May 2021

Autumn Tree Phenology in Northern Wisconsin: Humans Versus Photographs

Trevor Iglinski
University of Wisconsin-Milwaukee

AUTUMN TREE PHENOLOGY IN NORTHERN WISCONSIN:

HUMANS VERSUS PHOTOGRAPHS

by

Trevor J. Iglinski

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Geography

at

The University of Wisconsin-Milwaukee

May 2021

ABSTRACT

AUTUMN TREE PHENOLOGY IN NORTHERN WISCONSIN: HUMANS VERSUS PHOTOGRAPHS

by

Trevor Iglinski

The University of Wisconsin-Milwaukee, 2021
Under the Supervision of Professor Mark Schwartz

Ecosystem primary productivity halts when plants go dormant, and so the timing of dormancy as it relates to autumn phenology has been a focus of much interdisciplinary research. While monitoring plant phenology has its roots in directly observing specimens, digital sensors along with modern methods have also become a mainstay in phenology. Results from different methods often vary, so there is still a need to better understand how digital cameras record autumn phenology, especially in comparison with ground-based observations (Keenan et al. 2014). This study compared autumn phenology derived from direct ground observations with upward-facing fisheye photography, in the context of a larger research project (C.H.E.E.S.E.H.E.A.D.19), to precisely determine autumn tree phenology across 53 field sites in a heterogeneous temperate deciduous forest with over 220 individual trees and 1,000 digital photos sampled. Less-studied trees such as aspen (*Populus spp.)*, birch (*Betula spp.*), and basswood (*Tilia americana)* were included in the project, as well as sugar maple and red maple (*Acer spp.)*.

The results show that inflection points from sigmoid curves and change point detection are in close agreement for critical transition dates including the start of leaf coloration (bias of change points later at i= -0.47 days) and end (i= -0.6), but with slightly less agreement for the start of leaf fall (bias of change points earlier at i= 3.8) and the end of leaf fall (bias of change points later at i= -3.39).

While camera-derived transition dates correlated poorly with corresponding human-derived transition dates, the best relationship detected was between green chromatic coordinate (GCC) inflection points and leaf fall (when foliage is mostly absent from tree canopies). This work is intended as a pilot study for novel methodologies in the field of ground-based phenology.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge:

The Cheesehead19 phenology field team, Braden Westerhoff, Elicio Ruiz Guzman, Kelsey Flathers, and Amos Kaldor for their work in the woods collecting data;

My graduate committee members, Prof. Alison Donnelly and Gretchen Meyer, Ph.D. for their feedback before and during my thesis defense;

My advisor, Dist. Prof. Mark D. Schwartz for making graduate school an enriching endeavor, his patience, and insight along the way;

and

My fiancée, Rosemary for her understanding and love, as home became office from March 2020 into May 2021.

1 INTRODUCTION

Phenology is defined loosely as “nature’s calendar”, referring to the timing of life stage events in both plants and animals. Events in plants may include flowering, budbreak, or foliage abscission. In animals, phenological events such as bird migration, reproductive activity, or insect maturation play a role in the ecosystem as well. These recurring events are somewhat variable in nature’s calendar from year to year, and plant phenology is driven by a multitude of factors such as day length, latitude, temperatures, genetics, and complex physiological processes (Friedman et al. 2011). Climate change has been influencing phenology and has received much focus from interdisciplinary scholarly research (Fitchett et al. 2015) as phenology has potential as an indicator of adaptability in a warming world (Badeck et al. 2004; May & Montgomery 2015; Prevey et al. 2020). For those studying phenological phenomena, plants receive significant focus as primary producers for ecosystems, but there are various modern phenological observation methods (Tang et al. 2016) which are discussed in greater detail in the literature review section. The use of digital repeat photography to study plant phenology has become widespread, and the results are often boiled down to a myriad of discreet transition dates which loosely refer to an observed cutoff point such as the timing of “greenness rising” or “leaf fall”. These transition dates are human constructions which help to compare different years of phenological data. Results from previous research indicates transition dates for spring and autumn are well captured by near-surface cameras but calculated seasonal transition dates vary among other methods including ground observations and satellite remote sensing (Hufkens et al. 2012; Keenan et al. 2014; Toda & Richardson 2018). This study seeks to analyze the

link between transition dates from upward-facing digital repeat photography with direct ground observations of tree phenology embedded within a broader research study aimed at better understanding autumn land-atmosphere interactions in a heterogenous ecosystem. Drawing from novel and established methodologies, this study addresses the following research goals:

1. To characterize autumn 2019 within the study area. Specifically, what are the most precise autumn transition dates for nine deciduous broadleaf tree species at field sites throughout the study area?
2. How well does upward-facing digital repeat photography compare against direct ground observations for detecting autumn tree phenology at 53 field sites?

2 LITERATURE REVIEW

Clearly, phenological timing is plastic with regards to climate variability leading to earlier spring and later autumns (Reed 2006; Dragoni & Rahman 2012; Ge et al. 2015). To understand phenological changes over time requires both spatially and temporally extensive datasets, but often these are incomplete. Furthermore, data interoperability issues may arise such as when comparing results from one method to another. As stated previously, there are several methodological approaches used to study phenology, such as direct ground observation, digital repeat photography, satellite remote sensing, carbon-flux measurements, and leaf area index. These methods' advantages and limitations will be discussed in this section, with

particular focus given to direct ground observation and digital repeat photography as they are the two main methods employed by this study.

2.1 Direct Ground Observation

The traditional method to study phenology is to simply watch the visual cues change throughout the seasons and record the time of phenological events. This method has been used for centuries to track culturally and agriculturally significant plant species across the world. Records date back hundreds and even thousands of years ago in some locations (Mikami 2008; Schwartz 2013). Most modern data, however, emerge from phenological networks—often a loose organization which may draw from direct observations made from many individuals from within a geographic area. The first phenological networks were established in 18th century Europe (Dahl & Langvall 2008) and by the mid-1900's several countries had some form of phenological observation network for people to share their data.

In the USA, it was not until 2007 that the nationwide USA-National Phenology Network (USA-NPN) was founded, which has grown to include over 20,000 active observers tracking 1,057 plant species as of 2021 (Schwartz et al. 2012; USA-NPN 2021). With the help of volunteers and organizers, phenological networks have greatly expanded on historical datasets into the 21st century thereby allowing for long-term research at the regional or even continental scale (Schwartz 1994; Cayan et al. 2001; McCabe et al. 2011). Long-term datasets of phenological observations are rare but valuable; phenological changes are most meaningful when arranged over vast periods of time to observe significant differences at decadal scales. Direct human observation, being the oldest form of phenological monitoring, will always

have a major role in longitudinal studies, even as continuous satellite monitoring becomes more available.

Observation networks greatly improved the sample size and collaborative efforts for scholarly research (McDonough et al. 2020; full list at: USA-NPN 2021), and networks have also set standard methods for phenological observations, as described by Denny et al. (2014). In these USA-NPN sampling protocols, the authors highlight the problems detecting subtle shifts in species' phenology, as well as the need for long-term data with standard metrics geographically distributed while covering many taxa. These scholars worked closely with the USA-NPN to develop *in-situ* monitoring protocols standardized across taxonomic groups as well as ecosystem type. Protocols cover the visual human observation methods for terrestrial, freshwater, as well as marine plants and animals. The biodiversity covered in this standardization gets at quantifying phenophases in terms of onset, duration, and intensity. Simplification is one goal of this standardization by favoring "status monitoring" techniques which could be answered with "yes" or "no" by the observer. This may be considered an advantage in that it enables citizen science even without training (Feldman et al. 2018) but was not the protocol used in this study for several reasons. First, simplified status monitoring has less precision when it comes to quantifying autumn transition, because trees breakdown chlorophyll and absciss their leaves asynchronously for different parts of the canopy, so that the most sun-exposed tops of tree canopies may be further along in their autumn phenology than other parts. Second, autumn leaf coloration and leaf fall occur more like a continuous trend than a discreet event. Thus, status monitoring protocols are suited for volunteer-based citizen science but less so for more locally specific research studies such as this.

While advancements have been made by regional or continental phenology monitoring networks, Stucky et al. (2018) raise the need for a more global understanding of phenological phenomena, specifically the terms used and globally standardized methodologies. Some programs, such as the Global Phenological Monitoring Programme has shown limited success at building datasets in Europe and North America, with most of the data coming from central Europe. Stucky et al. call for a much broader coordinated effort both to connect the various phenological monitoring networks and get them to agree upon a standard set of definitions. This could better enable climate scientists to utilize phenology of plants across the globe to understand the effects of climate change at various latitudes and ecosystem types.

Beyond protocol standardization, issues persist with respect to direct human observations of phenology, and that comes down to human subjectivity with respect to identifying phenological phenomena. Empirical science tends to favor primary data which are accurate and replicable, but especially tries to minimize observer bias. When humans are the phenological sensors, there may be issues when it comes to things like estimating leaf area percentages, detecting subtle morphology, or even plant identification. There is no direct solution to this challenge as it impossible to count every leaf on a tree, rather the subjectivity is best recognized and mitigated with protocols fit to specific research questions, such as with the ground observations used in Schwartz et al. (2013). Given that there are a multitude of other methods to track phenology, human observations must not be considered “ground truth” because of observer bias which quickly becomes apparent when multiple people attempt to estimate relative fractions of tree canopies in different phenological stages.

Overall, direct observations of phenology in the field remains a powerful tool to study phenology despite its potential introduction of human subjectivity. These observations occur at the individual specimen level which requires upscaling for comparison with other monitoring methods as will be discussed in the following sections on digital repeat photography and satellite remote sensing.

2.2 Satellite Remote Sensing

The geographic scale required to answer some research questions at scale means that direct ground observation data, even from large monitoring networks, are not uniformly spatially distributed or numerous enough to understand plant dynamics across the continuous landscape-- so researchers often use remote sensing. Satellite remote sensing covers large areas at regular and predictable time intervals and traditionally has been used outside of phenology to measure vegetation cover. Remote sensing has been employed for decades to study land surface characteristics with radiometry on a variety of sensor-platform systems, each with their unique set of spatial, spectral, and temporal resolutions. In particular, vegetation indices like the normalized difference vegetation index (NDVI) can detect healthy vegetation by exploiting reflectance values of the near-infrared band which vegetation strongly reflects and red band which vegetation strongly absorbs. Spectral vegetation indices including NDVI and others such as enhanced vegetation index (EVI) have been widely used in phenology (Fischer 1994; Moody & Johnson 2001; Yu et al. 2003). Some other vegetation indices are described in Table 1 in the review by Zeng et al. (2020) which also include biophysical vegetation indices such as leaf area index (LAI) which can be directly measured by instruments (e.g., LAI-2000) on the ground or by satellite-derived data products. Satellite remote sensing is generally not capable of detecting specific phenological events, rather

what can be detected are general landscape dynamics usually referred to as land surface phenology (LSP, de Beurs & Henebry 2005). In this section, the advantages and disadvantages will be discussed to contrast the fundamental concept of scale in satellite remote sensing to track phenology.

The potential of satellite remote sensing for phenology was greatly improved upon by Zhang et al. (2003) in which MODIS vegetation index time-series data were fit to a logistic curve hypothesized to ideally represent four key transition dates: green-up, maturity, senescence, and dormancy. This was a groundbreaking publication and has influenced the field by setting up the theoretical foundation for curve-fitting time series, while allowing flexibility for different vegetation indices to be used, even from other types of remotely sensed data. Other models have been developed, but the basic premise of fitting a function to the time-series of vegetation indices remains a dominant method used in satellite remote sensing of phenology (Fisher et al. 2006; Elmore et al. 2012). While this method works well at understanding broad landscape patterns over large areas, a clear distinction must be noted that ecosystems are actually comprised of countless individual interactions between the environment and biological cycles which may have their nuance washed out by a single value covering an arbitrarily defined square pixel.

Schwartz & Liang (2009) and Liang et al. (2011) reconciled the different scales between field and space with hierarchical upscaling to produce a “landscape phenology” at meso-scales, a flexible term referring to the bridge between small- and large-scale observations. Individual phenological events directly observed were cross-compared with remotely sensed data to check for agreement and provide an upscaling method similar in scale to satellite data. This represents an epistemological shift to understanding *in-situ* phenology as too complex and

heterogenous for direct human observations to adequately capture, but with a logical spatial grid of ground observations the individualized data can be upscaled to be comparable with remotely sensed LSP. This study also helped bridge the gap between “landscape phenology” and “land surface phenology”, the latter being what is actually measured by MODIS and other satellite-derived phenology.

Overall, remote sensing methods continue to improve and will always have limited applicability for local studies but are well suited for many types of regional or continental scale research. In their review of phenological research using remote sensing, Zeng et al (2020) discusses why remote sensing with satellites matters in the context of science, agriculture, and environmental monitoring. They discuss more of the technical progress in the field as well as problems of scale, data contamination, and lacking ground observations for “truthing”. The case has been made that global phenological change studies are best situated to take advantage of satellite-derived phenology metrics. This methodology presents opportunities for interdisciplinary collaboration among physical geographers to answer the bigger questions about planet-wide ecological seasonality.

2.3 DIGITAL REPEAT PHOTOGRAPHY

Digital cameras are ubiquitous in everyday 21st century life, as well as research. They are relatively cheap, reliable, and can produce data continuously with basically no storage limitations. When applied to phenological monitoring, this is in contrast with human observations, which either rely on sparse volunteered data or hired seasonal observers (which for remote research plots can be quite a labor investment). Furthermore, satellite imagery from high-quality commercial satellites is expensive and still requires highly skilled technicians to process multispectral

data. However, digital cameras can be mounted to a platform to capture images remotely at regular intervals, or (as in the case of this study) carried with human observers to revisited study sites to capture semi-regular images from the ground. Here, a distinction must be made between upward-facing digital repeat photography (DRP) and above-canopy, oblique-view DRP which is more common with open-data networks such as in the PhenoCam program (PhenoCam Explorer 2019). An overview of DRP and some gaps in the literature will be discussed in this section, along with the connection between DRP and direct ground observations.

Digital repeat photography in phenology can be theoretically reduced to capturing a time-lapse of vegetation photos of the same view over the course of weeks or longer. As vegetation transitions through phenological stages, visual cues (examples include green leaves emerging, or autumn coloration) are detectable by the sensor and a phenological signal can be processed from those images. For example, Kato and Komiyama (2002) proved that the first leaf flush in spring can be detected by upward-facing fisheye cameras through the decreasing light levels which is useful for understanding forest understory dynamics. Since the early 2000s, several methods involving image brightness thresholding or color ratios with digital repeat photography have been deployed to monitor plant phenology.

Significant progress on this topic of using digital cameras to study forest phenology can be traced back to work by Andrew Richardson (Richardson et al. 2007; Richardson et al. 2009) in the Bartlett Experimental Forest. With an existing carbon flux tower in use at the site, a camera was mounted above the tree canopy looking down on a section of forest and multiple images were saved each day to construct a time-series. With the connection already made in the previous section of this paper between satellite radiometry and phenology, the signal captured by

traditional digital cameras are ratios of color, specifically red, green, and blue (RGB). Richardson and colleagues proved that plant phenology can be studied cheaply and reliably with cameras. However, some disadvantages were recognized right away, such as varying brightness conditions affecting the relative color ratios (the various RGB indices used to detect different phenological activity). As expected, greenness of images increases in spring and decreases in autumn, while redness (RCC) increases in autumn as chlorophyl levels in foliage decreases. Furthermore, Ryu et al. (2012) showed that upward-facing cameras can accurately measure leaf-area-index (LAI), or the area of foliage, when compared with data from LAI-2000 and litterfall traps. Research done in oak-savannah by Liu et al. (2018) agrees with this and show that image greenness (GCC) is negatively correlated with leaf fall rates detected by litterfall boxes. They also draw links between various foliar senescence rates with phenological events extracted with Bayesian change point detection. It seems reasonable to assume what works in oak savannah is different but not too dissimilar from other ecosystem types such as the forests of northern Wisconsin.

