2018

Remote Sensing Methods and Applications for Detecting Change in Forest Ecosystems

David James Gudex-Cross

University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/graddis

Part of the Ecology and Evolutionary Biology Commons, Forest Sciences Commons, and the Remote Sensing Commons

Recommended Citation

Gudex-Cross, David James, "Remote Sensing Methods and Applications for Detecting Change in Forest Ecosystems" (2018). Graduate College Dissertations and Theses. 966.

https://scholarworks.uvm.edu/graddis/966

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
REMOTE SENSING METHODS AND APPLICATIONS FOR DETECTING CHANGE IN FOREST ECOSYSTEMS

A Dissertation Presented

by

David J. Gudex-Cross

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy Specializing in Natural Resources

October, 2018

Defense Date: 24 August, 2018

Dissertation Examination Committee:

Jennifer Pontius, Ph.D., Advisor
Shelly Rayback, Ph.D., Chairperson
Anthony D’Amato, Ph.D.
Paul Schaberg, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College
ABSTRACT

Forest ecosystems are being altered by climate change, invasive species, and additional stressors. Our ability to detect these changes and quantify their impacts relies on detailed data across spatial and temporal scales. This dissertation expands the ecological utility of long-term satellite imagery by developing high quality forest mapping products and examining spatiotemporal changes in tree species abundance and phenology across the northeastern United States (US; the ‘Northeast’).

Species/genus-level forest composition maps were developed by integrating field data and Landsat images to model abundance at a sub-pixel scale. These abundance maps were then used to 1) produce a more detailed, accurate forest classification compared to similar products and 2) construct a 30-year time-series of abundance for eight common species/genera. Analyzing the time-series data revealed significant abundance trends in notable species, including increases in American beech (Fagus grandifolia) at the expense of sugar maple (Acer saccharum). Climate was the dominant predictor of abundance trends, indicating climate change may be altering competitive relationships.

Spatiotemporal trends in deciduous forest phenology – start and end of the growing season (SOS/EOS) – were examined based on MODIS imagery from 2001-2015. SOS exhibited a slight advancing trend across the Northeast, but with a distinct spatial pattern: eastern ecoregions showed advance and western ecoregions delay. EOS trended substantially later almost everywhere. SOS trends were linked to winter-spring temperature and precipitation trends; areas with higher elevation and fall precipitation anomalies had negative associations with EOS trends.

Together, this work demonstrates the value of remote sensing in furthering our understanding of long-term forest responses to changing environmental conditions. By highlighting potential changes in forest composition and function, the research presented here can be used to develop forest conservation and management strategies in the Northeast.
CITATIONS

Material from this dissertation has been published in Remote Sensing of Environment in the following form:


Material from this dissertation has been submitted to Landscape Ecology on June 28th, 2018 in the following form:

DEDICATION

In loving memory of my dad, Ronald G. Cross.

I miss you every day, old man. I know you are proud of me.
I owe the most gratitude to my mom, Beverley J. Gudex, for her unwavering support of my seemingly endless academic pursuits in life. So, thanks ma – I love ya!

I can safely say this dissertation would not have been possible without my advisor, mentor, and friend, Dr. Jennifer Pontius. She was an endless source of inspiration, laughs, and – let’s be honest – relative misery for me, none of which I could have done without on this journey to doctordom. I would also like to thank my committee members for their support along the way and for putting up with my repeated defense date rescheduling, missing of self-imposed deadlines, and subsequent frantic emails (sorry!); they are Drs. Shelly Rayback (chair), Paul Schaberg, and Anthony D’Amato. An extra ‘cheers!’ to Paul for his helpful edits and insights on my last two data chapters. Finally, I’d like to thank all the friends, colleagues, and undergraduate students that made my time at UVM enjoyable. A special shout-out goes to my good friend Jim Duncan, director of the Forest Ecosystem Monitoring Cooperative (F-E-M-C.), for sharing many a tasty beer with me – cheers! The conversation and times in the field together were okay too.

Funding for this work was graciously provided by UVM, the Northern States Research Cooperative (NSRC), and the USDA McIntire Stennis Cooperative Forest Research Program.
CHAPTER 4: RECENT PHENOLOGICAL CHANGE IN NORTHERN HARDWOOD FORESTS AND POTENTIAL DRIVERS ACROSS THE NORTHEASTERN US

1 Abstract .................................................................................................................. 110
2 Introduction ............................................................................................................... 111
3 Methods .................................................................................................................... 115
  3.1 Study area ............................................................................................................. 115
  3.2 Phenology data .................................................................................................... 116
  3.3 SOS/EOS trends and post-hoc filtering ................................................................. 118
  3.4 Forest cover, topography, and climate data ......................................................... 119
  3.5 Modeling significant predictors of SOS/EOS trends .......................................... 121
4 Results and Discussion ............................................................................................ 122
  4.1 Regional trends in SOS/EOS metrics ................................................................. 122
  4.2 SOS trends – spatial patterns and significant predictors ...................................... 126
  4.3 EOS trends – spatial patterns and significant predictors ...................................... 132
  4.4 Implications of phenological change in the Northeast ........................................ 137
5 Conclusions ............................................................................................................... 139
6 Acknowledgements .................................................................................................... 140
7 References ................................................................................................................ 140

CONCLUDING REMARKS AND RESEARCH SUMMARY ........................................ 146
COMPREHENSIVE BIBLIOGRAPHY ...................................................................... 150
LIST OF TABLES

Table 2.1. The composition of ground-reference plots used for development of percent basal area (%BA) models.......................................................... 36

Table 2.2. Principal components analysis eigenvectors highlight the input bands that account for the most spectral variability among forested pixels. ............... 45

Table 2.3. Percent basal area model fits derived from spectral unmixing. .................. 47

Table 2.4. Error matrix based on 50 ground-reference plots for the MTSU pixel-level forest classification. ................................................................. 57

Table 2.5. Comparison of specificity, spatial resolution, and accuracy of forest mapping products. ................................................................. 58

Table 3.1. Abiotic variables examined for potential relationships with species abundance trends. ................................................................. 78

Table 3.2. Mean regional abundance over time. Trend data is reported as the mean of the per-pixel regression slopes for a subset of buffered, random pixels. ......................................................................................... 81

Table 3.3. Abundance trends by elevation zone reported as the mean (% basal area) of the per-pixel regression slopes for a subset of buffered, random pixels. ......................................................................................... 82

Table 3.4. Significant abiotic factors and their effect on 30-year abundance trends. ......................................................................................... 86

Table 3.5. Final regression model selection and fit statistics. ................................. 87

Table 4.1. Summary of the statistical tests used to identify strong predictors of start of season (SOS) trends in northern hardwood forests. ................. 129

Table 4.2. Summary of the statistical tests used to identify strong predictors of end of season (EOS) trends in northern hardwood forests. ......................... 135
LIST OF FIGURES

Figure 1.1. Comparison of an image with a 0.5-meter spatial resolution resampled to represent different pixel sizes........................................... 5

Figure 1.2. The spectral band characteristics of five different multispectral/hyperspectral satellite sensors and the wavelengths they cover, depicted against the atmospheric transmissivity of the electromagnetic spectrum (graphic accessed 8/1/2018 from NASA’s Earth Observatory website at earthobservatory.nasa.gov, adapted from Casey et al. 2012)...................... 8

Figure 2.1. The study area, spanning northern New York and Vermont, and distribution of ground-reference plots (Landsat Path 14, Row 29). .................. 34

Figure 2.2. Landsat preprocessing and percent basal area modeling workflow. .................................................................................................. 38

Figure 2.3. Object-based image analysis workflow and hierarchical ruleset used to create a thematic forest cover map. ........................................... 43

Figure 2.4. The spatial distribution of percent basal area derived from spectral unmixing for four common species in northern New York and Vermont........ 51

Figure 2.5. Comparison of the object-based (top) and pixel-level (bottom) classifications in the Stowe region of Vermont. ........................................ 53

Figure 2.6. Forest cover map of northern New York and Vermont produced by integrating spectral unmixing of multi-temporal Landsat imagery (MTSU) and a hierarchical classification scheme........................................ 55

Figure 2.7. Comparison of the MTSU (top left), LANDFIRE EVT (bottom left), NLCD (top right), and National Forest Type Map (bottom right) forest cover maps near Stowe, Vermont........................................... 61

Figure 3.1. The study area across northern New York and Vermont (Landsat Path 14, Row 29).................................................................................. 75

Figure 3.2. Spatial distribution of 30-yr abundance increases (red) and decreases (blue) for eight tree species/genera across ecoregions. .......................... 84
Figure 3.3. Regional variation in geographically-weighted regression model fit and abiotic parameter significance for sugar maple (a) and birch species (b)........... 88-89

Figure 3.4. Regional variation in geographically-weighted regression model fit and abiotic parameter significance for red spruce (a) and balsam fir (b). .................. 91-92

Figure 4.1. Study area and spatial distribution of northern hardwood forests across New York and northern New England......................................................... 116

Figure 4.2. Trends in start and end of season (SOS/EOS) from 2001-2015 across the entire study area (left pane) and each ecoregion (right panes). ................. 124

Figure 4.3. Spatial distribution of the start of season (SOS) trend (top left) and significant explanatory climate variables................................................................. 127

Figure 4.4. Spatial distribution of the end of season (EOS) trend (top left) and significant explanatory variables. ................................................................. 133
CHAPTER 1: INTRODUCTION

1.1 Overview

The main goal of this dissertation is to expand the usefulness of satellite remote sensing for detecting and analyzing change in forest ecosystems. From a data perspective, remote sensing usually offers coarser detail and less accurate products compared to field observations. This is especially true for the most popular satellite imagery datasets (e.g., Landsat). Yet satellite imagery offers broad spatial and, to a lesser extent, temporal coverage, while field data are limited to discrete locations. Additionally, the existence of climate and other environmental spatial data products can facilitate examinations of forest responses to different environmental drivers of change across time and space.

The research presented here helps bridge the gap in ecological utility between satellite and field data by using satellite imagery to develop more accurate, detailed forest mapping products, examine spatiotemporal changes in forest structure (species composition) and function (phenology), and relate these changes to environmental drivers across forests of the northeastern United States (US; the ‘Northeast’). This work is timely as climate change, invasive species, and human land use practices continue to rapidly alter these ecosystems. By modeling changes in forest structure and function and their potential drivers, the information provided in this dissertation can be used to inform forest research and management in the Northeast. Conducting this research in a region with diverse forests and topography also enables replication in similarly complex (or less so) regions.
1.2 Specific objectives

The specific research objectives (first number = chapter) of this dissertation are to:

2.1: Integrate multi-temporal Landsat imagery and field inventory data using spectral unmixing to develop pixel-level percent basal area (% BA) coverages for 10 common tree species/genera in the Northeast.

2.2: Incorporate the percent basal area maps and ancillary data into an object-based, hierarchical ruleset to generate a forest classification (10 species/genera and 6 common species assemblages).

2.3: Compare the forest classification’s detail and accuracy with existing large-scale forest mapping products, including LANDFIRE, the National Land Cover Database (NLCD), and the National Forest Type Map.

3.1: Use the technique and products developed in Chapter 2 to construct a 30-year (1985-2015) time-series of abundance (% BA) for eight dominant tree species/genera across northern New York and Vermont.

3.2: Examine 30-year changes in mean abundance across the study area and by elevation.

3.3: Detect and quantify spatiotemporal patterns in pixel-level abundance trends.

3.4: Identify possible abiotic correlates (i.e., climate indices, topographical factors, acid deposition inputs, and soil characteristics) associated with abundance trends.

4.1: Using MODIS-derived annual phenology metrics from 2001-2015, quantify trends in the start and end of the growing season (SOS/EOS) across the Northeast.

4.2: Examine intraregional variation and spatiotemporal patterns in SOS/EOS trends.
4.3: Investigate relationships between SOS/EOS trends and numerous climatic and site characteristics.

1.3 Remote sensing of forests: a brief review

1.3.1 Remote sensing systems and concepts

From Lillesand et al. (2014), “Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in direct contact with the object, area, or phenomenon under investigation.” Remote sensing systems consist of three basic components: 1) the platform, 2) the sensor, and 3) the data. A useful metaphor in remote sensing is human sight – our body is the platform, our eyes are the sensors, and what we see is the data (processed and analyzed by our brains, a supercomputer equipped with all the necessary software).

Platforms are simply what the sensor is mounted on and are divided into two classes, airborne/aerial (airplanes, drones, and balloons) and spaceborne (satellites and spacecraft). The sensor – the crux of a remote sensing system – acquires the data and can be either active or passive. Active sensors emit energy and record information on the returns while passive sensors record energy (solar radiation) reflected or emitted from the earth (Wang and Weng 2013). A well-known example of an active sensor is Light Detection and Ranging (LiDAR). LiDAR sensors emit frequent laser pulses and measure their return time, thereby recording the height and shape of surface features (e.g., trees) (Reutebuch et al. 2005). Passive sensors are largely ‘electro-optical’, meaning they measure and record data (numeric ‘reflectance’ values) in portions of the electromagnetic
spectrum (Wang and Weng 2013). While the use of LiDAR and other active sensor (e.g., radar) data is increasingly common, most applications of remote sensing in forest research and management rely on digital imagery captured from passive sensors.

Choosing an appropriate remote sensing system depends on the objectives at hand, though important considerations between (and within) aerial and satellite systems include: 1) imagery resolution (spatial/spectral/temporal), 2) cost/availability, and 3) data processing/analysis/storage requirements (extensions of cost). Imagery resolution is arguably the most important factor because it dictates what forest features/phenomena can be detected and to what extent (see Sections 1.3.2-1.3.4). *Spatial resolution* refers to the spatial detail of an image and is defined by the smallest object it can detect. This is usually expressed as *pixel size*, though these terms are not directly interchangeable since the original resolution of an image can be resampled to different pixel sizes (Fig. 1.1) (Wang and Weng 2013).
Figure 1.1. Comparison of an image with a 0.5-meter spatial resolution resampled to represent different pixel sizes.

*Spectral resolution* refers to the amount of the electromagnetic spectrum that a sensor can detect, expressed as the number and width of ‘*spectral bands’*. For example, most aerial imagery – and pictures we take with everyday cameras – consist of three spectral bands in the visible light wavelengths (blue, green, and red). Most satellite imagery, on the other hand, contain several to many bands that capture segments of the ultraviolet, visible, infrared, and/or thermal wavelengths. Finally, *temporal resolution* is a measure of the frequency with which images are captured, usually called a ‘*return interval’*, and encompasses the *time-series* of imagery available for analysis (*temporal coverage*).
Generally, the major advantages of aerial imaging systems are: 1) they can be rapidly deployed to targeted locations; 2) they capture high *spatial* resolution imagery; and 3) they avoid problematic atmospheric effects like clouds (Paine and Kiser 2003). These advantages are why aerial imaging is commonly used in forestry to map local forest types, assess damages (e.g., from insect outbreaks and fire), and survey timber – a practice that dates back to the early 20th century (Duggin et al. 1990). Aerial images are also a good alternative to ground data for assessing the accuracy of other remotely-sensed forest data products (Paine and Kiser 2003).

However, these advantages come with substantial trade-offs that make aerial systems less appropriate for objectives that require large geographic coverage and high spectral/temporal resolution. For example, the combination of low flight paths and high spatial resolution sensors produces images that cover relatively small areas (low *scene size*) and capture few spectral bands due to limitations imposed on the sensor’s field of view (Lillesand et al. 2014). Aerial imaging also comes with many financial and labor-related costs for data acquisition, processing, analysis, and storage, and thus the imagery itself is usually expensive (Meneguzzo et al. 2013). For these reasons, the rest of this review (and dissertation) largely focuses on satellite remote sensing systems.

### 1.3.2 Satellite remote sensing

The rapid advancement of satellite and sensor technology since the Cold War era, beginning with the launch of the Landsat-1 satellite in 1972 (Woodcock et al. 2008), has led to a host of satellite remote sensing systems orbiting the earth today (see Table 1 in
Melesse et al. 2007 and the USGS’s list at eros.usgs.gov/satellite-imagery for example). They can be classified into four general categories – coarse (i.e., multispectral but with spatial resolution > 30-meters), multispectral, hyperspectral, and hyperspatial (Melesse et al. 2007) – and their utility in remotely-sensing forest attributes is again a function of their various resolutions, cost/availability, and processing requirements (see Table 1 in Xie et al. 2008 for a good summary).

Hyperspatial imagery like that from QuickBird, Ikonos, and World-View is an extremely useful alternative to aerial imaging for localized forest mapping and accuracy assessment, having high spatial and temporal resolution (all <4-meter spatial resolution, with sub-meter options for QuickBird and World-View, and ~5-day return intervals) (Melesse et al. 2007, Xie et al. 2008), yet these datasets have limited spectral resolution (1-5 bands) and are relatively underutilized due to their prohibitive financial cost (Boyle et al. 2014).

Arguably the most powerful applications of satellite remote sensing are those based on publicly-available imagery from the coarse (e.g., Moderate Resolution Imaging Spectrometer, MODIS), multispectral (e.g., Landsat), and hyperspectral (e.g., Hyperion) systems that can not only detect the unique spectral properties of vegetation, but also how these properties change over time and space. These systems detect energy in the same regions of the electromagnetic spectrum, since energy reflection to space is governed by atmospheric transmissivity – some wavelengths are fully absorbed (or nearly so) in the atmosphere before they reach the sensor while others pass through atmospheric ‘windows’ (Fig. 1.2). Where they differ is in the number of spectral bands measured and their width
(i.e., broad-band vs. narrow-band). For example, the Landsat Enhanced Thematic Mapper Plus (ETM+) sensor is a broad-band, multispectral sensor that captures 8 bands across the visible (3 bands), near-infrared (1 band), shortwave-infrared (2 bands), and thermal (1 band) regions, as well as a higher spatial resolution (15m vs. 30m) ‘panchromatic’ band that spans the visible-near infrared wavelengths and is used to ‘pan-sharpen’ other bands. In contrast, Hyperion is a narrow-band, hyperspectral sensor that captures 220 bands throughout the visible to shortwave infrared regions (Fig. 1.2). These spectral differences are a major factor, though not the only one, in determining the appropriate applications of multispectral and hyperspectral systems since their spatial resolutions are largely the same.

![Figure 1.2. The spectral band characteristics of five different multispectral/hyperspectral satellite sensors and the wavelengths they cover, depicted against the atmospheric transmissivity of the electromagnetic spectrum (graphic accessed 8/1/2018 from NASA’s Earth Observatory website at earthobservatory.nasa.gov, adapted from Casey et al. 2012).](image)

Typically, the most informative electromagnetic region for studying vegetation is from the visible to infrared wavelengths, since chlorophyll strongly absorbs visible light and reflects infrared (particularly near-infrared) during photosynthesis (Fig. 1.3). This demonstrates the power of hyperspectral imaging in vegetation studies (e.g., Treitz and Howarth 1999), which contain hundreds of bands in these regions (Fig 1.2). Unfortunately,
Hyperion is currently a ‘tasking satellite’ that must be requested for imaging and thus has poor spatial/temporal coverage and limited application (He and Weng 2018).

In addition to providing useful spectral information, the major advantages of multispectral systems like Landsat and MODIS over hyperspectral are their spatial coverage and temporal resolution. Both have wall-to-wall global coverage; Landsat has a 16-day return interval and an imagery archive dating back to 1972, while MODIS has a daily return interval and dates back to 2000. Between these systems, Landsat has better spatial resolution than MODIS (30m vs. 250-1,000m depending on the band), but its longer return interval is problematic for cloud cover and especially time-sensitive applications (e.g., phenology). A major limitation of both multispectral and hyperspectral systems is the ‘mixed pixel’ effect – a single pixel can contain multiple surface features or land cover types, complicating inferences regarding their true composition – that comes with reduced spatial resolution. Several imagery analysis techniques have been developed to address this issue, which are discussed in the context of forest mapping (Section 1.3.3) and change/trend detection (Section 1.3.4).
1.3.3 Mapping considerations, methods, and products

The three most important factors to consider for forest mapping are: 1) spatial extent (i.e., how much forest needs to be mapped), 2) imagery/imagery analysis techniques (i.e., how is forest cover going to be mapped), and 3) end users/uses (i.e., how detailed and accurate do the mapping products need to be). For example, mapping large spatial extents generally calls for matching imagery (coarser spatial resolution), which limits the detail and accuracy of the products. Conversely, highly localized mapping efforts can integrate multiple data sources (e.g., ground, satellite, and LiDAR) to obtain very accurate, tree species-level products (e.g., Ke et al. 2010, Alonzo et al. 2014). However, they require a high level of technical expertise and specialized (expensive) imagery analysis software packages (though open-source programming languages like ‘R’ are steadily easing this
requirement at the cost of requiring even more technical knowledge), in addition to the products having very specific end users/uses (e.g., local governments). Thus, there exists a need to develop novel methods of large-scale forest mapping that can integrate data sources and analysis methods in effective ways to achieve detailed, accurate products.

