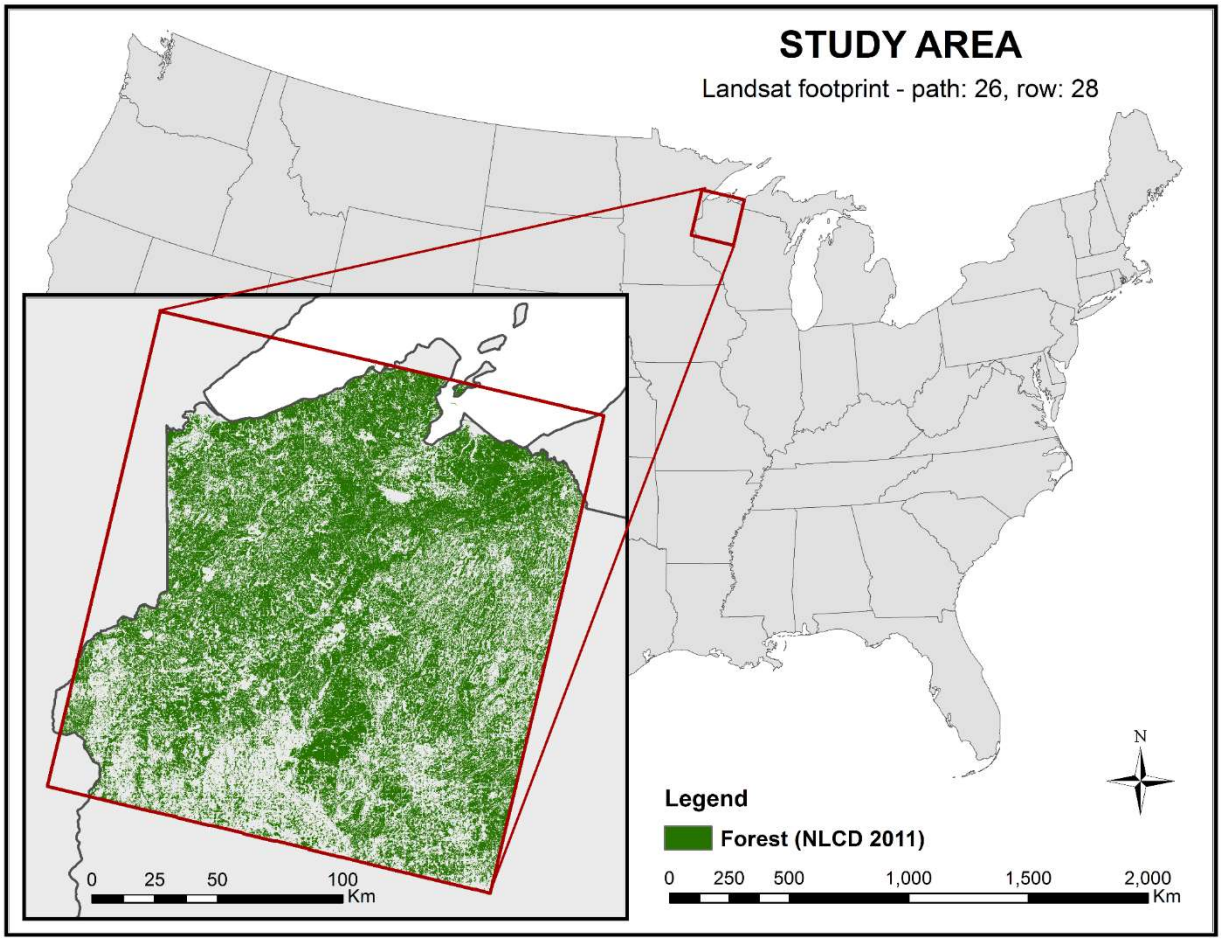

Figure 2. 1. Study area covering single Landsat footprint (path 26, row 28) located in Northern Wisconsin, USA.

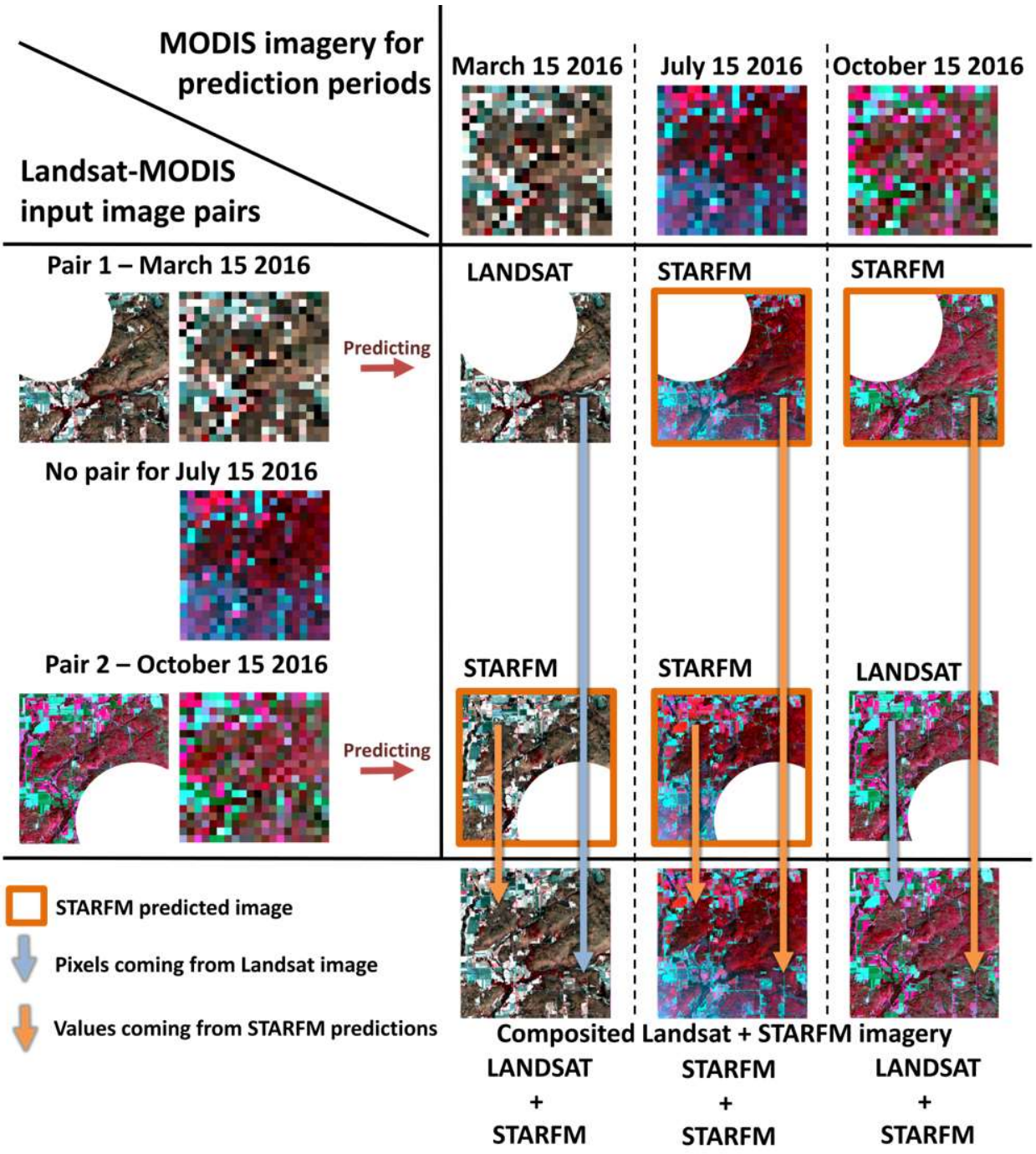

Figure 2. 2. General compositing scheme. The Landsat-MODIS input pairs are combined with the MODIS imagery for prediction periods in order to generate the STARFM predictions for periods where Landsat data is not available. Then, the original Landsat data is composited with STARFM creating the gap-free imagery.

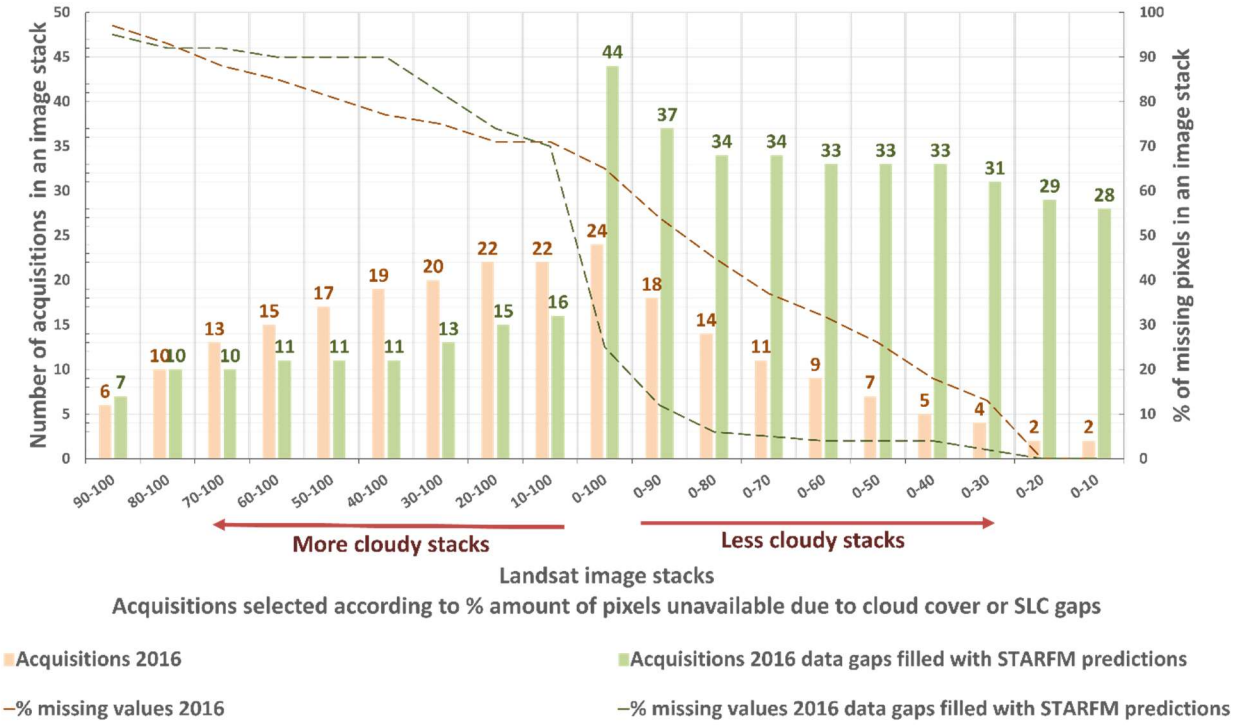

Figure 2. 3. Data augmentation resulted from filling gaps in Landsat data from 2016. Compositing original Landsat data with STARFM predictions allowed us to reduce the number of pixels missing due to cloud cover and SLC gaps and to increase the number of compleate acquisitions. This considerably decreased the level of missing pixels in the image stacks consisting imagery of various levels of missing pixels.

