

2019

## Crown-level mapping of tree species and health from remote sensing of rural and urban forests

Fang Fang  
fafang@mix.wvu.edu

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Forest Management Commons](#), [Geographic Information Sciences Commons](#), [Physical and Environmental Geography Commons](#), [Remote Sensing Commons](#), and the [Urban, Community and Regional Planning Commons](#)

---

### Recommended Citation

Fang, Fang, "Crown-level mapping of tree species and health from remote sensing of rural and urban forests" (2019). *Graduate Theses, Dissertations, and Problem Reports*. 4068.  
<https://researchrepository.wvu.edu/etd/4068>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact [researchrepository@mail.wvu.edu](mailto:researchrepository@mail.wvu.edu).

# Crown-level mapping of tree species and health from remote sensing of rural and urban forests

Fang Fang

Dissertation submitted  
to the Eberly College of Arts and Science  
at West Virginia University

in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in  
Department of Geology and Geography

Brenden McNeil, Ph.D.  
Timothy Warner, Ph.D.  
Trevor Harris, Ph.D.  
Jamison Conley, Ph.D.  
Gregory Dahle, Ph.D.

Department of Geology and Geography

Morgantown, West Virginia

2019

**Keywords:** tree species, machine learning, phenology, remote sensing, satellite, object-based, tree health, urban trees

Copyright 2019 Fang Fang

## ABSTRACT

### Crown-level mapping of tree species and health from remote sensing of rural and urban forests

Fang Fang

Tree species composition and health are key attributes for rural and urban forest biodiversity, and ecosystem services preservation. Remote sensing has facilitated extraordinary advances in estimating and mapping tree species composition and health. Yet previous sensors and algorithms were largely unable to resolve individual tree crowns and discriminate tree species or health classes at this essential spatial scale due to the low image spectral and spatial resolution. However, current available very high spatial resolution (VHR) remote sensing data can begin to resolve individual tree crowns and measure their spectral and structural qualities with unprecedented precision. Moreover, various machine learning algorithms are now available to apply these new data sources toward the discrimination and the mapping of tree species and health classes. The dissertation includes an introductory chapter, three stand-alone manuscripts, and a concluding chapter, each of which support the overarching theme of mapping tree species composition and health using remote sensing images.

The first manuscript, now published in the *International Journal of Remote Sensing*, confirms the utility of combining VHR multi-temporal satellite data with LiDAR datasets for tree species classification using machine learning classifiers at the crown level in a rural forest the Fernow Experimental Forest, West Virginia. This research also evaluates the contribution of each type of spectral, phenological and structural feature for discriminating four tree species: red oak (*Quercus rubra*), sugar maple (*Acer saccharum*), tulip poplar (*Liriodendron tulipifera*), and black cherry (*Prunus serotina*). The second manuscript investigates the performance of tree species classification in urban settings with three contributions: 1) 12 very high resolution WorldView-3 images (WV-3), whose image acquisition date covering the growing season from April to November; 2) a large forest inventory providing sufficient calibration/validation datasets in Washington D.C.; 3) object-based tree species classification using the RandomForest machine learning algorithm. This manuscript identifies the incremental losses in classification accuracy caused by iteratively expanding the classification to 19 species and 10 genera. It also identifies the optimum pheno-phases and spectral bands for discriminating trees species in urban settings. Building on these promising results from the second manuscript, the third manuscript detect a signal of statistical difference among individual tree health conditions using WorldView-3 images from June 11<sup>th</sup>, July 30<sup>th</sup> and August 30<sup>th</sup>, 2017 in Washington D.C.. It examines six vegetation indices calculated from WorldView-3 images to describe three health condition levels in good, fair and poor, and discusses the effects of green-down phenology for tree health analysis.

Overall, this dissertation research contributes to remote sensing research by combining data from both active and passive sensors to discriminate tree species in rural forest. For the species-rich urban settings, this dissertation illustrates the importance of phenology for tree species classification at crown level using VHR remote sensing images. Finally, this dissertation provides important insights on detecting statistical differences among tree health conditions at individual crown-level in the urban environment using VHR remote sensing images.

## Acknowledgments

It is hard to believe that five years go so quickly, and it is time to finish my Ph.D. journey in WVU. It is a pleasant, fruitful, and enjoyable journey. Even though I know, there were difficulties and pains during the entire process. However, there are two types of pains, one that hurts you, and the other changes you. My pleasant PhD journey in WVU changed my career path, my research opinions, and my view towards the world.

I still remember the first time I met Dr. Brenden McNeil, which happened during the new student orientation in 2014. He walked in the conference room with his water bottle and smiled to me. The story begins here: I am so lucky to work with Dr. Brenden McNeil, who is brilliant and smart as an outstanding scholar. At the same time, he is patient and positive as an academic advisor. I cannot list everything he has helped with during the past years. He is an excellent academic advisor, who not only has the broad knowledge in remote sensing to help me conducting research, but also has the insights foresight to help me attracting external funding. I am very lucky to be a student with Dr. Brenden McNeil, and he deserves my heartfelt thanks.

In the second place, I would like to thank my committee members: Dr. Timothy Warner, Dr. Trevor Harris, Dr. Jamison Conley, and Dr. Gregory Dahle. Thanks for their support and suggestions during the entire process. They were always available and willing to help no matter if I had questions about data collection, writing proposals and dissertation. I appreciate Dr. Warner for his detailed hand-written comments and suggestions on my remote sensing manuscripts. Thanks for Dr. Conley for his advices on quantitative analysis for my dissertation. I benefited a lot by Dr. Harris's courses such as ESDA, which expands my knowledge on spatial data visualization. I also appreciate the help from Dr. Dahle, who generously shared his resources on urban forestry studies especially the collaboration with DDOT. Dr. Dahle also introduced me to the ISA conference. I was awarded by ISA Student Travel Grant under his help.

I also would like to thank all the professors in WVU, especially Dr. Rick Landenberger, Dr. Aaron Maxwell, Dr. Mike Strager and Dr. William Peterjohn. Dr. Maxwell gave me rich suggestions on machine learning algorithms in R, which is a large portion for my dissertation research. I also appreciate the datasets of WV research forest from Dr. Strager and appreciate his suggestions on using UAS data to discriminate tree species. Thanks for Dr. Peterjohn who helped a lot for my field work in Fernow Experimental Forest.

At last, I also would like to thank all my friends and colleagues in WVU during the past five years. I really appreciate their help and support. Thanks for Brittany Casey who accompanies me during the past two years in my lab! Thanks to Dr. Matthew Purtill and Dr. Alex Dye for helping me with my proposal. Thanks to Dr. Yaqian He and Dr. Junli Li, who are always willing to help whenever I have questions. I do appreciate help from other graduate students: John Burkhart, Yixuan Zhu, Barbara MacLennan, Ruiqian Chen, Fei Shang, Jessica DeWitt, Christabel Devadoss, Jing Chen, Pragya Srivastava, Joshua Lohnes, Lee Ann Nolan, Jessica Brewer, Dave Kneiter, James Schindling, Park Muhonda, Haibin Di, Chris Ramezan, Liaoshao Song, Xinwei Bai, Audrey Millard, Zhi Zhong, and Tommy Brown. I also want to thank the staff in WV GIS Tech Center: Kurt Donaldson, Frank LaFone, Eric Hopkins, Kevin Kuhn, Maneesh Sharma, Xiannian Chen and Yibing Han. Special thanks to our department staff, who helped me with lots of technical issues and administration issues: Lisa Lively, Hope Stewart, Randy Crowe, and Donna Titus.

Funding for this dissertation was generously contributed by National Science Foundation (NSF) Doctoral Dissertation Research (DDIR) award entitled “Remote Sensing of Urban Tree Species and Tree Stress” in 2018 (#1758207). Thanks for all the comments and suggestions for this NSF grant from Dr. Brenden McNeil, Dr. Timothy Warner, Dr. Trevor Harris, Dr. Jamison Conley, Dr. Gregory Dahle and Earl Eutsler from Urban Forestry Administration District Department of Transportation.

# Contents

<b>1. Introduction</b> .....	1
<b>2. Combining high spatial resolution multi-temporal satellite data with leaf-on LiDAR to enhance tree species discrimination at the crown-level*</b> .....	4
2.1 Introduction.....	5
2.2 Methods .....	8
2.2.1 Study area .....	8
2.2.2 Methods overview .....	9
2.2.3 Field measurements .....	10
2.2.4 Imagery datasets .....	10
2.2.5 Image pre-processing .....	12
2.2.6 Crown delineation using field and imagery data.....	12
2.2.7 Variable extraction .....	13
2.2.8 Data Analysis.....	16
2.3 Results .....	18
2.3.1 Data combinations and random forest model accuracy .....	18
2.3.2 Useful variables identified by ANOVA and a decision tree .....	19
2.4 Discussion .....	23
2.5 Conclusion .....	25
<b>3. Discriminating tree species using the DDOT Street Tree Spatial Database and multi-temporal WorldView-3 imagery from Washington D.C., USA*</b> .....	34
3.1 Introduction.....	35
3.2 Methods .....	38
3.2.1 Study area .....	38
3.2.2 Field dataset .....	39
3.2.3 Image datasets.....	40
3.2.4 Tree crown delineation .....	41
3.2.5 Vegetation indices calculation and RandomForest classification .....	41

3.3 Results .....	44
3.3.1 The effect of number of classes on urban tree species classification .....	44
3.3.2 Variable importance for 19 species and 10 genera classification .....	45
3.4 Discussion .....	50
3.4.1 Effects from number of classes and taxonomic levels .....	50
3.4.2 Key pheno-phases for tree species mapping .....	52
3.5 Conclusion .....	53
<b>4. Street tree health from space? An evaluation using WorldView-3 data and the Washington D.C. Street Tree Spatial Database*</b> .....	55
4.1 Introduction .....	56
4.2 Methods .....	58
4.2.1 Study area .....	58
4.2.2 Image datasets .....	59
4.2.3 Field dataset .....	60
4.2.4 Tree crown delineation .....	61
4.2.5 Data analysis .....	62
4.3 Results .....	63
4.4 Discussion .....	65
4.4.1 Comparison of VIs for tree health condition class assessment .....	65
4.4.2 Effects of image acquisition date .....	66
4.4.3 Strategies for using remote sensing data in street tree health research and management .....	67
4.5 Conclusion .....	69
<b>5. Conclusion and future work</b> .....	74
5.1 Dissertation summary .....	74
5.2 Remaining challenges and future research .....	75
<b>6. Reference</b> .....	79

## List of Figures

<b>Figure 2.1</b> Study site at the Fernow Experimental Forest (a), WV, USA, showing a portion of the delineated tree crowns (White polygons) in the Watershed #4 (WS4) reference compartment (b). The background image illustrates tree crown differences in the autumn phenology by displaying Worldview-2 bands Red edge, Yellow, Blue in the red, green, and blue (RGB) colour channels. We use the yellow and red edge bands here because they are effective in capturing autumn phenology.....	8
<b>Figure 2.2</b> Flowchart of identifying tree species in FEF .....	9
<b>Figure 2.3</b> The Fernow Watershed #4 portion of the study area. (a) Spring image. (b) Summer image. (c) LiDAR canopy height model. (d) LiDAR intensity. (a) and (b) are standard false color infrared composites (bands NIR, Red, Green displayed as RGB). In each image, pixels with a canopy height less than 5 m are masked (and displayed in white) in order to highlight the variability among intact canopies. Compare to Figure 1, which shows the autumn phenology image.....	11
<b>Figure 2.4</b> Per cent of Kappa value lost by removing each group of variables.....	21
<b>Figure 2.5</b> Key results from the ANOVA tests, letters denote significant differences in means, as assessed by pair-wise t-tests.....	22
<b>Figure 2.6</b> Decision tree based on 50th percentile of variables. Affirmative logical decisions are in each case to the left.....	22
<b>Figure 3.1</b> a) Study area (within red boundary) in the District of Columbia. b) Portion of the digitized tree locations superimposed on a WorldView-3 true color image from September 28 2014.....	38
<b>Figure 3.2</b> Validation Kappa Accuracy from RF classification at species level (a) (from 2 to 19 most common species) and genus level (b) (from 2 to 10 most common genera). .....	46
<b>Figure 3.3</b> Top ten important variables at a) 19 species b) 10 genera level. Variable importance is defined by the RF algorithm. ....	49
<b>Figure 4.1</b> a) Study area (within red boundary) in the District of Columbia. b) Portion of the digitized tree locations superimposed on a WorldView-3 true color image from August 30, 2017. ....	59
<b>Figure 4.2</b> Example of the two-step crown delineation. Red circle: 4.55m radial tree crown buffer. a) True Color WV-3 image from August 30, 2017. b) NDVI1 calculated from the NIR1 and red bands. c) Application of the 0.5 NDVI1 threshold, showing the delineated tree crown as all the white pixels whose majority area falls within the red radial crown buffer. ....	61
<b>Figure 4.3</b> Mean (error bars represent +/- 1 standard error) vegetation index values for trees in three health condition classes. Based on webcam and Landsat phenology data for our study area (from Figure 7 in Elmore et al. 2012), we also show vertical dotted lines to mark the average DOY of full leaf expansion (DOY145), and leaf senescence (DOY255). ....	64