Most imagery analysis techniques used for forest mapping are designed to address the mixed pixel effect (Table 1 in Lu and Weng 2007 provides a ‘taxonomy’ of image classification techniques). As this effect is greatly reduced in hyperspatial imagery, more traditional pixel-based classifiers can often achieve satisfactory results. These include supervised and unsupervised classification algorithms based on the spectral properties of the image. In supervised classification, the analyst selects pixels that represent a certain forest type or species to be used as ‘training data’ for the algorithm, which then classifies the image by matching the spectral characteristics of the training set to the remaining pixels. Unsupervised classification works in the opposite direction, with the algorithm first grouping pixels into discrete ‘clusters’ based on spectral characteristics, then the analyst must combine or split the clusters to represent the desired forest classes. However, these techniques are prone to error caused by spectral overlap between forest classes and other types of vegetation and ignore the distinct shape and textural features contained in a hyperspatial image (e.g., tree crowns are often round and have a unique texture compared to other vegetation). To address this, a powerful advanced classification technique called object-based image analysis (OBIA) was developed that leverages all of the information contained in a hyperspatial image during the classification process.
OBIA techniques overcome pixel-based constraints by segmenting the image into groups of “objects”, whose characteristics are defined by the user based on shape/texture/spectral properties. The classification is then carried out on the objects (rather than individual pixels) in a hierarchical ‘ruleset’ developed by the analyst using threshold values of the same properties (ancillary data can also be incorporated, such as digital elevation models) (Chubey et al. 2006). Many studies have successfully employed OBIA to produce accurate, relatively detailed forest type-level classifications (see Pu 2013 for a review) and comparisons with pixel-based classifiers tend to favor OBIA (Dorren et al. 2003, Oruc et al. 2004, Agarwal et al. 2013). As these techniques are most useful on hyperspatial imagery, their applicability to large-scale forest mapping to date has been limited.

Obviously, the mixed pixel effect is more pronounced in imagery with lower spatial resolution, like the multispectral imagery required to map forests across large areas. In these cases, imagery analysis techniques rely heavily on the spectral and temporal characteristics of the image, as well as ancillary data, to classify forest cover. Hierarchical mapping is arguably the most common technique, where the image is sequentially classified into distinct landcover types (i.e., non-forest and forest, then deciduous and coniferous, etc.) based on threshold values of specific spectral bands or band ratios (e.g., the Normalized Difference Vegetation Index, NDVI). The classification is then carried out on the remaining, forest-stratified pixels (e.g., Wolter et al. 1995). Other studies have shown this process can be substantially improved by more than one image date (multi-
temporal imagery) to capture phenologically-significant differences driven by leaf-out and senescence (e.g., Mickelson et al. 1998, Dymond et al. 2002).

More advanced methods of mapping forest cover using low spatial-high spectral resolution imagery include various spectral mixture analyses (SMA), which decompose mixed pixels into linear or nonlinear combinations of ‘endmembers’ (i.e., spectra of a known cover type, such as a “pure” pixel of a particular tree species) to quantify the proportion of a tree species or forest type within a pixel (sub-pixel level classifiers) (Huguenin et al. 1997, Oki et al. 2002), and machine learning techniques (e.g., support vector machines and artificial neural networks – see Pu 2013 for a review of these). These techniques, and species-level mapping in general, have traditionally been reserved for hyperspectral imagery (e.g., Plourde et al. 2007).

Unfortunately, few large-scale forest cover maps exist due to the amount of data, labor, expertise, and money required to accurately characterize broad landscapes. Two popular products are the National Land Cover Database (NLCD; mrlc.gov) and LANDFIRE (LANDFIRE; landfire.gov). Both are developed using multiple data sources and hierarchical (or decision-tree) approaches, integrating field inventory data, multi-temporal Landsat imagery, and other environmental information (e.g., topography) to map broad forest classes at 30m resolution (Vogelmann et al. 2001, Rollins 2009). The main difference is NLCD produces a very coarse forest type classification (i.e., deciduous, evergreen, and mixed) while LANDFIRE refines an initial coarse classification (e.g., hardwood-dominated) into more specific forest types (e.g., oak-hickory) using regional expert opinion (Rollins 2009). Even at these broad classification levels, their accuracy in
heterogenous forests like those of the Northeast is limited (e.g., Wickham et al. 2013). The lack of detailed, accurate forest composition maps is one of the greatest limiting factors to the use of remote sensing in forest research and management (Frohn and Lopez 2017).

1.3.4 Change detection and trend analysis

Detecting spatiotemporal changes in landcover features is one of the most powerful environmental applications of remote sensing. In contrast to mapping where the spectral and spatial features of an image are most important, change detection and trend analysis are largely dictated by temporal resolution/coverage. This makes the long-running satellite systems like Landsat, the Advanced Very High Resolution Radiometer (AVHRR) and, to a lesser extent, MODIS particularly useful. However, change detection technically requires only two different image dates and thus all systems (aerial and satellite) remain relevant; meaningful trend detection, on the other hand, is largely limited to the aforementioned three satellite systems. Many imagery analysis techniques have been developed to detect change in forests using, 1) spectral reflectance values, from strictly visual assessments and simple algebra to complex data transformations (e.g., principal components analysis) and 2) advanced modeling of biophysical parameters derived from reflectance (e.g., biomass) (see Lu et al. 2004 for a comprehensive review).

Two examples of popular algebraic approaches are image differencing and regression. Image differencing is accomplished by simple subtraction between two or more image dates based on either raw reflectance values or more informative band ratios (e.g., NDVI). This technique is most appropriate for detecting abrupt changes in forest canopies,
such as those caused by insect defoliation, logging, or development (e.g., Muchoney and Haack 1994, Hayes and Sader 2001, Wilson and Sader 2002). Image regression allows for the detection of more detailed and subtle changes, with the regression slope calculated from multiple image dates representing the magnitude and nature of change at the pixel level. Since regression analysis of high temporal resolution satellite data (e.g., daily MODIS images) can be quite computationally-intensive, independent software programs have been specifically designed for this purpose (e.g., TIMESAT; Jöhnsson and Eklundh 2004).

Modeling trends in vegetation indices like NDVI can be used to approximate spatiotemporal changes in many important aspects of forests, including phenology (e.g., leaf-out/senescence and growing season length; Cleland et al. 2007), biomass (e.g., Powell et al. 2010), and growth/productivity (e.g., Keenan et al. 2014). However, direct modeling of trends in these and other detailed forest parameters (e.g., tree species abundance) is much more challenging, since deriving accurate characterizations of them from satellite imagery is a time-consuming and complex process (Lu et al. 2004). Nonetheless, more advanced techniques like those discussed in the context of forest mapping – namely SMA and machine learning algorithms (e.g., Random Forests) – have shown promise in recent years (e.g., Ali et al. 2015, Wang et al. 2016).

1.4 Forests of the northeastern United States

1.4.1 Historical and current composition patterns

Pre-European settlement forests of the Northeast (ca. 1600-1700s) were largely comprised of temperate forest types split along a tension zone – American beech (Fagus
grandifolia), maples (Acer spp.), and eastern hemlock (Tsuga canadensis) to the north and oak-hickory (Quercus-Carya spp.) to the south – with boreal spruce-fir confined to higher elevations and latitudes (Cogbill et al. 2002, Thompson et al. 2013). Following European settlement, this region was almost completely deforested for agriculture. As farms were abandoned from the early twentieth century on, forest regrowth led to a preponderance of early-mid successional species such as eastern white pine (Pinus strobus), red maple (Acer rubrum), and aspen-birch (Populus-Betula spp.) (Foster and Aber 2006).

Today, northeastern forests are again dominated by late-successional, shade-tolerant species, but with striking differences in abundance, largely due to the legacy effects of different land use histories (e.g., farming, logging, etc.) and altered disturbance regimes (Seymour and White 2002, Foster and Aber 2006, Nowacki and Abrams 2015). Notably, the once-dominant beech is now less abundant than maples (particularly sugar maple, A. saccharum); red spruce (Picea rubra) and hemlock are less abundant; and stands dominated by white pine, cherry (Prunus spp.), and aspen-birch remain scattered throughout the landscape (Cogbill et al. 2002, Thompson et al. 2013). However, recent research has suggested an ongoing shift towards increased beech and spruce abundance (e.g., Bose et al. 2017a, b, Wason and Dovciak 2017). Northeastern forests are also less diverse in structure and composition (Ducey et al. 2013), with most consisting of even-aged northern hardwoods versus the intricately-patterned, uneven-aged species mosaics that historically existed from low-intensity, localized natural disturbances (e.g., windthrows, insect outbreaks, Native American land use) (Abrams, 2005, Lorimer and White 2003).
1.4.2 Abiotic and biotic drivers of change

Recent important drivers of forest change across the northeastern US include: climate change, rising atmospheric carbon dioxide levels, climate-related disturbance (e.g., ice storms), atmospheric deposition of pollutants (e.g., nitrogen and sulfur), human land use (see Martinuzzi et al. 2015 for projected impacts) and land use history, exotic species/pests/pathogens, and elevated levels of herbivory by wildlife (e.g., deer). Disentangling the relative influence of each driver in affecting change across time and space is a significant challenge (e.g., Klepeis et al. 2013, Nowacki and Abrams 2015, Pederson et al. 2015). For example, while regional climate change has historically driven forest composition across long time periods (DeHayes et al. 2000, Shuman et al. 2009), some research suggests that species abundance and distribution patterns in the Northeast are currently decoupled from their historic climate controls due to land use legacies (e.g., Thompson et al. 2013, Nowacki and Abrams 2015; though see Pederson et al. 2015, Fei et al. 2017). Further, others have shown abiotic and biotic factors can interact in complex ways, often with species-specific responses, to influence landscape-scale forest composition and growth dynamics (e.g., Gandhi and Herms 2010, Pontius et al. 2016, Wason and Dovciak 2017, Wason et al. 2017b). Nonetheless, there are well-documented examples of each driver causing widespread effects on forest composition, structure, and function in the Northeast (note land use and anthropogenic climate change are mainly discussed in other sections, 1.4.1 and 1.4.3, respectively and thus not elaborated upon in detail here).
Atmospheric deposition of nitrogen and sulfur (i.e., acid deposition) has affected forest health, growth, and competitive dynamics through its effects on soil and tree physiology (Schaberg et al. 2001, Driscoll et al. 2003). In soils, acid deposition depletes calcium – an essential nutrient – resulting in nutritional deficiencies that predispose trees to decline following exposure to other stressors (e.g., extreme weather events) (Schaberg et al. 2001, Halman et al. 2011, Halman et al. 2013). These deleterious effects, however, are not consistent across species as those most heavily-impacted to date include red spruce (decreased cold tolerance and canopy health, mortality), sugar maple (calciphilic species, decreased canopy health, reduced competitive status), and birches (decreased canopy health, impaired growth) (Driscoll et al. 2001, Halman et al. 2011, Schaberg et al. 2011).

Exotic pathogens have caused (and continue to cause) significant declines and even near-extinction events in many northeastern tree species, with notable examples including American chestnut (*Castanea dentata*) and chestnut blight (Freinkel 2009), butternut (*Juglans cinerea*) and butternut canker (Ostry and Woeste 2004), and elm species (*Ulmus*) and Dutch elm disease (Strobel and Lanier 1981); exotic pests that will likely cause widespread future changes in northeastern forests include the recently-arrived emerald ash borer and hemlock woolly adelgid. Additionally, exotic earthworms have substantially altered soil composition and litterfall decomposition rates in the Northeast, affecting understory forest composition by negatively impacting native tree seedlings and favoring exotic shrubs (e.g., buckthorn) (Nuzzo et al. 2009); high rates of deer herbivory are influencing understory forest composition in similar ways (Horsley and Stout 2003).
Increased prevalence of exotic plants and pests/pathogens is also likely as anthropogenic climate change progresses (Dukes et al. 2009).

**1.4.3 Anthropogenic climate change impacts: observed and projected**

Over the past century, anthropogenic climate change has increased mean temperatures by ~2°F and altered precipitation regimes across the Northeast (Horton et al. 2014). Trombulak and Wolfson (2004) have shown that temperature change has varied in the region, with some areas experiencing more drastic increases (>3°F) and others less. While precipitation totals have generally increased, extreme flooding and drought events have become more common and winter precipitation is falling as rain rather than snow more often (Horton et al. 2014, Guilbert et al. 2015). Recent modeling efforts by Hayhoe et al. (2007) and Hayhoe et al. (2008) suggest that these climatic changes have been increasing in intensity since the 1970s and will continue to do so at similar or higher rates in the future depending on climate mitigation strategies.

Several recent studies suggest these climatic changes are significantly impacting northeastern forest composition via their effects on competition and survival, particularly in the understory. For example, Woodall et al. (2009) found the mean latitude of seedling abundance was higher than that of mature biomass for several northern hardwood species – an indication of poleward (latitudinal) migration (though see Zhu et al. 2012). Whereas Fei et al. (2017) found significant poleward and westward (longitudinal) shifts in sapling abundance over the past 30 years. In the dominant maple-beech-birch forests of the region, there have been documented increases in American beech abundance at the expense of
maple and birch (Duchesne and Ouimet 2009, Pontius et al. 2016, Bose et al. 2017a, b, Wason and Dovciak 2017). These shifts have been linked to both increased temperature and precipitation (Bose et al. 2017a, b). However, other investigations suggest spatially-complex dynamics involving climate, beech bark disease, deer herbivory, and the legacy effects of land use and acid deposition (e.g., Pontius et al. 2016, Wason and Dovciak 2017).

As climate change progresses, habitat suitability and associated competitive dynamics for many northeastern tree species are likely to change (Iverson and Prasad 1998; see Rustad et al. 2012 for a review). For example, increased temperatures and prolonged droughts will likely favor the northern expansion of warm-adapted oak and hickory species from the south, while the current cool-adapted dominants (e.g., maples) retreat to cooler climate refugia (Mohan et al. 2009, Tang et al. 2012). Changes in forest phenology associated with climatic changes are already occurring, with spring arriving earlier and autumn senescence later (lengthening the active growing season) (see Richardson et al. 2013 for a review). These changes are expected to further alter the important habitat components that govern species abundance and distribution patterns, like microclimate, nutrient availability, and soil moisture/temperature gradients (Chuine and Beaubien 2001, Huntington et al. 2009). Together, it is clear that anthropogenic climate change is having, and will continue to have, significant impacts on the northeastern forest dynamics.

1.5 Conclusions

Remote sensing is a powerful tool for mapping and detecting change in northeastern forest ecosystems. However, the Northeast also poses unique challenges to satellite
imagery-based research given its: 1) heterogeneous forest composition, 2) persistent cloud cover and diverse topography, and 3) numerous existing drivers of forest change across varying spatial scales. Current limitations include a lack of detailed, accurate large-scale forest maps and few studies that have comprehensively examined spatiotemporal changes in forest composition and function (e.g., tree species abundance and phenology) in response to climate change and other environmental drivers across the region. The research contained in this dissertation directly addresses these issues.

1.6 References


Huntington, T. G., A. D. Richardson, K. J. McGuire, and K. Hayhoe. 2009. Climate and hydrological changes in the northeastern United States: recent trends and


CHAPTER 2: ENHANCED FOREST COVER MAPPING USING SPECTRAL UNMIXING AND OBJECT-BASED CLASSIFICATION OF MULTI-TEMPORAL LANDSAT IMAGERY

David Gudex-Cross\(^1\),*\, Jennifer Pontius\(^1\), and Alison Adams\(^1\)

\(^1\) The Rubenstein School of Environment and Natural Resources, University of Vermont – Aiken Center, 81 Carrigan Drive, Burlington, Vermont 05405

* Corresponding Author: dgudexcr@uvm.edu (D. Gudex-Cross)
2.1 Abstract

Spatially-explicit tree species distribution maps are increasingly valuable to forest managers and researchers considering the effects of climate change and invasive pests on forest resources. Traditional forest classifications are limited to broad classes of forest types with variable accuracy. Advanced remote sensing techniques, such as spectral unmixing and object-based image analysis, offer novel forest mapping approaches by quantifying proportional species composition at the pixel level and utilizing ancillary environmental data for forest classifications. This is particularly useful in the northeastern region of the United States where species composition is often mixed.

Here, we employed a hierarchical forest mapping approach using spectral unmixing of multi-temporal Landsat imagery to quantify percent basal area for ten common tree species/genera across northern New York and Vermont. Basal area maps were then refined using an object-based ruleset to produce a thematic forest classification. Validation with 50 field inventory plots covering a range of species compositions indicated that the quality of percent basal area mapping largely reflected the number of “pure” (>80% BA) endmember plots available for calibration, with more common species mapped at a higher accuracy (i.e., *Acer saccharum*, adj. $r^2 = 0.44$, compared to *Populus* spp., adj. $r^2 = 0.24$).

The resulting thematic forest classification mapped 15 forest classes (nine species/genus level and six common species assemblages) with overall accuracy = 42%, KHAT = 33%, fuzzy accuracy = 86% at the pixel level, and 38%, KHAT = 29%, fuzzy accuracy = 84% at the object level. Using the validation plots to compare existing forest classification products, this hierarchical approach provided more class detail (11 represented classes) and higher accuracy than the National Forest Type Map (six represented classes, overall accuracy 18%, fuzzy accuracy 70%), LANDFIRE (five represented classes, overall accuracy 28%, fuzzy accuracy 80%) and National Land Cover Database (three represented classes, overall accuracy = 56%). These results show that more detailed and accurate forest mapping is possible using a combination of multi-temporal imagery, spectral unmixing, and rule-based classification techniques. Improved large-scale forest mapping has important implications for natural resource management and other modeling applications.

**Keywords:** remote sensing, basal area mapping, biophysical modeling, hierarchical classification, object-based image analysis
2.2 Introduction

Developing cost-effective methods to accurately classify forest cover is essential to inform sustainable forest management at local, regional, and national levels. These products are increasingly valuable considering the anticipated effects of climate change and invasive pests on forest resources. Warming temperatures and changing precipitation regimes are expected to cause shifts in tree species distributions (Iverson and Prasad 2001, Hamann and Wang 2006, Tang et al. 2012) and increases in the duration and severity of pest/pathogen outbreaks (Dale et al. 2001, Dukes et al. 2009). Yet our ability to direct management actions is limited by the coarse detail and relatively low accuracy of existing large-scale forest cover maps.

Existing forest cover maps include field inventory and remote sensing-based products, including those generated through the Forest Inventory and Analysis program (FIA; www.fia.fs.fed.us), the National Land Cover Database (NLCD; www.mrlc.gov) and LANDFIRE Existing Vegetation Type (LANDFIRE EVT; www.landfire.gov). More recently, the US Forest Service (USFS) used FIA data, multi-temporal Moderate Resolution Imaging Spectroradiometer (MODIS) data, vegetation indices, and other ancillary environmental data to produce the National Forest Type map (NFTM, data.fs.usda.gov/geodata/rastergateway/forest_type/index.php). The LANDFIRE and NLCD programs provide national forest type maps at a 30m x 30m spatial resolution, but in coarser forest type classes than FIA/USFS species-level products (250m x 250m).

Several remote sensing studies have successfully mapped species-level distributions, though largely at highly localized spatial scales (Martin et al. 1998, Carleer
and Wolff 2004, Plourde et al. 2007, Ke et al. 2010, Immitzer et al. 2012). These studies typically rely on data-intensive hyperspectral and/or hyperspatial resolution imagery (e.g., Ikonos, QuickBird, WorldView-2, Airborne Visible/Infrared Imaging Spectrometer – AVIRIS, Light Detection and Ranging – LiDAR), limiting their applicability to tree species/genus classification across larger regions.

Wolter et al. (1995), Mickelson et al. (1998), and Hill et al. (2010) achieved relatively accurate species-type classifications by utilizing multi-temporal Landsat imagery, demonstrating the usefulness of acquiring multiple image dates that capture phenologically-significant differences among species (e.g. green-up, senescence, etc.). Dymond et al. (2002) also found improved deciduous forest type discrimination when multi-temporal Landsat imagery was supplemented with Normalized Difference Vegetation Index (NDVI) and Tasseled Cap Transformation (TC) bands, as well as their respective differences among image dates.

Advanced remote sensing techniques, such as spectral unmixing and object-based image analysis (OBIA), utilize a wealth of spectral, spatial, and ancillary environmental data to enable more precise forest cover mapping (see Xie et al. 2008, Pu 2013 for reviews). Spectral unmixing has been shown to outperform traditional pixel-based classifiers by decomposing (“unmixing”) mixed pixels and assigning component proportions at the subpixel level (Huguenin et al. 1997, Oki et al. 2002). This is particularly useful in northeastern forests where species composition is often mixed at local scales. The resulting per-pixel proportions of each species obtained from the spectral unmixing process also facilitate the mapping of other forest attributes that are dependent upon the complexity of
species composition common in northeastern forests (e.g., carbon storage, basal area, productivity) (Hall et al. 1995, Sonnentag et al. 2007, Yan et al. 2015). OBIA techniques overcome individual pixel constraints by segmenting imagery into homogenous “objects” upon which classification is then carried out. This allows for the additional characterization of shape, size, and texture into classifications and minimizes impacts of canopy architecture-driven variability in spectral signatures (Chubey et al. 2006).