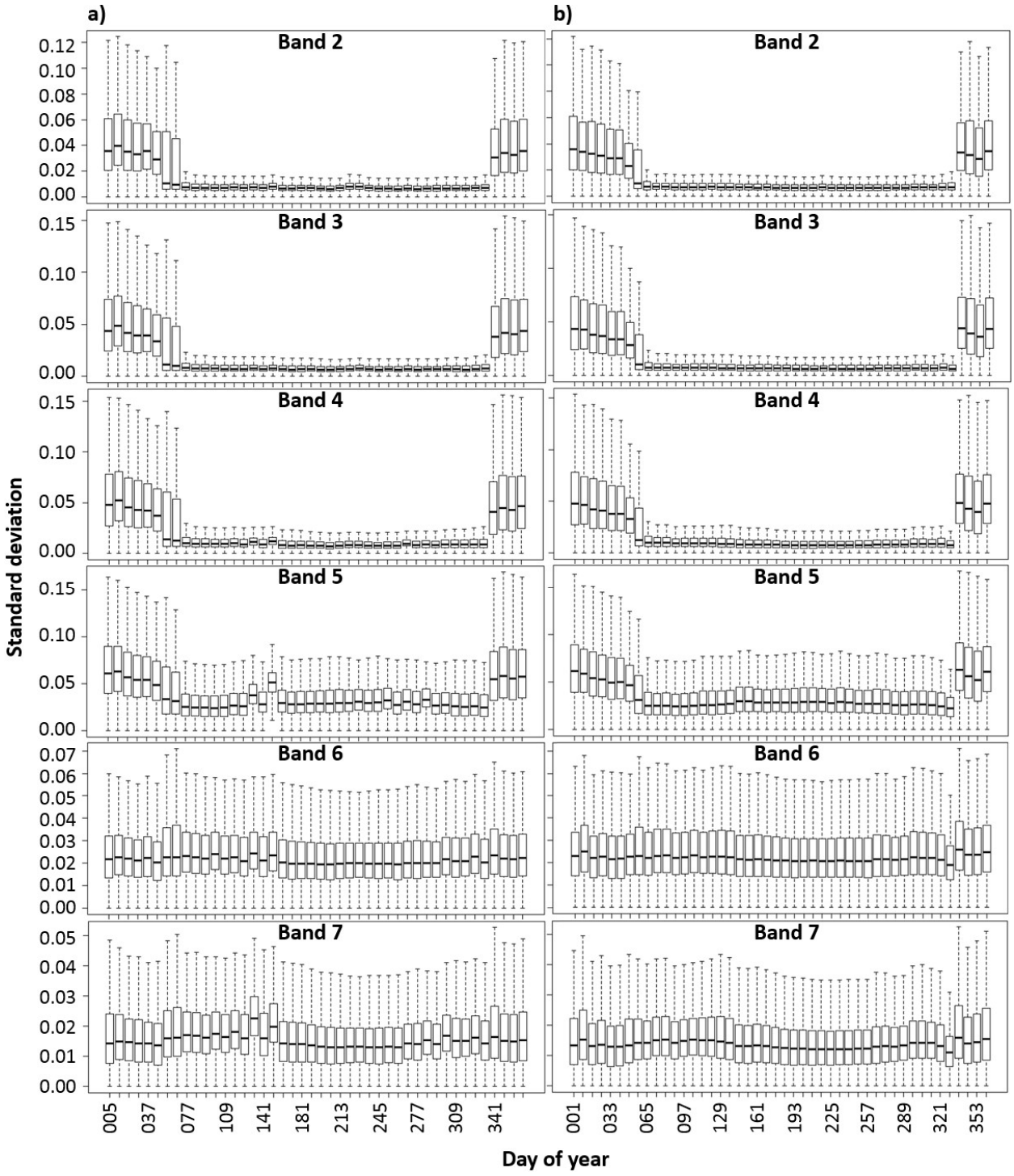

Figure 2. 4. Variability among STARFM predictions. Each box-plot represents values of standard deviation of predicted surface reflectance from 18 STARFM models based on Landsat-MODIS image pairs from all seasons; a) dates of Landsat acquisition scheme, b) MODIS periods.

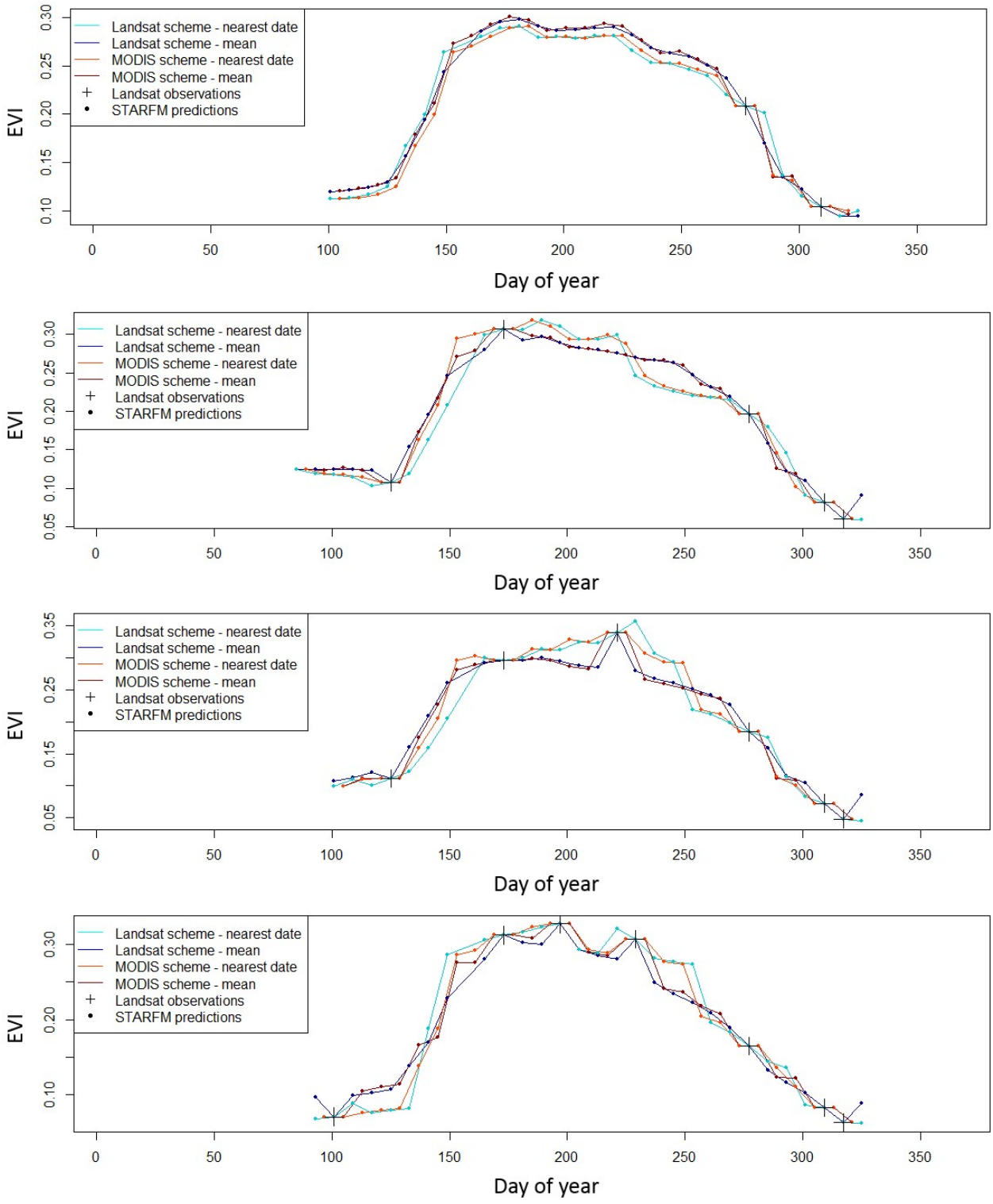

Figure 2. 5. The examples of phenology time series for 4 random forest pixels containing varying number of original Landsat observations. The EVI time series were generated for two temporal schemes: Landsat acquisition and MODIS product scheme, and with two compositing approaches: nearest date and mean value.

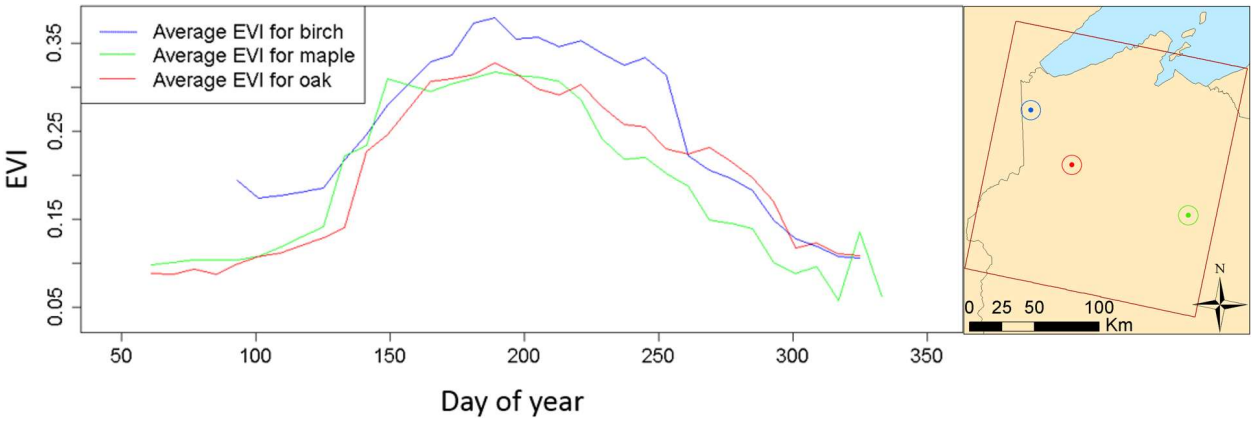

Figure 2. 6. Differences in the timing of phenology expressed with average EVI values for forest stands for three tree types: birch, maple, and oak.