<b>Figure 5.1</b> Species distribution for WV Research Forest in 2000.....	77
<b>Figure 5.2</b> Inconsistency of tree species reference map in WV Research Forest. Crown polygons were delineated by Dr. Landenberger in 2000 (Left). Tree stem locations were provided by Henry Liebermann in 2016 (Right). Dr. Michael Strager provided the UAV image in 2016 as base map. ....	78

## List of Tables

<b>Table 2.1</b> Metadata of the imagery used in the study .....	12
<b>Table 2.2</b> Image band wavelengths .....	12
<b>Table 2.3</b> Variables used in the analysis .....	15
<b>Table 2.4</b> Features selected by Random Forest, feature selection sorted by decreasing F-value obtained from ANOVA tests of species differences .....	20
<b>Table 2.5</b> Confusion matrix from RF classification of individual trees with top 50th percentile of variables. ....	21
<b>Table 3.1</b> The most common 19 species with genus information, ranked by sample size .....	40
<b>Table 3.2</b> 12 sets of WorldView-3 data .....	41
<b>Table 3.3</b> Vegetation indices calculated for tree species classification .....	43
<b>Table 3.4</b> Confusion matrix from the validation of the 19 species classification. ....	47
<b>Table 3.5</b> Confusion matrix for 10 genera classification .....	48
<b>Table 4.1</b> Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis tests of statistical difference in VI values among trees in good, fair and poor health condition classes. All tests were significant at $p < 0.001$ . The bold value denotes the most sensitive image date/VI combination. . ....	63

# 1. Introduction

Spatially explicit tree species and health information are valuable for answering research questions in forest inventory management, biodiversity, invasive species monitoring and habitat mapping (Fassnacht et al. 2016). With the rapid development of VHR remote sensing images during recent decades, it is increasingly possible to obtain tree species and health information from remote sensing imagery instead of traditional time-consuming field surveys especially over large areas. VHR imagery provide the opportunity to visualize individual tree crowns and motivate mapping forest types and tree species at individual crown-level in both urban and rural forests (Pu and Landry 2012, Fassnacht et al. 2016, Ferreira et al. 2019). Moreover, VHR imagery also help to establish linkages among spectral reflectance at certain band and tree health conditions (Huang et al. 2007, Eitel et al. 2011).

Both active and passive sensors provide useful information to measure species- and health-related traits (Key et al. 2001, Pu and Landry 2012, Vaughn et al. 2012, Ghosh et al. 2014, Fang et al. 2016, Fassnacht et al. 2016, Karlson et al. 2016, Madonsela et al. 2017). In general, these traits consist of three types of patterns: leaf-level structure and chemistry, crown level architecture/texture patterns, and phenological patterns expressed at the leaf and crown levels. First, images from passive optical sensors (i.e. multispectral and hyperspectral) can provide important wavelength regions to describe species-related spectral responses dominated by factors such as photosynthetic pigments, chlorophylls, and leaf water content. Researchers observed these species-related spectral responses via various vegetation indices (i.e. NDVI, RENDVI, Green ratio, etc.) calculated from important wavelength regions (Key et al. 2001, Pu and Landry 2012, Karlson et al. 2016, Madonsela et al. 2017). Second, past studies also articulated the utility of images from active sensor (LiDAR) to describe the crown level architecture/texture patterns of individual tree species (Heinzel and Koch 2011, Vaughn et al. 2012, Ghosh et al. 2014, Liu et al. 2015, Fang et al. 2016), especially the utility for discriminating conifers from broadleaf trees (Knyazikhin et al. 2013, Vauhkonen et al. 2014). Third, there is utility in sensing the phenological patterns, or temporal resolution of remote sensing images such as multispectral Geoeye,

WorldView-2, hyperspectral AVIRIS for tree species identification (Reed et al. 1994, Wolter et al. 1995, Key et al. 2001, Pu and Landry 2012, Li et al. 2015, Madonsela et al. 2017). However, for the explorations of species phenology, the image acquisition date of previous studies only covered part of the entire leaf-on period. For example, they only compared bi-temporal images from two seasons for tree species classification (Pu and Landry 2012, Li et al. 2015, Karlson et al. 2016, Madonsela et al. 2017). Exploring images covering pheno-phases from the entire growing season is of great necessity to extensively assess the importance of phenology for tree species mapping projects. In summary, there are two major challenges for tree species mapping: 1) the determination of optimum vegetation index to describe species-related spectral patterns; 2) the determination of optimum phenological period or image acquisition dates to classify tree species.

Rural and urban settings are two different types of environments for trees. Typically, rural forests consist of naturally grown trees with little interruption and maintenance from forest managers. Urban forests are more species-diverse with higher species variability, which are usually well organized and often maintained regularly by urban foresters. This dissertation systematically implemented tree species classification in both rural and urban environments. In urban settings, tree health condition is another significant factor for urban forest management. Previous researchers also adopted remote sensing images for tree health condition detection (Knipling 1970, Chaves et al. 2002, Xiao and McPherson 2005, Huang et al. 2007, Sankaran et al. 2010, Eitel et al. 2011, Asmaryan et al. 2013). Despite these successes using remote sensing imagery to describe health condition, substantial challenges remain for using satellite data to assess different levels of tree health at crown-level in urban environments.

The overall objective of this dissertation is to discriminate tree species in both rural and urban forests using remote sensing imagery with machine learning algorithms, as well as tree health condition analysis in urban settings. To address the overall objective, this dissertation answers the following questions:

- 1) For trees in rural forest, which are the optimum vegetation index and phenological periods to describe tree species-related spectral, structural and phenological traits?
- 2) For trees in species-rich urban settings, which are the optimum vegetation index and phenological periods to discriminate different tree species? What are the influences of the number of tree species and taxonomic levels on the performance of tree species mapping?
- 3) For trees in urban environments, is it possible to capture the statistical differences between tree health conditions at crown-level using multi-temporal VHR images?

## 2. Combining high spatial resolution multi-temporal satellite data with leaf-on LiDAR to enhance tree species discrimination at the crown-level\*

\*Published as Fang, F., McNeil, B. E., Warner, T. A., & Maxwell, A. E. (2018). Combining high spatial resolution multi-temporal satellite data with leaf-on LiDAR to enhance tree species discrimination at the crown level. *International Journal of Remote Sensing*, 39(23), 9054-9072.

### Abstract

The long-standing goal of discriminating tree species at the crown-level from high spatial resolution imagery remains challenging. The aim of this study is to evaluate whether combining (a) high spatial resolution multi-temporal images from different phenological periods (spring, summer and autumn), and (b) leaf-on LiDAR height and intensity data can enhance the ability to discriminate the species of individual tree crowns of red oak (*Quercus rubra*), sugar maple (*Acer saccharum*), tulip poplar (*Liriodendron tulipifera*), and black cherry (*Prunus serotina*) in the Fernow Experimental Forest, West Virginia, USA. We used RandomForest models to measure a loss of classification accuracy caused by iteratively removing from the classification one or more groups from six groups of variables: spectral reflectance from all multispectral bands in the (1) spring, (2) summer, and (3) autumn images, (4) vegetation indices derived from the three multispectral datasets, (5) canopy height and intensity from the LiDAR imagery, and (6) texture related variables from the panchromatic and LiDAR datasets. We also used ANOVA and decision tree analyses to elucidate how the multispectral and LiDAR datasets combine to help discriminate tree species based on their unique phenological, spectral, textural, and crown architectural traits. From these results, we conclude that combining high spatial resolution multi-temporal satellite data with LiDAR datasets can enhance the ability to discriminate tree species at the crown-level.

**Key words:** phenology, LiDAR, classification, multi-spectral, tree species, crown architecture, texture

## 2.1 Introduction

Obtaining spatially explicit data on tree species composition can lead to improved management of biodiversity and the provision of ecosystem services related to air and water quality (Dwyer, Nowak and Noble 2003, Karlson et al. 2016). Tree species information is traditionally derived via field survey or aerial interpretation (Pu and Landry 2012), which is time-consuming and quite limited for broad-scale tree species mapping. To date, with the growing availability of high-spatial resolution and high-spectral resolution images from various sensors, researchers are increasingly using imagery to classify individual trees in both rural and urban settings (Cochrane 2000, Key et al. 2001, Melgani and Bruzzone 2004, Colgan et al. 2012, Immitzer et al. 2012, Pu and Landry 2012, Zhang and Hu 2012, Dalponte et al. 2013, Alonzo et al. 2014, Van et al. 2014, Verlič et al. 2014, Waser et al. 2014, Cho et al. 2015, Omer et al. 2016). These studies not only illustrate the potential to discriminate species using phenological variation in their reflectance spectra, but also that tree species can be better differentiated by also using LiDAR-derived structural and intensity properties (for a detailed review, see Fassnacht et al. 2016). Here, we take advantage of both spectral and structural properties to help quantify and better understand how passive and active remote sensing data can improve discrimination of tree species at the individual crown-level.

Using multi-temporal data to capture different phenological patterns is a widely used approach for tree species discrimination (Reed et al. 1994, Wolter et al. 1995, Key et al. 2001, Madonsela. et al. 2017). By incorporating seasonal variation in the biochemical, physiological and structural properties of tree crowns, spectral variation linked to phenology has been adopted for many different remote sensing applications, including climate interactions (e.g. He et al. 2015), land cover assessment (Ganguly et al. 2010), crop observation (Sakamoto et al. 2005) and tree species mapping (Reed et al. 1994). However, at high spatial resolutions, it is challenging to obtain satisfactory images that capture phenological patterns for differentiating individual trees for an entire growing season (Reed et al. 2009). A number of studies have therefore used a variety of sensors to capture the most striking seasonal characteristics (Wolter et al. 1995, Voss and Sugumaran 2008). Studies involving multi-temporal datasets generally conclude that the

most informative image acquisition times are often in short windows surrounding the spring green-up period, peak summer, and autumn senescence phenology periods (Key et al. 2001, Voss and Sugumaran 2008, Hill et al. 2010, Wolter and Townsend 2011, Klosterman et al. 2014).

A key advantage of LiDAR for tree species discrimination is its ability to measure three-dimensional features useful for delineating tree crowns and describing species differences in crown architecture and structure (Brandtberg 2007, Dalponte et al. 2008, Heinzel and Koch 2011, Vaughn et al. 2012, Dalponte et al. 2013, Alonzo et al. 2014, Ghosh et al. 2014, Liu et al. 2015, Fang et al. 2016). For tree species classification, some studies have also recognized the value of combining LiDAR and multispectral data. Especially due to the complexity of tree crowns and similarity of spectral features between species, height information from LiDAR is often the most important factor to improve discrimination performance (Dalponte et al. 2008, Jones et al. 2010, Alonzo et al. 2014). LiDAR is also increasingly being used to link structural and spectral information in the form of intensity of returns, which is related to the reflecting object's spectral reflectance and the proportion of the laser beam cross-section that intersects with that object (Brandtberg 2007, Eitel et al. 2016, Fassnacht et al. 2016). LiDAR intensity provides invaluable information on species differences in crown architecture, thereby increasing differentiation of tree species (Brandtberg et al. 2003). For instance, broadleaf trees and conifers can be distinguished by LiDAR intensity values because coniferous tree architectures scatter more infrared light downward into their crowns, and thus provide a lower intensity value compared with broadleaved trees (Kim 2007, Knyazikhin et al. 2013, Vauhkonen et al. 2014). Intensity also differs within functional groups; for needleleaf evergreen trees, Holmgren and Persson (2004) found that the mean intensity is higher for Norway spruce (*Picea abies* L. Karst) trees, with higher standard deviation of intensity due to denser needle leaves, compared with Scots pines. Nevertheless, a full understanding of LiDAR intensity for species discrimination (Korpela et al. 2010, Fassnacht et al. 2016), and its potential, especially relative to other spectral and structural information, warrants further evaluation.