While OBIA is often more accurate than pixel-based methods for mapping forest cover at high spatial resolutions (Dorren et al. 2003, Oruc et al. 2004, Agarwal et al. 2013), comparative studies indicate that coupling pixel-based and OBIA techniques can improve the accuracy of forest type classifications (Wang et al. 2004, Aguirre-Gutiérrez et al. 2012). Using Ikonos imagery, Wang et al. (2004) achieved the highest mangrove classification accuracies when integrating a pixel-level classification to identify spectrally-distinct classes, then carrying out an object-based nearest neighbor analysis on spectrally-mixed classes. Similarly, Aguirre-Gutiérrez et al. (2012) obtained the highest accuracy in montane landscapes when merging the best pixel-based and object-based classes to produce the final thematic land cover classification.

Here, we test a novel approach to tree species mapping that integrates many of the successful approaches used in previous studies. This involves pixel-level spectral unmixing that integrates multi-temporal Landsat imagery and field inventory data to develop percent basal area coverages for 10 common species/genera. These percent basal area coverages are then incorporated into an object-based hierarchical ruleset to generate 16 forest classes (10 species/genera and 6 common assemblages). To evaluate the utility of
this integrated multi-temporal, spectral unmixing (MTSU) approach, we compare accuracy with existing large-scale forest mapping products, including LANDFIRE EVT, NFTM, and NLCD.

Achieving accurate, species-specific forest classifications is necessary to fill critical gaps in our knowledge of current tree species distributions and provide baselines for future comparisons. This integrated approach attempts to maximize the accuracy and detail possible from widely available Landsat imagery, allowing for improved, widespread mapping of important forest resources. Improved forest mapping also enables better parameterization of dynamic vegetation and climate models.

2.3 Methods

2.3.1 Study area and base imagery

This study was conducted on Landsat Row 29, Path 14, which spans much of northern New York and Vermont (Fig. 2.1). Forest composition across the region is highly heterogeneous with dominant canopy species including sugar maple (Acer saccharum), red maple (Acer rubrum), American beech (Fagus grandifolia), eastern hemlock (Tsuga canadensis), eastern white pine (Pinus strobus), and yellow birch (Betula alleghaniensis). Upper elevations are dominated by balsam fir (Abies balsamea), red spruce (Picea rubens), and birches (Betula spp.) (Widmann 2015, Morin and Widmann 2016).

Seasonal Landsat Operational Land Imager and Thermal Infrared Sensor (OLI-TIRS) and Enhanced Thematic Mapper Plus (ETM+) images (USGS level 1T products) were acquired for targeted, phenologically-representative dates: full snow cover (winter),
green up (spring), mid-growing season (summer), and peak fall color (fall). Because cloud cover is a common issue across this mountainous region, we included the lowest cloud cover image within a two-year buffer around the representative year (i.e. “2014” candidate images were chosen from 2012-2016). Cloud cover was masked, then backfilled via seamless mosaicking with another image acquired within two weeks of the base image. Even within these parameters, for this study we were unable to compile a spring image with sufficiently low cloud cover and hence excluded this season from further processing.

Figure 2.1. The study area, spanning northern New York and Vermont, and distribution of ground-reference plots (Landsat Path 14, Row 29).
2.3.2 Ground-reference data

“Pure” endmember spectra (plots with >80% basal area for a given species) for spectral unmixing algorithms were obtained from FIA plots distributed throughout the region, with an additional 20 variable-radius plots (collected with a 10-factor wedge prism) also used to improve representation of species underrepresented in the FIA data (Fig. 2.1). Aggregated to the plot level, this resulted in 54 plots containing >80% basal area to calibrate the unmixing models for ten common species or genera (Table 2.1).

For model validation, the FIA endmember plots were supplemented by mixed species composition plots from the Vermont Monitoring Cooperative (VMC; www.uvm.edu/vmc) for a total of 50 plots covering a range of species compositions (Table 2.1). Both programs employ the same sampling design, with four 1/24 acre subplots (see Bechtold and Patterson 2005) and measurement of all stems greater than five inches diameter at breast height.

Sugar maple, birches, American beech, red spruce, and red maple respectively occurred on the most ground-reference plots, while sugar maple, balsam fir, birches, and eastern hemlock had the highest percent basal area. It is important to note that percent basal area measurements did not differentiate between canopy dominant and understory trees, likely contributing to error in the resulting percent basal areas models that are based solely on reflectance signatures from the top of the canopy.
Table 2.1. The composition of ground-reference plots used for development of percent basal area (%BA) models. SD = standard deviation.

<table>
<thead>
<tr>
<th>Tree spp./genus</th>
<th>No. of pure endmember plots</th>
<th>Mean %BA (±SD)</th>
<th>Max %BA</th>
<th>No. of Plots w/Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balsam fir</td>
<td>8</td>
<td>14.3 (27.7)</td>
<td>92.5</td>
<td>14</td>
</tr>
<tr>
<td>Red maple</td>
<td>2</td>
<td>6.7 (14.2)</td>
<td>80.5</td>
<td>18</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>10</td>
<td>27.7 (36.6)</td>
<td>96.0</td>
<td>27</td>
</tr>
<tr>
<td>Birches</td>
<td>6</td>
<td>13.1 (20.4)</td>
<td>80.7</td>
<td>26</td>
</tr>
<tr>
<td>American beech</td>
<td>2</td>
<td>6.3 (13.2)</td>
<td>81.8</td>
<td>22</td>
</tr>
<tr>
<td>Red spruce</td>
<td>1</td>
<td>5.7 (14.6)</td>
<td>92.0</td>
<td>20</td>
</tr>
<tr>
<td>Eastern white pine</td>
<td>11</td>
<td>5.8 (21.4)</td>
<td>100.0</td>
<td>6</td>
</tr>
<tr>
<td>Aspens</td>
<td>1</td>
<td>3.3 (13.8)</td>
<td>86.5</td>
<td>5</td>
</tr>
<tr>
<td>Oaks</td>
<td>2</td>
<td>3.4 (13.1)</td>
<td>65.0</td>
<td>4</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>11</td>
<td>9.2 (25.0)</td>
<td>93.1</td>
<td>10</td>
</tr>
</tbody>
</table>

2.3.3 Preprocessing

Landsat Level 1T products come with basic radiometric calibration and topographically corrected georegistration. In-house preprocessing (Fig. 2.2) included atmospheric corrections to at-surface reflectance using a dark-object subtraction technique (Chavez Jr 1989). We then derived NDVI and TC (Crist and Cicone 1984) bands for each season and calculated seasonal TC differences between summer and fall. These indices have previously been shown to improve landscape-level forest type discrimination of multi-temporal Landsat imagery (Dymond et al. 2002).

Running a principal component analysis (PCA) on forested pixels only on the resulting 33 band imagery stack as a precursor to the Minimum Noise Fraction (MNF) transform (see section 2.3.4 below) allowed us to minimize autocorrelation among the full component of input bands. This step removed noise inherent in many of these bands due
to differences in illumination and atmospheric conditions across different image acquisition dates and isolated the spectral signal specific to distinguishing forested pixels.

The final stacked image for spectral unmixing included the first three PCA bands (accounting for >99% of the spectral variability in the full 33-band stack). Because these PCA bands were primarily distinguishing among species composition (see section 2.4.1), the final stacked image also included summer Landsat reflectance bands, NDVI, Tasseled Cap, and Tasseled Cap difference vegetation index products (Fig. 2.2) to capture information about canopy density for percent basal area modeling.
Figure 2.2. Landsat preprocessing and percent basal area modeling workflow.
2.3.4 Spectral unmixing

The spectral unmixing process outlined here largely follows that developed by Nielsen (2001) and Boardman and Kruse (2011), which has previously been used to classify tree species with hyperspectral imagery (see Plourde et al. 2007, Hallett et al. 2010). A MNF transform was first applied to the final imagery stack (17 bands) for data decorrelation and spectral noise reduction (Green et al. 1988) (Fig. 2.2). Endmember pixels were refined using a Pixel Purity Index to ensure spectral similarity of MNF bands among geographically distinct sites, with spectral outliers being excluded from further analysis. The resulting MNF image was then “unmixed” using a Mixture-tuned Matched Filtering (MTMF) algorithm (Boardman 1998) based on the target endmember spectra (i.e. tree species signatures). MTMF is a form of spectral mixture analysis that employs partial linear unmixing to map the abundance or fraction of target endmember spectra within each pixel (Boardman and Kruse 2011). The MTMF output consists of a matched filter and infeasibility score for each pixel, with the former reflecting how well the pixel matches the target spectra and the latter representing the likelihood of a false positive.

We considered several approaches to model percent basal area for input into the object-based classification ruleset based on the MTMF products. The traditional approach involves identifying thresholds for matched filter and infeasibility scores to maximize the binary accuracy of a species’ presence/absence. Because we were mapping heterogeneous forest cover dominated by mixed species composition, a binary classification scheme was ruled out for our purposes. Regression models have also been used to map species fractional basal area using hyperspectral imagery (Pontius et al. 2005). This study differed
from these previous single species efforts based on the large number of ground-reference plots across a range of forest species composition. The diverse plot network resulted in a variable number of plots where the target species was completely absent, as well as a suite of possible matched filter and infeasibility scores derived from the 10 species unmixing products.

Using linear regression models based only on plots that contained the species of interest produced more stable regression metrics but resulted in many false positives where particular species were absent. We also tested zero-inflated regression to account for the propensity of zero basal area plots in the calibration data. Results were generally lower model fit than the general linear models, with continued over-prediction of zero basal area plots. Further, zero-inflation p-values resulting from regression estimates were not significant, indicating that the presence of zero value data was not a significant contributor to overall model variability.

Our most consistently accurate results came from a stepwise linear regression model that included all ground-reference plots (including those where the target species was absent). Model terms were limited to matched filter and infeasibility variables significant at the 0.05 level, with a maximum variance inflation factor of 10 to avoid autocorrelation among parameters. We used the minimum Bayesian Information Criterion (Bhat and Kumar 2010) to select the best fit model. The resulting regression equation was then applied to the MTMF image via band math to create a percent basal area raster for each target species/genus.
It is important to note that the resulting fractional basal area products were not intended to be stand-alone products, but instead to be used as inputs to quantify the relative abundance of species within each pixel to inform classification. These relative abundances were not aggregated for all species but instead used as independent inputs to the object-based hierarchical ruleset (see section 2.3.5 below).

**2.3.5 Object-based classification**

Percent basal area rasters obtained from the pixel-based spectral unmixing were then incorporated into an object-based, hierarchical ruleset classification scheme (Fig. 2.3). This allowed us to refine the percent basal area products using ancillary environmental data (i.e. digital elevation data from the National Elevation Dataset available through the U.S. Geological Survey) and produce classifications on a stand- versus pixel-level.

Object-based classifications begin with segmentation to aggregate like pixels into larger image objects. Segmentation settings and input layer weightings were informed by knowledge of the image resolution, spatial characteristics of the landscape, and spectral nature of the feature objects. As is common in object-based classifications, iterations of various settings were evaluated to confirm selection of final segmentation settings. We used a multiresolution segmentation algorithm (see Chubey et al. 2006 for further explanation) based on layer inputs that highlighted differences in vegetation characteristics across our study area. This included weighting the first three MNF bands most, followed by summer and winter NDVI and seasonal TC differences. Given the moderate spatial resolution of Landsat imagery and heterogeneous nature of forest composition patterns across the landscape, a very low scale parameter (1) with no shape or compactness
weighting was used for object segmentation. To compare the pixel-level ruleset results to this object-based approach, a chessboard segmentation with a scale parameter of 1 and no band weighting was used to create pixel “objects”.

The ruleset started by differentiating forest from non-forest objects using thresholds for winter band 3 leveraging snow cover (non-forest > 0.60), and spring band 4, masking water (non-forest < 0.065). Forest classes were then assigned based on percent basal area rasters and elevational constraints outlined by Burns and Honkala (1990b) (Fig. 2.3) following a rule-based hierarchy. A species/genus class was assigned if the object contained greater than 40% basal area of that species or genus and did not exceed the specified elevation threshold (if there was one). Since rare species are spottier across the landscape and more likely to be smoothed out when averaged within image objects, forest type assignment in the hierarchical ruleset progressed from the least to most common species to maximize representation of rare species in the final thematic classification.

To capture regionally-common species assemblages where no species was greater than 40% basal area, we also classified six common forest assemblages by summing the percent basal area values for their respective component species (Fig. 2.3). The final thematic forest classification of 16 possible forest types was then exported as a 30-meter by 30-meter raster product.
Figure 2.3. Object-based image analysis workflow and hierarchical ruleset used to create a thematic forest cover map.
2.3.6 Accuracy assessment

Inventory data for the FIA and VMC plots described above (see section 2.3.2) were used to assign a forest class according to the same rule thresholds applied to the imagery. A confusion matrix of actual versus predicted forest classes was created to examine overall, kappa, User’s, and Producer’s accuracies. We also determined fuzzy accuracy by allowing misclassification between common species/species assemblages. For example, we considered sugar maple pixels that were classified as northern hardwoods to be “correct” at the fuzzy level.

We similarly calculated accuracy for three existing forest mapping products: the 2011 LANDFIRE EVT classification; the National Forest Type Map classification; and the 2011 NLCD classification. Only the LANDFIRE and National Forest Type Map classifications could be compared at the species-type level, with accuracy being determined following the same process outlined above with field plots assigned to match their respective classes. For the NLCD product, we classified the validation data as deciduous (>75% deciduous species), evergreen (>75% evergreen species), or mixed forest (a plot was considered mixed when both deciduous and evergreen species were present but neither exceeded 75% of the plot basal area).
2.4 Results and Discussion

2.4.1 Spectral decomposition

Our approach included the aggregation of a variety of image dates and vegetation index products in order to maximize the spectral information available to differentiate physiologically similar species. Eigenvectors from spectral decomposition were used to identify which bands accounted for the most variability among forested pixels. From the full 33 band multi-temporal stack, the largest eigenvectors came from the fall image (Table 2.2). It is important to keep in mind that this PCA was run on forested pixels only to isolate the potential spectral signal specific to differentiation among forest types (not forest/non-forest). The fall image was timed at the peak of physiological differentiation among species for our region, providing key spectral information to help separate otherwise spectrally similar species. Other studies have also cited the importance of using shoulder seasons with unique phenological information to assist in species classification (Dymond et al. 2002, Hill et al. 2010).

Table 2.2. Principal components analysis eigenvectors highlight the input bands that account for the most spectral variability among forested pixels.

<table>
<thead>
<tr>
<th>Input Band</th>
<th>PCA Band</th>
<th>Eigenvector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall Band 5 (Mid-IR)</td>
<td>1</td>
<td>0.233</td>
</tr>
<tr>
<td>Fall Band 2 (Green)</td>
<td>1</td>
<td>0.207</td>
</tr>
<tr>
<td>Fall Band 4 (NIR)</td>
<td>2</td>
<td>0.295</td>
</tr>
<tr>
<td>Fall NDVI</td>
<td>2</td>
<td>0.271</td>
</tr>
</tbody>
</table>
2.4.2 Percent basal area modeling

MTMF models of percent basal area resulted in significant but relatively weak (adj. $r^2 = 0.24$; RMSE = 0.04, *Populus* sp.) to moderate relationships (adj. $r^2 = 0.59$; RMSE = 0.06, American beech). These relatively low model fits likely result from several sources of known error. The sensor primarily records the spectral reflectance from the canopy surface, with a mix of canopy dominant trees that may differ from understory composition included in ground-reference inventories. Further, percent basal area is based on main trunk diameter at breast height with no accounting for variability in crown size, health, or geometry among species. This is reflected in lower fit statistics for species that are more common in the understory of northeastern forests (e.g., eastern hemlock) or with relatively small crown geometry relative to common co-occurring species.

The lack of fit is likely also driven by the preponderance of “pure” species plots included in the validation dataset. This resulted in plots with extreme high and extreme low (zero occurrence) values of each target species, levels where regression models are typically weakest. Species/genera with the lowest percent basal area fit were those with the fewest endmember calibration plots and lowest general abundance across the study area (per FIA forest demographic reports). For these target species, percent basal area was typically under-predicted (Table 2.3). The most accurate percent basal area models were associated with the dominant species in the region (e.g., American beech).
Table 2.3. Percent basal area model fits derived from spectral unmixing.

<table>
<thead>
<tr>
<th>Tree spp./genus</th>
<th>$r^2$</th>
<th>Adj. $r^2$</th>
<th>Mean %BA</th>
<th>RMSE</th>
<th>PRESS RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balsam Fir</td>
<td>0.34</td>
<td>0.32</td>
<td>0.15</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Red maple</td>
<td>0.47</td>
<td>0.46</td>
<td>0.08</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>0.46</td>
<td>0.44</td>
<td>0.28</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>Birches (<em>Betula</em> spp.)</td>
<td>0.32</td>
<td>0.30</td>
<td>0.13</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>American beech</td>
<td>0.60</td>
<td>0.59</td>
<td>0.07</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Red spruce</td>
<td>0.52</td>
<td>0.51</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Eastern white pine</td>
<td>0.3</td>
<td>0.29</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Aspens (<em>Populus</em> spp.)</td>
<td>0.25</td>
<td>0.24</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Oaks (<em>Quercus</em> spp.)</td>
<td>0.49</td>
<td>0.48</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>0.32</td>
<td>0.30</td>
<td>0.11</td>
<td>0.09</td>
<td>0.1</td>
</tr>
</tbody>
</table>

These results are similar to other species mapping efforts. Savage et al. (2015) used a zero-inflated regression model, based on a two-step process, to first predict the presence or absence of the target species and then species composition only where the target species was present. They modeled five different conifer species in heterogeneous forests of northwestern Montana using Landsat TM and OLI imagery, reporting independent accuracy assessment RMSE from 0.11 to 0.23 (no $r^2$ values were reported). These errors are slightly higher than the range of RMSE values reported for our ten target species (0.04 to 0.16).

Moisen et al. (2006) compared generalized additive regression modeling, classification and regression tree (CART) techniques, and stochastic gradient boosting for modeling live basal area from multi-temporal Landsat imagery for thirteen tree species in Utah. Basal area prediction results for all modeling techniques were poor for most species ($r^2$ less than 0.5 and RMS errors greater than 0.8). While the general approach employed by Moisen et al. (2006) is similar to that described here (multi-temporal Landsat imagery),
our range of model fit is higher, indicating that the additional image processing techniques and spectral unmixing approach employed here may improve abundance mapping using Landsat imagery.

Our percent basal area modeling results also compare favorably to those obtained in other studies using MTMF techniques. Hyperspectral imagery, with its wealth of narrow reflectance bands, is well suited to spectral unmixing and species abundance mapping. Hyperspectral instruments have reported comparable accuracy to that reported here for eastern hemlock abundance in the Catskills region ($r^2 = 0.65$; RMSE 0.12, Pontius et al. 2005). Plourde et al. (2007) used spectral unmixing to model percent sugar maple and American beech in New Hampshire using both hyperspectral AVIRIS imagery, as well as modifications of the hyperspectral imagery to match broadband sensors. They found weak relationships between field-measured and predicted percent basal area based on the broadband imagery but results similar to those reported here for spectral unmixing of the full hyperspectral data ($r^2 = 0.49$; RMSE = 0.09 for sugar maple and $r^2 = 0.36$; RMSE = 0.18 for beech).

These studies collectively underscore that modeling continuous variables, like individual tree species basal area, is a difficult task. Clearly the spatial resolution of Landsat imagery is limiting for mapping forest cover at the species level in highly mixed forests. Difficulties associated with scaling field data to the Landsat pixel level include: overlap in canopy dominant species (Plourde et al. 2007, Hallett et al. 2010); incongruities between field measurements (which include understory stems) and sensor-derived canopy reflectance (particularly for shade-tolerant species such as hemlock); and incorrect
registration between calibration field plots and pixel centers. Atmospheric and topographic shadow impacts on spectra are also particularly troublesome in mountainous regions. Within-species spectral variability due to differences in tree health can also confound unmixing algorithms (Carter 1993, Plourde et al. 2007).

While these errors impact the overall accuracy of the models, it is interesting to note that the multi-temporal, broadband, spectral unmixing approach described here reports similar accuracy to hyperspectral efforts (Pontius et al. 2005, Plourde et al. 2007) and improved accuracy compared to other broadband-based tree species abundance mapping (Moisen et al. 2006, Plourde et al. 2007, Savage et al. 2015). We attribute the improved performance of our MTSU integrative approach to a combination of factors: 1) the use of multi-temporal imagery to capture species-specific spectral characteristics during key phenological times; 2) the inclusion of vegetation indices derived from the multi-temporal images to isolate species-specific differences in vegetation characteristics across seasons; and 3) the use of MTMF products from multiple species components to model abundance of the target species. Previous broadband sensor-based studies have shown the utility of using multiple phenologically-important image dates and vegetation indices when classifying heterogeneous forest cover at the species-type level (e.g. Dymond et al. 2002, Hill et al. 2010). Others have highlighted that the use of multiple endmembers in spectral mixture analysis can improve assessments of forest structural attributes (Hall et al. 1995, Roberts et al. 1998).