One challenge to using digital repeat photography *in situ* is changing light conditions due to things like weather, solar angle, or technical issues. Anyone who has taken a photo of the bright daytime sky can tell you that too little camera exposure can wash out darker features which, in the case of phenological monitoring, can influence results. Foundational research on this topic in forest phenology research came from Zhang et al. (2005), who found that 16-71% of leaf area can be undetectable at automatic exposure settings on a DSLR camera. They recommended one-stop of underexposure and reference protocols which significantly reduced bias, when compared with ground measurements. Publications

by Macfarlane et al. (2011, 2014) also recommend under-exposure settings and provide standardization techniques for scene illumination levels for upward-facing photographs and above-canopy DRP. Canopy gap fraction increases for upward-facing image series in autumn as leaves fall off branches, thus transitioning pixels to detecting sky. This is an advantage, in that primary data can be processed in multiple ways to understand several forest foliar traits, namely leaf area index, gap fraction and GCC/RCC.

Phenology can be studied at multiple scales, typified by highly localized direct ground observation studies and with broad spatial extent best studied with satellite remote sensing. However, like ground observation networks, networks of near-surface digital sensors help bridge the different. These networks include “PhenoCam”, “European Phenology Network”, and “Phenological Eyes Network” which have been established at hundreds of sites including forest, grassland, and agriculture (Richardson et al. 2009; Motohka et al. 2010; Nasahara & Nagai 2015). These camera networks are capable of monitoring leaf, tree, and landscape phenology, but do require significant calibration to be fully interoperable (Wingate et al. 2015). Nonetheless, at least in North America the potential of the PhenoCam network has been expanded by Richardson et al. (2018) which processed imagery from over 130 cameras across dozens of vegetation classifications. The spatial extent for their data covered mostly the continental USA, and they provide phenological transition dates based on “greenness rising” and “greenness falling”. The number of images now publicly available is vast, with the PhenoCam program adding about one million images per month (“PhenoCam”, NAUTV 2019). Preprocessed repeat digital photography from these open-data platforms enables sophisticated analyses at broader geographic scale than could be done otherwise.

Reliable digital repeat photography does not have to come from professional academic camera installations, such as those mounted on human-made towers, but could someday come from citizen scientists in a similar fashion as the USA-NPN's Nature's Notebook (observation program). Access to high-quality digital cameras is not uncommon in the United States, which could be pointed toward trees and other plants during key transitional periods. Of course, volunteered geographic information can be extremely valuable for research but citizen scientists often lack formal training (See et al. 2016) which could present issues for data quality despite standardization protocols like those published by Denney et al. (2014). Still, volunteered information from citizen scientists using digital repeat photography could one day be an additional source of data for phenology research. The use of digital repeat photography continues to become more accessible as previous limitations like resolution and storage limits attenuate, which support the notion that citizen science could someday have a role in digital repeat photography to study phenology.

Digital repeat photography, both downward facing and upward facing, has been established as an objective way to monitor phenology, and has mostly done with fixed-camera installations. Scholars highlight the need for ground validation across different vegetation classifications (Tang et al. 2016) to allow for "big-data" breakthroughs. Also, there is still a need to test mobile upward facing photography which could enhance our ability to monitor phenology at the ground level. This study builds on ground-based methods –direct human observations alongside digital repeat photography in a heterogeneous temperate deciduous forest.

3 METHODS

3.1 DATA COLLECTION

3.1.1 Human observations

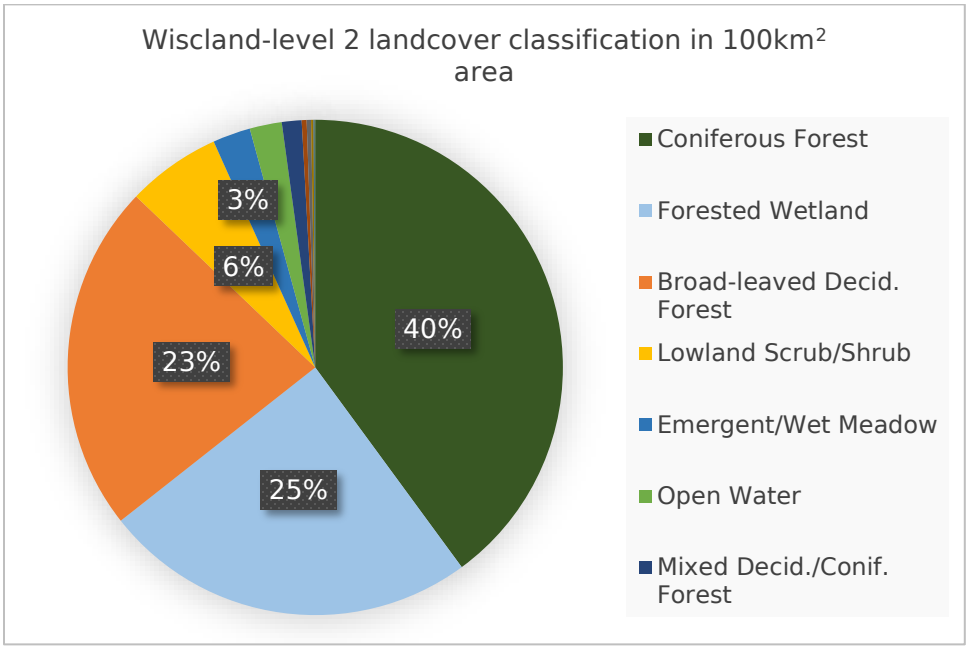
3.1.1.1 Study area

The Chequamegon Heterogeneous Ecosystem Energy-Balance Study Enabled by a High-Density Extensive Array of Detectors 2019 (C.H.E.E.S.E.H.E.A.D.19) project took place within a 10x10km square area of forested and aquatic landscapes in northern Wisconsin. The study domain location, centered on the existing Park Falls “tall” 447m tower AmeriFlux/NOAA supersite (U.S.-PFa/WLEF) was partly chosen due to its history of atmospheric and environmental research (Butterworth et al. 2020). Past research in the area includes previous years of phenological human observations, above-canopy Phenocam imagery, and under-canopy light sensors.

Wisconsin’s glacial history influenced the land within the study area, with 45% of the study area classified by the Wisconsin Department of Natural Resources (WDNR) Land Type Association as “Chequamagon Washed Till and Outwash”, 31% “Glidden Drumlins”, and 24% “Northern Highland Outwash Plains”. Before European colonization of the area, the land was territory of the Anishnaabe (Ojibwe) tribe. Nowadays, the 100km^2 domain is a mixture of private land and USFS National Forest. There is still some active forestry in the area, including management of even-aged pine plantations within the study area although this had no known influence on results of this study.

To broadly summarize the heterogeneous vegetation landcover in the area, the “Wiscland 2” classification data product was used (Wisconsin DNR, 2016). This dataset was developed by the WDNR using machine learning classification on Landsat data in conjunction with ground validation sites. The level-2 classifications by areal percentage of the study area are presented in *Fig. 3-1*.

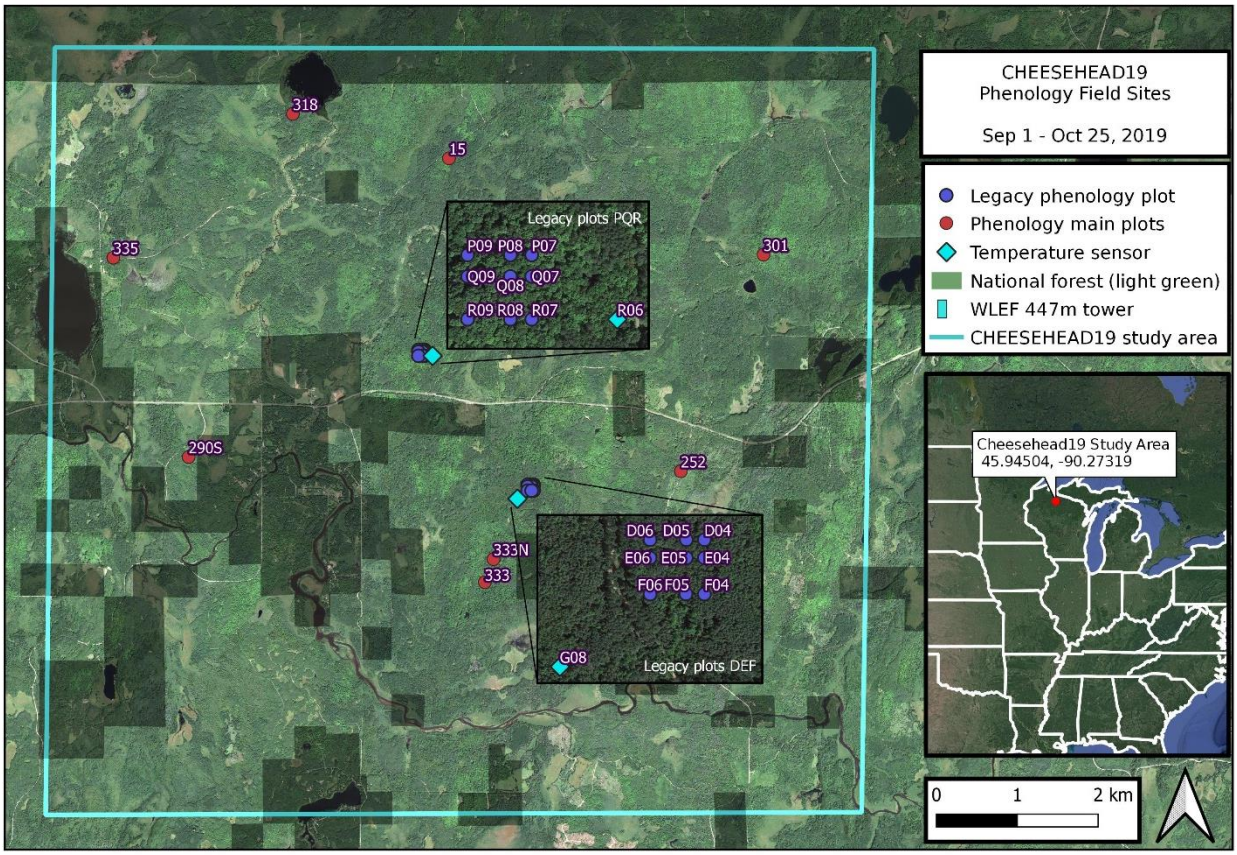
Figure 3-1: Landcover classification of study area



Coniferous forest, Forested wetland, and Broadleaf deciduous forest dominated the area, with several other broad land cover classifications present. In actuality, the landscape in the study was more complex than the broad landcover classifications used in Wiscland 2 which was useful for getting a general understanding of the heterogeneity of the study area.

3.1.1.2 Sampling design

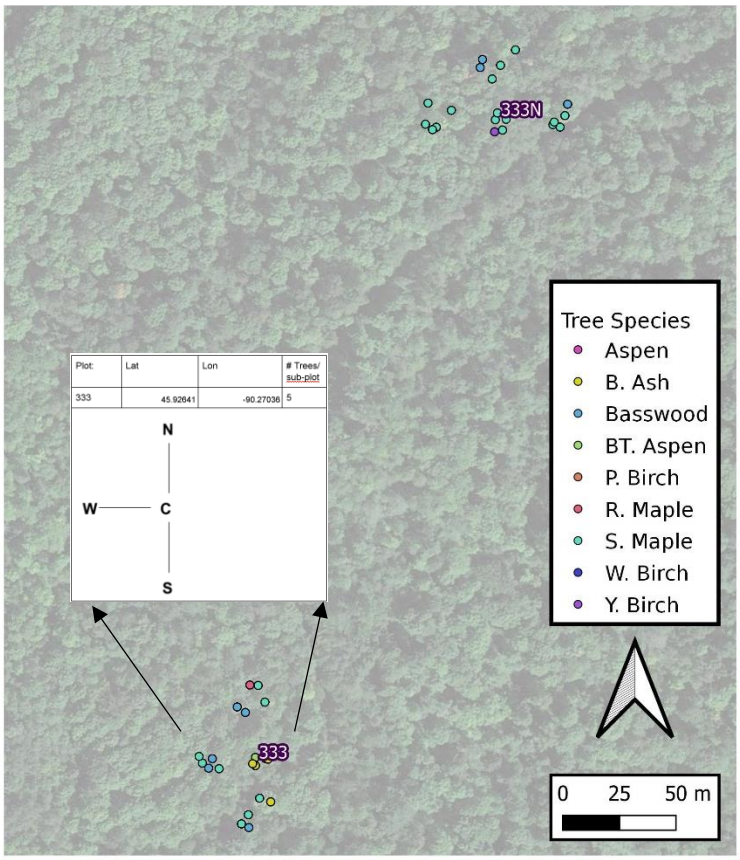
Figure 3-2: Map of phenology plots. Main plots are marked with red circles and legacy plots marked with blue circles in insets. Plots are labelled with their designation used in the Cheesehead19 study.



The unique situation of CHEESEHEAD19 allowed for studying many aspects of land-atmospheric interaction, and plant phenology monitoring was just one aspect of the broader research context. Field surveys aimed at detecting tree leaf characteristics such as pigment content and plant primary production were conducted within this study area throughout the growing season, although those data were not used in this study. Of the eight phenology mainplots added for this study, some were also field sites at which the leaf characteristics and plant growth were monitored. The eight phenology mainplots are shown in *Fig. 3-2* as red circles, and at each mainplot there were four or five subplots arranged with one subplot at

the mainplot center with other subplots at 25m distance in cardinal directions around the center. A map with tree locations for two mainplots dubbed 333 and

Figure 3-3: Example of subplot orientation around mainplots "333" and "333N". Tree Locations marked with colored circles indicating their species.



333N are shown in *Fig. 3-3*.

Two legacy phenology plots called DEF and PQR were reused from previous studies (Schwartz et al. 2013) which featured a grid system of subplots arranged along linear transects (shown in map insets within *Fig. 3-2*). These legacy plots had three to six trees observed at each location.

This plot design was used to increase observer efficiency because by keeping subplots clustered around mainplots then there was less travelling, while still spatially arranging trees apart from each other by 25m to reduce potential bias from over-clustering. The number of subplots varied per mainplot to avoid placing research plots at non-representative locations such as too close to a road or to avoid forest gaps. The total number of dominant canopy deciduous broadleaved trees at all the subplots for all eight mainplots was 20 individuals. Observed trees were selected based on proximity to subplot center and overall tree health. Each of the trees in this phenological study had their spatial relation in terms of distance and degrees to the sub-plot center point recorded, as well as their species. The diameter at breast height (DBH) was measured initially but was unused in the analysis. All 223 specimen trees were visually marked and had their GPS coordinates taken with high precision so that observers could find them during revisits when they recorded their phenological state and took images of the canopy.