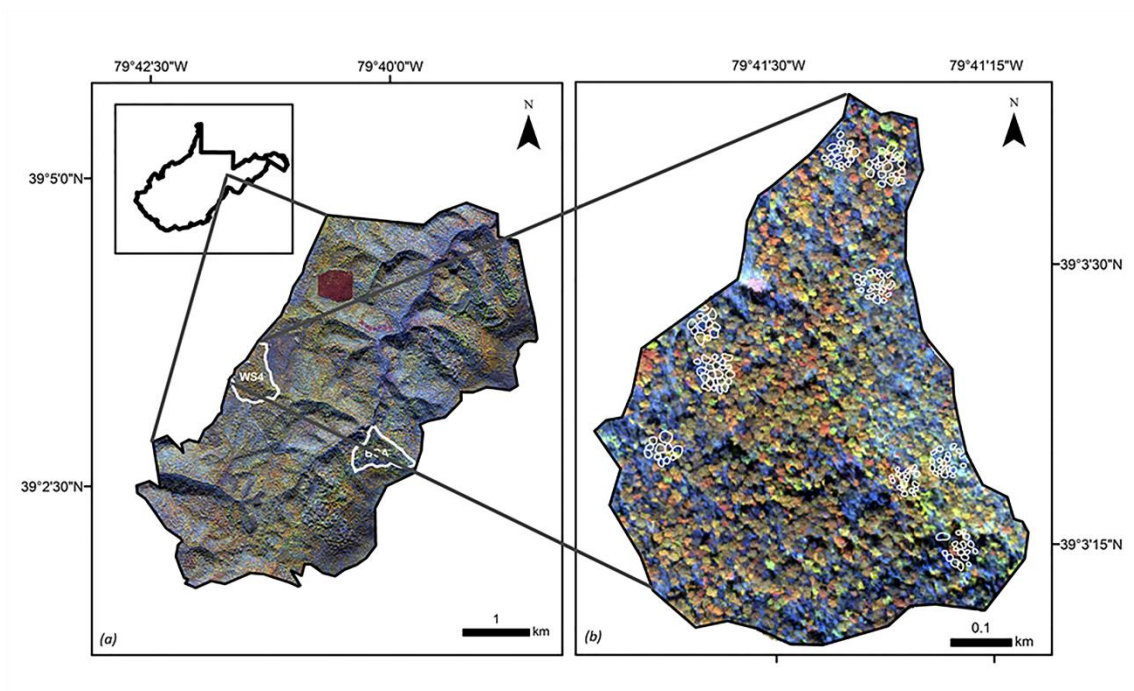
The primary goal for this study is to evaluate the utility of combining (a) high spatial resolution multispectral images from three different phenological stages, and (b) height and intensity data from leaf-on LiDAR data, to classify tree species at the individual crown level with a machine learning classifier at the Fernow Experimental Forest (FEF), West Virginia. We hypothesize that each of these image datasets will improve the separability of tree species, and add meaningful biophysical information, thereby enhancing the ability to discriminate tree species remotely.

Our approach assumes a geographic object-oriented (GEOBIA) framework for tree-species classification (Warner, Lee and McGraw, 1999). From each tree crown polygon, we extract six groups of variables, including spectral reflectance information from optical imagery that captures (1) spring, (2) summer, and (3) autumn phenology, (4) height and intensity data from leaf-on (peak summer) LiDAR, as well as (5) spectral indices and (6) texture information derived from the first four groups of variables. Then, we evaluate the relative performance of these groups of variables in Random Forest (RF) classification models. We also use analysis of variance (ANOVA) and decision trees to help understand how the LiDAR and phenological variables work individually, as well as work together to identify the distinct spectral and structural properties of each tree species. Together, these approaches are designed to help test how combinations of variables can provide a more robust method for measuring a fuller suite of spectral, phenological, and crown architectural differences useful for discriminating tree species with remotely sensed data.

## 2.2 Methods

### 2.2.1 Study area

Our study site includes two unmanaged reference compartments within the United States Department of Agriculture (USDA) Forest Service's Fernow Experimental Forest (FEF), located in Tucker County, West Virginia (**Figure 2.1** : 39°3'15''N, 79°41'15''W). The topography within each of these roughly 40 ha compartments is steeply sloped with elevations ranging between 762 m and 854 m (Burnham et al. 2017). The forests have been unmanaged following an almost complete harvest in the 1920's. Mean annual precipitation is approximately 1450 mm and mean annual temperature is 9.2 °C. The dominant tree species are red oak (*Quercus rubra*), sugar maple (*Acer saccharum*), tulip poplar (*Liriodendron tulipifera*), and black cherry (*Prunus serotina*). Other species include American basswood (*Tilia Americana*), black birch (*Betula lenta*), chestnut oak (*Quercus prinus*), cucumber magnolia (*Magnolia acuminata*), frasier magnolia (*Magnolia fraseri*), red maple (*Acer rubrum*), shagbark hickory (*Carya ovata*), white ash (*Fraxinus americana*) and white oak (*Quercus alba*).

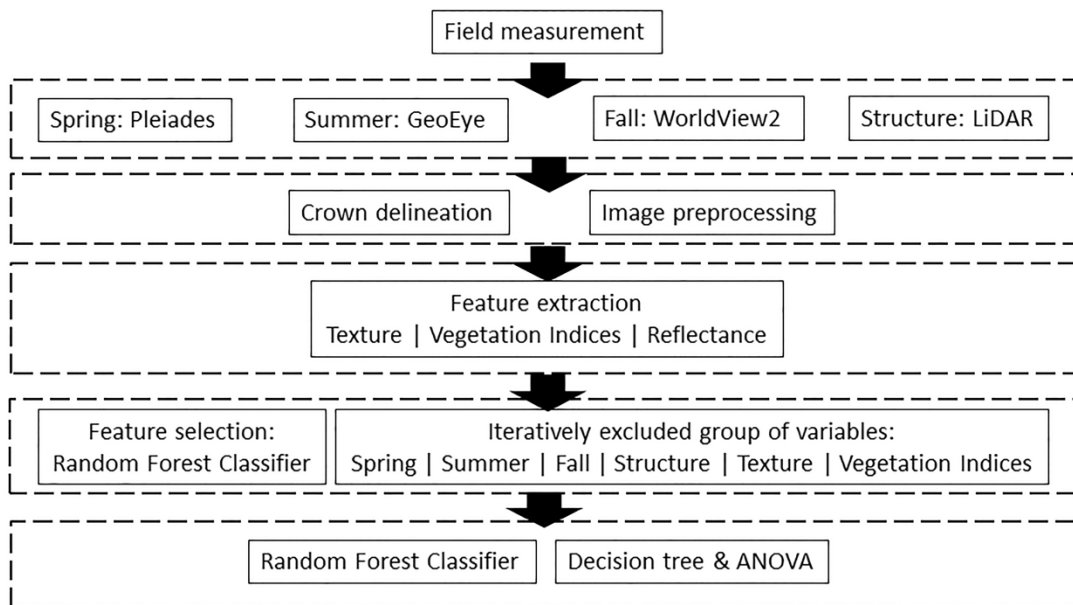


**Figure 2.1** Study site at the Fernow Experimental Forest (a), WV, USA, showing a portion of the delineated tree crowns (White polygons) in the Watershed #4 (WS4) reference compartment (b). The

background image illustrates tree crown differences in the autumn phenology by displaying Worldview-2 bands Red edge, Yellow, Blue in the red, green, and blue (RGB) colour channels. We use the yellow and red edge bands here because they are effective in capturing autumn phenology.

### 2.2.2 Methods overview

Our general workflow included a field survey, manual delineation of individual crown polygons, data preprocessing, feature extraction, feature selection and grouping using Random Forests to evaluate the relative performance of feature groups, and finally, using decision trees and ANOVA to evaluate data combinations (**Figure 2.2**).



**Figure 2.2** Flowchart of identifying tree species in FEF

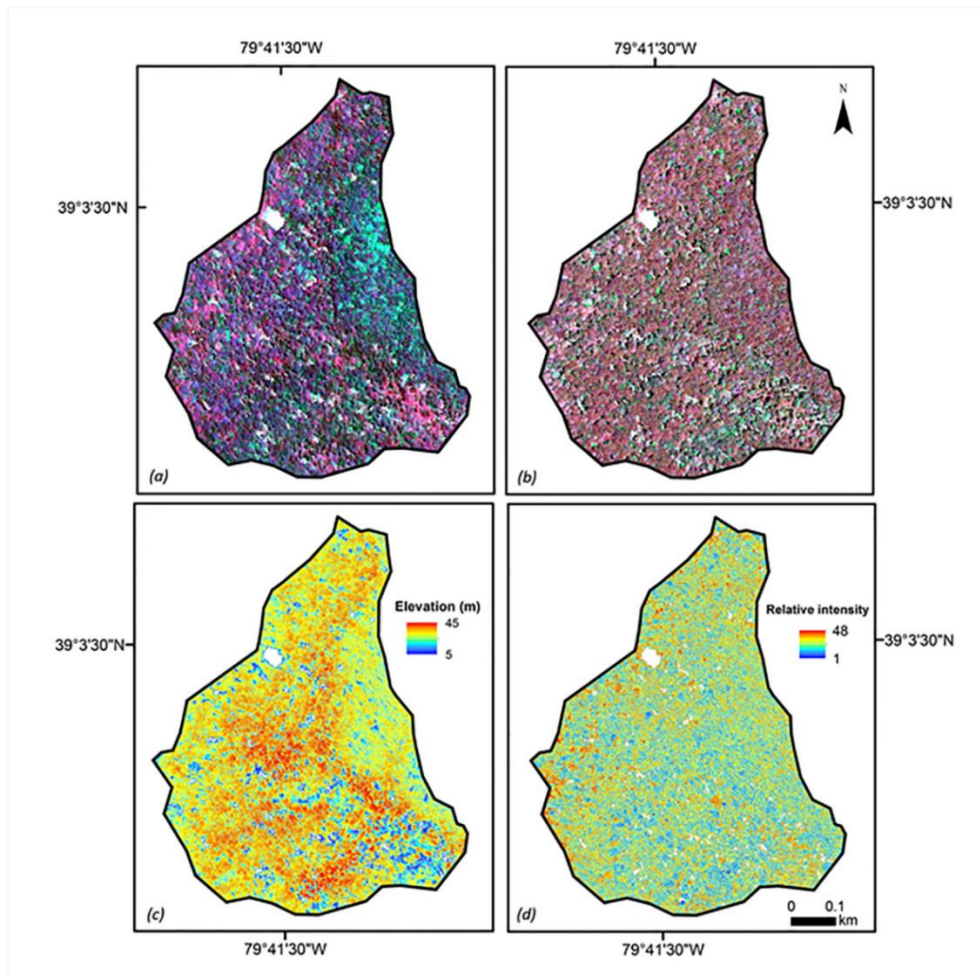
### 2.2.3 Field measurements

We completed all the field mapping of tree crowns during May 2015 using 0.2 hectare (0.5 acre) permanent forest plots within the two selected reference compartments at the Fernow: watershed 4 (WS 4) and the biological control area (BCA) (**Figure 2.1**). For each tree reaching the canopy in these plots, we recorded the species name, diameter at breast height (DBH), and tree height as measured with a Nikon Forestry Pro Laser Rangefinder. We also used the laser range finder to precisely map the distance and bearing of each stem from the plot centroids. We recorded the location of all centre points with a survey grade post-processed GPS unit ( $<0.5$  m error). Following these measurements, we hand-sketched the crown extents of each mapped tree to aid our manual crown delineation process conducted with the imagery datasets.

### 2.2.4 Imagery datasets

The LiDAR data were collected on 20 July 2014 using an aircraft flying at 915 m with an average speed of  $250 \text{ km hour}^{-1}$ , and an Optech ALTM 3100 LiDAR sensor with a pulse rate frequency of 100 kHz. Intensity data from this sensor's near-infrared laser beam has been found to correlate strongly with measured spectral reflectance at the laser's 1064 nm wavelength (Ahokas et al. 2006, Vain et al. 2009). The LiDAR data, collected with  $>50\%$  overlap between flight-lines, provided an average of 6.6 returns  $\text{m}^{-2}$ . We focused on LiDAR variables that are routinely calculated using standard GIS and remote sensing software in a raster environment. First-return data were rasterized to produce a Digital Surface Model (DSM), and last returns to produce a Digital Elevation Model (DEM). Rasterization was done in ArcMap 10.5 with a common 0.5 m cell size by assigning the average of all points within a cell as the elevation and linear interpolation as void fill method. We created a Canopy Height Model (CHM) from the difference between the DSM and the DEM. We rasterized the LiDAR intensity data using all the returns in each 0.5 m pixel.

We selected three high-resolution satellite images to best characterize phenological differences among species (**Figure 2.1** and **2.3** and **Table 2.1**). Visual inspection of the images indicates that the image dates correspond to three key phenology phases within broadleaf deciduous forests: the leaf emergence and flowering phase of spring (**Figure 2.3a**), the full-canopy development of summer (**Figure 2.3b**), and the leaf senescence and abscission of autumn (**Figure 2.3c**). All three sensors, GeoEye, Pleiades and WorldView2 provide approximately 2 m multispectral (MS) bands, as well as an approximately 0.5 m panchromatic band. Our spring and summer images have four MS bands, and the autumn image has eight bands (**Table 2.2**).