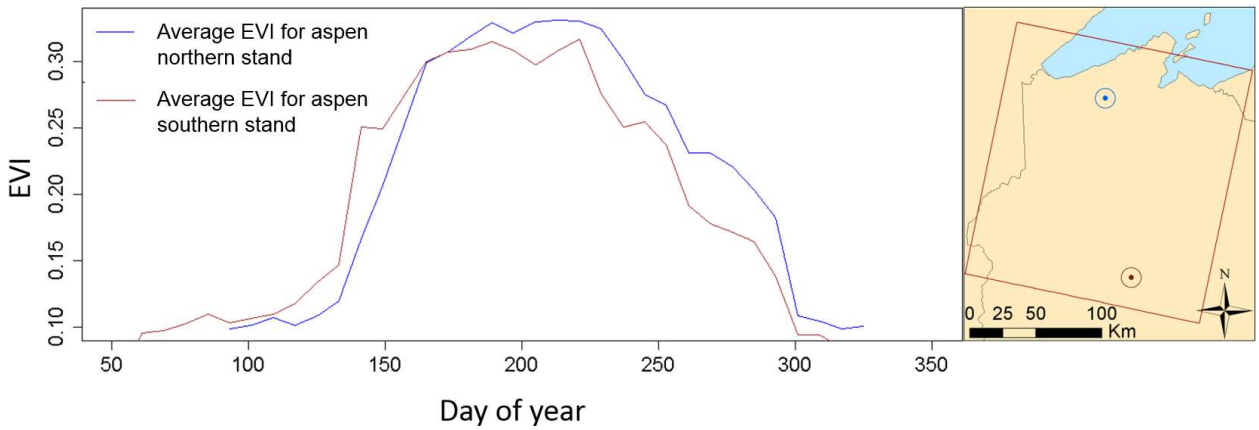

Figure 2. 7. Differences in the timing of phenology expressed with average EVI values for forest stands for two aspen stands located in the northern and in the southern part of the study area.

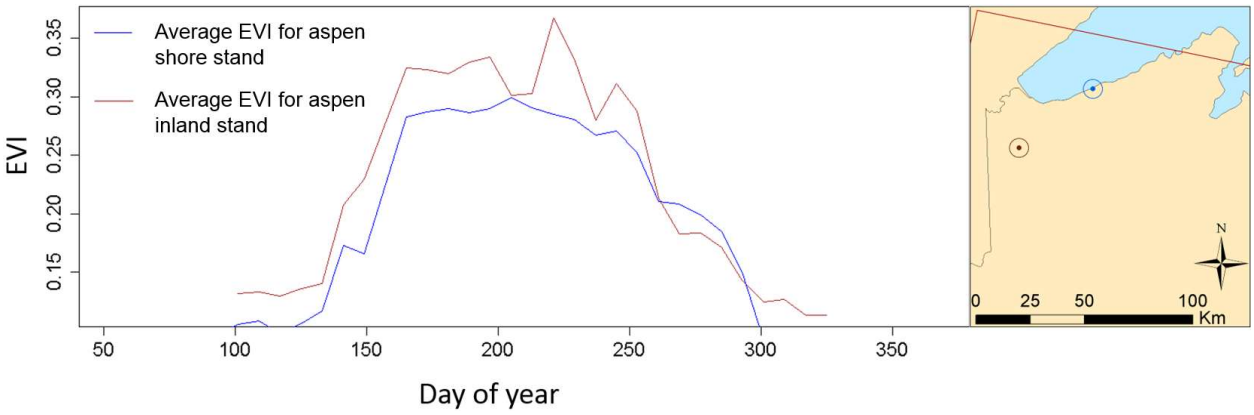

Figure 2. 8. Effects of phenology delay caused by neighborhood of the Lake Superior visible in the average values of EVI for two aspen forest stands located inland and near the shore.

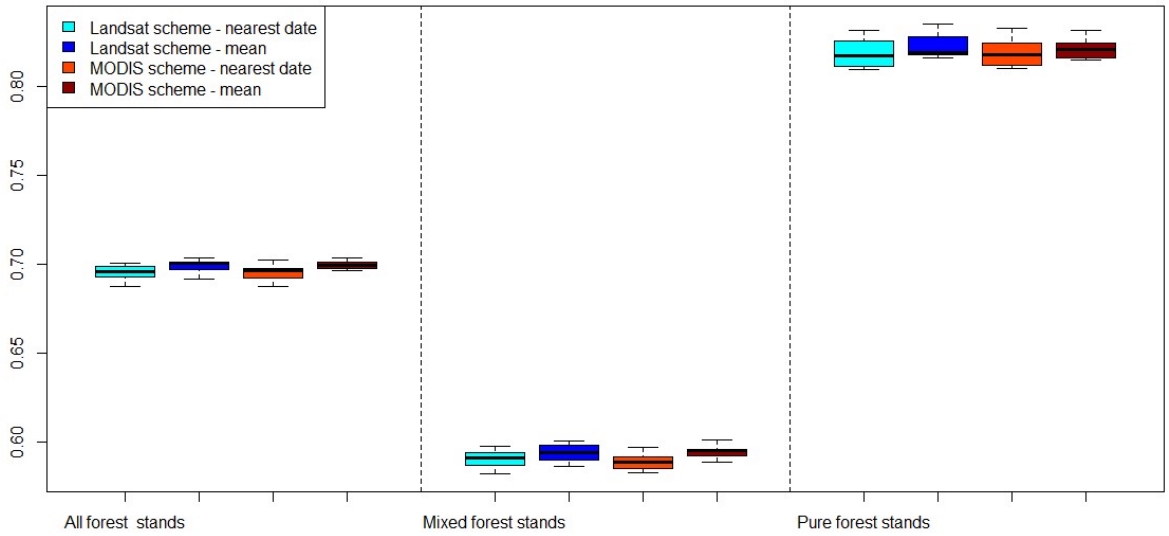

Figure 2. 9. Accuracy assessment: comparison of overall accuracy of forest type mapping performed with four image compositing algorithms and for three forest stand types. Ranges of accuracies resulted from multiple iterations of classification performed with randomly selected subsets of training and validation areas.

Chapter 3 - The value of Landsat-8 and Sentinel-2 Virtual Constellation for mapping forest types

Contributors: Konrad Turlej[1], Volker C. Radeloff[1]

[1] SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison

Abstract

The higher frequency of available remote sensing data with moderate spatial resolution provides new opportunities for more accurate mapping of forest tree species. The combination of recently launched satellites: European Sentinel-2 (S-2A in 2015, S-2B in 2017) and NASA-USGS Landsat-8 (2013) provides ≤30-m resolution imagery every 2-3 days. Such frequent observations may provide sufficient information to depict spectral differences among tree species at key phenological periods even in cloudy areas. Our goal was to evaluate how the combination of data from Landsat-8 and Sentinel-2 satellites affects the mapping of temperate tree species. We analyzed a single year of Harmonized Landsat and Sentinel-2 data (September 2017 – mid October 2018) and mapped 23 forest types present in northern Wisconsin, USA. Specifically, we tested: 1) combined HLS L30 and S30 imagery, 2) 8-day and 16-day composites generated from HLS data, 3) imagery from each system classified separately, and 4) data sets without any cloud-free imagery. We achieved an overall accuracy of 71.1% for maps based on either the full HLS data set, or 16-day composites, 70.6% for Sentinel-2 data only, and 67.2% for Landsat-8 only. For pure forest stands the accuracy ranged from 79.6% to 83.6%, for mixed stands it was <= 60.7%. However, a larger number of satellite images is only advantageous when no cloud free images exist. Compositing the HLS data in 8- or 16-day periods did not improve the classification accuracy, but accelerated the processing speed. While we did not find a major improvement in classification accuracy when analyzing Harmonized Landsat-8 and Sentinel-2

imagery, we still see the HLS data useful for forestry applications. The data from three sensors increases the probability of sensing a cloud free imagery in spring and fall, i.e. the periods of the year that are most important for mapping forest types.

Keywords

Forest type, remote sensing, Harmonized Landsat Sentinel-2, HLS, Landsat-8, Sentinel-2, gradient boosting decision trees, C50, intra-annual image composite

Introduction

The frequency of available remote sensing data with 30-m resolution has never been so high, thanks to concurrent operation of the recent launch of satellites: USGS/NASA Landsat-8 (L8) (launched in 2013), and European Union Copernicus: Sentinel-2 (S2) (platform A launched in 2015 and platform B in 2017). This Virtual Constellation of three similar satellites provides imagery across the globe every 2-3 days (Claverie et al., 2018). The triad of Landsat-8 Sentinel-2 is a powerful data source for numerous remote sensing applications including the monitoring of forests, which has been one strong focus of the Landsat program since its beginnings in 1972 (Iverson et al., 1989; Cohen et al., 2004; Wulder et al., 2012). Several key forest attributes are routinely monitored, including forest phenology (Fisher and Mustard, 2007), forest cover (Wulder et al., 2003), forest disturbance (Kennedy et al., 2010; Zhu et al., 2012; Griffiths et al., 2014), and tree species (Fassnacht et al., 2016). All of these applications can potentially benefit from the increased number of satellite observations available at ≤30-m resolution.

Combining data from Landsat-8 and Sentinel-2 is relatively straightforward because both sensors make similar measurements in regards to spectral, radiometric, spatial, and angular characteristics and timing (Wulder et al., 2015). Furthermore, both systems are placed in sun-

synchronous orbits setting the acquisition time to 10 am for Landsat-8 in a 16-day repeat cycle, and 10:30 am for Sentinel-2 in a 10-day repeat cycle (5-day for two satellites). The Multi-Spectral Instrument (MSI) installed on Sentinel-2 records the electromagnetic energy in 12 spectral bands out of which 7 are equivalent to the Landsat-8 Operational Land Imager (OLI) bands: a costal aerosol (L8/S2 band 1), optical bands blue, green, and red (L8/S2 bands 2-4), near infrared (L8 band 5; S2 band 8A), and two short wave infrared bands (L8 bands 6-7; S2 bands 11-12). The differences are that OLI collects 30-m resolution data for aerosol, all optical, and infrared bands, and 100-meter for thermal bands, whereas MSI provides 10-m data for optical and wide Infra-Red bands, 20-m for remaining NIR and SWIR bands, and 60-m for aerosol bands.

However, before combining images from different sensors, it is necessary to normalize the data, and resample it to common spatial reference (Wulder et al., 2015). Data harmonization and distribution of Landsat imagery has resulted in Analysis Ready Data (ARD) for the entire Landsat archive for every 8 days (Young et al., 2017). An equivalent of ARD for Landsat-8 and Sentinel-2 data is the Harmonized Landsat and Sentinel-2 (HLS) product (Claverie et al., 2018). The HLS provides surface reflectance data from Landsat-8 (L30) and Sentinel-2 (S30) satellites processed and resampled to a seamless moderate spatial resolution (<30-meters) product. HLS processing steps include atmospheric correction using LaSRC version 3.5.5 (Vermote et al., 2016), cloud and cloud-shadow masking, spectral bandpass adjustment, and Nadir BRDF adjustment (NBAR) according to the c-factor global 12 month fixed BRDF technique (Roy et al., 2016). The Sentinel-2 data included in the HLS are resampled to Landsat's 30-m spatial resolution via: 1) boxcar method (Sentinel-2 10-m bands), 2) area weighted averaging (Sentinel-

2 20-m bands), and 3) nearest neighbor split (Sentinel-2 60-m bands). The Landsat data are co-registered to common gridding system following Sentinel's 2 tailing scheme.