A baseline phenological observation for every one of the marked trees (n=223) was taken on two dates in early September 2019 at which point nearly all the study trees still had 100% of their green leaves. From September 17th onward, sub-plots were revisited every two or three days to record the “event number” which ranged from 800 to 890 for leaf color, and 900 to 990 for leaf fall (*Table 3-1*).

Table 3-1: Autumn phenological event protocol (from Schwartz et al. 2013)

Event number observation codes			
Leaf color *(including those on ground)*		*Leaf fall* *(based on total leaves initially)*	
800	<10% of leaves colored	900	<10% of leaves fallen
810	10-50% of leaves colored	910	10-50% of leaves fallen
850	50-90% of leaves colored	950	50-90% of leaves fallen
890	>90% of leaves colored	990	>90% of leaves fallen

At the eight main plots, there were a total of 33 subplots and there were a total of 20 legacy subplots in the “DEF” and “PQR” study areas. When these sub-plots were revisited, a digital fisheye-lens photograph was taken at zenith on the center of the sub-plots; this is discussed in greater detail in the following section.

Event numbers for both leaf color and leaf fall were intended as a category of the relative percentage of a tree’s canopy so there was some human subjectivity when it came to estimating fractions of canopy. To account for some of the observer subjectivity, field observations of event numbers and digital images were taken by a three-person team, which helped with data consistency by allowing observers to check for agreement on their estimated event numbers. Data were recorded up until October 25, 2019 (DOY 298), when all the study trees had reached near-maximum leaf color and leaf fall. At this point, over 90% of leaves had fallen from the study trees and plots were longer visited.

3.1.2 Digital Camera Images

3.1.2.1 Image capture

As observers visited sub-plots to record event numbers for trees, they carried with them a DSLR camera (Canon EOS Rebel T3i) with fisheye lens, also referred to as hemispherical lens, mounted on a folding tripod. A radial bubble level permanently affixed above the lens was used to point the lens at zenith directly above the subplot's center (marked with permanent stake). The radial orientation could vary each time an image was taken, but with care to make sure the aperture was level (with bubble level) and at sub-plot center (marked with a stake). At the end of each field day, the image files were moved to external storage and saved as the raw CR2 format and as JPEG. As not every plot was able to be visited on the same date, the total number of images for the plots varies, with the maximum number of images in a series twenty, and the minimum number sixteen. Horizontal images were captured to attempt to measure understory phenology but not used in this study.

3.1.2.2 Exposure setting

The level of brightness in an image captured with a DLSR camera is determined by several settings which can be set manually or automatically by the device. In this study, there were two images taken every time a site was visited: with manual settings and with automatic exposure settings. ISO speed determines how sensitive the camera's sensor is to light, aperture controls the lens diaphragm to allow more or less light reaching the sensor, and shutter speed affects how long light is allowed to reach the sensor.

Other studies have shown that phenological dates derived from upward-facing photographs can be sensitive to changing scene illumination (image brightness), which can vary day-to-day or even by time-of-day, but Macfarlane et al. (2014) developed a simple protocol to standardize photographic exposure to be used for sensitive phenological metrics. They recommend standardizing under-exposed images in their raw file format by applying a contrast stretch to the color image. This was shown to reduce the effect of image brightness variation on canopy gap fraction, leaf area index, and chromatic coordinate calculations.

In this study, to correct for varying brightness conditions washing out foliage pixels, an algorithm for color image histogram equalization was developed in Python for this purpose. This histogram equalization reduced the variability in brightness levels in the under-exposed images, which allowed for finer tree canopy details to be visible whereas in the non-standardized under-exposed and over-exposed images fewer fine details could be detected (*Fig. 3-4*).

Figure 3-4: Example of histogram equalization on photo

Automatic exposure	Manual exposure (f/3.5 stop; 1/4000 sec)	Manual exposure with Histogram equalization
		

In the image on the right, there are fine details at the tips of branches including branches themselves and foliage which is not visible in the automatic exposure image on the left (*Fig. 3-4*). Based on the advantages of applying contrast stretching to under-exposed images, this standardization method was applied to each image for all sub-plots and were the imagery data used for further analysis.

3.1.3 Temperature Data

Micrometeorological data was available at two legacy sites (G8 and R6 shown as cyan squares in *Fig. 3-2* insets), which were utilized to understand the inter-annual differences in microclimate temperature. This was useful because the legacy phenological observations were near to these temperature sites. Temperature data were collected with HOBO data loggers attached to the north side of large, mature trees and set to take air temperatures (at ~1.5m shelter height) and 20cm deep soil temperatures every ten minutes within the repurposed legacy plots. Data were available for 2010, 2012, 2013, and 2019. Other years' data were available but not analyzed because legacy phenological observations were not recorded in those years. Statistical analysis on these datasets included ANOVA and post-hoc tests were performed on temperature data at G8 and R6 across the different years.

3.2 ANALYSIS

3.2.1 Year 2019 temperature context

At the reused legacy plots (DEF and PQR) where historical ground observations and temperature series data were available, many of the same individual trees were still standing and resampled with the same observation method. This allowed for contrast with other years to understand general comparisons such as whether autumn was late or early in 2019. Every tree which

was sampled in 2019 and in past years was summarized with descriptive statistics grouped by species. Non-linear regression was performed on individual trees' data at legacy plots by fitting a sigmoid curve and calculating its inflection points. These inflection points were compared interannually using ANOVA on the sample means and post-hoc tests.

Finally, a basic comparison of the average first date of 850, 890, 950, and 990 event numbers was performed on the median values by species. This was considered supplementary to the post-hoc tests to generally characterize 2019 tree phenology.

3.2.2 Direct phenological observation

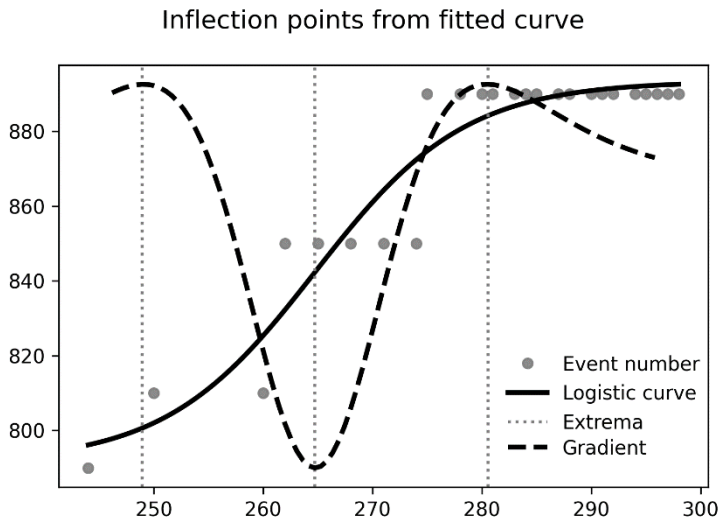
3.2.2.1 Sigmoid curve fitting

To understand the rate of change of event numbers for leaf color and leaf fall, it is pragmatic to look for distinct singular dates in time which could represent the "beginning", "middle", or "end" of autumn, which can be observed through time as leaves change color and absciss. A logistic curve was fitted by minimizing least squares of residuals between observed and expected curve values. Sigmoidal curve fitting was done using Scipy Using the "dogbox" method of minimizing least squares. The use of logistic curves to model phenology was best described theoretically in Zhang et al. (2003) to represent spring and autumn vegetation indices. The sigmoid function fitted to leaf color values and leaf fall values is defined as:

$$y = \frac{L}{1 + \exp(-k(x - x_0))} + b$$

For this sigmoidal logistic equation, y is the median event number, x is the independent date variable as day-of-year (DOY), and $k, L,$ and b are fitting parameters. This logistic sigmoid function is best understood as starting and ending periods of no y-value change, and a critical exponential relationship when the phenological event numbers were changing most rapidly. Therefore, x-values which define the beginning, middle, and end of the steep slope can be calculated using differentiation. In other phenological studies these critical dates calculated from vegetation indices are commonly used to describe the start of season (SOS), middle of season (MOS), and end of season (EOS). Spring or autumn could be understood in this fashion, as best described in Zhang et al. (2003). Autumn is the only season of interest for this study, and event numbers will be used for the vegetation indices; the first inflection point will be thought of as the start of autumn, second inflection point for middle of autumn, and third inflection point as the end of autumn following suit of other studies. An example of the intermediary curves and final inflection points (extrema) are shown in *Fig. 3-5.*

Figure 3-5: Example of subplot 15-C fitted sigmoid curve with curvature and inflection points



The three inflection points were identified using extrema on rate of sigmoid curvature (shown with the dashed “Gradient” line in *Fig. 3*-5); these inflection points were expected to represent the “beginning”, “middle”, and “end” of the phenological response for both leaf color and leaf fall.

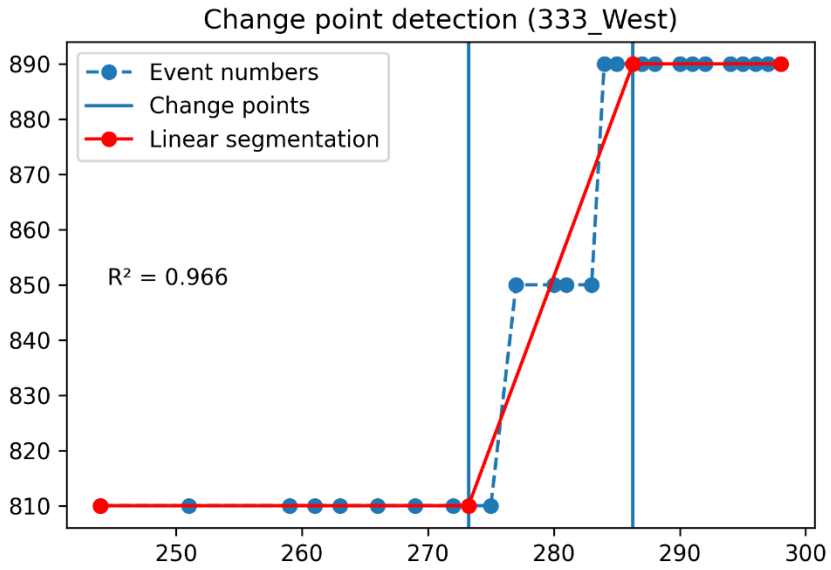
R-squared and MSE for each fitted logistic regression were calculated at the individual tree level and median sub-plot values. Different aggregation methods were used: median values by species, by subplot, and no aggregation (by tree). By species aggregation was useful for understanding the difference between the nine studied trees, whereas the by plot aggregation was useful to compare human-derived transition dates with camera-derived dates. By tree aggregation proved computationally demanding and resulted in some obviously fallacious results and were not used for statistical analysis. The calculated transition dates using sigmoid curve-fitting were stored in csv file format to be compared with transition dates from other methods.

3.2.2.2 Change point/linear segmentation

Recent studies support the use of linear segmented regression as an alternative to sigmoidal curves. Linear segmentation is conceptually more simplistic in that the vectors connecting change points indicate a rate of change as well as well as having a discreet beginning and end at the “change points”. This emerging method has proven more effective at linking human-observed ground observations with MODIS vegetation indices products (Xie & Wilson 2020) and with repeat digital photography relative greenness (Liu et al. 2018).

This study applies the linear segmentation method on event number data aggregated with median event number values by subplot to calculate the DOY at which change points fall (*Fig. 3-6*).

Figure 3-6: Example of linear segmentation and change points at subplot 333-W



The first and second change points, shown with dashed blue lines in the example above (*Fig. 3-6*) were then stored in csv file format to be compared with transition dates from other methods.

3.2.3 Digital Camera Images

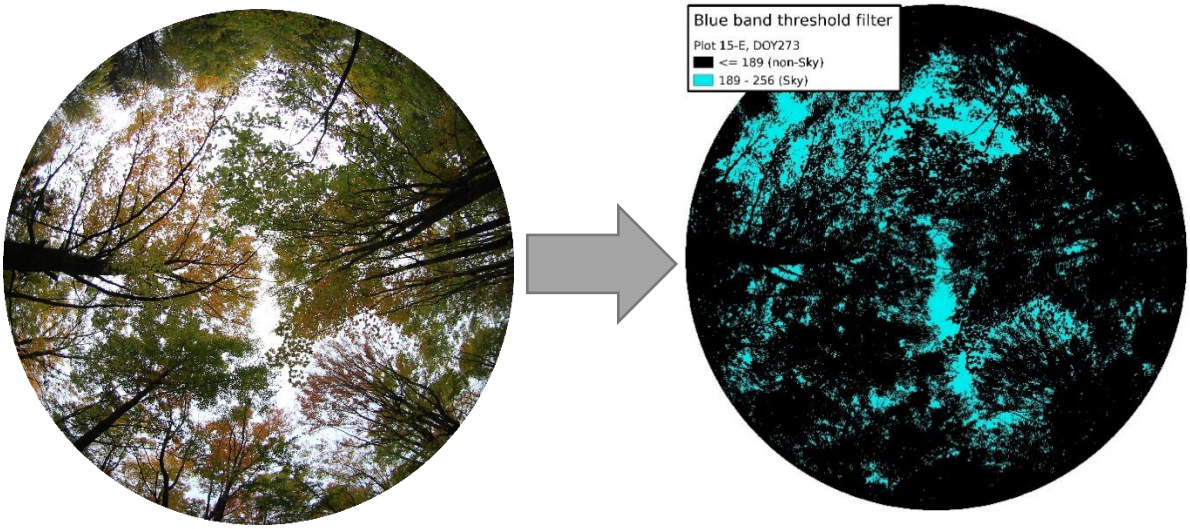
3.2.3.1 Preprocessing

With just under 100GB of data collected as fisheye lens photography and over 1,000 unique images, extracting meaningful phenological data proved to be a significant undertaking. The first step in pre-processing the images was to remove the black rectangular space around the circular fisheye view, resulting in a circle with a width x height of 2200x2200 pixels. This was done by adding a fourth band, “alpha”, which was binary and commonly used in digital images to make pixels

appear transparent. Next, histogram equalization following contrast stretching methods supported by Macfarlane et al. (2014) was performed on the manually exposed images. This histogram equalization method increased contrast between very bright pixels of the sky and darker non-sky pixels. The effect this had on image sharpness by reducing washing out effect is discussed in previous section 3.1.2.2 as it relates to image exposure setting.

Next, to remove non-sky pixels from the masked and contrast stretched images, a simple digital number (DN) threshold on the blue band was performed. MacFarlane (2011) advocates for the use of simple classification methods over more complex methods as it does not significantly affect calculated phenological transition dates. A threshold of Blue DN greater than 190 (DN>190) was selected based on data exploration as well as trial and error using different thresholds. In agreement with this, Leblanc et al. (2005) show that foliage have a much lower reflectivity and transmittance of blue light. This performed well for images at different plots across different light/sky conditions with little to no filtering-out of non-sky pixels. An example of this thresholding is shown in *Fig. 3-7*.

Figure 3-7: Example of sky pixels vs non-sky pixels using blue band thresholds



The preprocessing of the images was now complete at this point, leaving only non-sky pixels for every image. However, these pixels could contain not only the foliage of tree canopies, but also things like branches, tree stems, or other obstructions not being studied. Image thresholding using Red, green, or blue bands proved ineffective at removing these obstructions from the images as they shared too many similarities with foliage pixels, and their pixel locations changed from one image in a series to the next.