**Figure 2.3** The Fernow Watershed #4 portion of the study area. (a) Spring image. (b) Summer image. (c) LiDAR canopy height model. (d) LiDAR intensity. (a) and (b) are standard false color infrared composites

(bands NIR, Red, Green displayed as RGB). In each image, pixels with a canopy height less than 5 m are masked (and displayed in white) in order to highlight the variability among intact canopies. Compare to Figure 1, which shows the autumn phenology image.

**Table 2.1** Metadata of the imagery used in the study

Sensor	Acquisition date	Phenological period	Off-nadir view angle (°)	Pan band resolution (m)	Multispectral band resolution (m)
<b>LiDAR</b>	20-Jul-14	Summer	0-25		
<b>GeoEye</b>	6-Jun-14	Summer	8	0.42	1.67
<b>Pleiades</b>	15-May-15	Spring	20.6	0.7	2.8
<b>WorldView2</b>	26-Oct-14	Autumn	26.7	0.57	2.28

**Table 2.2** Image band wavelengths

Sensor	Wavelength (nm)									
	Pan	Coastal	Blue	Green	Yellow	Red	Red Edge	NIR1	NIR2	LiDAR laser
<b>LiDAR</b>										1064
<b>GeoEye</b>	450-800		450-510	510-580		655-690		780-920		
<b>Pleiades</b>	470-830		430-550	500-620		590-710		740-940		
<b>WorldView2</b>	450 - 800	400 - 450	450-510	510 - 580	585 - 625	630-690	705-745	770-895	860-1040	

## 2.2.5 Image pre-processing

We separately registered the multispectral and panchromatic images to the base layer of the LiDAR Canopy Height Model using Erdas Imagine 2016 and ArcMap 10.5. We applied a rational polynomial coefficient (RPC) sensor model (Toutin 2004) with elevation information and with ground control points (GCPs) to achieve a low RMSE (<1 pixel) for image registration (**Appendix A, Table S1**). Next, for each GeoEye, Pleiades and WorldView 2 multispectral and panchromatic band, we used the image metadata (IMD file) to convert digital numbers to ground reflectance using the CosT approach (Chávez 1996).

## 2.2.6 Crown delineation using field and imagery data

The focus of this work is on object-based classification. In this paper we only explore the issue of classification, and not errors due to segmentation issues (Liu and Xia 2010). Therefore, we used manually delineated crown polygons to ensure the highest accuracy in the segmentation. We visually compared tree crown field maps with high-resolution LiDAR height data, as well as panchromatic data

layers from the spring, summer, and autumn images. From these data sources, we manually delineated the non-overlapping polygons as 267 mapped tree crowns. In our analysis, we only attempt to discriminate crowns from the four most abundant tree species. Thus, our analysis dataset includes 57 red oak, 32 sugar maple, 26 tulip poplar, and 20 black cherry crowns for a total of 135 total tree crowns.

### **2.2.7 Variable extraction**

By computing the zonal average within each delineated tree crown, we extracted six groups of variables (**Table 2.3**): spectral reflectance from all multispectral bands in the (1) spring, (2) summer, and (3) autumn images, (4) vegetation indices derived from the three multispectral datasets, (5) canopy height and intensity from the LiDAR imagery, and (6) texture related variables from the panchromatic and LiDAR datasets. For vegetation indices, we computed the normalized difference vegetation index (NDVI) using the red and NIR multispectral bands from the spring, summer and autumn images. Using the eight spectral bands of WorldView 2 imagery, we also calculated additional vegetation indices on the autumn image: red edge and yellow band ratio (Waser et al. 2014), plant senescence index (Omer et al., 2016), NIR and yellow band ratio (Waser et al. 2014) and enhanced vegetation Index (EVI) (Omer et al. 2016) (**Table 2.3**).

Textural characteristics may vary based on several aspects of crown architecture that may differ among species, including internal shading, leaf orientation, as well as leaf density and size (L.Sayn-Wittgenstein 1978). We used the panchromatic reflectance of the spring, summer, and autumn images, as well as LiDAR height and intensity data layers to compute four types of texture variables (Haralick and Shanmugam 1973) based on the grey-level co-occurrence matrices (GLCM). The texture was calculated within eCognition. The eCognition GLCM texture averages the four directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$  at an offset of 1 pixel (Trimble 2011). The four texture features are as follows: (1) Homogeneity, which describes the similarity of pixel values within the local region; a high value indicates a more homogeneous region. (2) Contrast, which summarizes the variation of pixel values (exponentially) and a

high value indicates a locally more heterogeneous region. (3) Entropy, which is a measure of the disorder or randomness in the image values. (4) Dissimilarity, which is a measure of the heterogeneity of pixel values within the local region linearly (Warner 2011).



**Table 2.3** Variables used in the analysis

Group	Feature name	Description*	Sensor/source	Reference
Spectral reflectance	Reflectance 1-4	Average reflectance band 1-4	GeoEye	Pu and Landry 2012
	Reflectance 1-4	Average reflectance band 1-4	Pleiades	
	Reflectance 1-8	Average reflectance band 1-8	WorldView2	
	Brightness 1-3	Mean value of the means of all bands	GeoEye, Pleiades and WorldView 2	
LiDAR	Reflectance panchromatic band 1-3	Average panchromatic reflectance	GeoEye, Pleiades and WorldView 2	Waser et al. 2014
	Height	Average height	LiDAR CHM	
	Intensity	Average intensity	LiDAR intensity image	
	RedEdge yellow ratio	$(R_{\text{RedEdge}} - R_{\text{Yellow}}) / (R_{\text{RedEdge}} + R_{\text{Yellow}})$	WorldView 2	
Vegetation indices	Plant Senescence 1	$(R_{\text{Red}} - R_{\text{Blue}}) / R_{\text{NIR1}}$	WorldView 2	Omer et al. 2016
	Plant Senescence 2	$(R_{\text{Red}} - R_{\text{Blue}}) / R_{\text{RedEdge}}$	WorldView 2	Waser et al. 2014
	NIR yellow ratio	$(R_{\text{NIR2}} - R_{\text{Yellow}}) / (R_{\text{NIR2}} + R_{\text{Yellow}})$	WorldView 2	Waser et al. 2014
	NDVI WV2	$(R_{\text{NIR}} - R_{\text{Red}}) / (R_{\text{NIR}} + R_{\text{Red}})$	WorldView2	Karlson 2016
	NDVI P	$(R_{\text{NIR}} - R_{\text{Red}}) / (R_{\text{NIR}} + R_{\text{Red}})$	Pleiades	Omer et al. 2016
	NDVI G	$(R_{\text{NIR}} - R_{\text{Red}}) / (R_{\text{NIR}} + R_{\text{Red}})$	GeoEye	
	EVI	$2.5 \times ((R_{\text{NIR1}} - R_{\text{Red}}) / (R_{\text{NIR1}} + 6 \times R_{\text{Red}} - 7.5 \times R_{\text{Blue}} + 1))$	WorldView 2	
Texture	GLCM homogeneity	GLCM homogeneity	Panchromatic band of WorldView 2, Pleiades, GeoEye; and LiDAR CHM and intensity	Pu and Landry 2012
	GLCM contrast	GLCM contrast	Panchromatic band of WorldView 2, Pleiades, GeoEye; and LiDAR CHM and intensity	
	GLCM entropy	GLCM entropy	Panchromatic band of WorldView 2, Pleiades, GeoEye; and LiDAR CHM and intensity	
	GLCM dissimilarity	GLCM dissimilarity	Panchromatic band of WorldView 2, Pleiades, GeoEye; and LiDAR CHM and intensity	

In total, we extracted 52 variables for each individual crown. These include 11 variables from the spring Pleiades image, 11 variables from the summer GeoEye image, 20 variables from the autumn WorldView2 image, (Extra variables from WorldView2 were calculated from four additional bands compared with GeoEye) and 10 variables from LiDAR (**Table 2.3**).

### **2.2.8 Data Analysis**

To avoid problems due to the “curse of dimensionality” (Hughes 1968) associated with discriminating 4 species from 52 variables, we selected the best subset of variables before classification in R (R Core Development Team 2016). We used rfUtilities package (Evans and Murphy 2014) to select variables based on variable importance. We used caret package (Kuhn et al. 2016) to run a step-wise RF with the best 10%, 15%, 25%, 50%, 75%, and 90% of the variables retained. We generated 500 trees for each RF model. Based on prior experience which suggested any number greater than 100 was sufficient; we used 500 to be sure the number was large enough. The optimal number of variables available for splitting at each node (mtry) was selected using 10-fold cross-validation and ten different values were assessed. This 10-fold cross validation randomly split data into 10 subsets; 9 of them are used for training and 1 retained for testing each time. The average was calculated in the model as final accuracy (Maxwell, Warner and Fang 2018). From this, we selected the top 50% of the variables as the greatest per cent that retained the most signal while minimizing redundancy and noise. The cross-fold validation approach was used so that classification trains a classifier with a number of samples that is similar to the number of variables. Accuracy should be evaluated based on data entirely separate from that used in developing the model. RF does offer estimates of accuracy based on data withheld from subsets of the model used in developing the trees (so-called out of bag estimates), but we prefer to use entirely separate data for both overall accuracy assessment and the confusion matrices (Maxwell, Warner and Fang 2018). Using the six groups of variables described in section 2.6, we conducted a preliminary test to evaluate the accuracy of RF models that only drew

from an individual group of variables (see **Figure S1** and **Table S2** in the **Appendix A**). As we expected, the Kappa values obtainable from a single group were markedly lower than combining groups, so we focused the remainder of our analysis on testing our hypotheses concerning the combination of spectral features with structural features to enhance tree species discrimination.

To test how each group of variables enhanced model accuracy, we iteratively removed each of the six groups of variables and recorded the effect on the Kappa value derived from the 10-fold cross-validation of the RF model. These RF models used 31 features with spring image variables excluded, 31 features with summer image variables excluded, 21 features with autumn image variables excluded, 32 features with LiDAR variables excluded, 30 features with texture features excluded, and 32 features with vegetation indices excluded. Finally, in order to understand the patterns and mechanisms of how the variables used in the RF modelling are useful individually and in combination for discriminating tree species, we used ANOVA (assumptions tested in **Appendix A, Table S3**), paired *t*-tests, and unpruned decision trees generated in R with the *rpart* package (Therneau, Atkinson and Ripley 2017). An ANOVA test can help to test the difference among group means with the hypothesis that the group means are equal. *F*-ratio is the ratio of between groups to within-group variance. Group means are significantly different when the *p*-value (significance) is less than 0.05 and *F*-ratio is larger than 1.

A single decision tree classification tends to be less accurate than RF, which employs an ensemble of trees (Maxwell, Warner and Fang 2018). Nevertheless, single decision tree has the advantage of being simple and intuitive to interpret. Because our main aim in this study was to explore the importance of different data layers, a decision tree classification is therefore used as a supplement to the RF classification. The decision tree classification was carried out with the *rpart* package (Therneau, Atkinson and Ripley 2017). The tree was automatically pruned based on *cp*=0.01, *minsplit*=20, *maxcompete*=4.

## 2.3 Results

### 2.3.1 Data combinations and random forest model accuracy

The Random Forest model achieved a Kappa value of 42.4% (with an overall accuracy (OA) of 62%) when we ran it with the top 50<sup>th</sup> percentile of all the variables from all groups. Notably, the variables in this top 50<sup>th</sup> percentile were drawn from each of the six groups of variables (**Table 2.4**). When we iteratively excluded groups of variables, we found that removing each group caused a notable decrease in Kappa value compared with results from a full model (**Figure 2.4**). Removing variables from the summer GeoEye image caused the largest percentage (6.1%) drop in Kappa value, from 42.4% with the full model to 36.3%. The LiDAR variables were the next most useful group of variables, reducing Kappa by 5.1% when they were excluded. The four other groups of variables all caused smaller decreases in Kappa value, but still reduced model performance considerably. The overall magnitude of reduced performance in these four other groups was between 2.3%-3.9% in terms of Kappa value, corresponding to a 5-10% relative reduction in the model performance in comparison to the full model.

The confusion matrix from RF using the top 50<sup>th</sup> percentile of variables indicates that the sample size of each species likely influenced the results. Red oak ( $N = 57$  crowns) had the largest producers (PA) and users accuracy (UA), while black cherry ( $N = 20$  crowns) had the lowest accuracy (**Table 2.5**). Confusion matrices from other RF models highlight substantial drops in accuracy for individual species when certain of groups of variables were withheld (**Appendix A, Table S4**). This was especially evident for sugar maple, which fell from a 58% UA and 48% PA in the 50<sup>th</sup> percentile RF model to a 40% UA and 31% PA in the model with the LiDAR data withheld (**Appendix A, Table 4(b)**).