The HLS products provide the most frequent, openly available data set of surface reflectance at 30-m resolution. After the successful launch of Sentinel-2B satellite in 2017, the HLS data have reached the expected temporal resolution of 2-3-days. Such frequent observations make it possible to generate temporal composites of surface reflectance similar to the MODIS product MCD43A4 (Schaaf et al., 2002). The advantage of such 8- or 16-day composites is that cloud contamination is minimized, resulting in better spatial coverage and a more stable temporal scheme (Griffiths et al., 2019).

In the context of monitoring forest resources, the HLS data meets the requirements of large scale, spatially explicit mapping of forest types better than any other data set that is currently available (Fassnacht et al., 2016). The HLS data may overcome the problems of limited Landsat data availability, when the theoretical 8-day repeat cycle for two concurrent Landsat platforms was rarely achieved because of clouds and cloud shadows. HLS may thus be frequent enough to capture crucial aspects of forest phenology that are necessary to distinguish between tree species. First, the HLS data capture the species-specific surface reflectance in the most important visible and IR and SWIR parts of the spectrum that corresponds to tree type (coniferous, deciduous) and chemical structure (Shen et al., 1985; Fassnacht et al., 2014). Second, the imagery is acquired every 2-3 days, making it possible to monitor even slight changes in the reflectance characteristics due to phenology (Kodani et al., 2002). These changes in phenology differ among species and depend on particular adaptation of trees to late ground frost in spring, and leaf structure and longevity in the fall (Lechowicz, 1984) and can be observed with multi-temporal imagery (Townsend et al., 2001; Zhu and Liu, 2014). The improved ability to monitor phenology

thus may improve forest type mapping especially if there are acquisitions from periods when the trees appearance is most apparent (Wolter et al., 1995). Moreover, multi-date compositing can provide a more spatially complete view of differences in tree appearance in the most important green-up and senescence periods (Griffiths et al., 2013; Griffiths et al., 2019).

Our goal was to test how the combined imagery from Landsat-8 and Sentinel- in the HLS surface reflectance product performs in context of mapping temperate forests. Our objectives were to compare tree species classifications based on: 1) the full HLS dataset, 2) the temporal image composites (8- and 16- day), 3) the data from only Sentinel-2 and only Landsat-8, and 4) the dataset without cloud-free imagery.

Methods

Study area

Our study area was located in northern Wisconsin (**Figure 3. 1**) due to the highly forested landscapes and high spatial heterogeneity. The study area contains portions of three ecoregions: Lake Superior Lowland, Northern Highland, and the western part of the Central Plain (Martin, 1965). The local climate is temperate-continental and influenced by three air masses: 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 obstacles for air masses causing the zonal distribution of vegetation (Curtis, 1959). In addition to air masses, climate is influenced by Lake Superior in the north causing local temperature gradients.

Tree species composition of the forests is mixed consisting mostly aspen spp. (Populous), oak spp. (*Quercus*), pine spp. (*Pinus*), spruce spp. (*Picea*), maple spp. (*Acer*), and miscellaneous hardwoods, accompanied by tamarack (*Larix larcina*), eastern hemlock (*Tsuga canadensis*), and

white cedar (*Thuja occidentalis*). Contemporary tree species distributions differ substantially from those prior to European settlement in the mid-19th century due to land use (Curtis, 1959; Radeloff et al., 1999; Rhemtulla et al., 2007; Rhemtulla et al., 2009). In the past the forests on sandy outwash plains were dominated by pines (Curtis, 1959), sometimes accompanied by aspen and oaks in forest-savannas (Schulte et al., 2002), whereas the till plains and mesic moraines were dominated by hemlock and northern hardwood species (Curtis, 1959; Radeloff et al., 1999; Rhemtulla et al., 2007; Rhemtulla et al., 2009). Forest composition changes resulted from deforestation, i.e., land conversion from forest to croplands accompanied by rapid logging between 19th and 20th centuries, and from disturbance, i.e., insect outbreaks and return of severe fire events in middle of the 20th century (Schulete and Mladenoff, 2005). As a result, the forests that regrew since the 1930s differed considerably in species distribution from those prior to European settlement (Goring et al., 2015; Goring et al., 2016).

Data

We analyzed the Harmonized Landsat and Sentinel-2 (HLS) data version 1.4 (Skakun et al., 2018) for an area covering a single Landsat footprint (WRS2 path: 26, row: 28; MGRS tiles: 15TWL, 15TWM ,15TXL, 15TXM, 15TYM). We included a single year of HLS imagery starting from September 2017 to mid-October 2018 in order to include the data from two Sentinel-2 satellites (A and B). Contrary to version 1.3 (Claverie et al., 2018), the version 1.4 uses Landsat 8 Collection-1 data instead of the pre-Collection data set.

The HLS provides surface reflectance data from Landsat-8 (L30) and Sentinel-2 (S30) satellites processed and resampled in order to obtain a seamless moderate spatial resolution (30-meters) product. We included 6 bands of Landsat-8: 1) optical: Blue, Green, and Red, 2) Near Infra-Red, and 3) two Short Wave Infra-Red bands. In case of Sentinel-2 we also included the three narrow

Red-Edge spectra (bands 5-7), and the broad Near Infra-Red spectrum (band 8) that do not have Landsat equivalents. We masked out clouds, cloud-shadows, and snow using the outputs of Fmask (Zhu et al., 2015) embedded in the HLS quality bands.

One of our tests was to composite the HLS data to 8-day and 16-day periods in order to improve spatial coverage and temporal stability. We selected the maximal band values from all available acquisitions recorded both by the Landsat-8 and Sentinel-2 within each period. We performed this procedure for all 6 bands that are equivalent between the systems.

We used forest inventory data from the Wisconsin Forest Inventory & Reporting System (WisFIRS) generated by the Wisconsin Department of Natural Resources and County Forests (State of Wisconsin Department of Natural Resources, 2013). The data were collected in the field and provide stand-level information on forest types (**Appendix 1. B**): the primary tree type (>= 50% of the stand basal area), secondary tree type, understory type, and additional information including the tree height, tree density, total basal area. For our analyses, we selected two subclasses for forest stands: 1) pure stands (where primary, secondary, and understory vegetation type were all of the same tree species), and 2) mixed stands, all remaining stands. The WisFIRS dataset does not provide detailed information on the composition of tree species in mixed stands. The majority of the Reconnaissance forest stands were examined in the field after 2000, i.e., less than 16 years before the satellite data we used for this study were acquired (**Figure 1. 2**).

Mapping forest types

We tested four combinations of HLS data: 1) L30 + S30 with all optical and infrared bands of both Landsat-8 and Sentinel-2 systems, 2) the L30 + S30 for equivalent OLI and MSI bands only composited to 8- and 16- day periods, and 3) the L30 and S30 data separately. For each data

subset we repeated analyses two times: 1) all imagery, and 2) only imagery that were missing data with >=25% clouds and cloud shadows, to simulate a situation where no cloud-free imagery is available.

We randomly divided our reference forest stands in to 25% training and 75% testing subsets, and applied the Gradient Boosting Decision Trees, implemented in C5.0 package developed for R statistical software as our tool for classification. The method is described in detail in Quinlan (1993), Friedl and Brodley (1997), and DeFries and Chan (2000). We used this tool because it can handle incomplete datasets (Farhangfar et al., 2008), such as satellite imagery affected by clouds and cloud shadows. We built each model based on 100 boosting iterations, the maximum in C5.0, because higher numbers improved the classification accuracy. However, the actual number of iterations varied between the models depending on the amount of the data and the level of missing pixels in the training data set.

We assessed the accuracy of our maps, and compared the accuracy of the maps generated with the three subsets of HLS data and the two compositing schemes. We performed an area-weighted accuracy assessment and reported overall accuracy and class-specific producer's and user's accuracy (Congalton, 1991; Foody, 2002). For this assessment, we aggregated our results to the stand level labeling each stand's pixels with the value taken from majority vote from our classification.

Results

The combined HLS L30 and S30 imagery resulted in an overall classification accuracy of 73.1% for all types of forest stands, 83.5% for pure stands, and 60.7% for mixed stands **(Figure 3. 2**). Our resulting forest types map generally follows the patterns of species distribution visible in the

reference data set (**Figure 3. 3**). The most abundant tree classes: aspen, northern hardwoods, red pine, jack pine, and also the less abundant class 'lowland brush' are mapped in the correct locations. However, our map omits rare forest types, which occur throughout the study area, and this is reflected in the species-specific accuracy measurements (**Table 3. 1**). The user's accuracy was >80% for most of the classes, and exceeding 90% for jack pine (90.5%), and red pine (90.4%), but only 73.6% for black spruce. In terms of producer's accuracy, we achieved the highest accuracies for northern hardwoods (97.6%), aspen (95.3%), red pine (90.4%), black spruce (84.3%), brushes (82.6%), and jack pine (71.3%). The values for swamp hardwoods, scrub oak, and oak were considerably lower: 55.2%, 47.5%, 31.7%, respectively. The remaining types, which covered <2% of the total forest area, were mostly omitted (producer's accuracy equal to 19%) but had high user's accuracy. We present the full error matrix for all 23 forest types in **Appendix 3. A**.

The 8-day and 16-day temporal composites generated from the HLS data performed almost as well as the original HLS product (**Figure 3. 4**). However, the 8-day composites resulted in a minor decrease in classification accuracy for all forest stands (a 3.3 percentage point drop in overall accuracy). The results for the full HLS product and 16-day composites were virtually identical.

Analyzing only Landsat-8 and only Sentinel-2 data also resulted in only minor differences (**Figure 3. 5**). We recorded the highest overall accuracy for the full HLS product. Maps based only on Sentinel-2 imagery had 0.6, 0.4, and 0.8 percentage points lower accuracy for all forest stands, pure forest stands, and mixed forest stands. For data from Landsat-8 only, the drop in accuracy comparing to full HLS product was 3.9, 3.9, and 4 percentage points for all, pure, and mixed forest stands.