In some of the image series, obstructions such as shorter understory plants, major lower dead branches, or evergreen conifers (*Abies balsamea & Pinus strobus)* were blocking a clear view to the canopy of the tallest deciduous trees. Some series of images had significant obstructions, which were manually identified as an additional preprocessing step and noted for future considerations of overall data quality and accuracy.

Table 3-2: Image-series quality assessment

IMAGE SERIES QUALITY ANALYSIS	***Conifer presence***			
Obstruction level	**HIGH**	**LOW**	**NONE**	**Total number of plots**
HIGH	3	2	7	**12**
LOW	7	10	9	**26**
NONE	3	4	8	**15**
Total number of plots	**13**	**16**	**24**	**53**

Table 3-2 shows the count of image series which had evergreen conifers present in the camera view and, as a separate category, those which had >50% of field-of-view obstructed with non-foliar elements such as significant dead lower branches, or understory shrubs. 29 image series out of 53 had some level of evergreen conifer presence, while 38 image series out of 53 had some level of image obstruction. These obstructions were manually identified using a visual estimation since this data was not using statistically, only anecdotally.

Despite about half of subplots having some level non-removable obstructions, all 53 image series were included in the rest of image analyses.

The number of pixels relative to non-sky pixels can be considered the image gap fraction, but the sampling design limited its usefulness. To calculate leaf area index (the total area of foliage in a system), with this method would have required a more sophisticated photography method and thus was not suited to this study.

However, the relative number of sky pixels to non-sky pixels can be considered a validation of the blue-band DN threshold of >190, because the relative fraction of sky pixels generally increases throughout autumn. Generally, the relative fraction of sky pixels increased throughout autumn as expected.

Macfarlane et al. (2014) describe these principles of processing upward-facing digital images to study phenology and proved (using a variable gap screen and *in situ* monitoring) that ideal exposure settings would decrease over time in autumn as the gap fraction increases under similar light conditions. Without knowing the gap fraction *a priori*, two to three stops of under-exposure are recommended and was utilized in this study. Mixed pixels were not filtered out because a similar study by Macfarlane (2011) indicated that mixed sky-foliage pixels are at maximum 10% of non-sky pixels, which decreases over time and was deemed acceptable.

3.2.3.2 Chromatic coordinates

As trees progressed through their autumn phenology, the leaves exhibited a response observed in the red, green, and blue visible spectrum. Using the decreasing values of relative greenness in an image series has been validated against other phenological monitoring methods most recently by Liu et al. (2018, 2020) as an effective and simple method to use in repeat digital images in the visible electromagnetic spectrum. The term used is green chromatic coordinate, defined as:

$$GCC = \frac{G}{R + G + B}$$

For context, GCC is expected to increase in spring when more leaves are present to reflect or transmit green light and is expected to decrease sharply in autumn when chlorophyll breaks down and leaves eventually senesce. Other band ratios have shown success, such as RCC and VARI (Keenan et al. 2014). The red chromatic coordinate (RCC) is most used in autumn phenology and is understood as the relative redness of an image which would increase when leave display characteristic fall colors due to other leaf pigments. Another chromatic coordinate (VARI), though less widely used, shows promise at detecting anthocyanin content in leaves which becomes exposes in autumn as chlorophyll breaks down (Viña & Gitelson, 2010). These three chromatic coordinates, or RGB band ratios, are defined as:

$$GCC = \frac{G}{R+G+B} \qquad RCC = \frac{R}{R+G+B} \qquad VARI = \frac{G-R}{R+G-B}$$

In this study, these chromatic coordinates were calculated for every preprocessed image for all 53 sub-plots using the python imaging library (PIL). This method proved fast and effective at these rigorous algebraic calculations, by calculating the mean DN values for the red, green, and blue bands then calculating chromatic coordinates from that. However, VARI was calculated but not analyzed further due to time constraints. The series of chromatic coordinates were stored as csv format for future steps to smooth the data and then analysis.

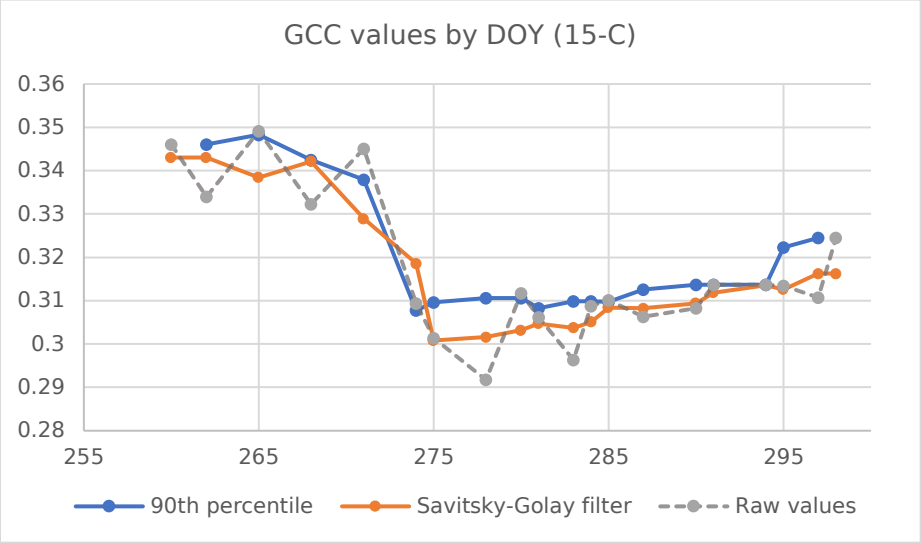
3.2.3.3 Smoothing the data

The calculated relative greenness (GCC) and redness (RCC) in their raw states were plagued with “noise”, but following procedures developed by Sonnentag et al. (2012) using the 90th percentile moving window algorithm (per90) the level of noise was reduced. This smoothing algorithm considers each value at a particular

date a function of itself and its neighbors and assigns the 90th percentile value to the center date. This algorithm is limited, in that it reduces the number of data points due to edge cases. While this is acceptable for other studies with high temporal frequency and larger datasets, the data in this study were sparser so the loss of samples from edge cases was not acceptable. The per90 smoothing method was tested but ultimately unused for this reason.

To preserve as much data points as possible while still smoothing GCC, RCC, and VARI values, a less-used algorithm called Savitsky-Golay (SG) signal filtering was employed. This method uses a flexible local polynomial function fitted to a subset of data points, then iteratively adjusts the dependent variable to reduce error. This method is flexible in terms of defining the window of sub-set points and the degree of smoothing. The window for this method was set to 3 points, so that a local polynomial function was fitted to each set of three adjacent data points. This is like the per90 method but can account for edge cases to not reduce sample size. A more detailed description of the SG filter, its parameters, and applications are discussed in Luo & Bai (2005). The smoothing effects on the green chromatic coordinate (GCC)are shown below comparing the raw GCC values, per90 smoothed GCC, and Savitsky-Golay smoothed GCC for a single sub-plot 15-C (*Fig. 3-8).*

Figure 3-8: Example showing the effect of two different smoothing algorithms



Aside from the loss of edge cases with the per90, the smoothing effect was only slightly different, and it was assumed that the effect on results would be negligible. Therefore, the Savitsky-Golay filter was performed on raw GCC and RCC values and resultant data were used for the next step in analysis, calculating transition dates.

3.2.4 Date extraction from chromatic coordinates

With a smoothed time-series of chromatic coordinates for each set of images (from the 53 subplots) it is visually apparent that GCC and RCC are sensitive to phenology. GCC decreases with leaf color and leaf fall, and RCC undergoes a peak of values, but then decreases to some minimum value. The next step in this analysis was to extract discreet dates (DOY) at critical points in the GCC and RCC series which could be compared with critical dates from human observations.

3.2.4.1 RCC Spline max

The first method involves the highest value of image redness, RCC, which was observed in the smoothed data. This is straightforward and proven effective by

a similar study (Xie et al. 2018). However, my study had special considerations in that due to project limitations, some days there were no images of the canopy taken. Those days in autumn without images could have potentially had greater image RCC values indicating greater abundance of leaf color right before leaf abscission occurs. So, to attempt to account for these data gaps, a novel method was applied called RCC Max-Spline. A spline curve was calculated as centered at the maximum observed RCC value plus the nine days prior and after the max value. This was designed so that the maximum y value of the spline may fall at a different date than the observed maximum value, thereby interpolating to fill in the days which had no image taken. The accuracy of this method was tested against the human observation critical values for both leaf color and leaf fall using Pearson correlation as well as the difference in days between corresponding transition dates from other methods.

3.2.4.2 GCC Sigmoid curve

The negative sigmoid curve is a staple for autumnal phenological studies and was fitted in this study to GCC values. This was done with the “dogbox” method of least squares minimization for each series of GCC values at every sub-plot, using the same sigmoid function fit to direct observation data. From the fitted sigmoid curve, the three critical inflection points were calculated and stored as csv file format for comparison with human observations. R-squared values were calculated for the non-linear regression during model fitting. The covariance of model parameters were stored for future error analysis, as well as the fitted curve parameters itself to avoid having to perform regression again and again. The inflection points were calculated using extrema from the rate of curvature, or gradient, in a very similar fashion as the example in *Fig. 3-5*: the only difference

being that the fitted sigmoid curve was downward sloping (negative) because the GCC values decreased throughout autumn. The three calculated inflection points from the GCC sigmoid curve were statistically compared with transition dates of other methods.

3.2.4.3 GCC Linear segmentation

Much like the linear segmentation method to calculate the change points for human observation data, the same principle was applied to digital camera images. Three linear segments were fitted to GCC values, then the first and second change points were stored in csv format to compare with corresponding human observations critical dates (using Pearson correlation coefficients as well as bias between methods).

3.2.5 Cross-method comparison

In total, there were ten calculated transition dates across the different methods based on human observation data (five from leaf color and five from leaf fall), and six transition dates based on photograph-derived data (five from GCC and one from RCC). A list of the transition dates are given in *Table 3-3*.

Table 3-3: List of Phenological Transition Dates

HUMAN OBSERVATIONS	DIGITAL CAMERA IMAGES
Fitted leaf color sigmoid curve (inflection points 1, 2, 3)	Fitted GCC sigmoid curve (inflection points 1, 2, 3)
Fitted leaf fall sigmoid curve (inflection points 1, 2, 3)	GCC Linear segmentation (change points 1, 2)
Leaf color linear segmentation (change points 1, 2)	RCC Spline (maximum)
Leaf fall linear segmentation (change points 1, 2)	
Total: 10 calculated transition dates	*Total: 6 calculated transition dates*

Methods were checked for validation such as between inflection points on leaf color and with change points on leaf color, but also compared between human observations and photographs. Pearson's correlation coefficients (with associated significance values) were calculated for each pair of data, as well as the difference, or bias between dates by subplot. In all, six critical dates from digital camera methods were compared with five critical dates for human observations of leaf color event numbers and five critical dates from leaf fall event numbers.

4 RESULTS

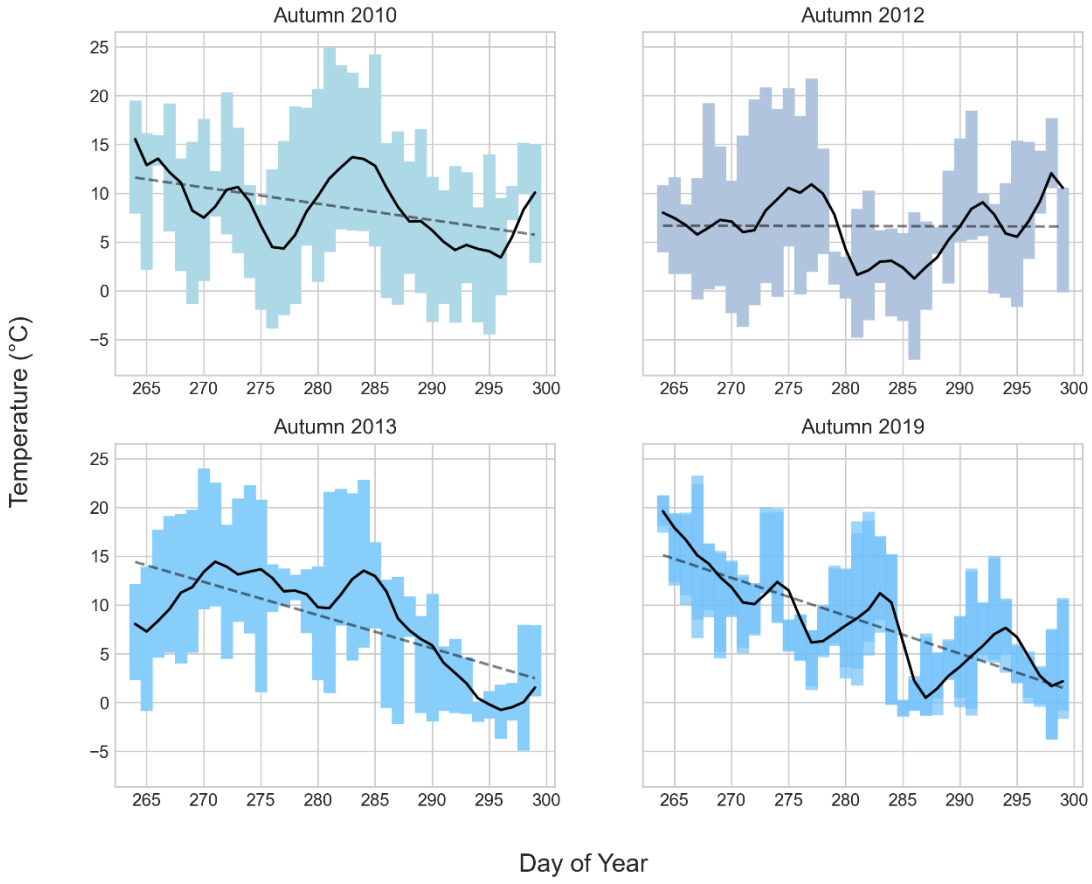
4.1 2019 AUTUMN PHENOLOGY

4.1.1 Inter-annual comparison

4.1.1.1 Temperature

Temperature has been shown to influence spring and autumn phenology, which suggests lengthened growing seasons as a result of accumulated growing degree days (Vitasse et al. 2009; Yu et al. 2016) Based on that simple assertion, daily temperatures throughout autumn (daily range in temperature, 3-day rolling average, and OLS regression line shown in *Fig. 4-1)* were explored to check whether 2019 was an anomalous year compared with 2010, 2012, and 2013.

Figure 4-1: Temperature trends in Autumn for 2010, 2012, 2013, and 2019. Daily temperare range shown with blue bars indicating maximum and minimum temperature, solid black line representing 3-day rolling average, and dashed line representing trendline from OLS regression.



Based on results from daily air minimum temperature ANOVA between groups of different years (p-value=0.0007) and between different sites (p-value=0.038), a main effect was detected between years and between sites, suggesting that at least one of the years was conclusively different. Following the initial two-way ANOVA, I performed a post-hoc test, the Tukey HSD test, to see which years of temperature were different. Results indicate 2019 was on average minimum air temperature 2.8°C warmer than 2012, but not significantly different from 2010 or 2013 based on a 99% confidence that the sample means could not be drawn from the same temperature series. Results of Tukey HSD tests on air mean

temperatures were less pronounced, with no statistically significant difference between means of different years, although autumn 2012 was on average 1.2°C warmer from DOY265 to DOY 300.