### 2.3.2 Useful variables identified by ANOVA and a decision tree

ANOVA analyses revealed that LiDAR intensity had the strongest individual ability to discriminate tree species (**Table 2.4**), and further analysis using pair-wise *t*-tests revealed that sugar maple had significantly higher intensity than other species (**Figure 2.5 (a)**). Reflectance in the blue wavelengths during the spring and summer was among the next strongest variables (**Table 2.4**), and *t*-tests revealed that red oak crowns had lower summer blue reflectance and higher spring blue reflectance (**Figures 2.5 (b) & (c)**). The vegetation indices also proved to be useful variables (**Table 2.4**), particularly the plant senescence index which indicated a notably higher index value for tulip poplar during the autumn phenology image acquisition (**Figure 2.5 (d)**).

**Table 2.4** Features selected by Random Forest, feature selection sorted by decreasing *F*-value obtained from ANOVA tests of species differences

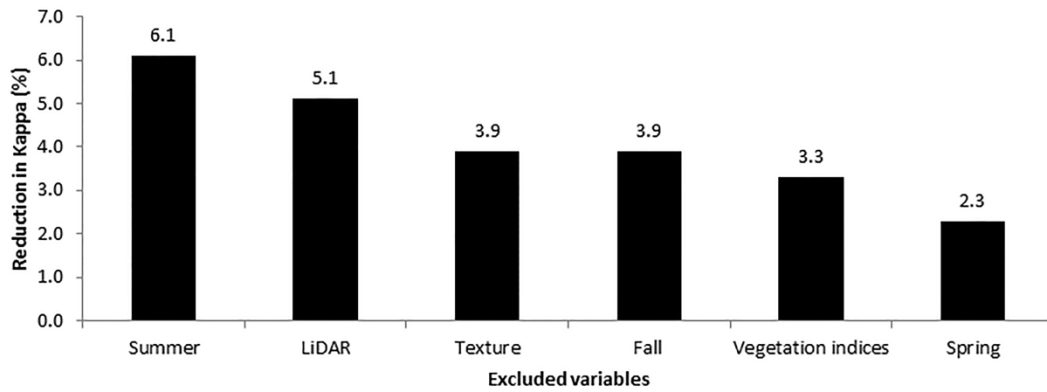
Feature name	Significance <sup>1</sup>	<i>F</i> -ratio	Variable groups <sup>2</sup>
Average intensity	<0.001	10.1	LI
Spring blue reflectance	<0.001	9.22	SP
Summer NDVI	<0.0001	9.02	SU&VI
Spring NDVI	<0.0001	8.2	SP&VI
Spring NIR reflectance	<0.0001	7.93	SP
Summer red reflectance	<0.001	7.75	SU
GLCM entropy of intensity	<0.0001	7.72	LI&TE
Summer blue reflectance	0.0002	7.2	SU
Autumn GLCM entropy	0.0002	7.16	AU&TE
Spring GLCM entropy	0.0002	7.11	SP&TE
Spring green reflectance	0.0003	6.79	SP
Spring Brightness	0.0003	6.79	SP
Summer GLCM entropy	0.0005	6.33	SU&TE
LiDAR GLCM entropy	0.0013	5.53	LI&TE
Autumn blue reflectance	0.002	5.21	AU
Autumn Plant Senescence	0.0022	5.13	AU &VI
Autumn NIR and yellow ratio	0.0031	4.84	AU &VI
Autumn coastal reflectance	0.0033	4.81	AU
Autumn Red Edge and yellow ratio	0.0075	4.16	AU &VI
Autumn NIR2 reflectance	0.0096	3.96	AU
Autumn red reflectance	0.0108	3.88	AU
NIR1 reflectance	0.013	3.73	AU
Spring red reflectance	0.014	3.67	SP
Autumn Brightness	0.0172	3.51	AU
Autumn NDVI	0.0218	3.33	AU &VI
Autumn Enhanced vegetation index	0.0269	3.16	AU &VI
Autumn Red Edge reflectance	0.0343	2.97	AU
LiDAR GLCM homogeneity	0.0346	2.96	LI&TE
Autumn Panchromatic reflectance	0.0406	2.84	AU
Spring Panchromatic reflectance	0.0427	2.8	SP
LiDAR Tree height	NS	2.58	LI
Intensity GLCM contrast	NS	1.76	LI&TE
Intensity GLCM dissimilarity	NS	1.59	LI&TE
Autumn yellow reflectance	NS	1.32	AU
Summer NIR reflectance	NS	1.23	SU
Summer GLCM dissimilarity	NS	0.79	SU&TE
Summer Brightness	NS	0.63	SU
Autumn green reflectance	NS	0.52	AU
Autumn GLCM homogeneity	NS	0.13	AU &TE

<sup>1</sup>NS: not significant

<sup>2</sup>SP: Spring; SU: Summer; AU: Autumn; LI: LiDAR; TE: Texture;; VI: Vegetation indices

Most of these useful variables were also selected by the decision tree analysis. The most parsimonious decision tree used 8 variables. As with the RF model, the decision tree notably

contained variables derived from each of the four image datasets (spring, summer, autumn, and leaf-on LiDAR) (**Figure 2.6**). Summer blue reflectance was the initial node, followed by blue and brightness bands from the spring image. Three spectral indices describing species differences in the autumn senescence, and two LiDAR variables describing differences in intensity and entropy of height provided the final nodes discriminating all four species (**Figure 2.6**).

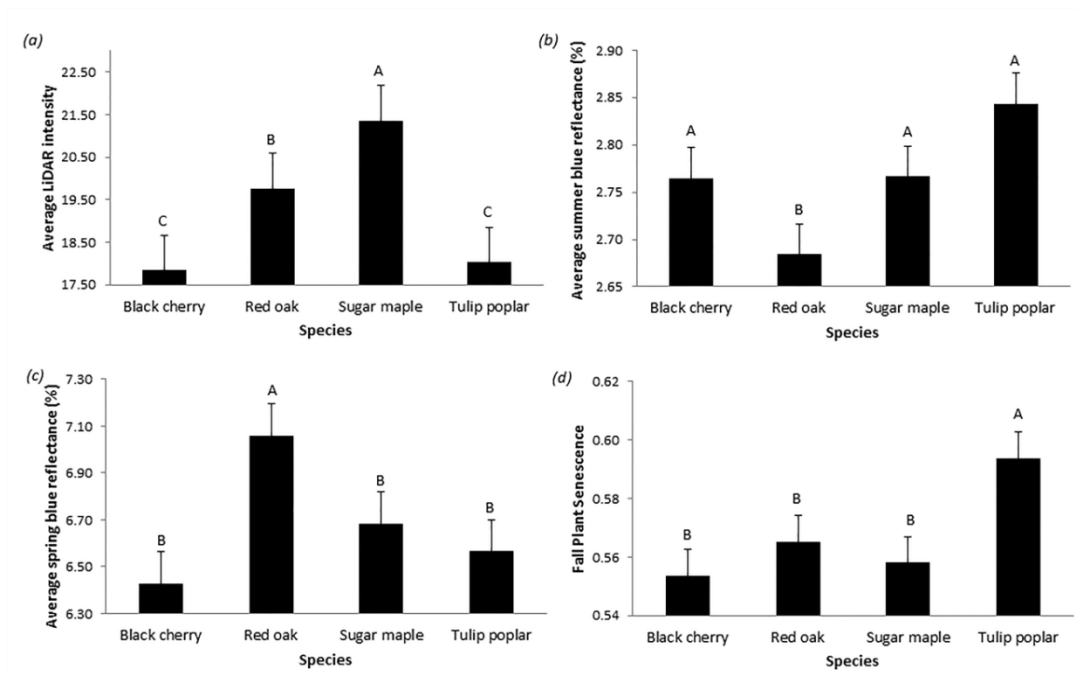


**Figure 2.4** Per cent of Kappa value lost by removing each group of variables

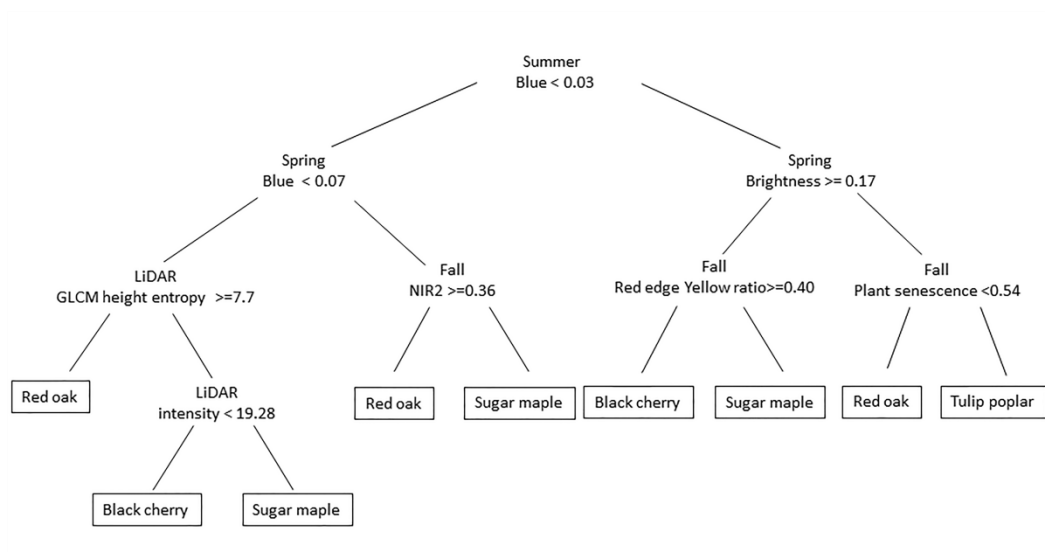
**Table 2.5** Confusion matrix from RF classification of individual trees with top 50<sup>th</sup> percentile of variables.

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	2	0	1	1	4	50
Red oak	10	50	14	5	79	63
Sugar maple	3	5	15	3	26	58
Tulip poplar	5	2	2	17	26	65
Total	20	57	32	26	135	
PA (%)	10	88	47	65		
						OA (%)=62

PA: Producer's accuracy; UA: User's accuracy; OA: Overall accuracy



**Figure 2.5** Key results from the ANOVA tests, letters denote significant differences in means, as assessed by pair-wise t-tests



**Figure 2.6** Decision tree based on 50th percentile of variables. Affirmative logical decisions are in each case to the left.



## 2.4 Discussion

Our results strongly support the hypothesis that the combination of phenology and structural information from individual tree crowns can enhance the discrimination of tree species. First, our RF results supported the hypothesis that each group of variables would enhance model performance (**Figure 2.4**). Our separate analysis using a decision tree supported this combination because all three phenology images and LiDAR were included in the most parsimonious decision tree (**Figure 2.6**). However, the relative degree to which the groups of variables affected model accuracy differed from some previous studies. For instance, whereas we found the summer and spring phenology periods to contain the most useful variables for species discrimination (**Figure 2.4**), others have found a peak autumn image to be most useful (e.g. Key et al. 2001). Even so, reflectance and spectral indices from the autumn image were selected by RF (**Table 2.3**), and also formed key parts of the decision tree analysis (**Figure 2.6**). Thus, our data still indicate strong support for including autumn phenology in species discrimination methods. The RF analysis also indicated that LiDAR-derived variables were the second most useful group of variables, which adds support to a growing number of studies emphasizing the importance of LiDAR information for tree species classification (Donoghue et al. 2007, Kim 2007, Dalponte et al. 2008, Voss and Sugumaran 2008, Jones et al. 2010, Korpela et al. 2010, Ghosh et al. 2014, Eitel et al. 2016).