Our resulting maps of species percent basal area match expected patterns across northern New York and Vermont (see Burns and Honkala 1990b for species-specific
descriptions), particularly for the five most abundant species (Fig. 2.4). Balsam fir (Fig. 2.4a) was predicted throughout the high elevations of both the Green and Adirondack mountains, in addition to the lowland boreal forest areas of northeastern Vermont. Birches (Fig. 2.4b) followed a spatial distribution similar to sugar maple but with greater occurrence at higher elevations. Since birches were modeled at the genus level, this reflects the northern hardwood type-driven distribution of yellow birch (*Betula alleghaniensis*) and that of paper birches (*Betula papyrifera/cordifolia*), which are common constituents of high elevation spruce-fir forests in the northeastern USA (Burns and Honkala 1990b, Leak et al. 2014). Sugar maple was predicted as prevalent throughout much of the region (Fig. 2.4c), which matches recent FIA field inventories in both states (Widmann 2015, Morin and Widmann 2016). Its highest estimated percent basal area densities were along the low to mid elevation slopes of the Adirondack and Green mountain ranges, where it is an integral component of the northern hardwoods forest type (Leak et al. 2014). Eastern white pine (Fig. 2.4d) was largely predicted to be found along the Lake Champlain valley lowlands of both states, but more prevalent in New York.
Figure 2.4. The spatial distribution of percent basal area derived from spectral unmixing for four common species in northern New York and Vermont.
2.4.3 Comparison of object-based and pixel-level thematic forest classifications

Rule-based, OBIA classification schemes are commonly used with high spatial resolution imagery that exhibits unique shape and texture features. Due to the relatively coarse pixels of Landsat, we compared the thematic results of the hierarchical ruleset applied to both individual pixels (pixel-level, PL) and image-segmented stand “objects” (object-level, OL) to determine if image segmentation was necessary to maximize accuracy of forest classifications. The relative abundance of the 16 forest classes was similar for both the pixel-level (PL) and object-based (OB) maps. The most striking difference was far fewer pixels classified as species-dominant in the OB map. This result is to be expected given the averaging of neighboring pixel values to create one common value for each stand-level object, which effectively washes out single-species dominant pixels. Spatial patterns for the PL and OB maps were indiscernible at the regional level. However, a localized, side-by-side comparison of both products revealed the PL map’s finer species-level detail and grainier appearance against the smoother, species assemblage-dominated OB map (Fig. 2.5). In the Stowe region of Vermont, for example, the PL map predicted more single-species dominant stands of balsam fir, red spruce, and eastern hemlock, largely in areas classified as mixed conifers on the OB map. Yet the general spatial distribution patterns of the predominant forest classes around Stowe were very similar, with both maps showing mixed classes around lowland and developed areas, mountain slopes dominated by northern hardwoods and sugar maple, and spruce-fir related classes at high elevations.
Figure 2.5. Comparison of the object-based (top) and pixel-level (bottom) classifications in the Stowe region of Vermont.
Based on ground-reference plots, overall classification accuracy among forest types was slightly higher for the PL (overall accuracy = 42%, KHAT = 33%, fuzzy accuracy = 86%) versus the OB classification (overall = 38%, KHAT = 29%, fuzzy = 84%). The increased detail of the PL classification also better matches the complex spatial heterogeneity of forests across the region. Given this, we consider the PL more appropriate for mapping forest types using Landsat imagery in the Northeast. For this reason, we include only a discussion of the PL results below.

2.4.4 Pixel-level thematic forest classification

Applying the classification ruleset across the study area shows a spatial distribution of forest classes that match expected patterns across northern New York and Vermont (Fig. 2.6). Mixed hardwoods dominate the lowlands, while sugar maple and northern hardwoods occupy the low to mid elevation slopes of the Adirondack and Green mountain ranges. Spruce-fir and spruce-fir-birch assemblages were classified throughout the high elevations of both the Green and Adirondack mountains. Eastern white pine and hemlock were classified primarily along the Lake Champlain valley corridor. Interesting anomalies include the near absence of pixels classified as oak- or birch-dominant, and complete absence of aspens. While this may simply reflect their relatively low abundance as pure stands across the region, it is also likely that the low number of calibration endmembers has limited our ability to capture a sufficient range of spectral signatures for these tree species.
Figure 2.6. Forest cover map of northern New York and Vermont produced by integrating spectral unmixing of multi-temporal Landsat imagery (MTSU) and a hierarchical classification scheme.
Accuracy assessment for the pixel-level classification resulted in 42% overall accuracy (KHAT = 33%) (Table 2.4). When allowing for confusion between pure target species and common assemblages that by definition contain a significant portion of the target species, overall accuracy doubled (fuzzy accuracy = 86%), indicating that a majority of error resulted from incorrectly predicting mixed species classes for plots that were dominated by one species (but likely also contained others). Typically, the actual dominant species was an important component of the incorrectly predicted mixed species class (e.g. sugar maple was often incorrectly classified as northern hardwoods, of which it is a major component).

The highest producer’s accuracies were obtained for the most common forest types across the study area (Table 2.4): sugar maple, northern hardwoods, and spruce-fir-birch. Lower user’s accuracies for northern hardwoods highlight the tendency of the ruleset to categorize single species-dominant validation plots into this species assemblage class. The lowest user’s accuracies were obtained for less common species with relatively low abundance across the study area. These included birches and the conifer species (balsam fir, eastern hemlock, and eastern white pine), all of which were often classified as mixed species assemblages. If identification of less abundant species is desired, the percent basal area thresholds of the ruleset could be lowered to denote “dominant stands”. However, we suggest that if the goal of using these forest maps is examining the spatial and structural distribution of a particular species, using the percent basal area maps themselves would be preferential to using the thematic classification.
Table 2.4. Error matrix based on 50 ground-reference plots for the MTSU pixel-level forest classification. Bold indicates correct at the species-type level; italic indicates correct at the fuzzy level.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar Maple</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>71%</td>
<td>93%</td>
</tr>
<tr>
<td>Red Maple</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>American Beech</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Northern Hardwoods Birches</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>25%</td>
<td>83%</td>
</tr>
<tr>
<td>Birches</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Spruce-Fir-Birch</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Spruce-Fir</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Red Spruce</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Balsam Fir</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Eastern Hemlock</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Eastern White Pine</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Mixed Conifers</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Mixed Hardwoods</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Aspens</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Oaks</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td>50</td>
<td>42%</td>
<td>86%</td>
</tr>
</tbody>
</table>

Producer's Accuracy 67% 0% 60% 0% 83% 0% 100% 0% 20% 33% 0% 0% 0% 0% 0% 0%
Fuzzy PA 93% 100% 100% 80% 100% 0% 100% 67% 100% 100% 100% 0% 0% 0% 0% 0%

Overall Accuracy 42%
KHAT 33%
Fuzzy Accuracy 86%
To evaluate how this integrated forest classification compared to other commonly used forest cover maps, we consider the specificity of forest classes (number and structure of distinct classes), the spatial resolution, and the mapping accuracy of each product (Table 2.5).

### Table 2.5. Comparison of specificity, spatial resolution, and accuracy of forest mapping products.

<table>
<thead>
<tr>
<th>Product</th>
<th># Forest Classes</th>
<th>Spatial Resolution (m)</th>
<th>Spp-Type Accuracy</th>
<th>Fuzzy Accuracy</th>
<th>NLCD Coarse Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTSU</td>
<td>15</td>
<td>30</td>
<td>42%</td>
<td>86%</td>
<td>76%</td>
</tr>
<tr>
<td>LANDFIRE</td>
<td>17</td>
<td>30</td>
<td>28%</td>
<td>80%</td>
<td>66%</td>
</tr>
<tr>
<td>National Forest Type Map</td>
<td>29</td>
<td>250</td>
<td>18%</td>
<td>70%</td>
<td>62%</td>
</tr>
<tr>
<td>NLCD</td>
<td>3</td>
<td>30</td>
<td>--</td>
<td>--</td>
<td>56%</td>
</tr>
</tbody>
</table>

Our forest classification resulted in 15 forest types (no aspen stands mapped) across the study area, based on the 10 most common genera/species in the region and six common assemblages of these species. The National Forest Type Map and LANDFIRE EVT forest class structures are most comparable to our MTSU integrated classification with 29 and 17 predicted across the study area. Both include common species assemblages such as spruce-fir and northern hardwoods. The National Forest Type Map also includes species-specific classes (e.g., balsam fir, eastern hemlock, eastern white pine, etc.). Where the LANDFIRE EVT classification diverges from ours is in its use of disturbance and geographic modifiers to describe certain forest types (e.g., ruderal forest, Atlantic swamp forest). Further, its mixed forest classes often cover a broader range of species assemblages, (e.g., pine-hemlock-hardwood and spruce-fir-hardwood). The NLCD product only classifies three broad forest types: deciduous, evergreen, and mixed.
Our 50 ground-reference plots represented 11 forest types for our MTSU integrated classification, five for the LANDFIRE EVT, and six for the National Forest Type Map (Table 2.5). Of the five LANDFIRE EVT classes, nearly all were predicted as belonging to one of three mixed forest types (pine-hemlock-hardwood, spruce-fir-hardwood, or yellow birch-sugar maple). Of the six National Forest Type Map classes, our ground-reference plots were predominantly categorized as one mixed forest type (sugar maple-beech-yellow birch). This simplification of the heterogeneity of species assemblages found across the Northern Forest region into broad categories resulted in a gross over-prediction of yellow birch-sugar maple (LANDFIRE EVT) and sugar maple-beech-yellow birch (National Forest Type Map) across the landscape, while missing other species entirely.

Focusing on the topographically diverse forests in the Stowe region of Vermont, a comparison of these forest classifications highlights the increased spatial detail and specificity of our MTSU product (Fig. 2.7). The MTSU predicts balsam fir, red spruce, spruce-fir, and spruce-fir-birch stands at high elevations, in addition to scattered balsam fir dominated stands in lowland swamp areas near suburban developments. Along mountain slopes, northern hardwoods and sugar maple stands are found throughout the low-mid elevations, with rare occurrences of birch and American beech dominated pixels. The valleys are largely dominated by the MTSU’s broadest species assemblages: mixed, mixed conifers, and mixed hardwoods. These results contrast those of the National Forest Type Map and LANDFIRE EVT, which both classify much of the region as a mixed northern hardwoods-type (maple/beech/birch and yellow birch-sugar maple, respectively). The National Forest Type Map also does poorly distinguishing forest from non-forest and has
a more pixelated appearance due to its lower spatial resolution. The spatial distribution of NLCD forest cover aligns most closely with that of the MTSU product, but at a much coarser forest type specificity.
Figure 2.7. Comparison of the MTSU (top left), LANDFIRE EVT (bottom left), NLCD (top right), and National Forest Type Map (bottom right) forest cover maps near Stowe, Vermont.
To compare accuracy among the mapping products, we used the same 50 ground-reference plots referenced throughout this study. Since there are inherent differences in how each product categorizes forest types, ground-reference plots were assigned to match the comparison product categories based on their species composition. Our results indicate that our MTSU classification was more accurate than the LANDFIRE EVT product (42% compared to 28% overall accuracy respectively) and more than twice as accurate as the National Forest Type Map (42% compared to 18% overall accuracy respectively) (Table 2.5). While fuzzy accuracies are improved for the National Forest Type Map and LANDFIRE EVT products, this is likely inflated by their broad class structure and near uniform assignment of plots into mixed forest type classes that include most of the common species/genera found within our ground-reference dataset.

When modifying all four classifications to match the coarser NLCD forest types (i.e. deciduous, evergreen, and mixed forest) for a more direct comparison of the general performance of these models, again the MTSU outperformed the LANDFIRE EVT, National Forest Type Map, and NLCD products (76%, 66%, 62% and 56% overall accuracy, respectively) (Table 5). Most of the error in the MTSU was due to an over-prediction of mixed forest in conifer dominated plots. Deciduous forest, by far the most common class in the ground-reference data, was also the most accurately predicted in each classification. The high deciduous class accuracies of the LANDFIRE EVT and National Forest Type Map were again driven by their propensity to predict yellow birch-sugar maple and sugar maple-beech-yellow birch across the landscape.
2.5 Conclusions

Our results indicate that the use of multi-temporal Landsat imagery, spectral unmixing, and a hierarchical ruleset classification (‘MTSU’ integrated approach) offers improved species specificity and accuracy relative to existing forest classification products. The key to this approach includes: 1) the use of multi-temporal imagery to capture species-specific differences during important phenological periods; 2) spectral unmixing to more accurately characterize the mixed composition of forests in the study area; and 3) integration of resulting percent basal area maps and ancillary environmental variables into a hierarchical, rule-based classification scheme.

Public availability of Landsat and FIA data enable the broad implementation, as well as scalable nature, of this approach. However, it is important to note that this approach hinges upon the user’s ability to obtain high quality (low cloud cover) multi-temporal imagery during key phenological periods, which is often difficult in temperate and mountainous regions. It also requires a robust set of “pure” species plots for use as endmembers in spectral unmixing and calibration of the percent basal area models. This can be difficult for rare and non-dominant species, or those that typically do not form homogeneous stands. A final limitation is the significant amount of remote sensing expertise required to implement the pixel- and object-based workflows.

Accurate, species-specific percent basal area and thematic forest maps provide forest researchers, managers, and policymakers with powerful demographic tools to inform management activities, identify potential ‘hotspots’ for invasive pest/pathogen outbreaks, and inform other large-scale modeling applications (e.g. carbon storage dynamics, forest
fragmentation/conversion, wildlife habitat/movements, etc.). That we were successful in mapping species distributions in the Northeast, given the high spatial heterogeneity of its forests, bodes well for applying this approach in other, less diverse regions. Further, the extensive Landsat archive lends itself to using this approach to investigate spatiotemporal trends in tree species composition, of particular interest given the anticipated effects of climate change on forest demographics.

2.6 Acknowledgements

This research was funded by the Northeastern States Research Cooperative (NSRC) through the US Forest Service Northern Research Station, and the McIntire-Stennis Cooperative Forestry Program through the USDA National Institute of Food and Agriculture. The authors would like to thank several people for their assistance with this project: Shelly Rayback, Anthony D’Amato, Terri Donovan, and Jennifer Santoro for reviewing the manuscript; Noah Ahles and Jarlath O’Neill-Dunne of the UVM Spatial Analysis Lab for eCognition support; and the McIntire-Stennis “Integrated Forest Ecosystem Assessment to Support Sustainable Management Decisions in a Changing Climate” research group for helpful feedback. We are also very grateful to our senior undergraduate field crew: Anna Smith, Benjamin Comai, Monica Johnson, Matthias Sirch, Harry Voelkel, Quinn Wilcox, and William Sherman.

2.7 References

Agarwal, S., L. S. Vailshery, M. Jaganmohan, and H. Nagendra. 2013. Mapping urban tree species using very high resolution satellite imagery: comparing pixel-based and


Bhat, H. S., and N. Kumar. 2010. On the derivation of the Bayesian Information Criterion. School of Natural Sciences, University of California.


CHAPTER 3: SPATIOTEMPORAL PATTERNS AND POTENTIAL DRIVERS OF CHANGING TREE SPECIES ABUNDANCE IN THE NORTHEASTERN US

David Gudex-Cross¹,*, Jennifer Pontius¹,², Paul G. Schaberg², and Alison Adams¹

¹ The Rubenstein School of Environment and Natural Resources, University of Vermont, Aiken Center, 81 Carrigan Drive, Burlington, Vermont 05405
² USDA Forest Service Northern Research Station, Aiken Center, 81 Carrigan Drive, Burlington, Vermont 05405

*Corresponding Author: dgudexcr@uvm.edu (D. Gudex-Cross)
3.1 Abstract

Climate change is projected to alter tree species abundance and distribution in the northeastern United States. To date, studies have focused on climate-sensitive areas (i.e., range boundaries and ecotones) or used site-specific data or coarse spatial models to characterize changes across broad landscapes. Leveraging a novel remote sensing product that estimates species-level abundance (percent basal area), we modeled fine-scale (30m x 30m) spatiotemporal trends in the abundance of eight northeastern US tree species between 1985-2015 and assessed associations with site, environmental, and climate factors across northern New York and Vermont. We detected significant decreases in sugar maple (Acer saccharum, -12%), eastern hemlock (Tsuga canadensis, -11%), balsam fir (Abies balsamea, -5%), and birches (Betula spp., -4%), and increases in American beech (Fagus grandifolia, +9%) and red maple (Acer rubrum, +3%). These changes varied by elevation and showed significant spatial clustering, with associated species often exhibiting opposing trends (e.g., increased red spruce and decreased balsam fir at high elevations). Climate-related metrics (primarily temperature variability and extremes) were commonly associated with abundance changes. The abundance changes we documented largely contradict traditional succession pathways, with significant implications for key species in the region. Though climate metrics were important predictors of abundance trends, their importance varied by species and across space, highlighting the problematic nature of generalizing the impacts of climate change across diverse forests and complex topographies.

**Keywords:** remote sensing, forest dynamics, climate change, northern hardwood forest, spruce-fir forest, landscape ecology, species distribution
3.2 Introduction

Changes in climate have well-documented effects on the structure, function, and resilience of forest ecosystems. Dendroecological and long-term forest inventory data show that warming temperatures, changing precipitation regimes, climate-related disturbances, and extreme climate events can affect tree fitness, mortality, and migratory pathways (Shuman et al. 2009, Pederson et al. 2015). Over the past century and especially since the 1970s, the northeastern region of the United States (US) (‘the Northeast’) has experienced significant seasonal temperature increases, increased heavy precipitation events/flooding, prolonged growing season dry spells, and less snowfall, outpacing predictions from global climate models (Hayhoe et al. 2007). These trends are projected to continue or intensify over the next century (Hayhoe et al. 2008).

Climatic changes are expected to significantly alter suitable habitat conditions and associated competitive dynamics for many northeastern tree species (Iverson and Prasad 1998). For example, increased temperatures and prolonged droughts will likely favor the northern expansion of warm-adapted species (e.g., oaks – *Quercus* spp.) from the south, while others (e.g., maples – *Acer* spp.) retreat to cooler climate refugia (Tang et al. 2012). Already, warmer temperatures have resulted in longer growing seasons and associated changes in leaf phenology (Richardson et al. 2013), which in turn, are expected to alter important components of habitat suitability, including soil moisture/temperature gradients, nutrient cycles, and microclimate (Huntington et al. 2009). Species composition in montane areas of the Northeast, where assemblages occur along elevational gradients strongly regulated by climate (Cogbill and White 1991), is also projected to shift as higher
elevation habitats become more suitable for hardwood species and less favorable for spruce-fir forests (Tang et al. 2012, Wason et al. 2017a). The effects of climate change on pest/pathogen, wildlife populations, and weather-related disturbances could further impact forest composition patterns, with some species more affected than others (Groffman et al. 2012).

Recent studies suggest climatic changes are already impacting northeastern forest composition, particularly in the understory. Woodall et al. (2009) found the mean latitude of seedling abundance was higher than that of mature biomass for several northern hardwood species – an indication of poleward (latitudinal) migration (though see Zhu et al. 2012). Similarly, Fei et al. (2017) found significant poleward and westward (longitudinal) shifts in sapling abundance over the past 30 years, with conifers, birches (Betula spp.), and aspens (Populus spp.) tending poleward and most other deciduous species westward. These westward shifts were more strongly associated with precipitation changes than temperature (Fei et al. 2017).

In the dominant maple-beech-birch forests of the region, several studies have documented recent increases in American beech (Fagus grandifolia) abundance at the expense of maple (Acer spp.) and birch (Duchesne and Ouimet 2009, Pontius et al. 2016, Bose et al. 2017a,b, Wason and Dovciak 2017). Regional studies linked shifts from maple to beech dominance with increased temperature and precipitation (Bose et al. 2017a,b). However, other investigations suggest spatially complex dynamics involving climate, beech bark disease, deer herbivory, and the legacy effects of logging and acid deposition (Pontius et al. 2016, Wason and Dovciak 2017).
While site-specific field studies provide invaluable information on forest composition and structure, their spatial constraints limit broader examinations of drivers of change across complex landscapes. Remote sensing offers a complementary approach that provides continuous, landscape coverage over several decades and has been used to examine shifts across broad forest types, particularly in montane ecotones (transition zones) where high-elevation red spruce (*Picea rubens*) and balsam fir (*Abies balsamea*) converge with northern hardwood stands dominated by sugar maple (*A. saccharum*), beech, and yellow birch (*B. alleghaniensis*). Beckage et al. (2008) reported upward shifts in northern hardwoods at the expense of spruce-fir on two peaks in Vermont (VT). Conversely, Vogelmann et al. (2012) and Foster and D’Amato (2015) found the spruce-fir ecotone moved predominantly downslope in VT and NH. Recent field studies by Wason and Dovciak (2017) and Wason et al. (2017a) indicated that these downslope shifts were due to increases in red spruce at mid-low elevations, likely driven by climate change and recovery from logging and acid deposition-induced declines.

Traditionally, remote sensing analyses of species composition patterns have been constrained by the detail and accuracy of forest mapping products (typically limited to broad forest type classifications). The recent development of an improved species-level abundance mapping technique by Gudex-Cross et al. (2017) allows for a more detailed examination of spatiotemporal changes across the landscape. Here, we utilize these products to examine 30-year (1985-2015) trends in the abundance of several dominant tree species across northern New York (NY) and VT. For each species, our objectives were to: 1) examine regional trends in mean abundance, 2) detect spatiotemporal patterns in
abundance trends, and 3) identify abiotic correlates (i.e., climate, soil, acid deposition, or topographic factors) associated with abundance trends. This information is essential to understanding if the Northeast is experiencing unexpected changes in species abundance, and to what extent climate and other factors are involved. Furthering our understanding of northeastern forest dynamics as climate changes is critical to informing adaptive management strategies for this ecologically, economically, and culturally valuable natural resource.