Overall, 2019 was not an outlier year based on these comparisons of temperature throughout autumn, but 2019 did have the latest dates for a light frost event (0°C) and hard frost event (-2.2°C). This could be expected to influence tree phenology to some degree, because of the temperature dependence of foliage as well as the susceptibility of frost damage to vegetative tissues.

4.1.1.2 Tree Phenology

At the legacy plots (“D,E,F” and “P,Q,R”) most trees were still alive into 2019 and had their phenological leaf response observed in the same fashion as in the autumns of 2010, 2012, and 2013. 49 trees in 2019 were part of this legacy sub-set. There was only minor mortality in these dominant canopy trees. However, in 2013 only the southern sites (“D,E,F”) were sampled, comprised of 18 individuals. Results from inter-annual, pairwise t-tests on the second inflection point of a sigmoid curve fitted to an individual tree’s leaf color event numbers indicate 2019 was statistically different from 2010 and 2012, but not different from 2013 at 95% confidence level. The same pairwise t-tests on leaf fall indicates the same sample central tendencies, that 2019 was statistically different from 2010 and 2012, but not from 2013. In both cases for leaf color and leaf fall, the second inflection point of the fitted sigmoid curve relates to the overall central tendency of autumn tree phenology.

The start of autumn, represented by the first inflection point on the fitted sigmoid curves for both leaf color and leaf fall were also tested with paired t-tests. Leaf color pairwise t-tests indicate that the start of autumn 2019 was statistically

different from 2010 and 2012 but not different from 2013. These results are the same for the second inflection point. However, 2019 was different than all three years for the third inflection point of leaf color.

Interestingly, the start of leaf fall was not significantly different in 2019 compared with all years of past data based on the first inflection point values. However, the end of leaf fall followed the same pattern as before, that 2019 has a significantly different distribution of data than 2010 and 2012 but not from 2013.

However, these t-tests do not adequately capture the nuances in the data, which would vary by species to some degree in any year. Aggregated by species, median event numbers for leaf color and leaf fall on every sampling date were used to show the central tendency of tree phenology in a season. The difference between the DOY of the median dates (*Table 4-1*) show 2019 was five to eight days later than median dates averaged in the past.

Table 4-1: Difference between 2019 and past years event numbers

2019 – (2010, 2012, 2013)	Leaf color		Leaf fall	
Difference (days)	850	890	950	990
Sugar Maple (A. saccharum)	6	8	8	9
Red Maple (A. rubrum)	5	9	8	6
Aspen (P. tremuloides)	9	8	5	5

However, the sampling window varied in 2019 versus past years, mostly due to the lack of early-autumn observations taken in 2010, 2012, and 2013 (no observations before DOY 264). For this reason, the difference in 810 values were excluded and only the values for 850, 890, 950, and 990 were deemed valid.

The results from inter-annual comparisons of human observations of phenology are confounding. T-tests of sample means of sigmoid inflection points indicate 2019 was statistically different from 2010 and 2012 and somewhat different from 2013. But when looking at median values by species, autumn 2019 happened later than the average of past years' median DOY. Discounting the notion of spurious analyses, these results could be explored further to unravel the nuanced inter-annual differences in autumn tree phenology.

4.1.2 Direct ground observations

Ten transition dates were calculated from the human observation data using sigmoid curve inflection points and linear segmentation change points. Aggregated using median event numbers by subplots, sigmoid curves generally seemed to perform better with a closer model fit (discussed along with *Fig. 4-5*).

For the calculated critical dates by subplots, which included from three to five sampled trees, the mean difference between the first and third inflection point of leaf color was 10.4 days, with a standard deviation of 8.4 days. This is in contrast with the mean difference between the first and third inflection points of leaf *fall*, which was 13.1 days with a standard deviation of 11.2 days. From sigmoid curve inflection points corresponding to the rate of phenological change, autumn lasted less than two weeks for the leaf color and leaf fall transitions.

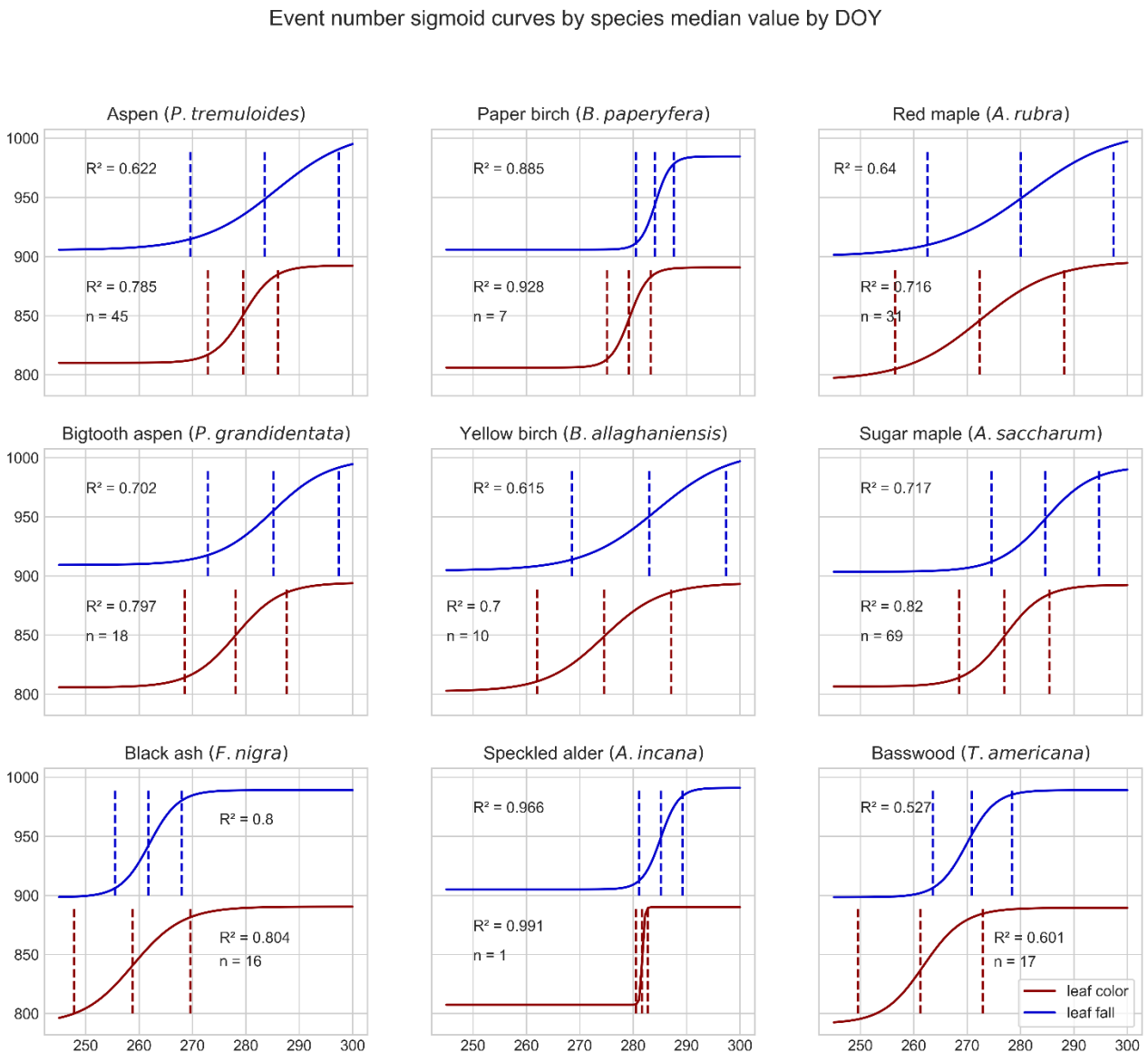
Change points from linear segmentation indicate a similar length of autumn based on the difference from the first and second change points. Leaf color change points were on average 10.3 days apart with standard deviation of 7.3 days. This is very similar to the sigmoid curve inflection point differences on leaf color. However, for leaf fall, change points one and two were on average 5.9 days apart with

standard deviation of 5.8 days. This does not align well with the difference between inflection points one and three on leaf fall, indicating possible issues with the calculation method or an undetected phenomenon.

Comparing leaf color with leaf fall calculated transition dates relates to the delay between chlorophyll breakdown and leaf abscission. From the first leaf color inflection points to the first leaf fall inflection point was on average 3.6 days with standard deviation of 9.9 days. The last leaf color inflection point was on average 6.3 days earlier than the last leaf fall inflection point with standard deviation of 6.1 days. These results indicate that a large portion leaves do not stick around on branches for much longer than a week when they are displaying autumn colors.

The observed phenological responses for trees studied in autumn 2019 were, as expected, varied among tree species. The fitted sigmoid curves with inflection points for the data (aggregated by species) are shown in *Fig. 4-2*.

Figure 4-2: Human observation sigmoid curves by species. Red line indicates fitted sigmoid curve for leaf color observations by tree species. Blue line indicates fitted sigmoid curve for leaf fall. Dashed lines represent inflection points 1, 2, and 3 for corresponding sigmoid curve.



It should be noted that only one individual speckled alder (*A. incana*) was studied, and so the calculated transition dates for that species should not be considered representative of the population due to small sample size. Black ash (*Fraxinus nigra*) showed the earliest autumn transition, with most trees having already dropped their leaves before other species had even hit their peak event numbers.

Basswood (*Tilia americana*) also showed a relatively early response but had greater variance between individual trees indicating a less uniform response for that species. Both species of Aspen *(P. tremuloides & P. grandidentata*) showed a slower leaf fall duration than their leaf color duration, based on the first and third dashed lines indicating sigmoid curve inflection points. The two species of birch were not all that similar to each other; *B. papyrifera* seemed to begin and end their autumn phenological transitions in a shorter timeframe than *B. alleghaniensis*. This is apparent visually by the greater period between the first and third inflections points for yellow birch.

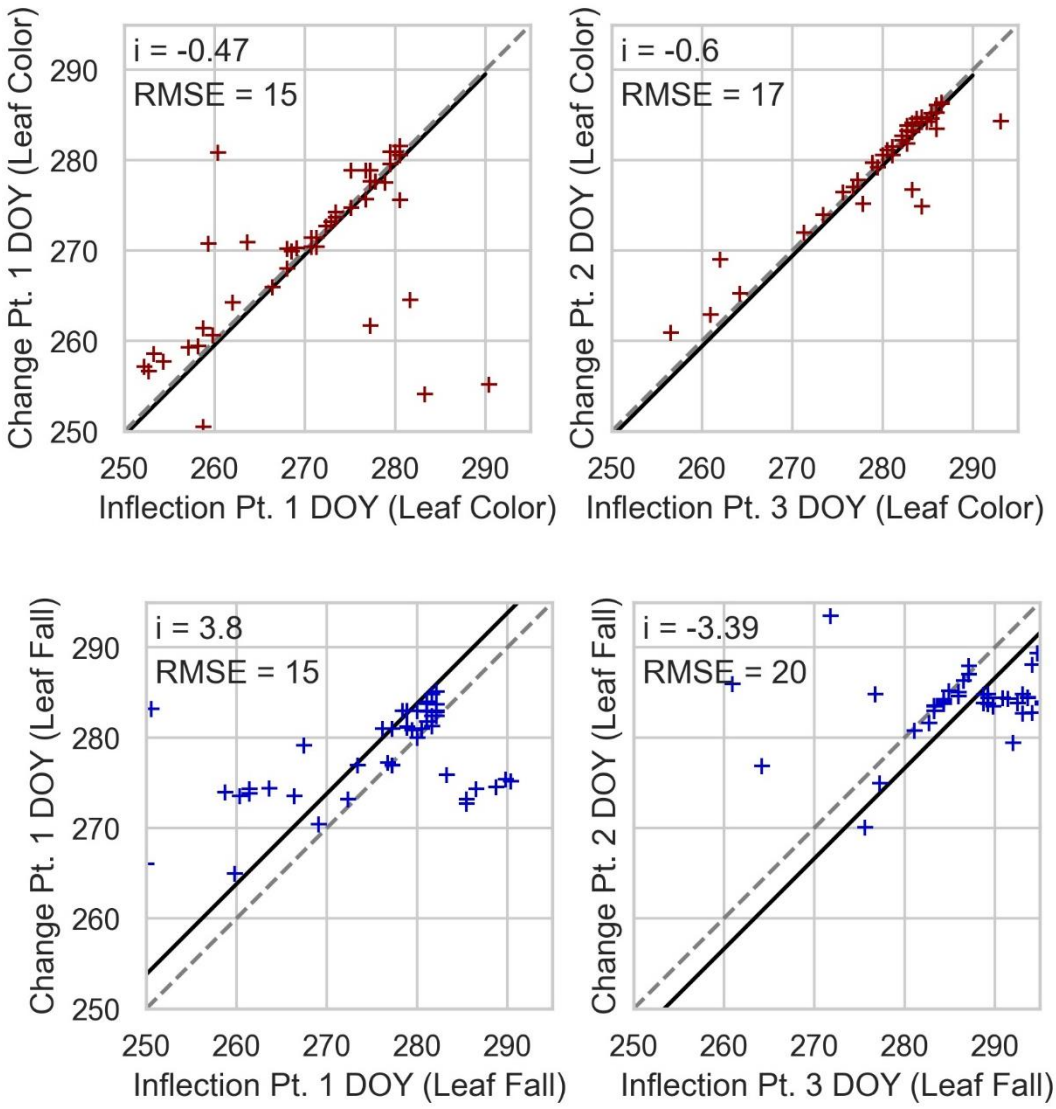
The genus of tree most studied (n=100) by these direct ground observations was *Acer*, which is best-known for its striking red and orange autumn leaf color. Red maple (*A. rubra*) began its autumn phenology earlier and ended later than sugar maple (*A. saccharum*). Red maple also showed a greater variability between individuals throughout the study domain. Red maple has been referred to as a "super-generalist" because of its adaptability to a wide range of habitats (Abrams, 1998) which speaks to its phenological flexibility observed in this study. The lower closeness of fit (R^2=0.716; R^2=0.64) for leaf color and leaf fall sigmoid curves support this, because some trees were so much earlier or later than others of the same species that the resultant fitted model for red maple was elongated. In contrast, sugar maple showed a much more uniform phenological response, and a tighter window for beginning and end of their leaf phenology.

Overall, deciduous broadleaf trees in the Cheesehead19 study area tended to be at peak fall colors around September 28th (DOY 271), and >90% of leaves on the vast majority had fallen by October 25th (DOY 298).

4.1.3 Sigmoid curves vs Linear segmentation validation

The statistical relationship between critical date calculation methods on human observations were tested, namely sigmoid curves vs linear segmentation. Pearson correlation coefficients (ρ) were expected to be 1.0 for a perfect validation between the two methods, but results were that between leaf color inflection point 1 and leaf color change point 1 there was ρ=0.61, and between leaf color inflection point 3 and leaf color change point 2 there was ρ=0.89 (closest to 1.0). There was a slight bias for the inflection points less than 1 day earlier than the change points (given as negative “i” values in *Fig. 4-3*). This represents average difference between calculated transition dates and is also the intercept of the fitted regression line with slope of 1.