Our ANOVA and decision tree analyses suggest several mechanisms by which spectral and structural information help discriminate broadleaf deciduous tree species based on their unique phenology and crown architectural traits. Specifically, the spectral information in the visible wavelengths helps to describe the unique phenology of leaf pigments and photosynthetic activity in each tree species, while the phenology of NIR spectral reflectance and LiDAR describe essential species differences in tree crown architecture (Gates et al. 1965, Asner 1998, Mohammed et al. 2000, Rautiainen et al. 2004, Ollinger 2011, Eamus et al. 2016). First, our observations that red oak has a high spring blue reflectance (**Figure 2.5 (c)**), low summer blue

reflectance (**Figure 2.5 (b)**), and a high autumn NIR reflectance (**Figure 2.6**) matches with field observations that red oak crowns tend to have a later leaf emergence in the spring, grow dense crowns of photosynthetically active leaves in the summer, and senesce later in the autumn relative to other species. High GLCM entropy of height as indicated by the decision tree differentiation of red oaks (**Figure 2.6**) suggests that red oak has relatively more within-crown gaps than other species. Next, a higher plant senescence index for tulip poplar (**Figure 2.6**) matches field observations of an earlier autumn leaf senescence in tulip poplar than other species (particularly red oak). Finally, coupled with observations of its lower entropy of height, observations of high summer LiDAR intensity (**Figure 2.5 (a) and 2.6**) indicate that sugar maple has a flatter and more horizontal crown architecture (Kim et al. 2009, Fassnacht et al. 2016, Budei et al. 2018). The importance of LiDAR for characterizing this unique crown architecture of sugar maple was borne out by the large drop in classification accuracy when LiDAR variables were removed from the RF model (**Table 2.4** and **Appendix A Table 3 (b)**).

Although the overall accuracies obtained in this study were lower than might be useful for mapping purposes, we emphasize that the ultimate purpose of this study was to provide guidance in data selection for future mapping, and not necessarily to produce a map directly from these data (see, for example, Key et al. 2001, for a similar research design). Indeed, the strong complementarity of spectral and structural information in our study illustrates how the combination of LiDAR and multi-temporal images can contribute to tree species discrimination, especially in forests where trees have similar functional types and reflectance characteristics (Dalponte et al. 2008, Jones et al. 2010, Alonzo et al. 2014). Yet, the overall accuracy of our discrimination algorithm (Kappa = 42%) highlights that substantial obstacles still remain toward an ultimate goal of automatic crown-level species discrimination from remotely sensed data. Our study exemplifies the challenge of these obstacles. First, our study design relied on available archived data. While we were able to match archive data to key phenology periods, it was

challenging to precisely co-register the images that had less than ideal viewing geometries (**Table 2.1**), and were collected over a topographically complex study site. Second, study areas like the Fernow Experimental Forest are especially challenging for discriminating tree species because the trees (1) have similar growth forms, leaf types and leaf habits, and (2) have high within-species variance in phenology and crown architecture linked to strong competition for light within a diverse forest on steeply sloping terrain. Given these unique characteristics of our study design and study area, we suggest that our classification accuracy and evidence of data combination are quite conservative results. As such, in more forgiving study contexts, we would expect substantially higher accuracies and even stronger evidence for the advantage of combinations of multispectral and LiDAR data.

## **2.5 Conclusion**

The significant contribution of this paper is that combining multi-temporal imagery with leaf-on LiDAR can enhance the ability to discriminating tree species based on their phenology and crown architectural characteristics. Specifically, our results reveal that the combination of multi-temporal high-resolution images with leaf-on LiDAR data improves the discrimination of four broadleaf deciduous tree species: red oak, sugar maple, black cherry and tulip poplar. We tested six groups of variables for their classification performance: (1) spring, (2) summer, and (3) autumn phenology, (4) height and intensity data from leaf-on LiDAR, (5) spectral indices and (6) texture information. When a specific group of variables was removed, the decrease in Random Forest classification accuracy demonstrates the importance of such variables. In addition, our ANOVA and decision tree results illustrate specific mechanisms to discriminate species based on their unique phenology and crown architectural properties. Continuing to build on these findings can provide a robust path toward the ultimate goal of automatic crown-level tree species discrimination.

The RF classification indicated that the Kappa value dropped each time a specific group of variables was removed. This strongly indicates all three seasons of high-spatial resolution multi-temporal satellite data and leaf-on LIDAR enhanced the ability to discriminate tree species. We find that summer phenology is most helpful to classify tree species since our model is most affected (declined in Kappa accuracy of 6.1%) by removing summer phenology features. Crown structural and architecture features from LiDAR were the second best variables for discriminating tree species. Next, texture features, fall phenology and vegetation indices showed similar importance for tree species classification. Variables from a spring image were the least helpful to classify tree species, but they still cause 2.3% drop in Kappa value when excluded.

The decision tree and ANOVA results provide additional evidence regarding how these imagery sources combine to identify the distinct spectral and structural properties for each tree species. We found blue band reflectance from multi-temporal remote sensing imagery to be a useful variable in describing unique phenological and structural features associated with leaf emergence. Autumn NIR may be important for capturing differences for red oak associated with leaf senescence. Moreover, for crown architectural properties, relative lower summer blue reflectance indicates dense crowns; higher entropy of height suggests more within-crown gaps in red oak. In another abundant species, sugar maple, we also found that higher summer LiDAR intensity corresponds with flatter crown architecture (Lower height entropy). Thus, for discriminating broadleaf deciduous tree species, this study supports the utility of high spatial multi-temporal satellite images in the spring, summer and fall to capture the most distinctive phenology patterns while also supporting the use of leaf-on LiDAR for capturing key differences in tree crown architecture.

## Appendix A.

**Table S1** RMSE for all the geo-referenced images in meters.

Sensor	MS RMSE (m)	PAN RMSE (m)
<b>GeoEye</b>	0.76	0.46
<b>Pleiades</b>	0.15	0.59
<b>WorldView2</b>	0.0056	0.86

**Table S2 (a)** Confusion matrix from RF classification with only summer variables

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	2	0	1	3	6	33
Red oak	9	50	17	9	85	59
Sugar maple	4	5	10	4	23	43
Tulip poplar	5	2	4	10	21	48
Total	20	57	32	26	135	
PA (%)	10	88	31	38		
						OA (%)=53

**Table S2 (b)** Confusion matrix from RF classification with only LiDAR variables

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	3	1	1	4	9	33
Red oak	9	39	16	11	75	52
Sugar maple	3	7	14	3	27	52
Tulip poplar	5	10	1	8	24	33
Total	20	57	32	26	135	
PA (%)	15	68	44	31		
						OA (%)=47

**Table S2 (c)** Confusion matrix from RF classification with only texture variables

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	3	3	1	1	8	38
Red oak	12	38	16	10	76	50
Sugar maple	3	6	8	5	22	36
Tulip poplar	2	10	7	10	29	34
Total	20	57	32	26	135	
PA (%)	15	67	25	38		
						OA (%)=44

**Table S2 (d)** Confusion matrix from RF classification with only autumn variables

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	4	0	0	1	5	80
Red oak	12	44	11	9	76	58
Sugar maple	1	8	13	6	28	46
Tulip poplar	3	5	8	10	26	38
Total	20	57	32	26	135	
PA (%)	20	77	41	38		
						OA (%)=56

**Table S2 (e)** Confusion matrix from RF classification with only vegetation indices

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	2	2	1	1	6	33
Red oak	15	47	17	10	89	53
Sugar maple	1	5	11	2	19	58
Tulip poplar	2	3	3	13	21	62
Total	20	57	32	26	135	
PA (%)	10	82	34	50		
						OA (%)=54

**Table S2 (f)** Confusion matrix from RF classification with only spring variables

Classification	Reference				Total	UA (%)
	Black cherry	Red oak	Sugar maple	Tulip poplar		
Black cherry	6	3	4	4	17	35
Red oak	6	45	12	9	72	63
Sugar maple	3	4	13	6	26	50
Tulip poplar	5	5	3	7	20	35
Total	20	57	32	26	135	
PA (%)	30	79	41	27		
						OA (%)=53

**Table S3.** ANOVA assumption test, based on errors normally distributed with equal variance.

<b>Feature name</b>	<b>Normality</b>	<b>Homogeneity of variance</b>	<b><i>p</i> -value</b>	<b><i>F</i> -ratio</b>
Average intensity	Yes	Yes	<0.001	10.1
Spring blue reflectance	No	Yes	<0.001	9.22
Summer NDVI	No	Yes	<0.0001	9.02
Spring NDVI	Yes	Yes	<0.0001	8.2
Spring NIR reflectance	No	Yes	<0.0001	7.93
Summer red reflectance	Yes	Yes	<0.001	7.75
Intensity GLCM entropy	Yes	Yes	<0.0001	7.72
Summer blue reflectance	Yes	Yes	0.0002	7.2
Autumn GLCM entropy	Yes	Yes	0.0002	7.16
Spring GLCM entropy	Yes	Yes	0.0002	7.11
Spring green reflectance	Yes	Yes	0.0003	6.79
Spring Brightness	No	Yes	0.0003	6.79
Summer GLCM entropy	Yes	Yes	0.0005	6.33
LiDAR GLCM entropy	Yes	Yes	0.0013	5.53
Autumn blue reflectance	Yes	Yes	0.002	5.21
Autumn Plant Senescence	Yes	Yes	0.0022	5.13
Autumn NIR and yellow ratio	Yes	No	0.0031	4.84
Autumn coastal reflectance	Yes	Yes	0.0033	4.81
Autumn Red Edge and yellow ratio	Yes	No	0.0075	4.16
Autumn NIR2 reflectance	Yes	Yes	0.0096	3.96
Autumn red reflectance	Yes	No	0.0108	3.88
NIR1 reflectance	Yes	Yes	0.013	3.73
Spring red reflectance	Yes	Yes	0.014	3.67
Autumn Brightness	Yes	No	0.0172	3.51
Autumn NDVI	Yes	Yes	0.0218	3.33
Autumn Enhanced vegetation index	Yes	Yes	0.0269	3.16
Autumn Red Edge reflectance	Yes	Yes	0.0343	2.97
LiDAR GLCM homogeneity	No	Yes	0.0346	2.96
Autumn Panchromatic reflectance	No	No	0.0406	2.84
Spring Panchromatic reflectance	No	Yes	0.0427	2.8
LiDAR Tree height	No	Yes	NS	2.58
Intensity GLCM contrast	No	Yes	NS	1.76
Intensity GLCM dissimilarity	Yes	Yes	NS	1.59
Autumn yellow reflectance	Yes	No	NS	1.32
Summer NIR reflectance	Yes	Yes	NS	1.23
Summer GLCM dissimilarity	Yes	Yes	NS	0.79
Summer Brightness	Yes	Yes	NS	0.63
Autumn green reflectance	No	No	NS	0.52
Autumn GLCM homogeneity	Yes	Yes	NS	0.13



**Table S4 (a)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables exclude summer variables

	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	3	0	1	1	5	60
Red oak	11	48	12	10	81	59
Sugar maple	2	7	16	3	28	57
Tulip poplar	4	2	3	12	21	57
Total	20	57	32	26	135	
PA (%)	15	84	50	46		
						OA (%)=59

**Table S4 (b)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables excludes LiDAR variables

	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	5	1	4	2	12	42
Red oak	8	45	16	5	74	61
Sugar maple	4	7	10	4	25	40
Tulip poplar	3	4	2	15	24	63
Total	20	57	32	26	135	
PA (%)	25	79	31	58		
						OA (%)=56

**Table S4 (c)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables excludes texture variables

	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	5	3	2	3	13	38
Red oak	8	44	13	5	70	63
Sugar maple	4	6	15	3	28	54
Tulip poplar	3	4	2	15	24	63
Total	20	57	32	26	135	
PA (%)	25	77	47	58		
						OA (%)=59

**Table S4 (d)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables excludes autumn variables

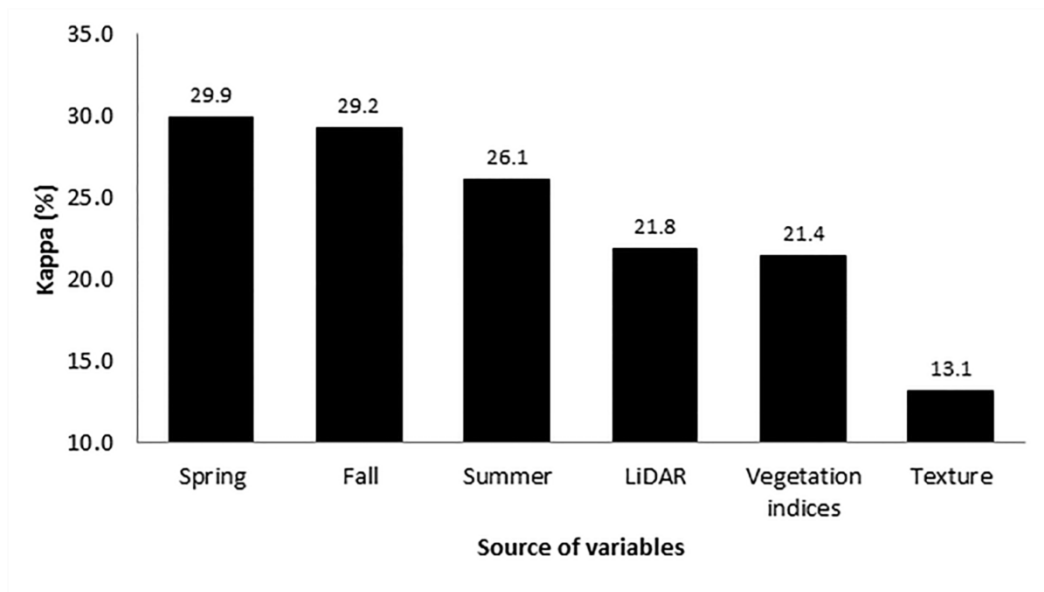
	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	4	2	2	2	10	40
Red oak	6	48	12	6	72	67
Sugar maple	5	3	14	5	27	52
Tulip poplar	5	4	4	13	26	50
Total	20	57	32	26	135	
PA (%)	20	84	44	50		
						OA (%)=59