3.3 Methods

3.3.1 Study area and focal species

This study was conducted across northern NY and VT using Landsat imagery that covers the Lake Champlain valley and significant portions of the Adirondack and Green mountain ranges (Fig. 3.1). Although the imagery extends into Canada, the study area was limited to the US due to the spatial constraints of ancillary data layers. This region features a wide range of forest composition, soil types, latitudinal and elevational climate gradients where shifts in species composition are expected in response to ongoing climate change (Iverson and Prasad 1998).
We examined eight common species/genera: balsam fir, red spruce, sugar maple, red maple (Acer rubrum), American beech, birches, eastern hemlock (Tsuga canadensis), and oaks. While additional species were of ecological interest, abundance maps were unavailable due to a lack of viable endmember training sites (e.g., ash and hickory species) or low overall mapping accuracy (white pine and aspens). Oaks and birches were modeled at the genus level because multiple species were combined to obtain adequate training sites.
### 3.3.2 Percent basal area (% BA) mapping

Our analysis is based on a remote sensing technique developed by Gudex-Cross et al. (2017) that uses forest inventory data and multitemporal (seasonal) Landsat imagery to model species-level abundance at four different 10-year timesteps (1985-2015). For each species, spectral unmixing techniques with endmembers comprised of pure stands (>80% BA) were used to quantify the likelihood of occurrence and infeasibility (likelihood of a false positive) for each 30m x 30m pixel. A % BA model based on these values was then calibrated using independent field data. Model accuracies rivaled those derived from hyperspectral sensors, with $r^2$ values ranging from 0.32 (hemlock) to 0.60 (beech) and root mean squared errors from 5 (oaks) to 17% BA (sugar maple).

Due to the inherent error in remote sensing estimates, several quality control steps were employed to minimize error in subsequent abundance change modeling. Pixels predicted with less than 5% BA were considered below “detection level” and masked. Any pixel with a single-year abundance estimate more than two standard deviations from its overall 30-year mean was also removed to minimize errors introduced by cloud effects, registration errors, or severe disturbance (e.g., logging, windthrow, etc.). This approach allowed us to focus on abundance changes resulting from natural succession.

### 3.3.3 Spatiotemporal patterns in species abundance trends

For each species, mean % BA was quantified across the full region and compared over the four timesteps to assess broad-scale changes in abundance. Abundance trends were then calculated on a per-pixel basis as the slope of the best fit linear regression line over
the 30-year period. One sample t-tests were used to determine if trends were significantly different from zero (p ≤ 0.01). Trends were evaluated by elevation class (low = <750 meters, transition zone = 750-900m, and high = >900m) based on work specific to the Adirondack and Green mountain ranges (Cogbill and White 1991). Most of the study area fell in the low elevation class (~90%), followed by high (~8%) and transition zone (~2%).

The Optimized Hot Spot Analysis tool in ArcGIS 10 (ESRI 2016) was used to identify spatiotemporal patterns in abundance trends. This tool uses the Getis-Ord statistic (Getis and Ord 1992) to identify areas with significant spatial clustering of increasing (hot) or decreasing (cold) abundance (evaluated at 95% confidence level), which was assessed for pixels with ≥5% BA change over the 30-year period.

3.3.4 Relationships between species abundance trends and abiotic factors

To examine potential drivers of abundance while minimizing spatial autocorrelation, collinearity, and heteroscedasticity in statistical analyses, we extracted a buffered (800m) random sample of pixels stratified by elevation and ecoregion (EPA Level IV, www.epa.gov/eco-research/ecoregions), resulting in 4,507 observations distributed across the study area. These random observations were used to examine several potential abiotic correlates with species abundance trends, including climate and soil metrics, acid deposition (assessed as nitrogen/sulfur critical load and exceedance), and topographic factors (Table 3.1). Climate data included 30-year monthly temperature and precipitation normals (averages from 1981-2010; 800m resolution across the region) obtained from the PRISM Climate Group (accessed 20 Jan 2017), from which we derived ecologically-
relevant indices to reduce collinearity and capture seasonal climate variability (O'Donnel and Ignizio 2012). Winter months included December, January, and February; summer months included June, July, and August. Factors related to landscape position included elevation and heat load index (HLI) (McCune and Keon 2002), which is a measure of potential direct incident solar radiation based on latitude, slope, and aspect.

Table 3.1. Abiotic variables examined for potential relationships with species abundance trends.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Resolution (m)</th>
<th>Raw Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Climate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual mean temperature</td>
<td></td>
<td>PRISM Climate Group: prism.oregonstate.edu</td>
</tr>
<tr>
<td>Annual temperature range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diurnal temperature range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature isothermality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature seasonality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum summer temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum winter temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean summer temperature</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>Mean winter temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual total precipitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total precip. of wettest month (July)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total precip. of driest month (Feb)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precipitation seasonality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total summer precipitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total winter precipitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Landscape Position</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>30</td>
<td>USGS National Map Viewer</td>
</tr>
<tr>
<td><strong>Soils</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth to Bedrock</td>
<td>30</td>
<td>NRCS Web Soil Survey</td>
</tr>
<tr>
<td>Water-holding Capacity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porosity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Acid Deposition</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical Load Threshold</td>
<td>30</td>
<td>Miller (2011)</td>
</tr>
<tr>
<td>Exceed. of Nitrogen and Sulfur 1984-1988</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exceed. of Nitrogen and Sulfur 1999-2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exceed. of Nitrogen and Sulfur 1984-2003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Spatial modeling efforts included several data reduction steps to minimize collinearity among potential input variables using stepwise partial least squares regression (PLS), which is more robust to collinearity than ordinary least squares (OLS) (Carrascal et al. 2009). We used a high variable importance in projection threshold ($\geq 1$) and held back one-third of the data for independent model validation, providing a conservative assessment of significance for input variables used in subsequent spatial regression models.

We employed a spatial regression with maximum likelihood estimation (MLE) workflow from Anselin (2004) to examine relationships between abundance trends and significant factors identified in the PLS. An exploratory ordinary least squares (OLS) global regression model was constructed to eliminate any variables that still exhibited significant collinearity (using a conservative variable inflation factor of $> 5$). Next, we calculated the spatial weights between data points in GeoDa (Anselin et al. 2006) using a minimum Euclidean distance threshold. GeoDa leverages these weights to provide useful diagnostic tests for spatial dependency and heteroscedasticity, which help determine the appropriate spatial regression technique: spatial autoregressive (SAR) or geographically-weighted regression (GWR) (Brunsdon et al. 1998). SAR models account for spatial dependencies by incorporating a spatial lag term ($Rho$) into the regression equation (Anselin et al. 2006). GWR deals with heteroscedastic data by conducting local regressions over a moving window and identifying where locally-weighted coefficients move away from their global values (Brunsdon et al. 1998).

Without significant spatial dependence or heteroscedasticity, the global OLS model was used. A SAR model was used if spatial dependence was significant, but
heteroscedasticity was not. If heteroscedasticity was significant, we ran GWR using GWR4 software (Nakaya et al. 2014). To better understand the regional variation in GWR model explanatory power and abiotic factor significance, we mapped the spatial distribution of local $r^2$ values and parameter coefficient t-statistics using two-tailed thresholds ($p < 0.05$). In all cases, the spatial regression model for a given species was compared against the global OLS model using the minimum corrected Akaike’s Information Criterion (Hurvich and Tsai 1993).

3.4 Results

3.4.1 Global trends

Sugar maple was by far the most abundant species across the study area (Table 3.2). Overall mean % BA estimates were relatively consistent for all species, with less than 2% change on average between each timestep (6% max change) and less than 3.4% change on average (6% max change) across the full period. The largest overall 30-year abundance increases were observed for beech, up 6% across the study area, and the largest decreases were in hemlock and sugar maple (down 5% and 6% respectively). These global changes in mean abundance largely matched the mean long-term trends (slope) obtained from the pixel-based random sample. Beech exhibited the highest positive 30-year trend (up 0.3% yr$^{-1}$; 9% net 30-yr gain), driven primarily by a large increase between 2005-2015 (Table 3.2). Red maple abundance also significantly increased (up 0.1% yr$^{-1}$; 3% net 30-yr gain). Negative trends were highest for birches, hemlock, and sugar maple (down 0.13%, 0.35%,
and 0.39% yr\(^{-1}\) respectively), resulting in estimated net 30-year losses of 3.9%, 10.5%, and 11.7% (Table 3.2).

Table 3.2. Mean regional abundance over time. Trend data is reported as the mean of the per-pixel regression slopes for a subset of buffered, random pixels. Significant values are in bold (p<0.01).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>1985</th>
<th>1995</th>
<th>2005</th>
<th>2015</th>
<th>Mean Ann. Change (%)</th>
<th>Mean 30-yr Change (%)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>American beech</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>13</td>
<td>0.30</td>
<td>9.0</td>
<td>349</td>
</tr>
<tr>
<td>Balsam fir</td>
<td>19</td>
<td>19</td>
<td>18</td>
<td>16</td>
<td>-0.18</td>
<td>-5.4</td>
<td>590</td>
</tr>
<tr>
<td>Birches</td>
<td>15</td>
<td>17</td>
<td>16</td>
<td>13</td>
<td>-0.13</td>
<td>-3.9</td>
<td>1,007</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>21</td>
<td>19</td>
<td>16</td>
<td>16</td>
<td>-0.35</td>
<td>-10.5</td>
<td>900</td>
</tr>
<tr>
<td>Oaks</td>
<td>12</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>-0.06</td>
<td>-1.8</td>
<td>243</td>
</tr>
<tr>
<td>Red maple</td>
<td>10</td>
<td>14</td>
<td>11</td>
<td>13</td>
<td>0.10</td>
<td>3.0</td>
<td>435</td>
</tr>
<tr>
<td>Red spruce</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>-0.05</td>
<td>-1.5</td>
<td>549</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>39</td>
<td>36</td>
<td>32</td>
<td>33</td>
<td>-0.39</td>
<td>-11.7</td>
<td>1,441</td>
</tr>
</tbody>
</table>

3.4.2 Elevational trends

Several species exhibited significant elevational abundance trends (p<0.01) that differed from their overall regional trend (Table 3.3). For example, red spruce significantly increased at upper elevations (0.28% yr\(^{-1}\)) with more modest increases in the transition zone (0.12% yr\(^{-1}\)). These increases were masked in the global trend by significant decreases in the more widespread lower elevations (-0.17% yr\(^{-1}\)). Similarly, significant decreases in birch abundance in transition zones and upper elevations (-0.46% and 0.50% yr\(^{-1}\) respectively) were tempered by moderate increases at lower elevations (0.09% yr\(^{-1}\)). The regional decrease in balsam fir abundance was driven primarily by significant losses at high elevation (-0.71% yr\(^{-1}\)), with insignificant abundance trends in the other two elevation classes (Table 3.3).
Table 3.3. Abundance trends by elevation zone reported as the mean (% basal area) of the per-pixel regression slopes for a subset of buffered, random pixels. Significant values are in bold (p<0.01).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Low (&lt;750m)</th>
<th>Transition (750-900m)</th>
<th>High (&gt;900m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Ann. Change (%)</td>
<td>Mean 30-yr Change (%)</td>
<td>N</td>
</tr>
<tr>
<td>American beech</td>
<td>0.29</td>
<td>8.7</td>
<td>293</td>
</tr>
<tr>
<td>Balsam fir</td>
<td>-0.08</td>
<td>-2.4</td>
<td>320</td>
</tr>
<tr>
<td>Birches</td>
<td>0.09</td>
<td>2.7</td>
<td>625</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>-0.35</td>
<td>-10.5</td>
<td>898</td>
</tr>
<tr>
<td>Oaks</td>
<td>0.01</td>
<td>0.3</td>
<td>180</td>
</tr>
<tr>
<td>Red maple</td>
<td>0.11</td>
<td>3.3</td>
<td>408</td>
</tr>
<tr>
<td>Red spruce</td>
<td>-0.17</td>
<td>-5.1</td>
<td>335</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>-0.41</td>
<td>-12.3</td>
<td>1,130</td>
</tr>
</tbody>
</table>

*N<50.
3.4.3 Spatiotemporal patterns

Exploring spatial patterns in the 30-year trends revealed significant clusters of increasing and decreasing abundance (Fig. 3.2). For each species, these clusters largely complemented patterns identified in the examination of elevation classes. Interestingly, compatriot species often exhibited opposing trends in the same areas. The most pronounced examples include co-located clusters of decreasing sugar maple (Fig. 3.2a) and increasing beech (Fig. 3.2c) and decreasing balsam fir (Fig. 3.2e) with increasing red spruce (Fig. 3.2g). Clusters for sugar maple-beech changes were concentrated throughout the Adirondacks, while virtually all the red spruce-balsam fir clusters occurred in two Adirondack ecoregions: Upper Montane/Alpine Zone and High Peaks. In contrast, both red spruce and balsam fir exhibited increasing abundance in the Green Mountain ecoregions of north-central VT.

The distribution of abundance clusters for sugar maple, red maple, and beech suggests a potential relationship with longitude (Fig. 3.2a, b, c). For maples, decreasing clusters were concentrated in the western portion of the region (NY) and increasing clusters in the east (VT). The opposite was true for beech. Given the significant decreasing trend of sugar maple across the study area, the NY losses clearly outweighed the VT gains. There were also stark differences in the distribution of birch abundance clusters throughout montane areas, with a clear pattern of decreasing to increasing abundance from south to north. Hemlock and balsam fir also showed abundance clustering with latitudinal relationships (Fig. 3.2e, f), with hot spots distributed throughout the northern portions of the study area and cold spots concentrated in the south.
Figure 3.2. Spatial distribution of 30-yr abundance increases (red) and decreases (blue) for eight tree species/genera across ecoregions. Confidence levels represent the significance of spatial clustering.
3.4.4 Correlates with global species abundance trends

Across the entire study area, climate-related indices were important predictors of 30-year abundance trends for seven of the eight species/genera (Table 3.4). These included minimum temperatures and measures of climate variability (e.g., temperature ranges). Temperature-related metrics were more common correlates than precipitation. Only the oak model excluded climate metrics, likely due to small sample size.

Larger temperature ranges were positively associated with beech and birch abundance trends and negatively associated with trends in red spruce abundance. Diurnal temperature range (an index of daily max-min fluctuations) was negatively associated with sugar maple abundance. Lower minimum winter temperatures were negatively associated with beech and balsam fir abundance trends but positively associated with red maple trends. Higher HLI values were associated with positive abundance trends for beech and sugar maple, indicating better growth conditions on warmer slopes, but negatively associated with balsam fir and eastern hemlock trends. Precipitation metrics generally reported increasing abundance trends with increasing precipitation for red spruce, sugar maple, and eastern hemlock. Elevation was positively associated with abundance trends for red spruce and negatively associated with those for balsam fir, oaks, and birches (Table 3.4).
Table 3.4. Significant abiotic factors and their effect on 30-year abundance trends. Rho is the spatial lag term in spatial autoregressive (SAR) models (OLS = global Ordinary Least Squares model).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Model Terms</th>
<th>Effect</th>
<th>OLS p-value</th>
<th>SAR p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>American beech</td>
<td>Rho</td>
<td>0.29</td>
<td>--</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>Minimum winter temperature</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>Annual temperature range</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Heat load index</td>
<td>+</td>
<td>0.005</td>
<td>0.006</td>
</tr>
<tr>
<td>Balsam fir</td>
<td>Heat load index</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Minimum winter temperature</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td>Birches</td>
<td>Elevation</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Annual temperature range</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>Rho</td>
<td>0.56</td>
<td>--</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Heat load index</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Total winter precipitation</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Total summer precipitation</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td>Oaks</td>
<td>Elevation</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>--</td>
</tr>
<tr>
<td>Red maple</td>
<td>Rho</td>
<td>0.37</td>
<td>--</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Minimum winter temperature</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Red spruce</td>
<td>Elevation</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Total annual precipitation</td>
<td>+</td>
<td>0.003</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Annual temperature range</td>
<td>-</td>
<td>0.01</td>
<td>*</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>Heat load index</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Diurnal temperature range</td>
<td>-</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Total summer precipitation</td>
<td>+</td>
<td>&lt;0.0001</td>
<td>*</td>
</tr>
</tbody>
</table>

3.4.5 Correlates with species abundance trends – spatial regressions

For seven of the eight species/genera, spatial regression models produced substantially better fit than global OLS (Table 3.5), indicating spatially-variable relationships between potential abiotic drivers and abundance trends. Due to high heteroscedasticity, GWR models were used for sugar maple, birches, red spruce, and balsam fir. SAR produced the best fit model for hemlock, beech, and red maple; global OLS was only sufficient for oaks (Table 3.5).
Table 3.5. Final regression model selection and fit statistics.

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Model Type</th>
<th>OLS* Adj. r²</th>
<th>SAR* Pseudo r²</th>
<th>GWR* Adj. r²</th>
<th>OLS AICc</th>
<th>Spatial AICc</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>American beech</td>
<td>SAR</td>
<td>0.12</td>
<td>0.17</td>
<td>--</td>
<td>146.1</td>
<td>136.6</td>
<td>338</td>
</tr>
<tr>
<td>Balsam fir</td>
<td>GWR</td>
<td>0.24</td>
<td>--</td>
<td>0.34</td>
<td>1,036.2</td>
<td>984.9</td>
<td>547</td>
</tr>
<tr>
<td>Birch</td>
<td>GWR</td>
<td>0.26</td>
<td>--</td>
<td>0.48</td>
<td>1,401.9</td>
<td>1,103.0</td>
<td>976</td>
</tr>
<tr>
<td>Eastern hemlock</td>
<td>SAR</td>
<td>0.09</td>
<td>0.18</td>
<td>--</td>
<td>1,468.5</td>
<td>1,401.4</td>
<td>805</td>
</tr>
<tr>
<td>Oaks</td>
<td>OLS</td>
<td>0.17</td>
<td>--</td>
<td>--</td>
<td>190.8</td>
<td>--</td>
<td>233</td>
</tr>
<tr>
<td>Red maple</td>
<td>SAR</td>
<td>0.08</td>
<td>0.13</td>
<td>--</td>
<td>115.4</td>
<td>100.1</td>
<td>435</td>
</tr>
<tr>
<td>Red spruce</td>
<td>GWR</td>
<td>0.28</td>
<td>--</td>
<td>0.36</td>
<td>374.0</td>
<td>322.3</td>
<td>502</td>
</tr>
<tr>
<td>Sugar maple</td>
<td>GWR</td>
<td>0.14</td>
<td>--</td>
<td>0.25</td>
<td>3,122.8</td>
<td>2,975.1</td>
<td>1,372</td>
</tr>
</tbody>
</table>

*OLS = Ordinary Least Squares; SAR = Spatial Autoregressive; GWR = Geographically-weighted Regression

GWR revealed significant spatial variation in model explanatory power (local r² values from ~0 to 0.75) and abiotic factor coefficient significance for sugar maple and birches (Fig. 3a, b). The sugar maple model performed best predicting positive abundance trends in the Green Mountains, but poorly when predicting negative abundance trends in the Adirondack Foothills. The birch model followed a similar spatial pattern, but yielded higher explanatory power across the landscape, primarily because of close associations with elevation. Clustering of significant positive and negative coefficients indicated temperature and precipitation have localized effects on sugar maple populations, while the influence of landscape position (HLI) was more spatially consistent. Similarly, birch model coefficients showed consistent negative relationships between elevation and abundance trends across the study area, but more localized negative relationships with annual temperature range.
(a) Sugar maple

Model Fit

Heat Load Index

Total Summer Precipitation

Diurnal Temperature Range

Local r

0.01 - 0.10

0.10 - 0.20

0.20 - 0.36

Significant neg. coefficient

Not significant

Significant pos. coefficient
Figure 3.3. Regional variation in geographically-weighted regression model fit and parameter significance for sugar maple (a) and birch species (b).
Red spruce (Fig. 3.4a) and balsam fir (Fig. 3.4b) GWR models also revealed substantial spatial variation in explanatory power and coefficient significance, with local $r^2$ values ranging from 0.03 to 0.60. The highest local $r^2$ values for red spruce were distributed throughout VT and most of the Adirondack Foothills regions, with the lowest in the Adirondack High Peaks/Alpine Zone and St. Lawrence Valley (Fig. 3.4a). Annual temperature range had a predominantly negative association with red spruce abundance trends in the northern portion of the study area, but a positive association in parts of the Adirondack Foothills and areas of the Lake Champlain Basin. Significant coefficients for annual precipitation and elevation were consistently positive, with precipitation being an important predictor of increased red spruce abundance in the Lake Champlain Basin and elevation in the southern and western parts of the study area. The highest local $r^2$ values for balsam fir were in the Adirondack High Peaks and Central Green Mountains; the lowest were found in the northern part of the Adirondack Foothills (Fig. 3.4b). The areas with the best model fit typically contained significant negative relationships between fir abundance, elevation, and HLI.
(a) Red spruce

Model Fit

Annual Total Precipitation

Annual Temperature Range

Elevation

Local r

-0.07 - 0.20
-0.20 - 0.40
-0.40 - 0.57

- Not significant
- Significant pos. coefficient
- Significant neg. coefficient
Figure 3.4. Regional variation in geographically-weighted regression model fit and abiotic parameter significance for red spruce (a) and balsam fir (b).
3.5 Discussion

3.5.1 Spatiotemporal patterns

Some of the species abundance changes we documented are consistent with changes expected in the Northeast from natural succession. These include a general increase in shade-tolerant species such as beech – historically dominant throughout the study area (Thompson et al. 2013) – and, to a lesser extent, red maple. Several recent northeastern studies have reported similar beech gains (e.g., Pontius et al. 2016, Bose et al. 2017a,b, Wason and Dovciak 2017). Our results show increases in beech were more pronounced over the last decade, which coincide with findings by Bose et al. (2017a, b) and Pontius et al. (2016). Also consistent with Bedison et al. (2007) and Bose et al. (2017b), spatial analysis of the beech trend data revealed hot spots of increasing abundance in the Adirondacks throughout its elevational range that were often co-located with sugar maple decreases, suggesting beech is directly benefiting from sugar maple decline. Red maple growth and abundance increases have been reported in VT (Kosiba et al. 2017), NH (Pontius et al. 2016), and across the broader eastern region (Fei and Steiner 2007).