Figure 4-3: Leaf color inflection points (red) and leaf fall (blue) vs corresponding change points. Black line represents regression line using intercept-only method. i values represent intercept, or bias in transition dates. Higher Root-mean-squared-error (RMSE) indicates greater average error between black line and subplot transition dates (+).

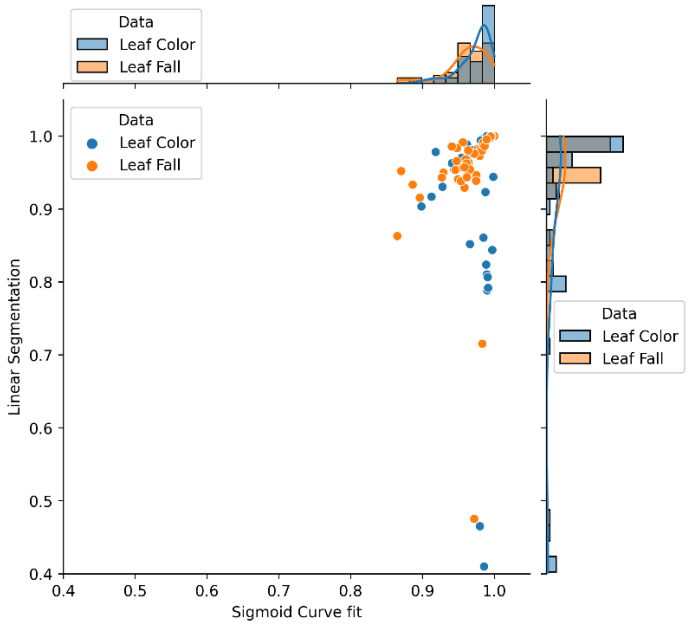


The higher correlation coefficient could indicate the end of the leaf coloration—meaning autumn leaf colors have hit their peak and the transition period has ended. On average, the difference between the first inflection point and the first change point for each field site was 0.47 days, with a standard deviation of

8.1 days for the differences. The difference between the third inflection point and the second change point was 0.60 days, with a standard deviation of 3.5.

Closeness of fit for models is better for the sigmoid curves as shown with R^2 distributions in *Fig. 4-4*.

Figure 4-4: R^2 distributions of sigmoid curve (x-axis) and linear segmentation (y-axis) for leaf color and leaf fall. Sigmoid curves had a greater overall R^2 distribution than linear segmentation.



This could be due to issues with errors when fitting the linear segments, rather than a superiority in the sigmoid curve fitting method because only a few samples using linear segmentation on leaf color values had $R^2<0.76$, with a similar relationship for leaf fallThe results from this study indicate that there is less agreement between sigmoid curve-fitting and linear segmentation in the beginning of autumn leaf color transition probably due to a few outlier data points contributing to the higher standard deviation in the beginning, but strong agreement for the end of leaf color transition.

4.1.4 Spatial autocorrelation

Spatial autocorrelation as it is applied in this study refers to how similar or dissimilar individual tree phenological event numbers are based on their proximity to their neighbors. The global Moran's I for all trees' first maximum leaf color is 0.232 indicating a positive spatial autocorrelation, and for all trees' first maximum leaf fall it was 0.200 also indicating a slight positive spatial autocorrelation. The spatial component of tree phenology is embedded within the local context of study trees, as the positive global Moran's I could be due to collinearity with other spatial variables such as microclimate.

This finding agrees with similar ground observation phenology research by Liang et al. (2011), who report that no consistent spatial autocorrelation exists for all trees, likely due to spatial heterogeneity from individual and species differences between trees. Local Moran's I analysis agrees with this assumption because several homogeneous species plots had a significant positive spatial autocorrelation for leaf color and leaf fall, but this could also be due to collinearity for trees of same species with similar phenological response. Although not the focus of this paper, a greater level of intra-species spatial autocorrelation would likely manifest at different levels of scale than the $10km^2$ study domain.

4.2 HUMANS VS PHOTOGRAPHS

Several phenological transition dates were calculated based on the human observation data and digital camera image series which were cross-compared using Pearson's correlation coefficient with associated p-values, as well as the average difference, or bias, between different methods. The bias is indicated by the "i"

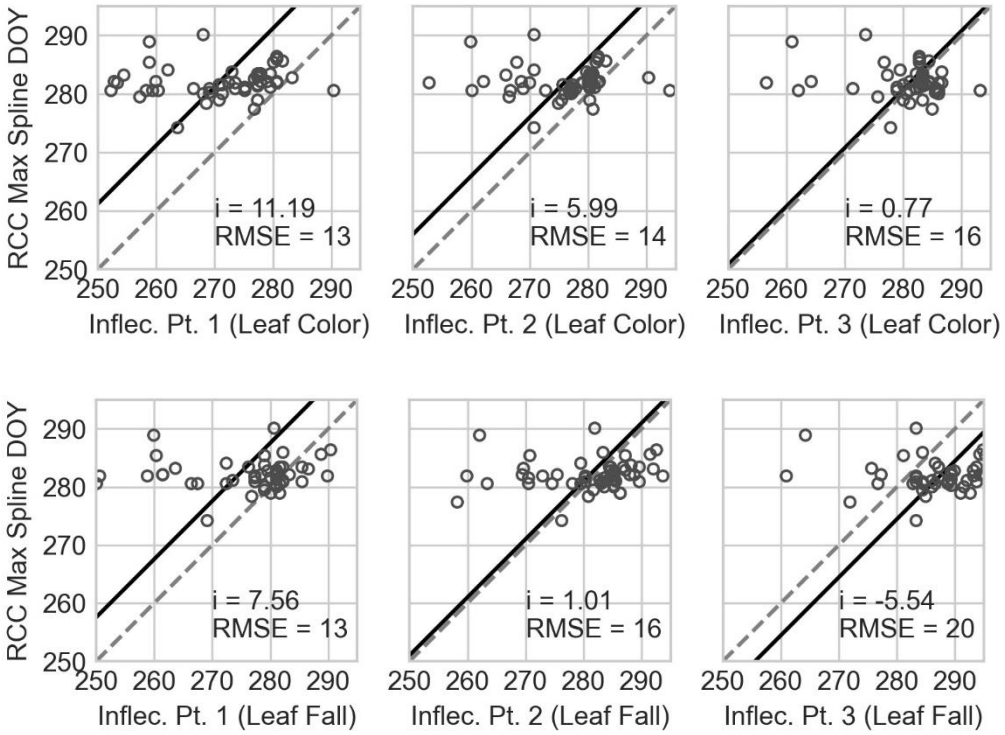
values on the following several figures. P-values were not generally statistically significant at >95% confidence level for the correlations, although calculated dates seemed to correlate better for the end of the autumn transition period rather than the beginning.

These results from cross-comparisons between human-derived transition dates and photo-derived transition dates varied in their correlations and biases. However, phenological photo-derived transition dates produced dubious results which seemed to vary widely from the human-derived transition dates. At 90% confidence level, GCC change point 2 and leaf color change point 2 were statistically correlated, which in accordance with results discussed toward the end of section 4.1.3 indicates that there is better agreement for the *ending* of autumn, but less so for the beginning of autumn transitions.

4.2.1 RCC Spline Maximum

In general, this method did not perform as expected, as seen from the flat results (circles) in *Fig. 4-5*. There was no significant correlation between RCC spline maximum method with any human-derived transition dates, which basically indicates that whatever this novel method was detecting was not tree phenology, or that there is another factor playing a significant role. The method comparison with intercept closest to zero was leaf color inflection point 3 (i=0.77, with SD=7.98), then leaf fall inflection point 2 (i=1.01, with SD=8.42).

Figure 4-5: Leaf Color (top row) and Leaf Fall inflection points (bottom row) vs RCC Max spline dates. Black line represents regression line using intercept-only method. i values represent intercept, or bias between transition dates. Higher Root-mean-squared-error (RMSE) indicates greater average error between black line and calculated transition date pairs shown with grey circles (n=53).

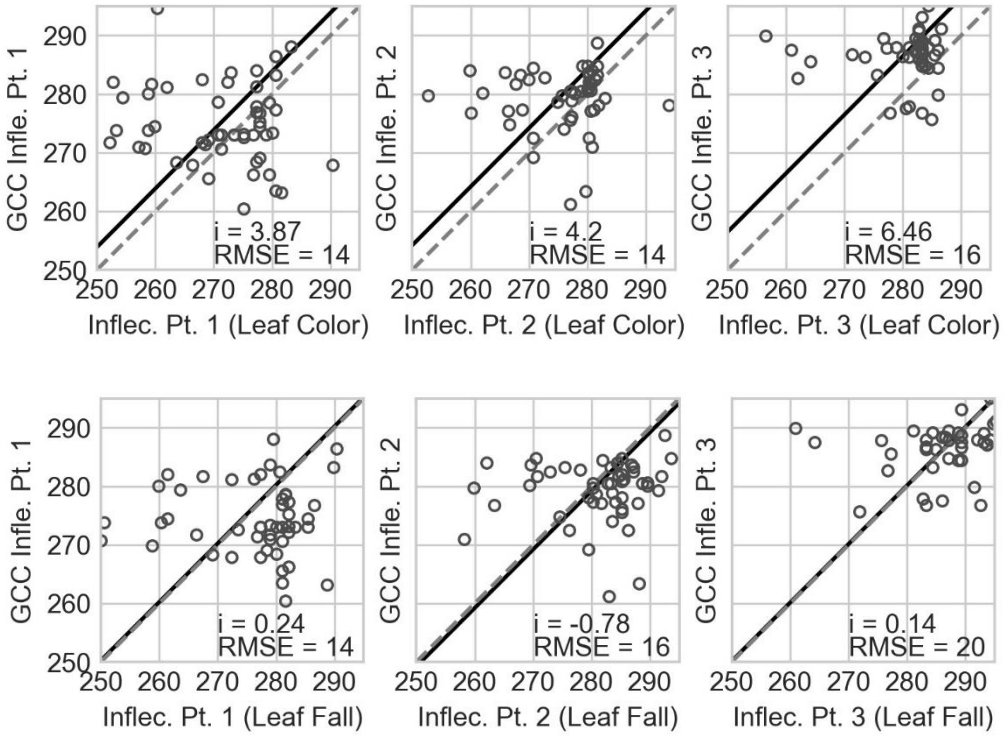


4.2.2 GCC Sigmoid inflection points

The GCC inflection points extracted from the fitted sigmoid curve had a weak, statistically insignificant correlations with inflection points from leaf color (ρ ranged from 0.13 to 0.18) and leaf fall (ρ ranged from 0.15 to 0.29). The positive correlation coefficients showed some promise but were all still not significant and nowhere near an ideal value of 1.0.

Surprisingly, GCC inflection points had a much lower bias when compared with leaf fall than with leaf color (*figure 4-7*). This suggest that leaf abscission is more closely coupled with relative canopy greenness inflection points.

Figure 4-6: GCC inflection points vs Leaf color (top) and vs leaf fall (bottom). Black line represents regression line using intercept-only method. i values represent intercept, or bias between transition dates. Higher Root-mean-squared-error (RMSE) indicates greater average error between black line and calculated transition date pairs shown with grey circles (n=53).

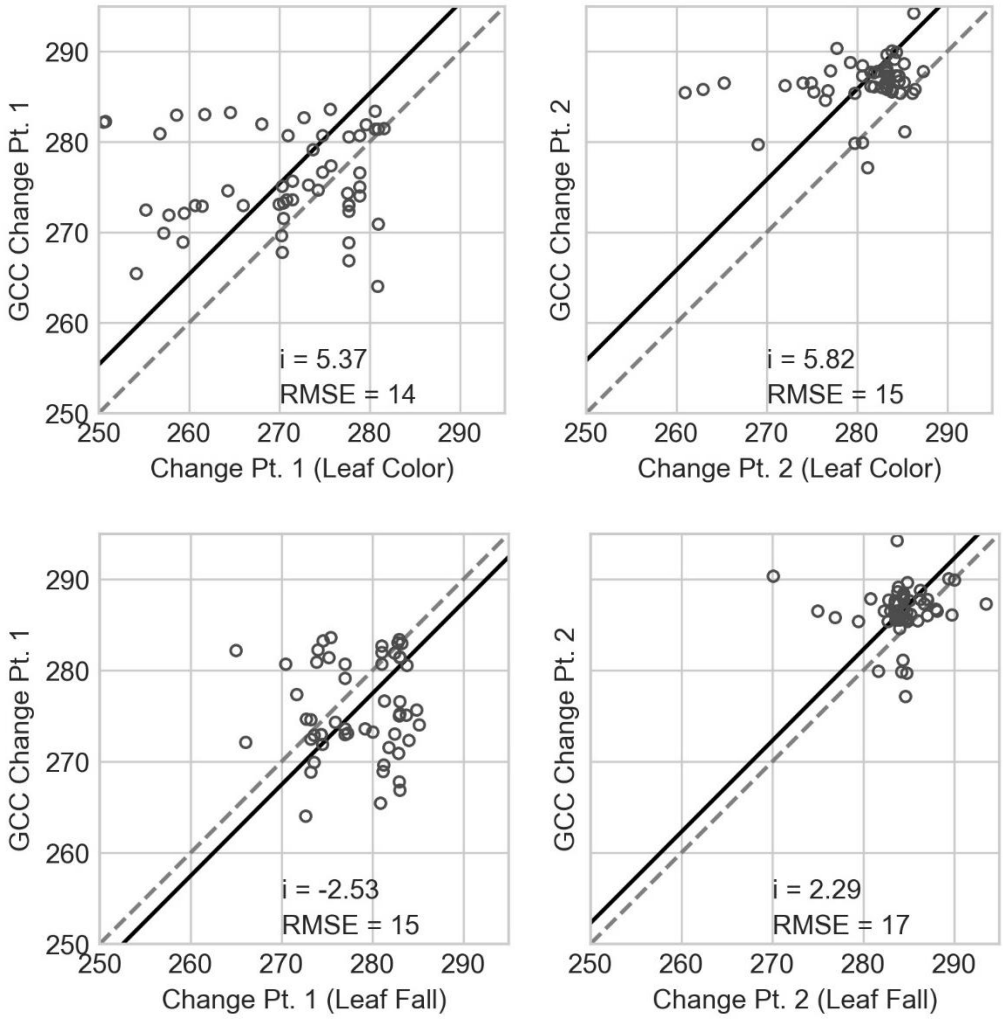


4.2.3 GCC linear segmentation change points

Compared with GCC inflection points, linear segmentation has relatively small differences in the corresponding transition date on a per plot basis. The GCC change point 1 was on average 1.04 days earlier than the first GCC inflection point, with a standard deviation of 8.5 days. The GCC change point 2 was on average 1.2 days later than the third GCC inflection point, with a standard deviation of 5.0 days.

Linear segmentation of leaf color human observations and GCC from digital images showed the most promise of the tested methods, with GCC change point 2 and leaf color change point 2 statistically significant at confidence level >90% (*Fig. 4-7*).