**Table S4 (e)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables excludes vegetation indices

	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	6	2	2	2	12	50
Red oak	5	47	12	4	68	69
Sugar maple	5	5	15	9	34	44
Tulip poplar	4	3	3	11	21	52
Total	20	57	32	26	135	
PA (%)	30	82	47	42		
						OA (%)=59

**Table S4 (f)** Confusion matrix from RF classification with top 50<sup>th</sup> percentile of variables excludes spring variables

	Reference					
Classification	Black cherry	Red oak	Sugar maple	Tulip poplar	Total	UA (%)
Black cherry	4	2	0	1	7	57
Red oak	9	47	12	6	74	64
Sugar maple	4	5	16	5	30	53
Tulip poplar	3	3	4	14	24	58
Total	20	57	32	26	135	
PA (%)	20	82	50	54		
						OA (%)=60



**Figure S1.** Kappa value by using each group of variables as a single source

### **3. Discriminating tree species using the DDOT Street Tree Spatial Database and multi-temporal WorldView-3 imagery from Washington D.C., USA\***

Fang Fang<sup>a</sup>, Brenden McNeil<sup>a</sup>, Timothy Warner<sup>a</sup>, Aaron Maxwell<sup>a</sup>, Gregory Dahle<sup>b</sup>, Earl Eutsler<sup>c</sup>

<sup>a</sup> Department of Geology and Geography, West Virginia University, Morgantown, WV, 26505, USA

<sup>b</sup> Davis College, Division of Forestry & Natural Resources, West Virginia University, Morgantown, WV, 26505, USA

<sup>c</sup> Urban Forestry Division, District Department of Transportation, Washington, DC, 20003, USA

<sup>d</sup> Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China

#### **Abstract**

With the promise of transformative changes for the management of rural and urban forests, the discrimination of tree species from satellite imagery has been a long-standing goal of remote sensing. For the species-rich urban setting of Washington, D.C. USA, we evaluate current prospects toward this goal by combining a RandomForest object-based tree species classification method with two large datasets 1) A suite of 12 very high resolution (VHR) WorldView-3 images (WV-3), whose image acquisition date cover each pheno-phase of the growing season from April to November; and 2) the Washington D.C. Department of Transportation's (DDOT) field inventory of almost 50,000 street trees. We classify the 19 most abundant tree species with an overall accuracy of 63.7% and classify the ten most abundant genera with an overall accuracy of 75.6%. We observe (1) a decline in classification accuracy with each added species (2) larger declines in accuracy when attempting to classify species in the same genus, and (3) images taken from spring leaf emergence and fall senescence phenological period have the most potential to discriminate trees at both species and genus level. Especially if satellite data can be matched to these key pheno-phases, our study highlights that current VHR satellite sensors now have the radiometric, spectral, and spatial resolution to begin serving in their long-awaited role as powerful tools for managing species-rich urban forests.

### 3.1 Introduction

Obtaining spatially accurate information of tree species is critical for sustainable forest management in rural and urban settings, where the large variety of trees provides economic, aesthetic, as well as ecosystem values (Miller 1996, Nowark et al. 2006). With the increased availability of high temporal, spatial and spectral resolution satellite data during the past two decades, researchers have made large efforts on collecting tree species information via remote sensing imagery (Key et al. 2001, Holmgren et al. 2008, Immitzer et al. 2012, Jensen et al. 2012, Pu and Landry 2012, Alonzo et al. 2014, Karlson et al. 2016, Wang et al. 2016, Madonsela et al. 2017, Fang et al. 2018). Yet, three key challenges remain: (1) a comprehensive assessment of phenology and its impact on spectral response of different species in tree species classification, (2) the availability of a sufficient forest reference inventory with a wide range of tree species in different taxonomic levels; and (3) a robust analysis method to discriminate a large numbers of tree species.

Researchers have investigated the effects of temporal resolution, or phenology, for tree species mapping using high spatial resolution images in both rural and urban settings (Key et al. 2001, Immitzer et al. 2012, Pu and Landry 2012, Karlson et al. 2016, Fang et al. 2018, Pu, Landry and Yu 2018, Hubert et al. 2019). They demonstrated the advantages of using multi-seasonal or multi-temporal images over using a single date of very high resolution (VHR) image for tree species classification (Key et al. 2001, Xiao Ustin and McPherson 2004, Li et al. 2015, Karlson et al. 2016, Madonsela et al. 2017, Fang et al. 2018, Pu, Landry and Yu 2018, Yan et al. 2018). Multi-seasonal VHR images are crucial for capturing key phenological development features of tree species during different transition periods. In urban settings, Li et al. (2015) reported an overall accuracy around 92% using bi-temporal VHR images from September and October to classify four urban tree species in parts of Beijing, China. Beside tree species mapping using bi-temporal data, Key et al. (2001) evaluated the potential of VHR images acquired in more than two seasons for tree species classification. By comparing images from four dates in May, June

and October, they recommended an autumn image (October) to reach the highest tree species classification accuracy. However, Pu, Landry and Yu (2018) found that an image from late spring (April) produced the highest classification accuracy for seven urban tree species compared with images from February, May, August and November, since April image captured the seasonal transition from dry-spring to wet-summer in Tampa, FL. These previous studies were limited to their image acquisition time, each of which only covered part of the phenophases from the growing season period from April to November. A more systematic study is therefore needed to fill the gap for a comprehensive understanding of the importance of phenology by using images taken from each of the phenophases covering the entire growing season for urban tree species classification.

Perhaps due to the limited sample size for calibration/validation in the forest inventory database, there is a wide range in the number of tree species and taxonomic levels included in remote sensing-based tree species mapping projects. (Xiao, Ustin and McPherson 2004, Holmgren et al. 2008, Immitzer et al. 2012, Ghosh et al. 2014, Karlson et al. 2016, Pu, Landry and Yu 2018). For example, Holmgren (2008) classified three types of trees (pine, spruce, and other deciduous trees) with an overall accuracy around 96%. Karlson et al. (2016) classified five dominant tree species in a mixed forest including both deciduous and evergreen trees via RandomForest (RF) classifier and the overall accuracy was 83%. Immitzer et al. (2012) identified ten tree species in a mixed forest and derived the overall accuracy around 82% from RF. Xiao, Ustin and McPherson (2004) classified three forest tree types (Conifer, broadleaf deciduous and broadleaf evergreen) and sixteen tree species using spectral mixture analysis with overall accuracy as 90% and 70% respectively; they also argued that the accuracy was highly affected by the crown size and the density of leaves. Among these studies, high accuracies were derived, but there is limited discussion on how the number or types of classes affect the accuracy of tree species mapping. In addition, the classification taxonomic level varies due to the complexity of tree species composition in urban settings. Some studies focused on classifying trees within the

deciduous forest (Pu and Landry 2012) while some focused on discriminating deciduous trees from conifers (Immitzer et al. 2012, Karlson et al. 2016). In order to avoid a bias due to the different number of tree species classified, in this study we concentrate on exploring how different taxonomic levels and number of classes affect the performance of tree species classification in the urban forest setting.

The recent advancement in machine learning algorithms has presented new prospects for object-based remote sensing image classification, as well as robust linkages leveraging remote sensing imagery and forest inventory database for tree species identification (Xiao, Ustin and McPherson 2004, Holmgren et al. 2008, Immitzer et al. 2012, Ghosh et al. 2014, Liu et al. 2015, Karlson et al. 2016, Fang et al. 2018, Pu, Landry and Yu 2018). Non-parametric classifiers such as RF are often preferred for tree species mapping because of their typically higher prediction accuracy (Maxwell, Warner and Fang 2018). Moreover, the RF classifier does not require normally distributed input data. Our study uses RF classification to leverage 1) high-resolution remote sensing images covering the entire growing season and 2) a large forest inventory database for object-based tree species classification. We chose object-based image analysis since it significantly improved tree species classification accuracy by incorporating the inherent spectral variability within individual tree crowns (Im et al. 2014, Li et al. 2015, Karlson et al. 2016, Fang et al. 2018).

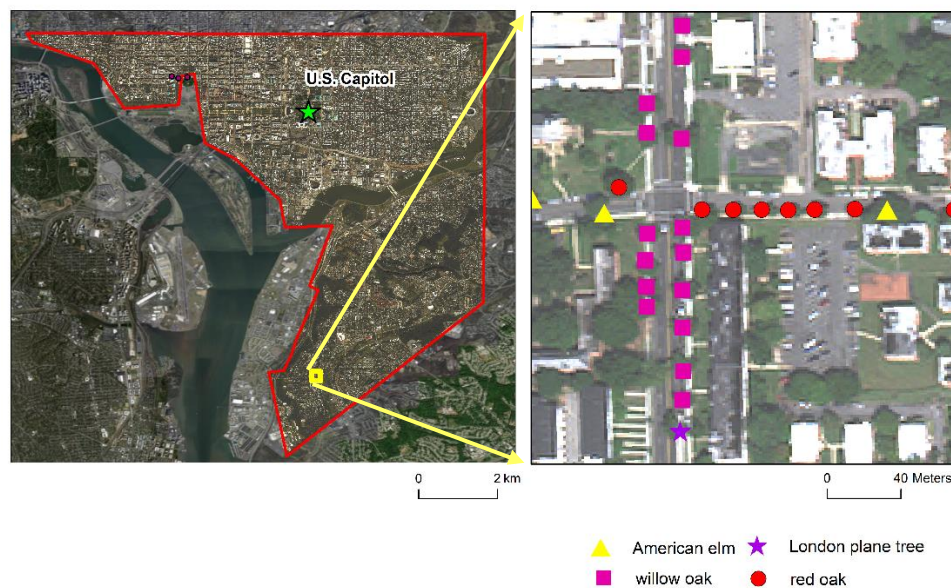
In this study, we aim to 1) examine the potential of using VHR images to identify urban street tree species in terms of the number of classes and taxonomic levels; 2) identify the key pheno-phases that can govern urban tree species classification accuracy, by using a time series of WV-3 imagery which has been adopted for tree species classification and health assessment (Li et al, 2015, Au 2018, Ferreira et al. 2019), and a large forest inventory reporting tree species information over 160 species from District Department of Transportation (DDOT). We hypothesize that 1) the overall accuracy will decrease with each added class because the potential for spectral similarity between classes increased; 2) leaf pigment-induced changes in visible-light

spectral bands during spring leaf emergence or fall senescence phenology periods will be the most predictive components of spectral variability linked to species differences.

## 3.2 Methods

### 3.2.1 Study area

Washington D.C. has extensive urban forests with tree coverage around 30% of the land area (Andrada II et al. 2015). The city landscape has a large variety of tree species, including native trees such as elms, maples, oaks, and exotic trees such as cherries and ginkgos. Besides trees under private ownership or maintained by various local agencies (e.g. National Capital Planning commission, National Park Services, District of Columbia's Department of Parks and Recreation), the Urban Forestry Administration (UFA) is the District's governmental agency which responsible for planting, pruning, removal and other routine maintenance of the remaining street trees in the public spaces. These street trees comprise approximately 9% of the entire Washington D.C. urban forest (Government of the District of Columbia 2010). In order to maximize the overlap of this street tree database with archived WV-3 images, we set our study site within central Washington D.C., covering approximately 45 km<sup>2</sup> (**Figure 3.1**).





**Figure 3.1** a) Study area (within red boundary) in the District of Columbia. b) Portion of the digitized tree locations superimposed on a WorldView-3 true color image from September 28 2014.

### 3.2.2 Field dataset

We obtained the tree inventory data from the DDOT Street Spatial Database (District of Columbia Government 2018). The field survey dataset contains attributes of street trees such as the tree stem coordinates, species common name, tree health condition class (excellent, good, fair, poor and dead), and diameter at breast height (DBH) in U.S feet. In total, there are 49,303 tree stems within our 45 km<sup>2</sup> study areas. To avoid difficulties from small tree crowns, we only considered mature trees with DBH greater than 10.16 cm (4 inches) (Nascimebene et al. 2009, Memiaghe et al. 2016). In addition, we also excluded trees labeled as “dead”. Finally, to eliminate bias from small sample sizes, we only considered tree species with more than 200 individuals (~1% of the full dataset). After applying these criteria, our final tree inventory data consists of the top nineteen species. These belong to ten different genera and account for 16,496 individual trees (**Table 3.1**) which is 33.5% of the total database.