Contrary to traditional patterns towards shade-tolerant, mid-late successional species, our results show significant regional decreases in sugar maple and hemlock abundance. Given its importance to the region, sugar maple has been intensively studied in the Northeast, with reported declines beginning in the mid-20th century (e.g., Bishop et al. 2015, Horsley et al. 2002). Our spatiotemporal analyses show decreasing sugar maple abundance concentrated in low-mid elevation forests of the Adirondacks (Fig. 2a), where others have reported unexpected growth declines across age and diameter classes despite
favorable growing conditions (warmer temperatures and increased precipitation) and reduced acid deposition rates over the recent past (Bedison et al. 2007, Sullivan et al. 2013, Bishop et al. 2015). However, these areas experienced significant canopy defoliation events from a combination of drought-like conditions and forest tent caterpillar outbreaks in the early-mid 2000s (Wood et al. 2009), which were likely reflected in our results (sugar maple mean % BA was lowest in 2005; Table 3.2). For hemlock, the decline reported here is a novel result for northern NY and VT, where the hemlock woolly adelgid – a widespread invasive pest that has recently invaded New England – has not yet been detected. The distinct north-south spatial distribution of hemlock trends and correlation with climate metrics suggest a relationship mediated by climate (see “Eastern hemlock” section below).

Our spatiotemporal analyses showed that shifting compositional patterns in high elevation spruce-fir-birch forests were more pronounced in the transition zone and higher elevation ecoregions of the Adirondacks (Fig. 2e, g). Birch declines in these locations likely reflect losses for two species, yellow birch (transition zone) and montane paper birch (transition and high zones). Other studies in the region have recorded similar birch losses and attributed them to either natural succession (i.e., age-related mortality) (Van Doorn et al. 2011) or a predisposition to freeze/thaw-related injuries caused by acid deposition (e.g., Halman et al. 2015). The high elevation balsam fir declines may be due to a combination of wind exposure and climate events known to cause mortality (Sprugel and Bormann 1981), though other nearby studies have documented losses to spruce budworm in previous decades (Filion et al. 2006).
Transition zone dynamics

The spatiotemporal abundance patterns in the transition zone provide additional insight into conflicting reports of altitudinal shifts where northern hardwood and spruce-fir forests converge. Given temperature is one of the main constraints on species distributions along elevation gradients, continued warming is expected to promote the upslope movement of northern hardwoods at the expense of more cold-adapted conifers (Tang et al. 2012). In contrast, our results show significant declines for several northern hardwood species and hemlock, with only beech and red spruce increasing in abundance in the transition zone. However, it is important to note these abundance shifts varied across the study area, particularly in the northern Green Mountains where some northern hardwood species exhibited clusters of increasing abundance.

Many of our findings are corroborated by recent studies of species demography. Using remote sensing, both Vogelmann et al. (2012) and Foster and D’Amato (2015) observed recent downslope movement in spruce-fir forests, likely driven by red spruce expansion. Similar to our results, the latter study also documented areas exhibiting upward shifts or no apparent change in northern hardwoods. Across montane northeastern forests, Wason and Dovciak (2017) compared the basal area distributions of several species at field sites spanning from low elevation northern hardwood to high elevation spruce-fir and found: 1) expansion of beech (upslope) and red spruce (downslope); 2) contraction of sugar maple (downslope); and 3) no change in balsam fir. Like our study, they found these shifts in abundance were predominantly correlated with climatic factors, though historical logging practices were also an important factor. Collectively, these studies indicate that
while shifts in species abundance have occurred in the transition zone, responses have been species-specific and variable across the landscape.

### 3.5.2 Abiotic factors

Though we analyzed many potential site, soil, and deposition factors, climate indices were most commonly associated with species abundance trends. Temperature variability, minimum winter temperatures, and HLI were consistent correlates across almost all species/genera, whereas precipitation metrics were important for known drought-sensitive species (e.g., hemlock). Elevation was a significant predictor of abundance for many species, perhaps capturing the interacting effects of climate, soil, and deposition levels known to vary with elevation. Climate metrics typically accounted for ~10-25% of the total variability in species abundance trends, supporting other recent studies that found climate to be an important driver of northeastern forest composition (Bose et al. 2017a,b, Fei et al. 2017, Wason and Dovciak 2017).

Nonetheless, climatic factors alone largely failed to predict the widespread decreases in sugar maple abundance and shifts from balsam fir to red spruce concentrated in the Adirondacks. We suspect these localized decreases may be due to a complex set of factors that include climatic change, poorer soils, and the legacy effects of higher acid deposition. For sugar maple, different land use history (less conversion to agriculture vs. sugar maple-dominated stands in VT) and pest/pathogen outbreaks (e.g., Wood et al. 2009) have also likely affected competition with beech (see Bishop et al. 2015, Lawrence et al. 2017 for further discussion). Although deposition and soil pH were not associated with
abundance changes, the low range of values in these data layers likely limited the statistical power of our analyses. Indeed, considerable research has shown that acid deposition has strongly influenced forest composition in our study area (Schaberg et al. 2001, Driscoll et al. 2003). Nitrogen and sulfur deposition in forest soils depletes calcium, resulting in deficiencies that impair tree stress and carbon regulation pathways, predisposing trees to decline following exposure to other stressors (Schaberg et al. 2001). Sugar maple is particularly prone to calcium deficiency-related declines in health, growth, and competitive status that are often associated with synchronous increases in beech (Bishop et al. 2015, Halman et al. 2015). However, like Bishop et al. (2015), our results indicate sugar maple declines (and associated beech increases) in the Adirondacks span deposition and elevational gradients, indicating additional factors (e.g., climate, pests/pathogens, etc.) may be influencing this relationship across broader spatial scales.

That climate-related factors were predominant predictors of species abundance trends is expected in a region where forest productivity is primarily limited by climate. However, given the diversity of landscapes, forest types, and land use histories contained in our study area, spatial modeling also revealed complex spatial variation in the significance of these factors. Next, we examine these relationships on a species-specific basis, excluding the potentially unreliable models for red maple (adj. $r^2 < 0.15$) and red oak ($N < 300$).

**American beech**

Constrained carbon relations and extreme cold damage likely explain the negative association with minimum winter temperatures (Table 3.4). Both impair physiological
function and growth (thereby affecting an individual’s competitive status): low winter temperatures induce deep physiological dormancy, leading to a shorter functional growing season, and extreme cold can cause significant canopy dieback events (Melancon and Lechowicz 1987, Sakai and Weiser 1973). Dendroecological studies have shown minimum winter temperature is also the main limitation on beech’s relative abundance and northern range extent (e.g., Huntley et al. 1989). This likely explains the positive association between beech abundance trends and HLI (Table 3.4), since carbon relations and competitive success would benefit in warmer areas. In these areas, favorable growth conditions may also counteract the deleterious effects of beech bark disease (BBD), which is widespread throughout the study area (e.g., Munck and Manion 2006, Pontius et al. 2016). The positive association with annual temperature range may be an extension of this effect and was corroborated by a recent study of beech sapling abundance relative to other hardwoods in the Northeast (Bose et al. 2017b).

**Sugar maple**

Sugar maple is also at the northern extent of its geographic range, which likely explains the positive association between HLI and sugar maple abundance (Table 3.4). Similarly, positive associations with total summer precipitation are consistent with its preference for moist habitats and drought sensitivity (Horsley et al. 2002). Negative associations with increased temperature fluctuations (diurnal range) likely reflects locations with the highest probability of spring freeze/thaw damage (Halman et al. 2013). However, mapping the diurnal temperature coefficients revealed local significance that often ran counter to the global coefficient (Fig. 3.3a), suggesting an interaction with other
factors. Dendrochronological analyses of the relationship between sugar maple growth and climate variability in NY (Bishop et al. 2015) and VT (Gavin et al. 2008) have recorded similar non-stationarities, confirming differential climate responses among local populations.

**Red spruce**

The positive association between red spruce abundance and elevation (Table 3.4) likely reflects a combination of recovery from historic declines from acid deposition and logging, in addition to an increasing competitive advantage over balsam fir with warming temperatures. The competitive advantage of red spruce is related to its ability to photosynthesize throughout the year, including the fall, winter, and spring when cold-hardy balsam fir is functionally dormant (DeHayes et al. 2001). Red spruce abundance was also positively associated with total annual precipitation, though this trend mainly applied to spruce populations at low elevations near Lake Champlain – the warmest part of our study area (Fig. 3.4a). We suspect this localized effect may reflect an interaction between warmer temperatures and increased precipitation that have been shown to favor spruce growth elsewhere (Burns and Honkala 1990a).

Red spruce’s negative relationship with annual temperature range (Table 3.4) likely builds upon its documented sensitivity to freezing injury that is exacerbated by acid deposition (DeHayes et al. 2001). Precocious dehardening during winter/spring freeze-thaw cycles (reflected in annual temperature ranges) increases the risk of freezing injury, crown dieback, and mortality (Schaberg and DeHayes 2000). Further supporting this link,
the temperature range-spruce abundance relationship we observed was concentrated in the north where winter cold extremes are more likely (Fig. 3.4a).

**Balsam fir**

We suspect that the negative associations between abiotic factors and balsam fir abundance relate to their effects on spruce-fir competition. Balsam fir is a cold-adapted species that occupies the more southerly portions of its geographic range in our study area. Thus, it is expected that balsam fir abundance would be negatively associated with greater heat stress (i.e., HLI; Table 3.4). For cold-adapted species like balsam fir, net photosynthesis goes down above certain threshold temperatures as respiration rates increase but photosynthesis plateaus/declines. This puts trees at a competitive disadvantage versus more warm-adapted species (e.g., red spruce) (Schaberg et al. 1996).

The negative association between balsam fir abundance and minimum winter temperatures (Table 3.4) is surprising because of its extreme cold hardiness (Strimbeck and Schaberg 2009). However, a consequence of achieving deep cold-hardiness is reduced photosynthetic function. Perhaps especially in the high elevations of the Adirondacks where balsam fir is pushed to be most cold tolerant, it would have a constrained carbon budget relative to its more photosynthetically opportunistic competitor red spruce (Schaberg et al. 1996). We suspect the negative association of balsam fir abundance and elevation is an extension of this dynamic.

**Eastern hemlock**
Because hemlock is generally restricted to regions with cool, humid climates (Burns and Honkala 1990a), the negative association between hemlock abundance and higher HLI (Table 3.4) was expected. The threshold temperatures above which respiration exceeds photosynthetic gain is particularly low for this species (Adams and Loucks 1971), and substantial temperature increases over the past century across the study area (Trombulak and Wolfson 2004) have likely created additional stress for it. Comparisons of the spatial distribution of hemlock abundance clusters (Fig. 3.2f) to temperature gradients (Trombulak and Wolfson 2004) shows a close association between higher temperature increases in the south and decreased hemlock abundance, with an opposite pattern in the north.

That precipitation metrics were significant predictors of hemlock abundance (Table 3.4) was also expected given its drought sensitivity (Burns and Honkala 1990a) and epispodic drought that occurred over the study period (e.g., early 2000s). Along with less dramatic shifts in temperature, the positive association between hemlock abundance and higher summer precipitation levels likely reflect favorable growing conditions remaining constant in the northern part of our study area. In contrast, the physiological connection(s) for the negative association with higher winter precipitation is less intuitive but could reflect indirect influences of snowpacks that govern soil temperature cues, triggering the start of the growing season (Groffman et al. 2012).

**Birches**

Our interpretation of abiotic factor effects on birch abundance is limited by assessment at the genus level, particularly since the three primary birch species (paper,
yellow, and mountain paper) occupy somewhat overlapping habitats along elevational gradients. The only significant correlates included elevation and annual temperature range (Table 3.4). Other regional work has documented declines in mountain paper birch at high elevations, paper birch at mid elevations in VT (Halman et al. 2011), and yellow birch at mid elevations in NH (Halman et al. 2015), which may reflect the legacy effects of acid deposition on growth and mortality. Significant negative coefficients for the relationship between birch abundance and increases in annual temperature range (Table 3.4) were confined to northern VT, which also suggests an interaction with temperature-associated injury (e.g., spring freeze/thaw events; Halman et al. 2011).

3.5.3 Links to Climate Change Tree Atlas projections

To better understand the potential effects of climate change on northeastern forests, the USDA Forest Service developed models of projected changes in suitable habitat for 134 tree species under several climate scenarios (‘Climate Change Tree Atlas’; Prasad et al. 2007, Iverson et al. 2008). These models incorporate abiotic factors related to climate, elevation, soil, and land use practices and link them to climate projections to create maps of favorable habitat under various climate scenarios. According to these projections, almost every species we evaluated is expected to decline in the Northeast except for oaks and red maple (under low emissions scenarios only).

Our data indicate that the abundance declines projected for sugar maple, hemlock, balsam fir, and birch species may already be underway. For these species, important abiotic predictors consistent between Tree Atlas models and ours include: landscape position and
summer precipitation totals for sugar maple; warmer temperatures and precipitation for hemlock; warmer temperature and elevation for balsam fir; and elevation for birches. Our results also provide some evidence in support of: 1) projected increases in red maple abundance; 2) projected poleward shifts in the abundance of hemlock and balsam fir; and 3) projected westward shifts in beech (see Fei et al. 2017 for further evidence of 2 and 3).

3.5.4 Study limitations

As a remote sensing assessment, the results of our study only reflect compositional changes in the overstory forest canopy. The inability to characterize understory species composition patterns is particularly limiting considering early changes are often first evident in juvenile age classes. Remote sensing assessments are also prone to image registration errors, cloud edge effects, and confusion among similar species. However, our confidence in the abundance trends we report is bolstered by our rigorous QAQC protocols and the consistency between our results and those reported by recent field studies.

Our spatial analyses of abiotic correlates were limited by the coarse resolution of some of the input layers. While our species abundance maps were modeled at a 30m resolution, inputs such as climate (800m), soils (low value ranges), and pollution critical load exceedance (low temporal resolution) were much less detailed, masking the fine-scale variability inherent across the study area. Though our results can be used to inform which climatic factors may be most closely linked with species demographics, it is important to note that we did not test 30-year changes in the climate metrics. Nonetheless, our results consistently associated climatic factors with species abundance changes in northeastern
forests, and by extension suggest future climatic changes will continue to shift competitive dynamics.

3.6 Conclusions

This study shows that some changes in northeastern tree species abundance have been consistent with expected successional pathways, but for others climate, sometimes mediated by other abiotic factors (e.g., landscape position), may have altered competitive relationships among dominant species. Overall, four species/genera (sugar maple, birches, balsam fir and hemlock) showed significant declines over time, and three (beech, red maple and red spruce) experienced general or localized increases in abundance. Local abundance patterns among compatriot species often contrasted one another, with notable examples including beech increases co-located with sugar maple decreases, and red spruce increases co-located with balsam fir decreases.

Species abundance trends were most closely aligned with temperature and precipitation gradients, indicating climate change will likely influence future compositional patterns. However, our hot spot analyses and spatial regression models also suggest that species responses are typified by localized relationships that vary across the landscape. These results demonstrate that generalizing climate impacts across species and large regions with complex topographies can be problematic, providing support for fine-scale niche mapping of species as climate continues to alter regional forest dynamics.
3.7 Acknowledgements

This research was funded by the Northeastern States Research Cooperative through the USDA Forest Service Northern Research Station (Grant #14-DG-11242307-087), and the McIntire-Stennis Cooperative Forestry Program through the USDA National Institute of Food and Agriculture (Grant #2017-32100-06050). The authors would like to thank the McIntire-Stennis “Integrated Forest Ecosystem Assessment to Support Sustainable Management Decisions in a Changing Climate” research group at UVM for helpful feedback on the manuscript.

3.8 References


CHAPTER 4: RECENT PHENOLOGICAL CHANGE IN NORTHERN HARDWOOD FORESTS AND POTENTIAL DRIVERS ACROSS THE NORTHEASTERN US

David Gudex-Cross1, *, Jennifer Pontius1, 2, and Paul G. Schaberg2

The Rubenstein School of Environment and Natural Resources, University of Vermont,
Aiken Center, 81 Carrigan Drive, Burlington, Vermont 05405

2USDA Forest Service Northern Research Station, Aiken Center, 81 Carrigan Drive,
Burlington, Vermont 05405

*COrresponding Author: dgudexcr@uvm.edu (D. Gudex-Cross)
4.1 Abstract

Climate change has been linked to well-documented shifts in important phenological events like the start and end of the growing season (SOS/EOS) in deciduous forests across the Northern Hemisphere. Yet regional variability and the role of climate at this scale are less understood. We addressed these issues by examining spatiotemporal trends in SOS/EOS and their potential drivers across northern hardwood forests in the northeastern United States (US). Using MODIS-derived phenology metrics from 2001-2015 to quantify trends over time at a 250-meter pixel level, we found slight regional advances in SOS that were characterized by a clear longitudinal pattern: eastern ecoregions experienced earlier SOS and western ecoregions had delayed SOS. Conversely, EOS trended significantly later across most of the study area. Modeling the SOS/EOS trends using spatial regression, we identified several important climatic drivers of phenological change. For SOS, these included significantly warmer mean Dec-May temperatures (linked to earlier SOS) and Feb-Mar precipitation totals (variable effect on SOS). Important predictors of EOS trends included elevation and Sep-Nov precipitation anomalies, with both having a negative association (i.e., areas with higher elevation and precipitation exhibited either lower rates of EOS delay or earlier EOS). This research further validates climate change-driven changes in phenology across northern hardwood forests and elucidates their complex spatial relationships in the northeastern US. Our spatial regression models also highlight the importance of considering autocorrelation and heteroscedasticity when evaluating important drivers of phenological change across topographically-diverse regions.

Keywords: remote sensing, MODIS, climate change, landscape ecology, start of season, end of season
4.2 Introduction

Long-term data from field and remote sensing studies show pronounced trends toward earlier spring green-up and later autumn senescence in the northeastern US and much of the Northern Hemisphere (Richardson et al. 2013, Gill et al. 2015), resulting in a shift toward longer growing seasons and changes in forest productivity (Richardson et al. 2009b, Richardson et al. 2010, Keenan et al. 2014; but see Wu et al. 2016). These trends have largely been linked to warming temperatures, though photoperiod cues and extreme precipitation events (e.g., drought) can mediate phenological responses to temperature (Polgar and Primack 2011, Melaas et al. 2016a, Wu et al. 2016). However, several recent studies have revealed that broad-scale changes in phenology exhibit considerable variability at a regional scale, particularly over the past few decades (e.g., Jeong et al. 2011, Xie et al. 2015b, Yue et al. 2015). Given the important role of seasonal canopy dynamics in regulating biogeochemical processes (carbon, nutrient, and water cycling) and physical properties (atmospheric and surface conditions – e.g., albedo) that affect forest structure, composition, ecosystem services, and wildlife habitat (see reviews by Peñuelas and Filella 2009, Polgar and Primack 2011, Richardson et al. 2013), more detailed regional examinations of forest phenology are needed to understand spatial patterns of phenological change and their potential drivers.

The northeastern United States (the ‘Northeast’) is an ideal region for examining the relationship between climate change and key events like the start and end of the growing season (SOS/EOS) because of its relatively dense, largely deciduous forest cover and diverse topography. Here, previous studies using remote sensing methods to estimate
SOS/EOS trends have found inconsistent results (Dragoni and Rahman 2012, Yang et al. 2012, Yue et al. 2015, Melaas et al. 2018), likely due to the varying temporal and spatial scales evaluated. Recent work by Yue et al. (2015) and Keenan et al. (2014) incorporating multiple sources of phenological information, including field data, carbon flux measurements, and satellite imagery, have reconciled some of these inconsistencies: both show later EOS across most deciduous forests of the Northeast and, to a lesser extent, earlier SOS. One major commonality among these and other studies (e.g., Friedl et al. 2014) is that temperature change is the dominant driver of phenological trends. However, the importance of when those temperature changes occur and the influence of other factors like photoperiod, precipitation, and topography remain somewhat unresolved. For example, while recent modeling efforts found spring warming alone can explain most of the variability in SOS trends, others have favored the inclusion of winter temperatures (related to chilling requirements for budburst) and photoperiod limitations (Migliavacca et al. 2012, Yue et al. 2015, Melaas et al. 2016a). Further, while EOS can be impacted by heavy rains and drought (Xie et al. 2015b, Xie et al. 2018), precipitation is generally not considered a significant predictor of SOS (Polgar and Primack 2011).