Type-specific accuracies followed the same trends as overall accuracies for most of the dominant tree species. In terms of producer's accuracy, the general trend of performance were: lowest classification accuracy for only Landsat-8 data, slightly better for 8-day HLS composites, almost equal for data from Sentinel-2 and 16-day composites, and the highest species-level accuracies for the full HLS data set (**Figure 3. 6a**). For oak, scrub oak, and swamp hardwoods differences were greater than for the other tree species.

In terms of user's accuracy our classifications were reliable for the most abundant tree species irrespective to the input satellite data that we used (**Figure 3. 6b**). We achieved around 80.0% accuracy for aspen, black spruce, northern hardwoods, scrub oak, and swamp hardwoods, and around 90.0% accuracy for jack pine, and red pine. The differences in user's accuracy among the data sets were minor in most cases. We found the most pronounced differences for oak, for which we achieved 95.9% user's accuracy using images from Landsat-8 only, but only 76.6% using 8-day composites. These value are opposite to the measures of producer's accuracy (**Figure 3. 6a**) suggesting that improvements in mapping with Landsat-8 data oak resulted in the increase of omission error.

Finally, the removal of cloud free acquisitions from the classified image stack resulted in a substantial drop in overall classification accuracy for all data subsets (**Figure 3. 7**). Limiting the data coverage to maximally 75% of the study area (i.e. simulating the lack of cloud-free images) reduced overall classification accuracies by about 9 percentage points for the full HLS data set, 11-12 percentage points for Sentinel-2data only, and 10 percentage points for Landsat-8 data only.

Discussion

We assessed the value of combining imagery from Landsat-8 and two Sentinel-2 satellites, represented by HLS (Harmonized Landsat-8 and Sentinel-2) surface reflectance product, for mapping tree species in Northern Wisconsin, USA. Our results proved the usefulness of the HLS data for distinguishing temperate tree species based on annual characteristics of their reflectance (Kodani et al., 2002). The combination of the three satellites provided more cloud-free image acquisitions than separate systems, making it possible to capture the stages of the tree phenology with greater frequency. In total, we had access to 270 images compared to 73 for Landsat-8, and 197 for Sentinel-2. Generally, a single year of HLS data provided the similar amount of useful imagery comparable to three years of Landsat data (**Chapter 1**).

While the HLS dataset resulted in the highest overall classification accuracy, the advantage of frequent 2-, 3-day repeat cycle of the HLS data was minor compared the data from the two Sentinel-2 satellites alone, and even a single Landsat-8. Overall accuracy ranged from 67.2% to 71.1% for all forest stands, and 79.6% to 83.5% for pure forest stands, and <= 60.7% for mixed stands. We found that a large number of satellite images did not improve classification accuracy very much, and what mattered most was if acquisitions added an unique information capturing spectral heterogeneity among tree species at certain times of the year (**Chapter 1;** Wolter et al., 1995; Zhu and Liu, 2014). Although data from three satellites increases the chances of more cloud free images, our results show that more imagery by itself does not necessarily improve the mapping accuracy. This result is similar to what has been reported for hyperspectral imagery (Hesketh and Sánchez-Azofeifa, 2012).

Compositing the HLS data in to 8-day and 16-day composites did not improve the classification accuracy, but reduced computing time. Our 16-day image composites performed similarly to the

full HLS data, whereas the 8-day composites resulted in a minor drop in mapping accuracy. This was unexpected considering the advantages of intra-annual image compositing (Griffiths et al., 2019). However, our study was the first to focus on mapping tree forest types, not just general land cover classes (Griffiths et al., 2013) or agriculture (Griffiths et al., 2019). Therefore, the lack of improvement in our case may reflect the fact that the available cloud-free imagery contained already most of the heterogeneity necessary to distinguish forest types.

The considerable differences in mapping accuracies for pure and mixed forest stands accurately may be partially due to our reference data. According to the WisFIRS methodology, each forest stand is labeled according to the: 1) dominating tree type which covers at least 50% of the given stand's basal area, 2) secondary tree species (no strict definition), and 3) understory tree type (no strict definition) (State of Wisconsin Department of Natural Resources, 2013). In contrast, we assigned the dominant tree species to each pixel, and labeled the forest stands according to the species that was most abundant. Thus, our classification may not reflect the on-the-ground measurements perfectly for mixed stands. However, considering that we captured pure forest stands correctly, i.e. the stands where primary, secondary, and understory tree types are identical, we see the results of our classification as reliable at the pixel level, but with the current reference data we were not able to quantify pixel-level accuracy.

Considering that mapping forest resources has been a strong focus of satellite remote sensing since the launch of the first Landsat satellite in 1972 (Iverson et al., 1989; Cohen et al., 2004; Wulder et al., 2012), our results have important implications for environmental science, resource management, and policy making. Our evaluation of the HLS data showed that it is a valid source of information to map forest type, and thereby update stand-level forest inventories and support

forest management (Portoghesi, 2006), and support the conservation of wildlife species (Loiselle et al., 2003).

Finally, we conclude that our findings provide insights into the process of mapping forest types using 30-m resolution optical data. We highlight the benefit of HLS imagery for forest type and its increased possibility for acquiring cloud-free data due to the frequent repeat cycle of 2-3 days. The harmonization of the data allows for 1) combining the data from Landsat-8 and Sentinel-2 seamlessly, and 2) generating 16-day compositions that can accelerate the mapping speed without compromising its accuracy.

Acknowledgments

The study was supported by the USDA McIntire-Stennis Program. The authors thank Eric Kruger, Annemarie Schneider, Mutlu Ozdogan, Philip Townsend from University of Wisconsin-Madison for ideas and constructive feedback at various stages of work on this manuscript; Jeff Masek and Junchang Ju for providing the HLS data set in version 1.4.

References

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., Goward, S.N., Cohen, Warren, B., Goward, S.N., Cohen, W.B., Goward, S.N., 2004. Landsat's Role in Ecological Applications of Remote Sensing. Bioscience 54, 535–545. https://doi.org/DOI: 10.1641/0006-3568(2004)054[0535:LRIEAO]2.0.CO;2

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

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

DeFries, R.S., Chan, J.C.W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515. https://doi.org/10.1016/S0034-4257(00)00142-5

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

Fassnacht, F.E., Lati, H., Sterenczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classi fi cation from remotely sensed data. Remote Sens. Environ. 186, 64–87.

Fassnacht, F.E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Kochb, 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, 2547–2561.

Fisher, J.I., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sens. Environ. 109, 261–273. https://doi.org/10.1016/j.rse.2007.01.004

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

Friedl, M. a. M.A., Brodley, C.E.C.E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sens. Environ. 61, 399–409. https://doi.org/10.1016/S0034-4257(97)00049-7

Goring, S., Mladenoff, D.J., Cogbill, C. V, Record, S., Paciorek, C.J., Jackson, S.T., Dietze, M.C.,

Dawson, A., Matthes, J., McLachlan, J.S., Williams, J.W., 2015. Changes in Forest Composition, Stem Density, and Biomass from the Settlement Era (1800s) to Present in the Upper Midwestern United States. bioRxiv 1–34. https://doi.org/10.1101/026575

Goring, S.J., Mladenoff, D.J., Cogbill, C. V., Record, S., Paciorek, C.J., Jackson, S.T., Dietze, M.C., Dawson, A., Matthes, J.H., McLachlan, J.S., Williams, J.W., 2016. Novel and lost forests in the upper Midwestern United States, from new estimates of settlement-era composition, stem density, and biomass. PLoS One 11, 1–34. https://doi.org/10.1371/journal.pone.0151935

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V.C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., Hostert, P., 2014. Forest disturbances, forest recovery, and changes in forest types across the carpathian ecoregion from 1985 to 2010 based on landsat image composites. Remote Sens. Environ. 151, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. Environ. 220, 135–151. https://doi.org/10.1016/J.RSE.2018.10.031

Griffiths, P., Van Der Linden, S., Kuemmerle, T., Hostert, P., 2013. Erratum: A pixel-based landsat compositing algorithm for large area land cover mapping (IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing). IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 2088–2101. https://doi.org/10.1109/JSTARS.2012.2228167

Hesketh, M., Sánchez-Azofeifa, G.A., 2012. The effect of seasonal spectral variation on species classification in the Panamanian tropical forest. Remote Sens. Environ. 118, 73–82. https://doi.org/10.1016/j.rse.2011.11.005

Iverson, L.R., Graham, R.L., Cook, E.A., 1989. Applications of satellite remote sensing to

forested ecosystems. Landsc. Ecol. 3, 131–143. https://doi.org/10.1007/BF00131175

Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Kodani, E., Awaya, Y., Tanaka, K., Matsumura, N., 2002. Seasonal patterns of canopy structure, biochemistry and spectral reflectance in a broad-leaved deciduous Fagus crenata canopy. For. Ecol. Manage. 167, 233–249. https://doi.org/10.1016/S0378-1127(01)00701-0

Lechowicz, M.J., 1984. Why do Temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities. Am. Nat. 124, 821–842.

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

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

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

Portoghesi, L., 2006. European Forest Types. Categories and types for sustainable forest management reporting and policy, Forest@ - Rivista di Selvicoltura ed Ecologia Forestale. https://doi.org/10.3832/efor0425-003

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

Radeloff, V.C., Mladenoff, D.J., He, H.S., Boyce, M.S., 1999. Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. Can. J. For. Res. 29, 1649–1659. https://doi.org/10.1139/x99-089

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, 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, 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, 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., Kovalskyy, V., 2016. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. Remote Sens. Environ. 176, 255–271. https://doi.org/10.1016/j.rse.2016.01.023

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., D'Entremont, R.P., Hu, B., Liang, S., Privette, J.L., Roy, D., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135–148. https://doi.org/10.1016/S0034-4257(02)00091-3

Schulete, L.A., Mladenoff, D.J., 2005. Severe Wind and Fire Regimes in Northern Forests :

Historical Variability At the Regional Scale. Ecology 86, 431–445.

Schulte, L.A., Mladenoff, D.J., Nordheim, E. V, 2002. Quantitative classification of a historic northern Wisconsin (U.S.A.) landscape: mapping forests at regional scales. Can. J. For. Res. 32, 1616–1638. https://doi.org/10.1139/x02-082

Shen, S.S., Badhwar, G.D., Carnes, J.G., 1985. Separability of boral fores species in the Lake Jennette Area, Minnesota. Photogramm. Eng. Remote Sensing 51, 1775–1783.

Skakun, P.S., Ju, J., Claverie, M., Roger, J., Vermote, E., Franch, B., Dungan, J.L., Masek, J., 2018. Harmonized Landsat Sentinel-2 (HLS) Product User ' s Guide 2, 1–30.

State of Wisconsin Department of Natural Resources, 2013. Public Forest Lands Handbook.