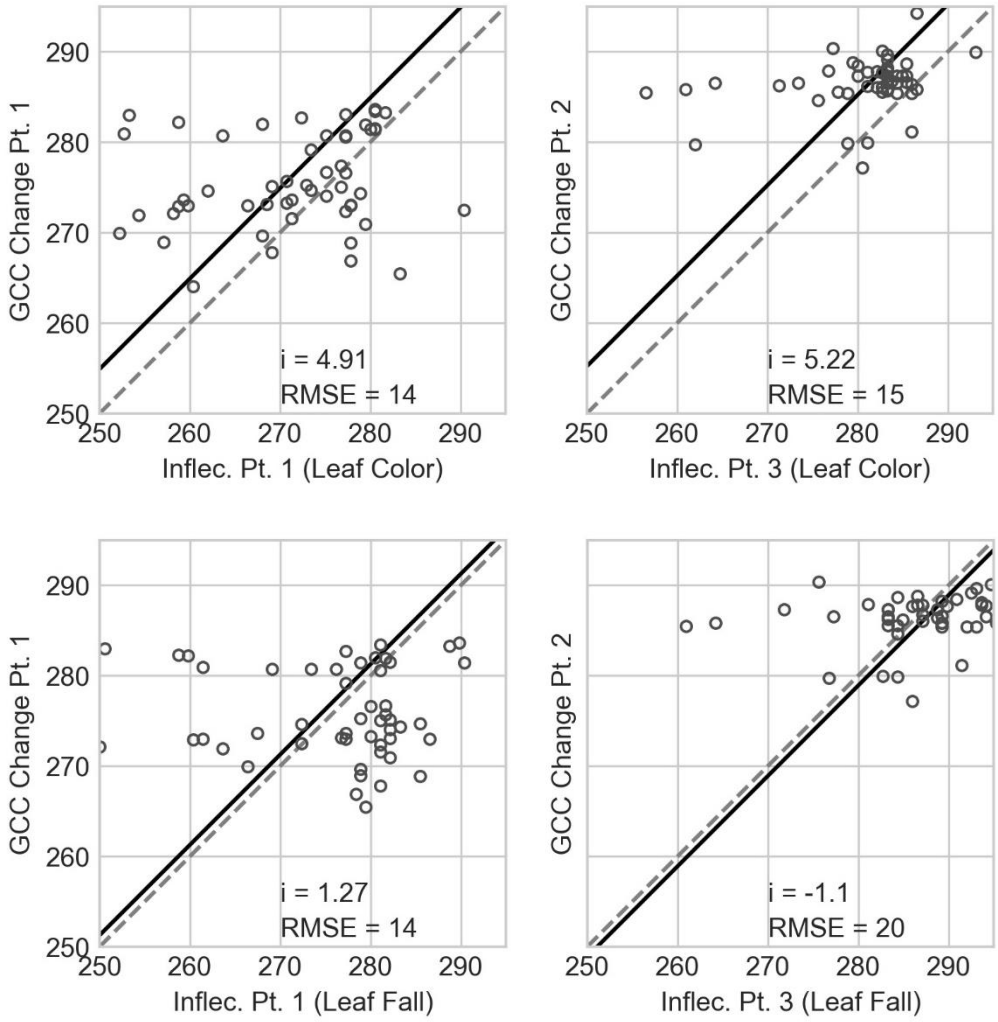
Figure 4-7: GCC change points vs leaf color change points (top) and leaf fall (bottom). Black line represents regression line using intercept-only method. i values represent intercept, or bias between transition dates. Higher Root-mean-squared-error (RMSE) indicates greater average error between black line and calculated transition date pairs shown with grey circles (n=53).



This indicates that human observations and digital camera images are in closest agreement with the timing of the end of autumn but are in less agreement detecting the beginning of autumn color transition. This is further supported by comparing the GCC change points with leaf color event number inflection points, with a statistically significant (>95% confidence) relationship between the GCC

change point 2 and leaf color inflection point 3, again related to the end of autumn leaf color transition. (*Fig. 4-8*)

Figure 4-8: GCC change points vs leaf color inflection points (top) and leaf fall (bottom). Black line represents regression line using intercept-only method. i values represent intercept, or bias between transition dates. Higher Root-mean-squared-error (RMSE) indicates greater average error between black line and calculated transition date pairs shown with grey circles (n=53).



These results are like the previously discussed results that linear segmentation change points agree well with the third inflection point of the fitted sigmoid curve on leaf color event numbers (described in section 4.1.3) again indicating that change point analysis detects the end of autumn better than the beginning, when considering this cross-validation between phenological date calculations.

The results from comparing human-derived transition dates with photo-derived transition dates are that they do not agree well for all plots, although there is some promise in GCC inflection being coupled to leaf fall, and an overall better agreement across methods for the ending of autumn.

5 DISCUSSION

5.1 CHARACTERIZING AUTUMN 2019 TREE PHENOLOGY

Considering inter-annual differences presented in section 4.1.1, the timing of 2019 was not a conclusively extreme-outlier year compared to three past years in terms of autumn temperature but comparing tree phenology with past years suggests that tree phenology was later in 2019 than what was usual in past years.

This conclusion is somewhat confounding, but perhaps could be explained by other factors known to influence autumn tree phenology such as first frost dates (2019 was latest as discussed in section 4.1.1.1), photoperiod, cooling-degree days, and total monthly precipitation as described in a meta-analysis of autumn phenology research by Gill et al. (2015). Photoperiod would not explain the differences in this study, because data were compared for the same trees at the same exact geographic location and thus would not vary in their day length from year-to-year. However, cooling-degree days or monthly precipitation were not studied in this thesis and could be analyzed further to perhaps explain the exaggerated differences in timing of phenology against the lesser inter-annual differences in daily autumn temperatures. If those factors were found to not be the cause of different phenological response in 2019, then the most reasonable explanation would be that the differing temporal frequency, lack of early autumn

observations, or other experimental design factors contributed to the different response.

5.2 CHANGE POINT DETECTION

Change point detection on vegetation indices to track phenology may have potential to shift the paradigm from sigmoid fitted curves (Liu et al. 2018; Xie & Wilson, 2020), but the results from this study show greater closeness of fit (R^2) for sigmoid curve fitting (*Fig. 4-4*). Several factors could have led to these results, such as the calculation method itself, noisy dataset, or the study design itself.

This study used a function with naïve logic to fit a variable number of points, which when interpolated was tested until its error was minimized. The permutations were based on the "Nelder-Mead" method of transforming the change point locations each computational iteration to settle on a satisfactory resultant set. This is in contrast with other methods used by other studies, such as those using Bayesian phenology models (Schleip et al. 2008; Thomson et al. 2010; Yang et al. 2014). Liu et al. (2018) advocates for a Bayesian approach for change point detection over fitted sigmoid curves on GCC values because it eliminates biases from strong, short term fluctuations which could influence the sigmoid logistic curve. In this study, temporal frequency made it so that short term fluctuations could create false positive change points as well as causing a significant effect on the fitted sigmoid curves, so Bayesian change points may help reduce biases if it were to be tested in the future.

A noisy dataset is not uncommon in scientific research, and to some degree it should be expected. However, some image series may have so many obstructions or varying light conditions that a particular change point could be undetectable

against outlier erroneous GCC data points. This would not apply to the direct observations which did not have sampling errors like in the digital repeat photography. Therefore, change point analysis cannot be disproven as a useful tool for analyzing phenological events for digital repeat photography, but is inferior to sigmoid inflection points for direct ground observation data in this study.

The study draws primary data from two sources: direction ground observations and digital repeat photography, to which change point analysis was applied. On days when a sub-plot was not sampled, trees would continue to progress through their autumnal phenology. The event number scheme for direct ground observations were based on categories of relative percentage so gaps in the data would be less likely to be greatly impacted by data gaps. On the other hand, data gaps in GCC, a scalar value, could be more drastic and thus could influence the change point detection algorithm to a false positive. With the number of change points vis-à-vis linear segments user-determined to be three linear segments, one change point may be spurious but the other change point accurate. So, it could be assumed that the lack of complete daily image series in the early part of autumn decreased model performance. These unexpected results in the change point detection being better suited for the end of fall could at least partially be explained by differing number of days between samples points to an understudied limitation to this model by the literature and could be studied in the future.

Overall, several factors contributed to the slight underperformance of change point detection on both direct ground observations and digital repeat photography. Nonetheless, these results maintain the efficacy of sigmoid curve fitting as well as change point detection on phenological data as standard analytical approaches.

5.3 HUMANS VS PHOTOGRAPHS

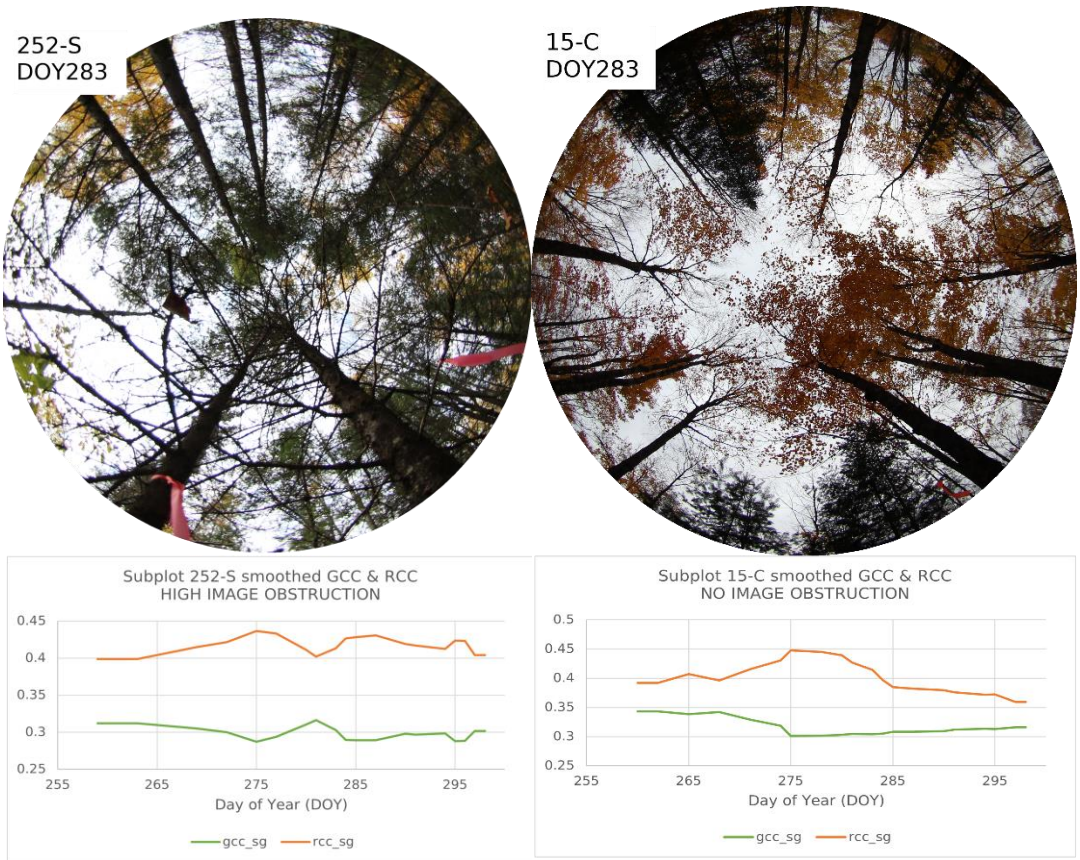
Calculating transition dates for autumn phenology proved challenging to match up from human observation and digital camera image series. The lack of major agreement between direct ground observations and digital repeat photography derived critical dates could be explained by the differing scale, or by photographic issues.

Direct event number observations were taken at the individual tree level, but there was no way to tell for sure which trees were captured in the camera field of view, which included upwards of 20 tree canopies in a single image. Some of the individually studied trees were in the camera images, due to the presence of tree markers being visible in some images, but it is likely that most of the tree canopy in images were not trees individually tracked with event numbers. Not only were different trees sampled by cameras than by human observers, but also the two sampling techniques operated at different spatial scales. There could be upward of 25 or more individual trees with their canopies at least partially captured in an image series, whereas four to five individual trees were sampled with event number observations. It is possible that with careful manual parsing, a count of tree species in each image could bridge the spatial divide. The role of scale in these differing methodologies was significant enough to distort results and was not easily remedied with over 1,000 images in this heterogenous forest study domain, mostly due to the radial orientation of camera view changing from day to day.

The lack of statistically significant validation between images and direct observations of phenology could also be partially explained by obstructions in the images. Upward-facing photographs were taken at the plot center, but the plot locations were selected for direct ground observations without considering

implications for digital repeat photography. The forest structure in the study domain was uneven-aged mixed northern deciduous-conifer secondary forest, so naturally things like understory, low dead branches, and evergreen conifers were present at some sub-plots which were not excluded from analysis (discussed in section 3.2.3.1). The effect this had on the two image series quality is shown in *Fig. 5-1*, showing the expected GCC and RCC smoothed data on the right for subplot 15-C with minor conifer presence in the field of view, and bad image series on the left for subplot 252-S, with high obstruction from branches, stems, and high evergreen conifer presence. The smoothed GCC and RCC data for that faulty image series is relatively flat but with several extreme values creating peaks, and so calculated critical dates were spurious and likely due to the poor image quality.

Figure 5-1: Examples of bad and good images, and their chromatic coordinates



This study succeeded at detecting phenological phase transition dates using human-derived samples at over 200 trees, but less so using the 53 upward-facing fisheye photography. The CHEESHEAD19 domain which was comprised of mixed and diverse northern deciduous forest, and digital repeat photography has been shown to perform poorly with mixed-species images (Nagai et al. 2015). Scholars recommend species-homogenous image series which agree more closely with direct ground observations, but that is sometimes not available or feasible with project constraints. This phenological study covered nine typical native tree species, which could potentially be analyzed further to separate the affect from different species phenological habits.

Heterogeneous landscapes including mixed-species forest are challenging to study across this spatial extent (100km^2) using handheld cameras especially because of the discreet nature of field plots leading to spatial gaps between sites. This could be true for direct observations at field sites as well, which could have led to important areas in the study area not sampled from. Upscaling toward landscape phenology (Liang & Schwartz 2009) by fusing with satellite-derived phenology products could alleviate some of the potential spatial gaps in the data. Nonetheless, direct observations proved more precise at calculating the phenological transition dates, in alignment with previous ground-based studies from the same area.

6 CONCLUSION

Handheld digital cameras are not yet capable of detecting phenological responses of trees in a mixed northern forest setting to the same level of precision as direct human observation, but there is still enormous potential for this type of

monitoring. Static installations of Phenocam-like setups account for several limitations in this study. Between two transition date calculation methods on human-derived phenology, there is a good agreement between sigmoid curve inflection points and change point detection, agreeing with previous studies (de Beurs & Hennebry 2010, Liu et al. 2018, Xie & Wilson 2020). The potential coupling shown between GCC and both leaf color and leaf fall point to the conclusion that repeat digital photography could prove highly useful as another technique to track individual plant phenology on the ground; the mobile aspect of mobile cameras could expand spatial constraints typically experienced by ground observations. To fully realize that potential, this study highlights several issues which can arise from improper site selection, image frequency, camera orientation, or possibly processing methods.

This study connected prior phenology data in a way which could prove useful for researchers in the context of Cheesehead19, as well as the broader phenology community, because of the unique intensive monitoring history at this location in Northern Wisconsin covering several common native broadleaf deciduous trees in their typical forest settings.

Links could be made between other, less understood factors which might influence or be influenced by autumn tree phenology, such as highly specific weather conditions studied by Cheesehead19 not limited to near-surface atmospheric eddy-covariance data products, forest canopy characteristics, and summer primary productivity.