A persistent challenge in determining significant drivers of phenological change in the Northeast is the high spatial heterogeneity in SOS/EOS dates across the region. This variability, which can occur even over very short distances, is due to diverse patterns of species composition and strong climate gradients governed by latitude, longitude, topography, large bodies of water, and urban heat island effects from large cities surrounded by sparsely populated areas (Zhang et al. 2004, Fisher et al. 2006, Xie et al.
Heterogeneous forests further complicate broad generalizations regarding climatic drivers because phenological responses to climate vary 1) among species with differing physiologies (i.e., ring- versus diffuse-porous wood), 2) among intraspecific populations adapted to local conditions in different parts of a species’ range, and 3) by canopy position and age (Polgar and Primack 2011, Migliavacca et al. 2012). Thus, it is not surprising that climate-based phenology models developed using site-specific data (e.g., at Harvard Forest in Massachusetts and the Hubbard Brook Experimental Forest in New Hampshire) to predict interannual variability in SOS/EOS day of year (DOY) often do not scale well when applied across the broader region; though Melaas et al. (2016a) found stratifying by forest type and incorporating data from other sources (e.g., digital cameras) substantially improved model performance.

An alternative approach to identifying key climate parameters affecting forest phenology that has yet to be used in the Northeast is to explicitly model the linear trend in SOS/EOS dates over time rather than specific DOY (e.g., Wang et al. 2017). Coupling temporal rates of change in phenology with those in climate metrics over the same period can elucidate more direct relationships between them. Also, since climate, soil, and other site characteristics vary across the region, relationships between phenology trends and potential drivers may also vary spatially. Therefore, utilizing spatial statistics can inform how the importance of these drivers varies across the landscape. Here, we employ spatial modeling of phenological trends using MODIS-derived annual estimates of SOS/EOS from 2001-2015 across the dominant forest type within the Northeast – northern hardwood forest
– and address existing gaps in our understanding of landscape-scale climate-phenology relationships with three specific objectives:

1) To quantify regional trends in SOS/EOS from 2001-2015,
2) Examine intraregional variation and spatiotemporal patterns in SOS/EOS trends, and
3) Investigate relationships between SOS/EOS trends and numerous climatic and topographic variables using spatial regression models to explore the possible causes of variations in phenology across the landscape.

4.3 Methods

4.3.1 Study area

This study was conducted across the northern hardwood forests of New York and northern New England, spanning four dominant ecoregions (level III; Omernik 1995): the Eastern Great Lakes Lowlands and Northern Allegheny Plateau in the west, and the Northeastern Highlands and Acadian Plains and Hills in the east (Fig. 4.1). Focusing on the northern hardwood forest type allowed us to examine phenological responses in forests with mixed species compositions (dominated by maple (Acer spp.), beech (Fagus grandifolia), and birches (Betula spp.)) across a diverse range of environmental conditions (i.e., differing edaphic, topographic, and climatic gradients). These include a wide latitudinal range (~40-47°N), several prominent mountain ranges (the Adirondacks and Catskills in New York, Greens in Vermont, and Whites in New Hampshire), and the transition zone between temperate and boreal forests.
4.3.2 Phenology data

To quantify phenology, we acquired annual SOS and EOS DOY estimates from 2001 (first year available) to 2015 from the “expedited” Moderate Resolution Imaging Spectrometer (eMODIS) database provided by the US Geological Survey’s Earth Resources Observation and Science Center (phenology.cr.usgs.gov/index.php; Brown et al. 2015). eMODIS SOS/EOS estimates are based on weekly composites of maximum
Normalized Difference Vegetation Index (NDVI) values and have been processed and standardized for time-series analysis (Jenkerson et al. 2010). Temporal smoothing using weighted least-squares is applied to the NDVI data to reduce unusual spikes caused by cloud contamination, atmospheric effects, and other artifacts; then a delayed moving average approach is used to predict NDVI values from other observations in the time-series. The timing of phenological events is represented by the point when actual NDVI values significantly depart from the trend predicted by the moving window (Reed et al. 1994). The range of possible DOY estimates provided by eMODIS are Julian day 60-183 (March 1st-July 1st) for SOS and 244-450 (September 1st through March of the following year) for EOS. We suspect the wide range of EOS dates is designed to capture evergreen forest phenology, which can be especially prolonged compared to deciduous forests. We did not constrain the EOS dates, as doing so would have resulted in the loss of most pixels across the region (most had at least one year where EOS was estimated as occurring after DOY 365). Instead, individual pixels were removed prior to trend analysis if: 1) an SOS/EOS DOY estimate for any given year was greater than two standard deviations from its overall 15-year mean and 2) both SOS and EOS estimates were not available for every year over the 15-year study period. This step was also designed to address potential errors from mixed pixel and atmospheric effects as well as disturbance events (e.g., logging, insect defoliation, etc.).

In the heterogenous forests of the Northeast, a major advantage of the eMODIS phenology metrics is their combination of high temporal frequency (weekly) and finer spatial resolution (250-meter pixels) compared to those derived from other sensors. For
example, Landsat has better spatial resolution (30m) but a longer return interval, and the Advanced Very High Resolution Spectroradiometer (AVHRR) has a longer data record but coarser spatial resolution (~1 kilometer). MODIS-derived NDVI estimates have also been shown to provide accurate predictions of green-up and senescence in northeastern deciduous forests based on ground validation and agreement with finer-scale remote sensing methods (e.g., LANDSAT and digital cameras) (Fisher and Mustard 2007, Hufkens et al. 2012b, Hmimina et al. 2013).

4.3.3 SOS/EOS trends and post-hoc filtering

To capture changes in phenology, we calculated pixel-level trends in SOS/EOS dates using ordinary least squares (OLS) linear regression, with the slope of the trendline representing the annual rate of change (days/year\(^{-1}\)). One-sample t-tests were used to assess if trends were significantly different from zero (p ≤ 0.05). Intraregional differences in phenology trends were tested using analysis of variance (ANOVA) with Tukey’s procedure for comparing means between ecoregions.

Post-hoc examination of the trend data revealed numerous ‘anomalous’ pixels with unrealistic trends, almost exclusively for EOS (i.e., delayed by more than a month per decade). Examination of these locations indicate that they were likely dominated by oaks and conifers, species with phenology typically distinct from northern hardwoods. To ensure that pixels included in our modeling represent true changes in phenology across northern hardwood forests, we further limited the geographic extent of our analyses to pixels with trends between -2 and 2 days/year\(^{-1}\), or a net change of approximately one month over the
15-year period, based on the upper limit of rates reported in other regional studies (e.g., Dragoni and Rahman 2012).

4.3.4 Forest cover, topography, and climate data

Three different LANDFIRE (LF; www.landfire.gov) 30-meter spatial resolution forest cover products (percent forest cover datasets for 2001 and 2014, and the existing vegetation type (EVT) classification for 2014) were used to constrain eMODIS phenology data to northern hardwoods-dominated pixels only. After aggregating the individual LF products to match the spatial resolution of our phenology data (250m), we retained only those eMODIS pixels classified as northern hardwoods with ≥75% forest cover.

A regional digital elevation model (DEM; 250m) was obtained from the Global Multi-resolution Terrain Elevation Dataset 2010 (lta.cr.usgs.gov/GMTED2010) and used to calculate three indices that describe local site characteristics: site exposure index (SEI; Balice et al. 2000), heat load index (HLI; McCune and Keon 2002), and compound topographic index (CTI; Gessler et al. 1995). The first two indices transform slope-aspect relationships to represent warmer (higher values) and cooler (lower values) sites while accounting for slope steepness, with the main difference being HLI also incorporates latitude. The CTI is a steady-state wetness index representing soil moisture gradients, with wet, depressed areas having the highest values and dry ridges the lowest.

Monthly total precipitation and maximum, mean, and minimum temperature data for each year in the study period (2001-2015), along with 30-year normals (1981-2010), were acquired from the PRISM Climate Group (4km resolution;
Raw climate averages in the Northeast exhibit strong spatial patterns that can complicate inferences regarding important regional drivers. For example, lower latitudes, elevations, and the western ecoregions are typically warmer on average than higher latitudes and elevation in eastern ecoregions. Therefore, our analyses focused on two sets of unique climate metrics: trends and anomalies calculated at the pixel level. Temporal trends were calculated using the slope from an ordinary least-squares (OLS) linear regression (representing the rate of change); anomalies were calculated by subtracting the 30-year normal from the 15-year average, representing departures from historical norms (1981-2010) over the entire study period.

Additionally, climate metrics inherently contain a high degree of autocorrelation (spatial and temporal) and collinearity. To minimize these associations, we examined the correlation structure between climate trends and anomalies within SOS- and EOS-specific timeframes (Dec-May and Jun-Nov, respectively) and combined highly-correlated months. The grouping of “like-months” was informed by cluster analysis and our understanding of important phenological timeframes in the Northeast (e.g., max temperature anomalies formed two distinct clusters that retained phenological relevance: Dec-Mar and Apr-May).

In order to further reduce collinearity and autocorrelation, we used a Pearson’s correlation coefficient threshold of $< 0.7$ (between variables) to identify the final set of SOS/EOS-specific climate variables to include in spatial modeling (Dormann et al. 2013; see Tables 4.1 and 4.2).
4.3.5 Modeling significant predictors of SOS/EOS trends

Our approach to identifying significant predictors of phenological trends was designed to reduce the effects of spatial autocorrelation and potential multicollinearity on statistical inferences. We based our analyses on a random sample of pixels, stratified by ecoregion and buffered by at least 4 kilometers (N = 982; see Fig. 4.3). Modeling the SOS/EOS trend as the response variable, our predictors included a set of ecologically-relevant climate and topography variables with reduced correlations (see Tables 4.1 and 4.2), with latitude added as a proxy for photoperiod. Additionally, the SOS trend was included as a predictor variable for the EOS trend, as previous studies have shown SOS DOY can influence EOS DOY in the Northeast (Keenan and Richardson 2015, Liu et al. 2016).

Variable reduction prior to spatial modeling was based on Spearman’s rho and partial correlations using only significant correlates (p < 0.05) with a SOS/EOS partial coefficient > 0.05 or < -0.05. This reduced set of predictor variables were then entered into a stepwise multiple linear regression model using the Bayesian Information Criterion (BIC) to identify the best predictive model and assess variable significance (p < 0.05 for retention), autocorrelation (variance inflation factor < 2 for retention; Graham (2003), and relative importance (using scaled and centered parameter estimates).

Prior to spatial modeling, significant, independent variables (see Tables 4.1 and 4.2) were assessed for spatial autocorrelation (Moran’s I test) and heteroscedasticity (Breusch-Pagan test), which are rarely evaluated in similar landscape-scale phenology studies but can have strong impacts on parameter estimates and hence perceived
importance (Anselin et al. 1996). In the presence of significant spatial autocorrelation but no heteroscedasticity, we constructed a spatial lag model that accounts for spatial dependencies by applying a ‘lag term’ to the dependent variable (DV). The lag term was a weighted average of neighboring values that smooths the DV, with the spatial weights determined by a minimum Euclidean distance between points (Anselin 2013). If heteroscedasticity was significant in the multiple regression model, we used geographically-weighted regression (GWR) to examine predictor significance across the study area. GWR uses an adaptive moving window based on an optimized number of neighbors to fit local regressions and identify areas where parameter estimates significantly depart from their global values (Brunsdon et al. 1998). As GWR does not output p-values, significance was assessed using t-value thresholds associated with 95% confidence (< -1.96 and > 1.96).

4.4 Results and Discussion

4.4.1 Regional trends in SOS/EOS metrics

At a regional scale, we detected a slight trend toward earlier SOS (mean = -0.05 ± 0.006 SE days/year⁻¹) and strong trend toward later EOS (0.86 ± 0.01 days/year⁻¹) (Fig. 4.2). While these region-wide trends were highly significant, SOS and EOS DOY estimates exhibited substantial interannual and within-year variability (Fig. 4.2), resulting in poor regression fit (p < 0.0001 and r² < 0.10). Within-year variability was more pronounced in the EOS data, where standard deviations for individual years ranged from ~10-28 days (~11 days on average overall), compared to SOS (range of ~7-14 days, ~5 days on average.
overall). In the heterogenous forests of the Northeast, high within-year variability is not surprising given the diverse latitudinal, elevational, and climate conditions, as well as differences in senescence timing among common deciduous species (Fisher et al. 2006, Richardson et al. 2006, Richardson et al. 2009a). For example, Fisher et al. 2006 found cold air drainage patterns in relatively flat uplands strongly influence phenology over distances of less than 500 meters in southern New England. In northern New England where more dramatic elevational gradients exist, others have shown trees at lower elevations leaf out earlier and tend to keep their leaves longer than those at higher elevations (e.g., Richardson et al. 2006, Xie et al. 2015b).
Figure 4.2. Trends in start and end of season (SOS/EOS) from 2001-2015 across the entire study area (left pane) and each ecoregion (right panes). All trends were highly significant (p<0.0001) but with poor model fit ($r^2 < 0.10$). The position of each ecoregion pane generally corresponds to its geographic location (e.g., the Eastern Great Lakes Lowlands occupy the northwest portion of the study area). CI = confidence interval.
High between-year variability is also typical considering the large area and relatively short timeframe we examined, as well as the strong link between year-to-year variability in weather patterns and phenological events. Examination of the temporal variability highlighted extreme differences between years that were consistent across the region: 1) SOS in 2009 and 2010 was particularly early relative to other years and later in 2011 and 2012; 2) EOS occurred later in 2001, 2012, and 2015 and earlier in 2011, with very high variability in 2011 and 2012; and 3) 2011 was characterized by a late SOS and early EOS (shortest growing season length; especially apparent in the Northern Allegheny Plateau), while the late SOS in 2012 was followed by a late EOS (Fig. 4.2).

The 2010 and 2012 SOS phenomena are well-documented in the Northeast (see Friedl et al. 2014). While both years had exceptionally warm springs, the early arrival of SOS in 2010 was due to high temperatures directly preceding and during budburst (April-May), whereas those in 2012 peaked earlier (March). Keenan and Richardson (2015) have shown that later SOS usually leads to later EOS in the Northeast, which could explain the 2012 patterns observed here (late SOS/late EOS). The extremely short growing season in 2011 (late SOS/early EOS) relative to other years coincides with an exceedingly wet late summer and early autumn, largely due to Hurricane Irene. Recent studies have shown heavy rainfall can lead to earlier senescence in northeastern deciduous forests (Xie et al. 2015b, Xie et al. 2018). The late EOS dates in 2001 and 2015 were likely temperature-driven: both years were exceptionally warm (2015 had the highest autumn temperatures on record), while 2001 also had low rainfall (yearly climate information obtained from NOAA’s State of the Climate reports, accessed at www.ncdc.noaa.gov/sotc/).
4.4.2 SOS trends – spatial patterns and significant predictors

The relatively low regional trend toward earlier SOS (-0.05 days/year\(^{-1}\)) in northern hardwood forests was the product of a significant longitudinal pattern: eastern ecoregions exhibited earlier SOS while western ecoregions had delayed SOS (Fig. 4.2 and 4.3). In the east, the rate of spring advancement was highest in the Northeastern Highlands (-0.12 ± 0.01 days/year\(^{-1}\)), whereas in the west the Northern Allegheny Plateau exhibited the strongest delaying trend (0.17 ± 0.01 days/year\(^{-1}\)) (Figs. 4.2 and 4.3). These trends generally showed good agreement with prior remote sensing and site-based studies in northeastern deciduous forests. The earlier SOS rates for the eastern ecoregions (combined mean = -0.10 days/year\(^{-1}\)), which span most of New England, are slightly lower than that found by Yang et al. (2012) (-0.14 days/year\(^{-1}\)) based on retrospective analysis of climate data over the past 50 years for that region. This rate was also lower than localized estimates by Keenan et al. (2014) and Yue et al. (2015) at Harvard Forest (-0.4 – -0.5 days/year\(^{-1}\)) and the Hubbard Brook Experimental Forest (-0.2 – -0.3 days/year\(^{-1}\)). However, the spatial distribution of our SOS trend data does show stronger advances in these locations (i.e., the south-central Northeastern Highlands; Fig. 4.3). Jeong et al. (2011) and Park et al. (2016) noted a slowing of spring advancement in large parts of the Northern Hemisphere since the turn of the 21\(^{st}\) century, which could help explain the lower magnitude of our trends versus those that include data from the 1980s-1990s.
Figure 4.3. Spatial distribution of the start of season (SOS) trend (top left) and significant explanatory climate variables. The SOS trend map was interpolated using inverse-distance weighting for visualization purposes only due to low sample pixel densities. The interpolation was based on the random points shown in the SOS panel (black dots), stratified by ecoregion and buffered by a minimum of 4 kilometers, which were also used for identifying the important predictors of SOS trends.
Initial correlations (Spearman's \( \rho \)) revealed several significant relationships between SOS trends in northern hardwood forests and climate, as well as latitude (Table 4.1). Higher precipitation always exhibited a delaying effect on SOS trends, while warmer temperatures were usually associated with earlier SOS (Table 4.1). A counterintuitive exception was higher spring-only temperatures (e.g., April-May maximum temperature anomaly) showed a delaying effect on SOS trends. Visually inspecting spatial patterns in the climate metrics suggests this was because of dissociation with the delayed SOS trends in western ecoregions – spring temperatures were warmer than normal across most the study area (e.g., Apr-May \( T_{\text{max}} \) anomaly mean = 0.27°C).

Subsequent partial correlation and regression analyses on significant correlates converged on a similar set of climate factors as having the most explanatory power in predicting SOS trends (Table 4.1). The final ‘global’ multiple regression model for the full region contained four climate variables (ordered by magnitude of importance): Dec-May mean temperature trends, Feb-Mar precipitation trends, Dec-Feb minimum temperature anomalies, and Dec-May precipitation anomalies (Table 4.1). Because this model’s residuals exhibited strong spatial autocorrelation (Moran’s I p-value < 0.0001) but no heteroscedasticity, we constructed a spatial lag model. After accounting for spatial dependencies, Dec-May precipitation anomalies were no longer significant and Dec-Feb min temperature anomalies only marginally so (Table 4.1), indicating these two parameters suffered from strong spatial autocorrelation effects (i.e., pseudo-replication) while the Dec-May mean temperature and Feb-Mar precipitation relationships held true across the study area.
Table 4.1. Summary of the statistical tests used to identify strong predictors of start of season (SOS) trends in northern hardwood forests. Tests are ordered sequentially from left (variable reduction) to right (final spatial model).

\[ p < 0.001 \] and \[ p < 0.05 \] for SPEAR and OLS significance tests

<table>
<thead>
<tr>
<th>Precipitation</th>
<th>SPEAR</th>
<th>PCOR1</th>
<th>PCOR2</th>
<th>STEP BIC</th>
<th>OLS EST</th>
<th>SLAG p-v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec-May precip. anomaly</td>
<td><strong>0.29</strong></td>
<td>0.08</td>
<td>0.09</td>
<td>✔</td>
<td>0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>Dec-Jan precip. trend</td>
<td><strong>0.21</strong></td>
<td>0.01</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Feb-Mar precip. trend</td>
<td>0.14</td>
<td>0.17</td>
<td>0.16</td>
<td>✔</td>
<td><strong>0.30</strong></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Apr-May precip. trend</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temperature</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec-May mean temperature trend</td>
<td>-0.26</td>
<td>-0.12</td>
<td>-0.17</td>
<td>✔</td>
<td>-0.47</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dec-Mar max temperature anomaly</td>
<td>-0.07</td>
<td>-0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Apr-May max temperature anomaly</td>
<td>0.17</td>
<td>0.07</td>
<td>0.07</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Dec-Feb min temperature anomaly</td>
<td>-0.20</td>
<td>-0.07</td>
<td>-0.11</td>
<td>✔</td>
<td>-0.18</td>
<td>0.05</td>
</tr>
<tr>
<td>Mar-May min temperature anomaly</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Site characteristics</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>-0.19</td>
<td>-0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Site exposure index</td>
<td>-0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Heat load index</td>
<td>-0.02</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Compound topographic index</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

SPEAR = Spearman’s \( \rho \); PCOR = partial correlations; STEP BIC = best stepwise regression model validated by lowest Bayesian Information Criterion value; OLS EST = multiple regression parameter estimates (scaled and centered for comparison); SLAG p-v = parameter p-value after removing spatial autocorrelation using a spatial lag model. The checkmarks for STEP BIC denote the terms in the best-fit model.