Townsend, P. a, Walsh, S.J., Ecology, S.P., Dec, N., Townsend, a, Stephen, J., J, S., 2001. Remote Sensing of Forested Wetlands : Application of Multitemporal and Multispectral Satellite Imagery to Determine Plant Community Composition and Structure in Southeastern Stable URL : http://www.jstor.org/stable/20051167 and sensing of forested wetland. Plant Ecol. 157, 129–149. https://doi.org/10.1023/A:1013999513172

Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sens. Environ. 185, 46–56. https://doi.org/10.1016/j.rse.2016.04.008

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.

Wulder, M.A., Dechka, J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudoin, A., Franklin, S.E., 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. For. Chron. 79, 1075–1083. https://doi.org/10.5558/tfc791075-6

Wulder, M.A., Hilker, T., White, J.C., Coops, N.C., Masek, J.G., Pflugmacher, D., Crevier, Y., 2015. Virtual constellations for global terrestrial monitoring. Remote Sens. Environ. 170, 62–76. https://doi.org/10.1016/j.rse.2015.09.001

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

Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R., Evangelista, P.H., 2017. A survival guide to Landsat preprocessing. Ecology 98, 920–932. https://doi.org/10.1002/ecy.1730

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., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. Remote Sens. Environ. 159, 269–277. https://doi.org/10.1016/j.rse.2014.12.014

Zhu, Z., Woodcock, C.E., Olofsson, P., 2012. 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

List of tables

Table 3. 1 Accuracy assessment for all HLS data for pure forest stands.

	ASPEN	BLACK SPRUCE	JACK PINE	NORTHERN HARDWOODS	OAK	RED PINE	SCRUB OAK	SWAMP HARDWOODS	BRUSH	OTHER	***R.sum***	***User's accuracy (%)***
ASPEN	**321900**	1080	4203	9332	8930	3405	7923	13845	15862	14762	401242	**80.2**
BLACK SPRUCE	112	**34735**	8	100	1	11	9	133	1718	10372	47199	**73.6**
JACK PINE	246	94	**31082**	16	2	2490	245	-	34	147	34356	**90.5**
NORTHERN HARDWOODS	13529	3803	586	**441184**	15886	813	1031	8433	2980	14609	502854	**87.7**
OAK	143	1	-	1259	**11578**	15	160	2	21	9	13188	**87.8**
RED PINE	285	93	5792	17	21	**74905**	136	30	157	1468	82904	**90.4**
SCRUB OAK	106	-	358	9	131	158	**9522**	-	1019	2	11305	**84.2**
SWAMP HARDWOODS	217	372	-	31	-	6	-	**34671**	3139	2506	40942	**84.7**
BRUSH	1109	477	1555	225	3	1023	1163	5118	**120821**	3372	134866	**89.6**
OTHER	44	572	-	35	-	13	-	531	486	**8042**	9723	**82.7**
C.sum	337691	41227	43584	452208	36552	82839	20189	62763	146237	55289		
Producer's accuracy (%)	**95.3**	**84.3**	**71.3**	**97.6**	**31.7**	**90.4**	**47.2**	**55.2**	**82.6**	**14.5**		

Appendix 3. A The full error matrix for pure forest stands. Classification based on all HLS images.

	ASPEN	BALSAM FIR	BLACK SPRUCE	BOTTOMLAND HARDWOODS	FIR SPRUCE	HEMLOCK	JACK PINE	LOWLAND BRUSH	LOWLAND BRUSH ALDER
ASPEN	**321900**	239	1080	949	615	461	4203	6390	6462
BALSAM FIR	1	**-**	-	-	-	-	-	-	-
BLACK SPRUCE	112	14	**34735**	-	60	3	8	543	1167
BOTTOMLAND HARDWOODS	-	-	-	**-**	-	-	-	-	-
FIR SPRUCE	1	-	2	-	**-**	-	-	-	-
HEMLOCK	-	-	-	-	-	**-**	-	-	-
JACK PINE	246	99	94	-	28	-	**31082**	19	10
LOWLAND BRUSH	688	50	359	-	29	-	6	**46767**	12387
LOWLAND BRUSH ALDER	419	17	118	-	6	1	19	5998	**42024**
LOWLAND BRUSH WILLOW	-	-	-	-	-	-	-	-	-
NORTHERN HARDWOODS	13529	56	3803	42	485	466	586	753	1833
OAK	143	-	1	-	-	-	-	19	2
RED MAPLE	1	-	-	-	-	-	-	-	-
RED PINE	285	3	93	-	11	9	5792	55	42
SCRUB OAK	106	-	-	1	-	-	358	-	-
SWAMP CONIFER	5	25	218	-	-	-	-	21	174
SWAMP HARDWOODS	217	54	372	4	4	34	-	2036	1013
TAMARACK	4	-	207	-	3	-	-	177	20
UPLAND BRUSH	2	-	-	-	-	-	1530	1	1
WHITE BIRCH	2	-	-	-	-	-	-	-	-
WHITE CEDAR	23	-	144	-	38	-	-	-	92
WHITE PINE	6	-	1	-	-	-	-	-	1
WHITE SPRUCE	1	-	-	-	-	-	-	-	-
C.sum	337691	557	41227	996	1279	974	43584	62779	65228
Producer's accuracy (%)	**95.3**	**-**	**84.3**	**-**	**-**	**-**	**71.3**	**74.5**	**64.4**

	LOWLAND BRUSH WILLOW	NORTHERN HARDWOODS	OAK	RED MAPLE	RED PINE	SCRUB OAK	SWAMP CONIFER	SWAMP HARDWOODS	TAMARACK
ASPEN	82	9332	8930	4449	3405	7923	480	13845	903
BALSAM FIR	-	-	-	-	-	-	-	-	-
BLACK SPRUCE	-	100	1	19	11	9	4270	133	5434
BOTTOMLAND HARDWOODS	-	-	-	-	-	-	-	-	-
FIR SPRUCE	-	-	-	-	-	-	-	-	-
HEMLOCK	-	-	-	-	-	-	-	-	-
JACK PINE	-	16	2	-	2490	245	-	-	1
LOWLAND BRUSH	339	104	2	12	179	120	122	4016	1024
LOWLAND BRUSH ALDER	56	118	1	47	22	5	153	1102	1044
LOWLAND BRUSH WILLOW	**-**	-	-	-	-	-	-	-	-
NORTHERN HARDWOODS	1	**441184**	15886	9016	813	1031	2222	8433	252
OAK	-	1259	**11578**	1	15	160	-	2	1
RED MAPLE	-	-	-	**-**	-	-	-	-	1
RED PINE	-	17	21	-	**74905**	136	265	30	19
SCRUB OAK	-	9	131	-	158	**9522**	-	-	-
SWAMP CONIFER	-	15	-	-	4	-	**2348**	375	3
SWAMP HARDWOODS	-	31	-	41	6	-	818	**34671**	32
TAMARACK	-	7	-	4	1	-	-	-	**232**
UPLAND BRUSH	-	3	-	-	822	1038	-	-	-
WHITE BIRCH	-	-	-	-	-	-	-	-	-
WHITE CEDAR	-	13	-	-	-	-	688	156	-
WHITE PINE	-	-	-	-	8	-	-	-	-
WHITE SPRUCE	-	-	-	-	-	-	1	-	-
C.sum	478	452208	36552	13589	82839	20189	11367	62763	8946
Producer's accuracy (%)	**-**	**97.6**	**31.7**	**-**	**90.4**	**47.2**	**20.7**	**55.2**	**2.6**

	UPLAND BRUSH	WHITE BIRCH	WHITE CEDAR	WHITE PINE	WHITE SPRUCE	***R.sum***	***User's accuracy (%)***
ASPEN	2928	2117	770	2948	831	393653	**80.2**
BALSAM FIR	-	-	-	-	-	0	**-**
BLACK SPRUCE	8	32	392	-	148	44559	**73.6**
BOTTOMLAND HARDWOODS	-	-	-	-	-	0	**-**
FIR SPRUCE	-	-	-	-	5	0	**-**
HEMLOCK	-	-	-	-	-	0	**-**
JACK PINE	5	-	-	8	11	40181	**90.5**
LOWLAND BRUSH	99	2	334	-	48	61714	**70.1**
LOWLAND BRUSH ALDER	157	-	452	2	29	60483	**81.1**
LOWLAND BRUSH WILLOW	-	-	-	-	-	0	**-**
NORTHERN HARDWOODS	393	931	964	85	90	482668	**87.7**
OAK	-	2	-	5	-	19071	**87.8**
RED MAPLE	-	-	-	-	-	106	**-**
RED PINE	60	25	154	909	73	79590	**90.4**
SCRUB OAK	1019	-	-	1	-	15477	**84.2**
SWAMP CONIFER	-	-	1416	-	-	4566	**51.0**
SWAMP HARDWOODS	90	5	1500	-	14	46621	**84.7**
TAMARACK	-	-	-	-	108	6510	**30.4**
UPLAND BRUSH	**12992**	-	-	-	-	18110	**79.3**
WHITE BIRCH	-	**-**	-	-	-	4	**-**
WHITE CEDAR	-	-	**3007**	-	-	4573	**72.3**
WHITE PINE	-	-	-	**95**	-	296	**85.6**
WHITE SPRUCE	1	-	-	-	**68**	15	**95.8**
C.sum	17752	3114	8989	4053	1425	***Overall***	**83.5**
Producer's accuracy	**73.2**	**-**	**33.5**	**2.3**	**4.8**	***Accuracy***	