6.1 Future Recommendations

The phenological data collected during autumn 2019 stemmed from a unique opportunity to study the land and atmosphere concurrently (which contributed to its somewhat hasty project design). It is recommended that future autumn phenological research consider insight learned post-hoc: site selection should favor site more suited to upward-facing photography, and sampling protocol decisions should attempt to control for more variables such as varying brightness, camera orientation, and image frequency.

As with this study and recommended by other research, underexposure of images is better than automatic settings (Zhang et al. 2005, MacFarlane et al. 2014). Even with this, brightness levels will vary due to weather or solar angle; the effects of this can be alleviated with maintaining the same camera rotation in each image, or with a reference image with a clear view of the sky. Fixed camera installations do not have the issue of varying orientation, but in this study the camera was not always oriented the same way which could have introduced error in the results. Additional steps to manually identify the species present in each image field of view could help validate results with direct phenological observation.

Furthermore, accurate site selection is paramount to maintain a clear, unobstructed view of the tree canopy. Previous studies showed that a disadvantage of this viewing geometry is that it emphasizes nearest vegetation because it is closest to the camera. Keenan et al. (2014) suggests oblique view to avoid this problem. The use of upward-facing fisheye photography employed in this study could have resulted, especially when understory shrubs were in view, in a loss of the ability to detect subtle changes in the dominant forest canopy.

Finally, temporal frequency plays a major role in the accuracy of data in situations like digital repeat photography where it is pretty much guaranteed there will be poor-quality images for any number of reasons. Like in the case of studies where a remote camera can be set up to take an image at set time intervals, the best image per day around the same time makes the most sense. This would require more technical setup which limits the ability to bring the camera around to different field sites. In future studies, I would suggest visiting plots every day by taking the same path so that images are taken around the same hour. This could also alleviate the effect solar angle would have on differing light-canopy interaction, and thus control for more variables. This type of mobile upward-facing fisheye photography would work best for situations where research is conducted near to base to reduce travel times.

Due to Cheesehead19 project logistics, the fisheye lens camera became available for research use not long before September 2019, and thus field methodological decisions needed to be make quickly. In future phenological monitoring projects, practices and principles related to illumination, view angle, obstructions, and frequency discussed in this study can be fully implemented.

7 REFERENCES

Abrams, M.1998. The Red Maple Paradox. *BioScience* **48**(5):355-364.

Badeck, F., A. Bondeau, K. Bottcher, D. Doktor, W. Lucht, J. Schaber, S. Sitch. 2004. Responses of spring phenology to climate change. *New Phytologist* **162**(2), 295-309.

Butterworth, B. J., Desai, A. R., Townsend, P. A., Petty, G. W., Andresen, C. G., Bertram, T. H., Kruger, E. L., Mineau, J. K., Olson, E. R., Paleri, S., Pertzborn, R. A., Pettersen, C., Stoy, P. C., Thom, J. E., Vermeuel, M. P., Wagner, T. J., Wright, D. B., Zheng, T., Metzger, S., Schwartz, M. D., Iglinski, T. J., Mauder, M., Speidel, J., Vogelmann, H., Wanner, L., Augustine, T. J., Brown, W. O. J., Oncley, S. P., Buban, M., Lee, T. R., Cleary, P., Durden, D. J., Florian, C. R., Lantz, K., Riihimaki, L. D., Sedlar, J., Meyers, T. P., Plummer, D. M., Guzman, E. R., Smith, E. N., Sühring, M., Turner, D. D., Wang, Z., White, L. D., & Wilczak, J. M. 2021. Connecting Land–Atmosphere Interactions to Surface Heterogeneity in CHEESEHEAD19, *Bulletin of the American Meteorological Society*, **102**(2), E421-E445. doi:10.1175/BAMS-D-19-0346.1

Dahl, A., & Langvall, O. 2008. Observations on phenology in Sweden - past and present. COST action 725: the history and current states of plant phenology in Europe, 161-165.

de Beurs, K.M., & G. M. Henebry. 2005. Land surface phenology and temperature variation in the International Geosphere–Biosphere Program high-latitude transects. *Global Change Biology* **11**:779–790.

Denny, E., K. L. Gerst, Ab. J. Miller-Rushing, G. L. Tierney, T. M. Crimmins, C. Enquist, P. Guertin, A. H. Rosemartin, M. D. Schwartz, K. A. Thomas, & J. F. Weltzin. 2014. Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. *International Journal of Biometeorology* **58**:591-601. doi:10.1007/s00484-014-0789-5

Dragoni, D. & A. F. Rahman. 2012. Trends in fall phenology across the deciduous forests of the eastern USA. *Agricultural and Forest Meteorology* **157**:96-105. doi:10.1016/j.agrformet.2012.01.019.

Feild, T.S., D. W. Lee, & M. N. Holbrook. 2001. Why Leaves Turn Red in Autumn. The Role of Anthocyanins in Senescing Leaves of Red-Osier Dogwood. *Plant Physiology* **127**:2 566-574. doi:10.1104/pp.010063

Fischer, A. 1994. A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Environment* **48**:220–230.

Fitchett, J. M., S. W. Grab, & D. L. Thompson. 2015. Plant phenology and climate change. *Progress in Physical Geography* **39**(4):460-482. doi:10.1177/0309133315578940

Friedman, J.M., J. E. Roelle, & B. S. Cade. 2011. Genetic and environmental influences on leaf phenology and cold hardiness of native and introduced

riparian trees. *International Journal of Biometeorology* **55**, 775–787 . doi:10.1007/s00484-011-0494-6

Gill, A., A. Gallinat, R. Sanders-DeMott, A. Rigden, D. Gianotti, J. Mantooth, & P. Templer. 2015. Changes in autumn senescence in northern hemisphere deciduous trees: a meta-analysis of autumn phenological studies. *Annals of Botany* **116**(6), 875-888. doi: 10.1093/aob/mcv055

Ge, Q., H. Wang, T. Rutishauser, & J. Dai. 2015. Phenological response to climate change in china: A meta-analysis. *Global Change Biology* **21**(1), 265-274. doi:10.1111/gcb.12648

Kato, S., & A. Komiyama. 2002. Spatial and seasonal heterogeneity in understory light conditions caused by differential leaf flushing of deciduous overstory trees. *Ecological Research* **17**, 687-693.

Keenan, T. F., B. Darby, E. Felts, O. Sonnentag, M. A. Freidl, K. Hufkens, J. O'Keefe, S. Klosterman, J. W. Munger, M. Toomey, & A. D. Richardson. 2014. Tracking forest phenology and seasonal physiology using digital repeat photography: a critical assessment. *Ecological Applications* **24**:1478—1489. doi:10.1890/13-0652.1

Liang, L., & M. D. Schwartz. 2009. Landscape phenology: An integrative approach to seasonal vegetation dynamics. *Landscape Ecology* **24**(4), 465-472. doi:10.1007/s10980-009-9328-x.

Liu, L., L. Liang, M. D. Schwartz, A. Donnelly, Z. Wang, C. B. Schaaf, & L. Liu. 2015. Evaluating the potential of MODIS satellite data to track temporal dynamics of autumn phenology in a temperate mixed forest. *Remote Sensing of Environment* **160**:156-165.

Liu, Z., S. An, X. Lu, H. Hu, & J. Tang. 2018. Using canopy greenness index to identify leaf ecophysiological traits during the foliar senescence in an oak forest. *Ecosphere* **9**(7):e02337. doi:10.1002/ecs2.2337

Luo, J., Y., Kui & B., Lijing. 2005. Savitzky–Golay smoothing and differentiation filter for even number data. *Signal Processing* **85**:1429-1434. doi:10.1016/j.sigpro.2005.02.002.

Macfarlane, C. 2011. Classification method of mixed pixels does not affect canopy metrics from digital images of forest overstorey. *Agricultural and Forest Meteorology* **151**:7 833-840. doi:10.1016/j.agrformet.2011.01.019.

Macfarlane, C., Y. Ryu, G. N. Ogden, & O. Sonnentag. 2014. Digital canopy photography: Exposed and in the raw. *Agricultural and Forest Meteorology* **197**:244-253. doi:10.1016/j.agrformet.2014.05.014.

May, D. A., & R. A. Montgomery. 2015. Photoperiod constraints on tree phenology, performance and migration in a warming world. *Plant, Cell, and environment* **35**:1725-1736. doi:10.1111/pce.12431

McDonough M., C., A. S. Gallinat, & L. Zipf. 2020. Low-cost observations and experiments return a high value in plant phenology research. *Applications in Plant Sciences* **8**(4): e11338. doi:10.1002/aps3.11338

Moody, A., & D. M. Johnson. 2001. Land-surface phenologies from AVHRR using the discrete fourier transform. *Remote Sensing of Environment* **75**:305–323. doi: 10.1016/S0034-4257(00)00175-9

Motohka, T., K. N. Nasahara, H. Oguma, & S. Tsuchida. 2010. Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing* **2**:2369–2387. doi:10.3390/rs2102369

Nagai, S., T. Inoue, T. Ohtsuka, S. Yoshitake, K. N. Nasahara, & T. M. Saitoh. 2015. Uncertainties involved in leaf fall phenology detected by digital camera. *Ecological Informatics* **30**:124–132.

Nasahara, K. N., & S. Nagai. 2015. Review: Development of an in situ observation network for terrestrial ecological remote sensing: the Phenological Eyes Network (PEN). *Ecological Research* **30**:211-223. doi:10.1007/s11284-014-1239-x

NAUTV, 2019: PhenoCam, https://www.youtube.com/watch?v=6BC3kj8FlCs accessed March 31, 2021.

PhenoCam Explorer, University of New Hampshire, 2019: PhenoCam Explorer, https://phenocam.sr.unh.edu/phenocam_explorer/, accessed March 31, 2021.

Prevéy, J.S., L. E. Parker, C. A. Harrington. 2020. Projected impacts of climate change on the range and phenology of three culturally-important shrub species. *PLOS ONE* **15**(5): e0232537. doi:10.1371/journal.pone.0232537

Reed, B. C. 2006. Trend analysis of time-series phenology of north america derived from satellite data. *GIScience & Remote Sensing* **43**(1), 24-38. doi:10.2747/1548-1603.43.1.24.

Richardson, A. D., B. H. Braswell, D. Y. Hollinger, J. P. Jenkins, & S. V. Ollinger. 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications* **19**:1417–1428. doi:10.1890/08-2022.1

Richardson, A. D., K. Hufkens, T. Milliman, D. M. Aubrecht, M. Chen, J. Gray, M. R. Johnston, T. F. Keenan, S. Klosterman, M. Kosmala, E. K. Melaas, M. A. Freidl, & S. Frolking. 2018. Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. *Scientific Data* **5**:180028. doi:10.1038/sdata.2018.28.

Schleip, C., A. Menzel, N. Estrella, & V. Dose. 2006. The use of Bayesian analysis to detect recent change in phenological events throughout the year. *Agricultural and Forest Meteorology* **141**:2-4, 179-191. doi:10.1016/j.agrformet.2006.09.013

Schwartz, M. D., J. M. Hanes, & L. Liang. 2013. Comparing carbon flux and high-resolution spring phenological measurements in a northern mixed forest. *Agriculture and Forest Meteorology* **169**:136–147. doi:10.1016/j.agrformet.2012.10.014

Schwartz, M. D. (editor), 2013: *Phenology: An Integrative Environmental Science*, 2nd edition. Springer, Netherlands, 610 pp.

See, L., P. Mooney, G. Foody, L. Bastin, A. Comber, J. Estima, Fritz, et al. 2016. Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. *ISPRS International Journal of Geo-Information* **5**(5), 55. MDPI AG. doi:10.3390/ijgi5050055

Sonnentag, O., K. Hufkens, C. Teshera-Sterne, A. M. Young, M. Friedl, B. H. Braswell, T. Milliman, J. O'Keefe, & A. D. Richardson. 2012. Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology* **152**:159–177. doi:10.1016/j.agrformet.2011.09.009

Toda, M., & A. D. Richardson. 2018. Estimation of plant area index and phenological transition dates from digital repeat photography and radiometric approaches in a hardwood forest in the Northeastern United States. *Agricultural and Forest Meteorology* **249**:457-466. doi: 10.1016/j.agrformet.2017.09.004

USA-NPN, 2021: USA-NPN Data Dashboard. http://www.usanpn.org/data/dashboard, accessed March 31, 2021.

Viña, A., & A. Gitelson. 2010. Sensitivity to foliar anthocyanin content of vegetation indices using green reflectance. *IEEE Geoscience and Remote Sensing Letters* **8**:464–468.

Vitasse, Y., A. J. Porte, A. Kremer, R. Michalet, & S. Delzon. 2009. Responses of canopy duration to temperature changes in four temperature tree species: relative contributions of spring and autumn leaf phenology. *Global Change Biology* **161**:187-198. doi:10.1007/s00442-009-1363-4

Wingate, L., J. Ogee, E. Cremonese, G. Filippa, T. Mizunuma, M. Migliavacca, C. Moisy, M. Wilkinson, C. Moureaux, G. Wohlfahrt, A. Hammerle, L. Hortnagle, C. Gimeno, A. Porcar-Castell, M. Galvagno, T. Nakaji, J. Morison, O. Kolle, A. Knol, W. Kutsch, P. Kolai, E. Nikinmaa, A. ibrom, B. Gielen, W. Eugster, M. Balzarolo, D. Papale, K. Klumpp, B. Koster, T. Grunwald, R. Joffre, J.M. Ourcival, M. Hellstorm, A. Lindroth, G. Charles, B. Longdoz, B. Genty, J. Levula, B. Heinesch, M. Sprintsin, D. Yakir, T. Manise, D. Guyon, H. Ahrends, A. Plaza-Aguilar, J. H. Guan, & J. Grace. 2015. Interpreting canopy development and physiology using the EUROPhen camera network at flux sites. *Biogeosciences Discussion* **12**:7979–8034. doi: 10.5194/bgd-12-7979-2015

Wisconsin Department of Natural Resources, 2016. Wisconsin Wiscland 2 Land Cover Database Level 2, 2016.

Xie, Y., & A. M. Wilson. 2020. Change point estimation of deciduous forest land surface phenology. *Remote Sensing of Environment* **240**: 34-4257. doi: 10.1016/j.rse.2020.111698

Xie, Y., D. L. Civco, & J. A. Silander Jr. 2018. Species-specific spring and autumn leaf phenology captured by time-lapse digital cameras. *Ecosphere* **9**(1):e02089. doi:10.1002/ecs2.2089

Yu, F., K. P. Price, J. Ellis, & P. Shi. 2003. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment* **87**:42–54. doi:10.1016/S0034-4257(03)00144-5

Yu, R., M. D. Schwartz, A. Donnelly, & L. Liang. 2016. An observation-based progression modeling approach to spring and autumn deciduous tree phenology. *International Journal of Biometeorology* **60**:335-349. doi:10.1007/s00484-015-1031-9

Zeng, L., B. D. Wardlow, D. Xiang, S. Hu, & D. Lis. 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment* **237**, 111511. doi:10.1016/j.rse.2019.111511.

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

Zhang, Y., J. M. Chen, & J. R. Miller. 2005. Determining digital hemispherical photograph exposure for leaf area index estimation*. Agricultural and Forest Meteorology* **133**:166-282. doi:10.1016/j.agrformet.2005.09.009.