In the global model, warming throughout the winter-spring months (Dec-May) and higher max temperatures in winter (Dec-Feb) were associated with earlier SOS. Conversely, increasing precipitation in late winter-early spring, likely representing late snowfall, and wetter than normal conditions from Dec-May delayed SOS. Though we detected a significant negative association with latitude (coeff. = -0.06° ± 0.01°, \( p < 0.0001 \)), further examination of the spatial data (Fig. 4.3) showed this was mainly a product of the stark longitudinal pattern: on average, the western ecoregions (delayed SOS) were
lower than the eastern (earlier SOS), and similar latitudes in the east versus west exhibited this divergent pattern. Latitude was also dropped in the first iteration of partial correlations, highlighting that climate trends had a stronger influence on SOS than photoperiod across the study area. Notably, we did not detect a regional relationship between SOS and elevation.

The significant relationships between the important climate metrics and SOS trends were spatially-consistent (Fig. 4.3). Both temperature variables captured the longitudinal split in the SOS data. Based on the climate data from 2001-2015, Dec-May mean temperatures trended higher in the east and lower in the west and Dec-Feb min temperatures were hotter/colder than normal in the same pattern. The relationship between Feb-Mar precipitation and SOS trends was less clear, though areas trending towards less winter precipitation generally exhibited earlier SOS, particularly in the Northeastern Highlands and eastern edge of the Northern Allegheny Plateau (Fig. 4.3).

Precipitation is not commonly believed to play a significant role in SOS in deciduous forests of the Northeast (Polgar and Primack 2011, Klosterman et al. 2014). However, Fu et al. (2014a) recently found significant positive correlations between winter precipitation and the amount of growing-degree days (GDD) required for green-up across temperate and boreal forests (mid-high latitudes). They postulated two mechanistic links related to soil temperature and solar radiation requirements for vegetative growth: 1) that heavier snowpacks and greater snow melt may require more GDD to warm the soil sufficiently and 2) higher winter precipitation may simply reflect more cloudy days and thus require more GDD to pass a certain absorbed radiation threshold (which had less
empirical support). The data of Groffman et al. (2012) supported the first of these proposed mechanisms – they noted that leaf expansion was associated with the rapid increase of soil temperatures from solar warming immediately following the loss of snowpack. Although these are plausible explanations for the precipitation-SOS trend relationship in our study area (more winter precipitation/snow = delayed SOS and vice-versa), more field studies and controlled experiments are needed to better define causal relationships between precipitation inputs and SOS.

The role of winter temperatures in mediating SOS responses in northeastern deciduous forests is still debated in the literature (see discussions in Migliavacca et al. 2012, Melaas et al. 2016b). A large source of uncertainty is the physiological underpinnings of chilling requirements (i.e., prolonged exposure to cold temperatures that prevents premature budburst during abnormally warm winter days) remain poorly understood both within and across species. Yet most process-based models of spring phenology fall into two categories: *spring warming* (where budburst is a function of GDD accumulation only) and *chilling* (where chilling requirements must be met before GDD can begin to accumulate) (Migliavacca et al. 2012). Both Migliavacca et al. (2012) and Melaas et al. (2016a) found strong support for the use of simple spring warming models (where SOS is a function of spring temperatures only) across most forest types of the Northeast, though with significant differences among species. Our results show winter temperatures play a prominent role in explaining regional variations in SOS trends exhibited by northern hardwood forests. As the previous studies are based on predicting SOS DOY, it is possible that winter temperatures are primarily significant in examining long-term trends rather than
annual DOY. Regional climate models suggest winter is warming more rapidly than other seasons in the Northeast (Hayhoe et al. 2007, Burakowski et al. 2008), which could explain the link we observed here. Nonetheless, these results indicate the need to examine the role of both winter and spring temperatures to better understand how climate change may impact regional phenology.

4.4.3 EOS trends – spatial patterns and significant predictors

In contrast to SOS, the overall trend toward delayed EOS (0.86 days/year\(^{-1}\)) was nearly universal across the study area. The only statistically significant difference between ecoregions was a slightly lower trend in the Northeastern Highlands (0.71 ± 0.01 days/year\(^{-1}\)) compared to the other three ecoregions, had a mean EOS trend of just over 1 day/year\(^{-1}\) (Fig. 4.2). Spatial patterns showed this was largely due to trends toward earlier EOS in parts of the southern Green and eastern Adirondack mountain ranges (Fig. 4.4; see below for further discussion). The regional trend of 0.86 days/year\(^{-1}\) is almost identical to that estimated by Dragoni and Rahman (2012) across temperate forests of the eastern US and Keenan et al. (2014) at Harvard Forest (both ~0.8 days/year\(^{-1}\)), with the former based on AVHRR data from 1989-2008 and the latter on ground observations from 1991-2013.
Figure 4.4. Spatial distribution of the end of season (EOS) trend (top left) and significant explanatory variables. The points on the explanatory variable maps represent those identified by geographically-weighted regression as significant (95% confidence level). Note the Jun-Nov mean temperature anomaly map is included for comparison to the EOS trend, though it was not statistically significant. The EOS trend map was interpolated using inverse-distance weighting for visualization purposes only due to low sample pixel densities. The interpolation was based on the random points shown in the EOS panel (black dots), stratified by ecoregion and buffered by a minimum of 4 kilometers, which were also used for identifying the important predictors of EOS trends.
Determining significant predictors of EOS trends was substantially more straightforward than SOS – only four variables were significant in our initial correlations (Table 4.2). Higher elevations, latitudes, and Sep-Nov precipitation anomalies were associated with earlier EOS trends (and sometimes earlier EOS in the case of elevation; Fig. 4.4), while the correlation with CTI suggested that higher moisture availability delayed EOS trends further (Table 4.2). As with SOS, latitude was a significant negative correlate with EOS trends, but the first partial correlation test indicated elevation exhibited a stronger control on EOS – though both are likely related higher elevations and latitudes in the study area having shorter growing seasons than their counterparts elsewhere. The final stepwise BIC and multiple regression model contained only elevation (strongest predictor) and Sep-Nov precipitation anomalies. However, this model exhibited significant heteroscedasticity (Breusch-Pagan test p-value < 0.0001), so GWR was used to examine how the significance of each parameter varied across the study area.
Table 4.2. Summary of the statistical tests used to identify strong predictors of end of season (EOS) trends in northern hardwood forests. Tests are ordered sequentially from left (variable reduction) to right (final ‘global’ model). Note geographically-weighted regression was used as the final spatial model, which does not provide global significance testing of individual predictors.

<table>
<thead>
<tr>
<th>Precipitation</th>
<th>SPEAR</th>
<th>PCOR1</th>
<th>PCOR2</th>
<th>STEP BIC</th>
<th>OLS EST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun-Aug precipitation anomaly</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sep-Nov precipitation anomaly</td>
<td><strong>-0.10</strong></td>
<td>-0.06</td>
<td>-0.08</td>
<td>✔</td>
<td>-0.21</td>
</tr>
<tr>
<td>Jun-Aug precipitation trend</td>
<td>0.02</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sep-Nov precipitation trend</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temperature</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun-Sep mean temperature trend</td>
<td>0.03</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Oct-Nov mean temperature trend</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Jun-Nov max temperature anomaly</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Jun-Nov min temperature anomaly</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Site characteristics</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>-0.07</td>
<td>-0.02</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Start of season trend</td>
<td>0.01</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Elevation</td>
<td><strong>-0.09</strong></td>
<td>-0.10</td>
<td>-0.10</td>
<td>✔</td>
<td><strong>-0.25</strong></td>
</tr>
<tr>
<td>Site exposure index</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Heat load index</td>
<td>-0.02</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Compound topographic index</td>
<td><strong>0.09</strong></td>
<td>0.09</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

SPEAR = Spearman’s rho; PCOR = partial correlations; STEP BIC = best stepwise regression model validated by lowest Bayesian Information Criterion value; OLS EST = multiple regression parameter estimates (scaled and centered for comparison). The checkmarks for STEP BIC denote the terms in the best-fit model.

The GWR results revealed a complex spatial relationship between EOS trends, elevation, and Sep-Nov precipitation. Significant negative coefficients (earlier or less delayed EOS) for elevation were found in parts of almost every mountain range in the study area, as well as the northernmost corner of the Acadian Plains and Hills and Northeastern Highlands; significant negative coefficients for Sep-Nov precipitation anomalies were exclusively in the northern Adirondacks and Green Mountains, where they largely overlapped with significant elevation points (Fig. 4.4). These relationships are likely a reflection of strong elevational climate gradients – particularly differences in the timing of low temperatures that generally occur earlier at higher elevations. Although many
environmental stresses can initiate leaf senescence, low temperatures that instigate changes in leaf pigment concentrations (Schaberg et al. 2017) and leaf abscission (Michaeli et al. 1999) are the most common triggers of autumnal leaf senescence and loss in northern regions (Heide and Prestrud 2005, Schaberg et al. 2017). Other studies (e.g., Xie et al. 2015b, Xie et al. 2018) found similar relationships between EOS DOY, elevation, and heavy autumn precipitation to those documented here. Although the physiological basis for an earlier EOS-heavy precipitation relationship remains poorly understood, it could involve improved conditions for foliar fungal diseases that shorten leaf lifespans, reduced light exposures that constrain photosynthetic gains and speed leaf senescence, or the physical damage and loss of leaves from intense precipitation and associated wind events.

We suspect that the lack of a significant statistical relationship between EOS trends and temperature was due to the limited variation in summer-fall temperatures (Fig. 4.4). Coincident with the near-universal regional trend towards later EOS, Jun-Nov mean temperatures were warmer than normal in over 99% of the study area. Further, other recent studies have related delayed EOS in the Northeast to warmer summer and autumn temperatures (e.g., Dragoni and Rahman 2012, Xie et al. 2015b, Xie et al. 2018). Additionally, the delaying effect of CTI on EOS suggests that sites with better growing conditions (i.e. higher soil moisture availability and warmer temperatures) retain leaves longer.

It is interesting to note that we did not detect a significant relationship between SOS and EOS trends, contrary to previous studies based on DOY in the Northeast (Keenan and Richardson 2015) and controlled experiments (Fu et al. 2014b). Our results indicate that
while EOS may be influenced by SOS within a given year, how these phenological events are changing over time may not be as closely linked. While EOS is consistently trending later, SOS is more variable across the landscape.

### 4.4.4 Implications of phenological change in the Northeast

While interannual variability in SOS/EOS remains high, the trends (earlier/later) documented here support previous studies that show growing season length is increasing in the Northeast primarily as a result of delayed senescence (Piao et al. 2007, Jeong et al. 2011, Dragoni and Rahman 2012). The apparent greater sensitivity of EOS to environmental stimuli is consistent with later EOS having few physiological risks but real potential gains as protracted leaf longevity may lead to greater seasonal carbon capture. In contrast, physiological constraints on a very early SOS likely reflect strong selection pressures against precocious leaf expansion at a time when frost is likely, and the risks of freezing injury and associated carbon loss are high.

Even the western ecoregions that exhibited delayed SOS had associated delays in EOS at rates that were much higher. At the rate estimated in this study (0.86 days/year⁻¹), EOS will occur later by ~8-10 days per decade or almost a month over the next 30 years. Such strong shifts in the seasonality of northeastern forests will likely have significant impacts on their function and structure, as phenology helps to regulate the abiotic properties and biogeochemical processes of these ecosystems and, in turn, their biotic communities (Polgar and Primack 2011, Richardson et al. 2013). The vast ecological implications of phenological change have been the subject of several reviews (e.g., Cleland
et al. 2007, Polgar and Primack 2011, Richardson et al. 2013), from which we highlight a few major points specific to forest structure and function in the Northeast:

1) Seasonal canopy dynamics directly influence forest composition by regulating the core habitat requirements of tree species for reproductive success and survival (e.g., microclimates and light/water/nutrient dynamics) and altering competitive interactions. Indeed, phenology has been shown to be a powerful predictor of species’ ranges and distributions (Chuine and Beaubien 2001), and current predictions suggest the coupled effects of climatic and phenological change (e.g., warmer temperatures and longer growing seasons) may favor the expansion of hardwood-dominated forest over spruce-fir ecosystems in the future (Iverson et al. 2008, Huntington et al. 2009).

2) Longer growing seasons may result in increased forest ecosystem productivity and carbon storage, thereby serving to mitigate global climate change. Results from recent studies suggest this is already occurring in the Northeast (e.g., Richardson et al. 2010, Keenan et al. 2014), though others have noted high variability in the growing season length-productivity relationship (Wu et al. 2016) and lower net ecosystem productivity gains across the broader Northern Hemisphere due to concomitant losses in soil carbon (Piao et al. 2007). Furthermore, reductions in wood density may at least partially offset carbon accumulations associated with increased growth (Pretzsch et al. 2018).

3) The potential detrimental effects of phenological change largely relate to increased mismatches in phenological timing among species that react
differently to changing climate cues. Examples of this include trees responding
to warmer spring temperatures with earlier budburst that are then more
susceptible to late spring frost damage (e.g. Gu et al. 2008, Hufkens et al.
2012a) and pollinators becoming asynchronous with flowering dates (Miller-
Rushing and Primack 2008).

4.5 Conclusions

Based on eMODIS data from 2001-2015, these results show that the magnitude and
direction of phenological trends vary considerably, particularly for SOS, across northern
hardwood forests of the Northeast. Phenological responses to climate drivers within this
forest type can also differ depending on geographic location, likely reflecting physiological
adaptations to specific environmental conditions across a given species’ range (Polgar and
Primack 2011). As others (e.g., Vitasse et al. 2009, Melaas et al. 2016a) have noted, this
highlights the importance of considering localized site characteristics, including species
composition, climate, and soils, when examining the potential long-term impacts of climate
change on forest phenology. Additional important results documented in this study include:

1) The modest regional trend towards earlier SOS can be explained by substantial
intraregional variability, with a distinct opposing pattern between eastern (earlier)
and western (later) ecoregions.

2) While warming temperatures are the main driver of delayed SOS trends, our results
suggest temperature and precipitation changes in winter outweighed influences in
spring for explaining long-term trends.
3) EOS is trending significantly later across nearly the entire region, but higher elevations, latitudes, and autumn precipitation totals can mitigate this response.

4) Explicitly modeling phenological trends (rather than DOY) and relating these to climate trends and anomalies for the same time period can identify and validate important parameters in phenology models.

5) Spatial regression techniques provide a powerful tool for addressing autocorrelation and heteroscedasticity issues in phenology data, thereby improving model parameterization and the understanding of spatiotemporal patterns.

These results expand our existing knowledge on the current and potential future impacts of climate change on deciduous forests. Given the heterogeneity in phenological trends and responses to environmental conditions we documented, future remote sensing-based phenology research would benefit greatly if higher resolution climate and soils data, as well as more detailed and accurate forest cover maps, were made available.

4.6 Acknowledgements

This research was funded by the Northeastern States Research Cooperative through the USDA Forest Service Northern Research Station (Grant #14-DG-11242307-087), and the McIntire-Stennis Cooperative Forestry Program through the USDA National Institute of Food and Agriculture (Grant #2017-32100-06050).

4.7 References


CONCLUDING REMARKS AND RESEARCH SUMMARY

This dissertation provides valuable remote sensing-based products and research related to changes in forest composition, structure, and function in the Northeast. This information can be used by forest researchers, managers, and policymakers to inform management and conservation strategies. Additionally, the methods employed here for mapping forests and forest responses to potential drivers of change offer a blueprint for similar remote sensing studies in other regions of the world. Together, this work is a powerful example of the important role satellite remote sensing plays in broadening our understanding of forest change and enabling predictions into the future.

The more detailed, accurate forest cover maps produced in Chapter 2 have a wide range of applications as inputs to improve regional land use/land use change models, dynamic vegetation models, and other large-scale modeling efforts (e.g., wildlife occupancy). For example, these maps have already been utilized to improve our understanding of carbon storage in northeastern forests (Adams et al. 2018). Chapter 3 demonstrated their usefulness in modeling spatiotemporal changes in tree species abundance, providing important insights into ongoing demographic shifts in northeastern forests and their potential climate/site-related drivers. Chapter 4 added further evidence that climate change is shifting forest phenology across the Northeast, identified significant climate parameters, and elucidated their complex spatial relationships with phenological responses at the start and end of the growing season. Some important limitations on this work include the error inherent in mixed pixels of moderate-coarse resolution (30m and 250m), general lack of downscaled climate data available at the regional level (800m and
4km), and relatively short timeframes examined in the context of forest change (15-30 years).

The outcomes of each specific research objective are summarized below:

**Obj. 2.1:** Integrate multi-temporal Landsat imagery and field inventory data using spectral unmixing to develop pixel-level percent basal area (% BA) coverages for 10 common tree species/genera in the Northeast.

**Outcome:** Our spectral unmixing technique showed multi-temporal Landsat imagery can be successfully used to derive species/genus-level abundance at a subpixel level, with model accuracies typically being higher for more dominant species (directly related to canopy composition patterns and the availability of calibration/validation ground data). Another informative outcome of this process was that fall imagery is particularly powerful for discriminating between species and forest types.

**Obj. 2.2:** Incorporate the percent basal area maps and ancillary data into an object-based, hierarchical ruleset to generate a forest classification (10 species/genera and 6 common species assemblages).

**Outcome:** We developed a detailed forest map across the diverse forests and topography of northern New York and Vermont using both pixel- and object-based classification schemes, with the former being more accurate at a species/genus-level and the latter more appropriate for broader forest types.

**Obj. 2.3:** Compare the forest classification’s detail and accuracy with existing large-scale forest mapping products, including LANDFIRE, the National Land Cover Database (NLCD), and the National Forest Type Map.
**Outcome**: Our forest classification was more accurate across all levels (i.e., species, forest type, and coarse) compared to existing large-scale mapping products, including NLCD, LANDFIRE, and the National Forest Type Map.

**3.1**: Using the technique and products developed in Chapter 2, construct a 30-year (1985-2015) time-series of abundance (% BA) for eight dominant tree species/genera across northern New York and Vermont.

**Outcome**: Abundance time-series were successfully developed in four timesteps (1985, 1995, 2005, and 2015) for sugar maple, red maple, American beech, eastern hemlock, balsam fir, and red spruce, in addition to oaks and birches at the genus level. Modeling accuracies were relatively consistent across years, again varying by species/genus (cf. Obj. 2.1 outcome). Time-series for eastern white pine and aspens could not be reliably constructed due to low modeling accuracies in specific timesteps.

**3.2**: Examine 30-year changes in mean abundance across the study area and by elevation.

**Outcome**: While overall mean abundance remained relatively constant, abundance trends indicated significant regional declines in sugar maple, eastern hemlock, and birches and increases in American beech and red maple. Notable elevational abundance trends included: sugar maple losses throughout low-mid elevations with increases in American beech in these same zones; red spruce increases at mid-high elevations, with balsam fir decreases at high elevation; and substantial losses in birch species from mid-high elevations.

**3.3**: Detect and quantify spatiotemporal patterns in pixel-level abundance trends.
Outcome: Compatriot species often exhibited opposing abundance trends in the same areas: beech increases-sugar maple losses were found in throughout New York and red spruce increases-balsam fir losses were concentrated in mid-high elevation ecoregions of the Adirondacks. Across species, most losses were concentrated in New York and gains in Vermont; birches and eastern hemlock were exceptions to this.

3.4: Identify possible abiotic correlates (i.e., climate indices, topographical factors, acid deposition inputs, and soil characteristics) associated with abundance trends.

Outcome: Climate-related indices, particularly those associated with low winter temperatures and high heat loading, were the dominant predictors of abundance change. However, responses to climate were largely species-specific and exhibited high intraspecific variation.

4.1: Using MODIS-derived annual phenology metrics from 2001-2015, quantify trends in the start and end of the growing season (SOS/EOS) across the Northeast.

Outcome: At the regional scale, SOS trended slightly earlier and EOS significantly later. Interannual and within-year variability was high, likely driven by differing responses among species and years with unique climate phenomena (e.g., Hurricane Irene).

4.2: Examine intraregional variation and spatiotemporal patterns in SOS/EOS trends.

Outcome: The modest regional trend towards earlier SOS was driven by a clear longitudinal pattern, with western ecoregions experiencing delayed SOS and eastern ecoregions earlier SOS. EOS trended later throughout the majority of the region, though slightly less so in the Northeastern Highlands due to some areas of high elevation exhibiting weaker delays or even trends toward earlier EOS.
4.3: Investigate relationships between SOS/EOS trends and numerous climatic and topographic variables.

**Outcome:** SOS trends were driven by warmer mean winter-spring temperature trends and winter temperature anomalies, as well as trends in Feb-Mar precipitation. Temperature changes were associated with earlier SOS while precipitation had a delaying effect. Elevation and autumn precipitation anomalies were associated with weaker delays and sometimes trends toward earlier EOS. However, these relationships were non-stationary across the region, with their effects being most apparent in parts of the eastern Adirondack and southern Green mountain ranges.

**COMPREHENSIVE BIBLIOGRAPHY**


Bhat, H. S., and N. Kumar. 2010. On the derivation of the Bayesian Information Criterion. School of Natural Sciences, University of California.


Klosterman, S., K. Hufkens, J. Gray, E. Melaas, O. Sonnentag, I. Lavine, L. Mitchell, R. Norman, M. Friedl, and A. Richardson. 2014. Evaluating remote sensing of
deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11:4305–4320.