List of figures

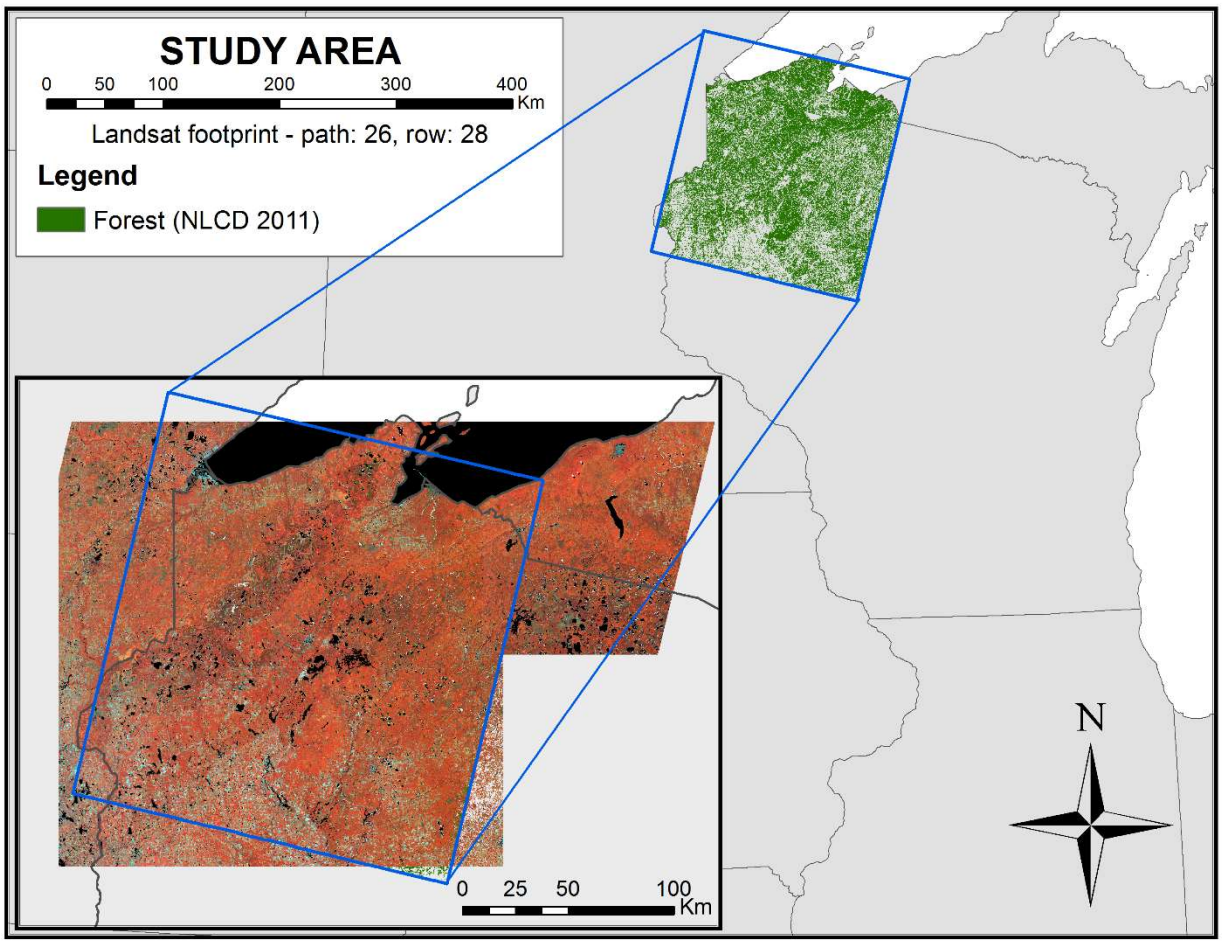

Figure 3. 1. The location of the study area in Northern Wisconsin, USA. The area covers a single Landsat WRS2 footprint (path: 26, row: 28) and 5 HLS UTM MGRS tiles (15TWL, 15TWM, 15TXL, 15TXM, 15TYM)

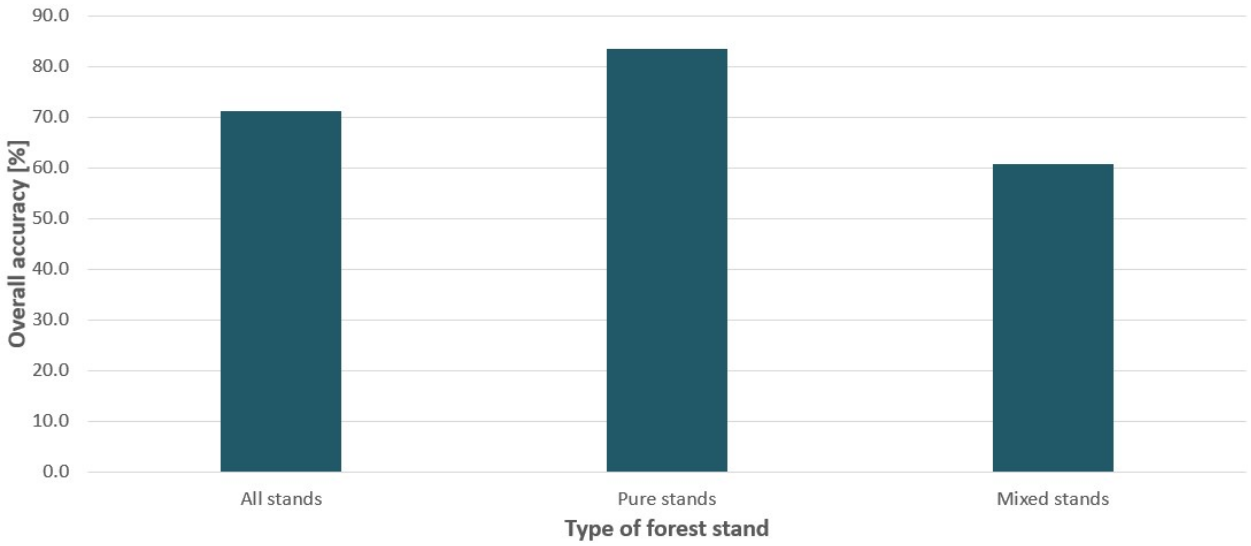

Figure 3. 2. Overall classification accuracy for the full HLS data for all forest stands, pure forest stands, and mixed forest stands.

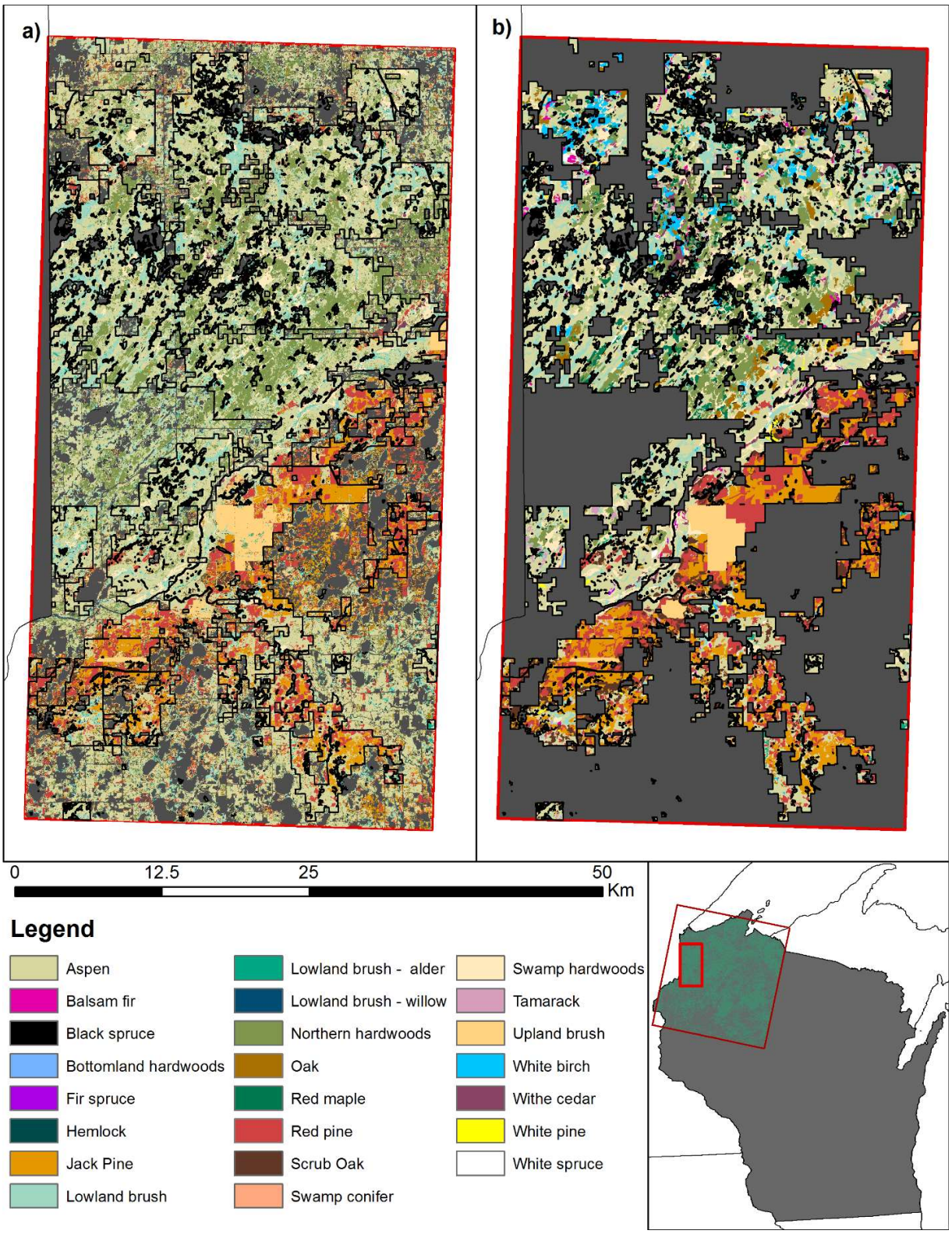

Figure 3. 3 Maps of forest forest typess for a portion of the study area for a) the best classification generated with the full HLS data set, b) the reference data set.

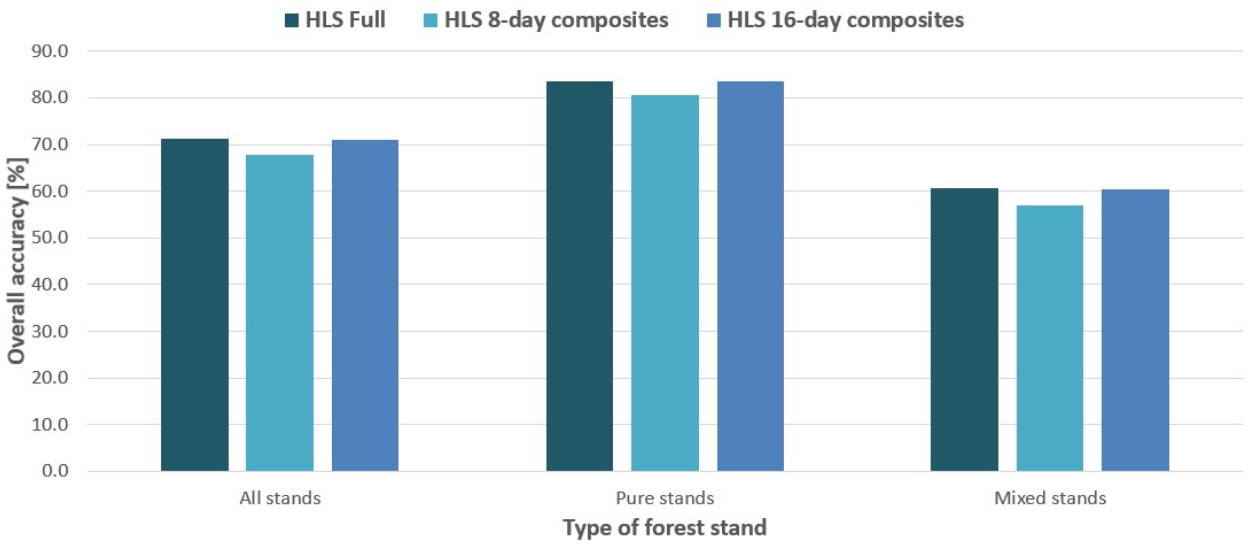

Figure 3. 4. Overall classification accuracy for the full HLS data set, and 8-day, and 16-day temporal compositions.

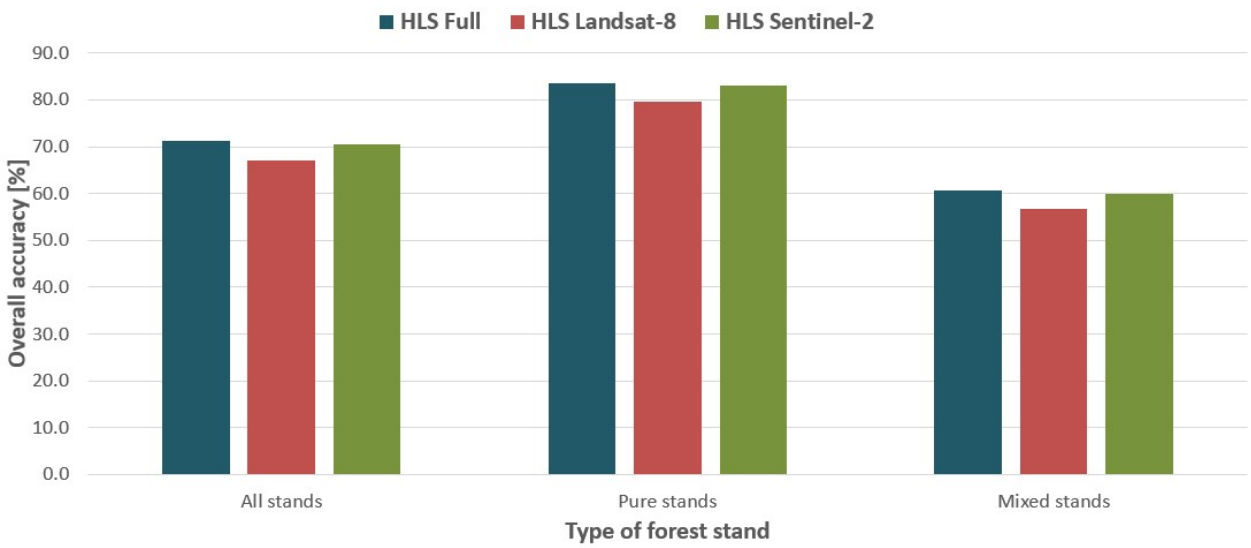

Figure 3. 5. Overall classification accuracy for the full HLS dataset, only Landsat-8 imagery, and only Sentinel-2 imagery.

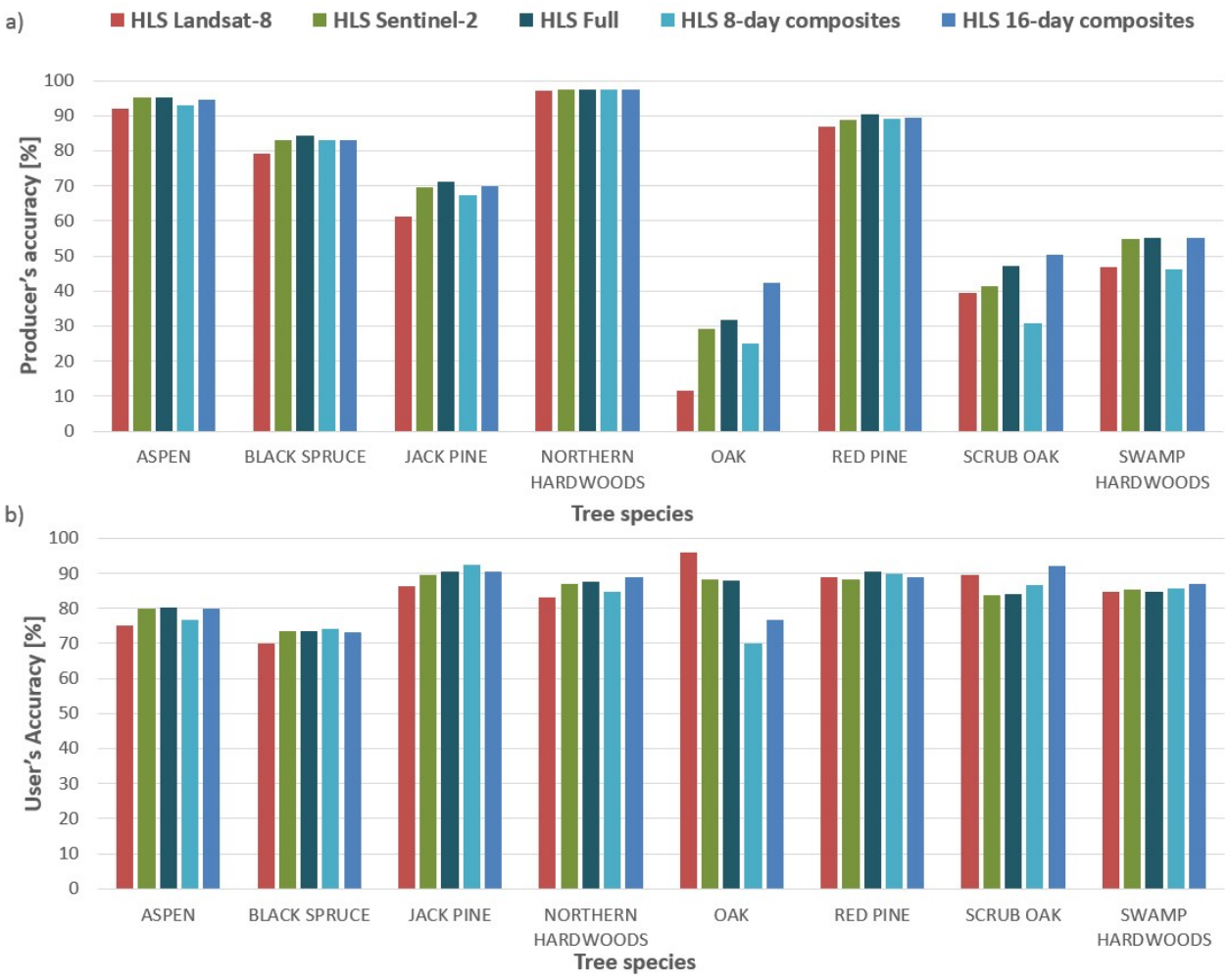

Figure 3. 6. Type-specific classification accuracies for the most abundant forest typess for different types of input data: a) producer's accuracy, b) user's accuracy.

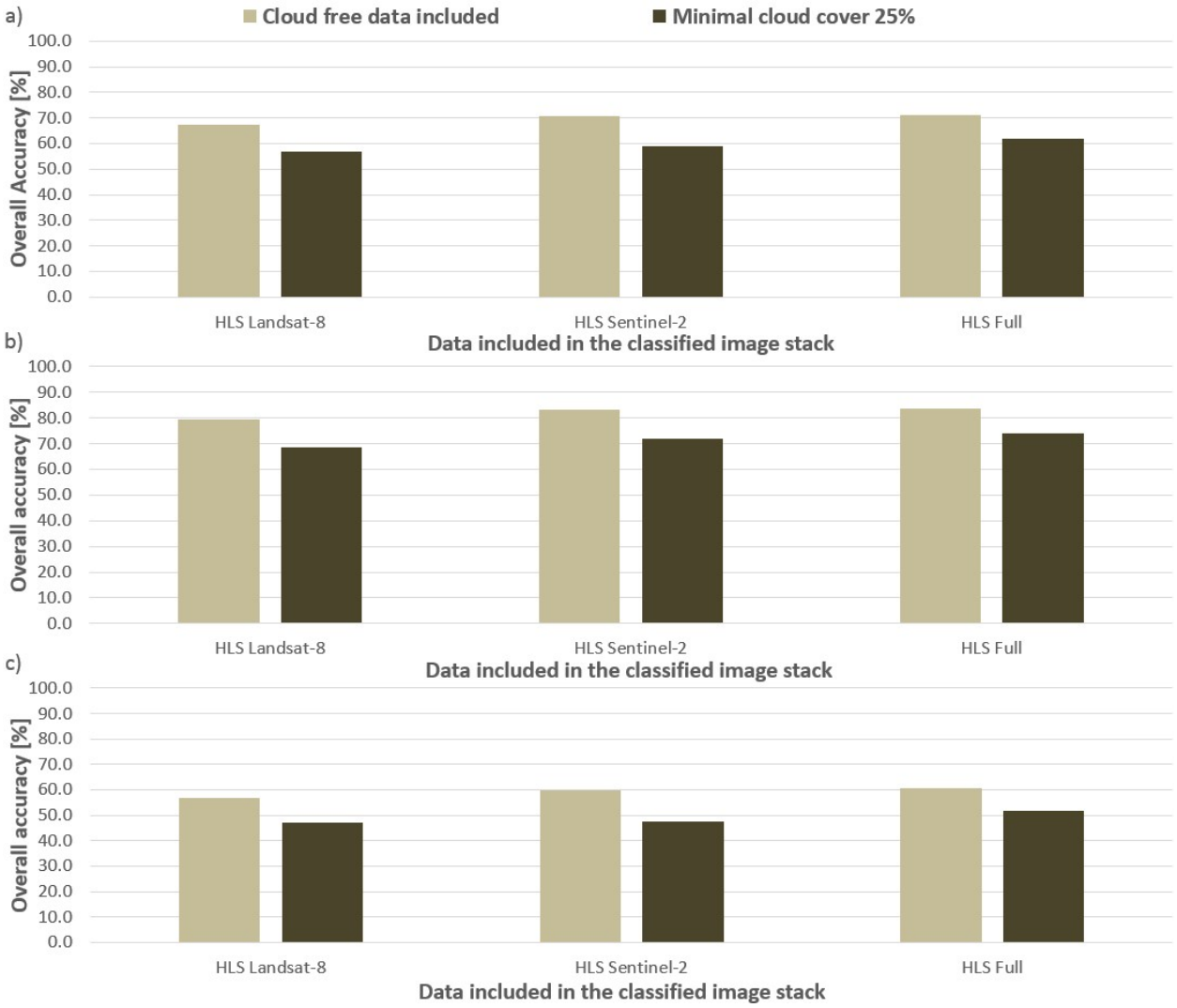

Figure 3. 7 Overall classification accuracy obtained with and without cloud-free imagery. Data presented for a) all forest stands, b) pure forest stands, and c) mixed forest stands.