**Table 3.1** The most common 19 species with genus information, ranked by sample size

	Common Name	Scientific name	Symbol*	Number of trees sampled
1	Red maple	<i>Acer rubrum</i>	ACRU	2854
2	Willow oak	<i>Quercus phellos</i>	QUPH	1909
3	American elm	<i>Ulmus americana</i>	ULAM	1728
4	Pin oak	<i>Quercus palustris</i>	QUPA2	1703
5	Japanese zelkova	<i>Zelkova serrata</i>	ZESE80	1226
6	Red oak	<i>Quercus rubra</i>	QURU	1065
7	Sugar maple	<i>Acer saccharum</i>	ACSA	891
8	London plane tree	<i>Platanus acerifolia</i>	PLHI	814
9	Little leaf linden	<i>Tilia cordata</i>	TICO2	700
10	Ginkgo	<i>Ginkgo biloba</i>	GINKG	682
11	Norway maple	<i>Acer platanoides</i>	ACPL	590
12	Chinese elm	<i>Ulmus parvifolia</i>	ULCH	469
13	Scarlet oak	<i>Quercus coccinea</i>	QUCO2	395
14	American linden	<i>Tilia americana</i>	TIAM	295
15	Sawtooth oak	<i>Quercus acutissima</i>	QUAC80	293
16	Cherry	<i>Prunus avium</i>	PRAV	231
17	American sycamore	<i>Platanus occidentalis</i>	PLOC	225
18	Sweetgum	<i>Liquidambar styraciflua</i>	LIST2	220
19	Bradford callery pear	<i>Pyrus calleryana</i>	PYCA80	206
	<b>Total</b>			16,496

\* Symbols adopted from USDA Plants Database

### 3.2.3 Image datasets

We purchased twelve WV-3 cloud-free Atmospheric Compensation (ACOMP) images from a DigitalGlobe vendor. WV-3 images comprise eight multispectral bands ( $B_{1-8}$ ) (four traditional multispectral bands (Blue ( $B_2$ : 450 – 510 nm), Green ( $B_3$ : 510 – 580 nm), NIR-1 ( $B_7$ : 770 – 895 nm), Red ( $B_5$ : 630 – 690 nm), four additional bands (Coastal ( $B_1$ : 400 – 450 nm), Yellow ( $B_4$ : 585 – 625 nm), Red edge ( $B_6$ : 705 – 745 nm), NIR-2 ( $B_8$ : 860 – 1040 nm)) at a spatial resolution of 1.2 m and one panchromatic band at a spatial resolution of 0.3m: Panchromatic band (450 – 800 nm). Our high temporal resolution WV-3 images from 2014 to 2017 cover the entire leaf-on period from April to November with different pheno-phases (**Table 3.2**). We applied Rational Polynomial Coefficient (RPC) orthorectification for WV-3 images using National Agriculture Imagery Program (NAIP) images as the reference raster (Maxwell et al. 2017) and LiDAR data

(District of Columbia Government 2009) as DEM raster in ENVI 5.4. The resulting orthorectification had an RMSE less than 1 pixel.

**Table 3.2** 12 sets of WorldView-3 data

Time	Season	Day Of Year	Mean Off nadir °
18-Apr-15	early leaf emergence	108	16.4
16-May-16	late leaf emergence	137	23.9
11-Jun-17	full leaf expansion	162	22.2
16-Jul-15	full leaf expansion	197	19.4
24-Jul-16	full leaf expansion	206	24.3
30-Jul-17	full leaf expansion	211	14.0
30-Aug-17	full leaf expansion	242	14.7
17-Sep-15	early leaf senescence	260	17.0
28-Sep-14	early leaf senescence	271	26.6
11-Oct-16	late leaf senescence	285	22.0
31-Oct-15	late leaf senescence	304	24.0
18-Nov-16	late leaf senescence	323	17.5

### 3.2.4 Tree crown delineation

We applied a two-step delineation procedure to delineate tree crowns as individual objects. First, we determined a common crown buffer size of 4.5 m, based on the average diameter at breast height DBH (33.8cm=13.3 inches) and an allometric equation (Lamson 1987). Then we applied these radial buffers for the GPS coordinates of all tree stems location in the DDOT database in ArcMap 10.6. Next, we checked the histogram of each NDVI image and masked pixels within these radial buffers that had NDVI values less than 0.25 in order to remove confounding influence of non-vegetated portions of possible tree crowns. We visually confirmed the performance of our crown delineation procedure by checking a random subset of 100 crowns.

### 3.2.5 Vegetation indices calculation and RandomForest classification

We first calculated 11 vegetation indices in ENVI 5.4 (**Table 3.3**) by combining both visible and near infrared bands from each of the twelve WV-3 images (12 dates \* 11 indices = 132 vegetation indices in total). These typical vegetation indices measure the general quantity of

vegetation greenness, which are also frequently mentioned for other tree species discrimination studies (Pu and Landry 2012, Li et al. 2015, Karlson et al. 2016, Fang et al. 2018, Pu, Landry and Yu 2018). Second, we calculated four types of texture features using panchromatic bands for each image based on the grey-level co-occurrence matrices (GLCM) in eCognition. These texture variables can help describe differences in crown architecture (e.g. within crown gaps) between different tree species. The four types of texture features are (1) Homogeneity. Within the delineated crown object, homogeneity describes the similarity of pixel values. (2) Contrast. Contrast summarizes the variation of pixel values exponentially. (3) Entropy. Within the delineated crown object, entropy measures the disorder or randomness in the image values. (4) Dissimilarity. It measures the heterogeneity of pixel values linearly within the tree crown. We used all 132 vegetation indices and 48 texture variables as predictors for RF classification. Since we tested the performance of RF by the number of classes, we didn't apply feature selection and feature reduction process to maintain consistency in using 180 predictors between each model.

We applied RandomForest (RF) machine learning algorithm for classifying tree species using the 180 predictors calculated above. Using large numbers of decision trees with a bootstrapped sample of training data each time, a RF classifier is an ensemble classifier which avoids the shortage of regular decision trees (Maxwell, Warner and Fang 2018). For each binary split within a tree, RF uses maximum node purity (Karlson et al. 2016). One of the benefits of RF classification is that, since the classification uses a subset of the data in each tree, it produces a so-called out-of-bag error. To evaluate how discrimination accuracy is sensitive to the number of classes included in the sample, we ranked our tree inventory data by sample size from high to low for top nineteen species (10 genera). Next, we grouped our tree inventory into 1) eighteen subsets at the species level: we selected the two most dominant species as the first subset, and we added one more species (the next largest class) each time iterately into the next subset until we reached nineteen species for the 18<sup>th</sup> subset; 2) 9 subsets at genus level: we selected the two most dominant genera as the first subset, and we added one more genus (the next largest class) each

time iterately into the next subset until we reached ten genera for the 9<sup>th</sup> subset. For each of these subsets, we randomly selected 75% of the dataset as training data and treated the remaining 25% as validation. We applied the up-sampling method for each of training set (Caret package) prior to RF classification in order to avoid bias induced by class imbalance (Maxwell, Warner and Fang 2018). Then we implemented a 10-fold cross validation RF classifier (mtry=100) for each training dataset using Caret package in R (Kuhn 2012, Maxwell, Warner and Fang 2018) independently. We used the validation data for prediction and confusion matrix generation. We evaluated the potential of WorldView-3 images to classify tree species by recording validation kappa and overall accuracy calculated from confusion matrices. Based on these evaluations, we tested the effect of the number of classes on urban tree species classification accuracy by adding one more class (the next most abundant class) each time at both species and genus level. We assessed the most optimum pheno-phases of WV-3 data to classify nineteen urban tree species and ten genera by analyzing variable importance provided by the RF classifier.

**Table 3.3** Vegetation indices calculated for tree species classification

	Feature name	Description*	Reference
Vegetation Index	NDVI1	$(R_{NIR-1} - R_{Red}) / (R_{NIR-1} + R_{Red})$	Rouse et al. 1973
	NDVI2	$(R_{NIR-2} - R_{Red}) / (R_{NIR-2} + R_{Red})$	Rouse et al. 1973
	RENDVI1	$(R_{NIR-2} - R_{RedEdge}) / (R_{NIR-1} + R_{RedEdge})$	Karlson et al. 2016
	RENDVI2	$(R_{NIR-2} - R_{RedEdge}) / (R_{NIR-2} + R_{RedEdge})$	Karlson et al. 2016
	Green Red Ratio (GR)	$(R_{Green} - R_{Red}) / (R_{Green} + R_{Red})$	Karlson et al. 2016
	Green Yellow Ratio (GY)	$(R_{Green} - R_{Yellow}) / (R_{Green} + R_{Yellow})$	
	Green Normalized Difference Vegetation Index 1 (GNDVI1)	$(R_{NIR-1} - R_{Green}) / (R_{NIR-1} + R_{Green})$	Gitelson, A., and M. Merzlyak, 1998
	Green Normalized Difference Vegetation Index 2 (GNDVI2)	$(R_{NIR-1} - R_{Green}) / (R_{NIR-1} + R_{Green})$	Gitelson, A., and M. Merzlyak, 1998
	Green Leaf Index (GLI)	$((R_{Green} - R_{Red}) + (R_{Green} - R_{Blue})) / (2 * R_{Green} + R_{Red} + R_{Blue})$	Louhaichi et al. 2001
	EVI1	$2.5 \times ((R_{NIR-1} - R_{Red}) / (R_{NIR-1} + 6 \times R_{Red} - 7.5 \times R_{Blue} + 1))$	Omer et al. 2016
Texture	EVI2	$2.5 \times ((R_{NIR-2} - R_{Red}) / (R_{NIR-2} + 6 \times R_{Red} - 7.5 \times R_{Blue} + 1))$	Omer et al. 2016
	GLCM homogeneity	GLCM homogeneity	Pu and Landry 2012
	GLCM contrast	GLCM contrast	
	GLCM entropy	GLCM entropy	
	GLCM dissimilarity	GLCM dissimilarity	

\*Where *R* is the reflectance and the subscript refers to the spectral band (e.g. Blue, Red, etc).

### 3.3 Results

#### 3.3.1 The effect of number of classes on urban tree species classification

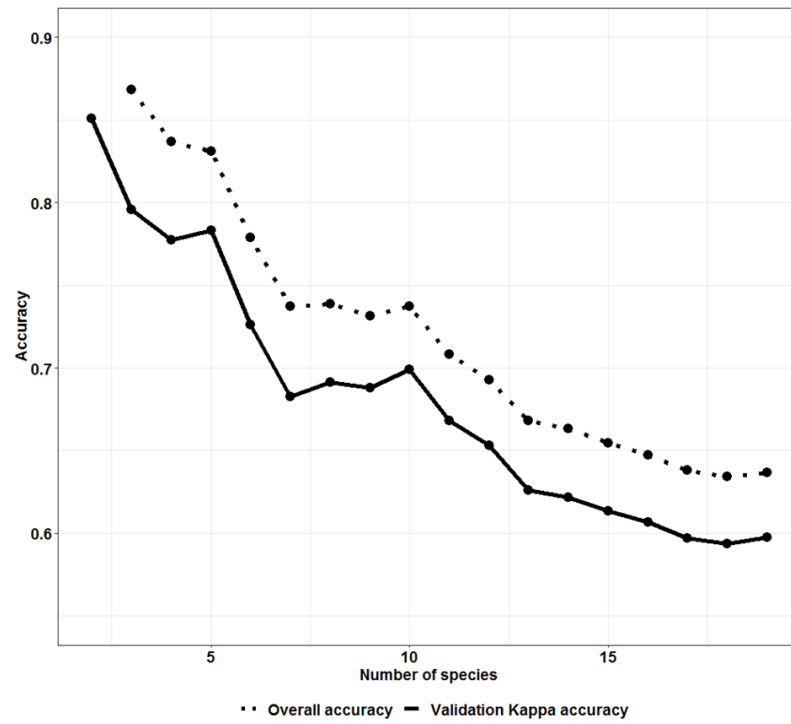
Generally, by adding one more class each time, validation kappa accuracy and overall accuracy exhibited similar trends at both species and genus level (**Figure 3.2 a, b**). With each added class, the validation kappa accuracy gradually decreased with each added class. When classifying nineteen classes at the species level, we derived the lowest validation kappa accuracy as 59.8% (Overall accuracy =63.7%). Similarly, we derived the lowest validation kappa accuracy as 68.5% (Overall accuracy =75.6%) when classifying ten genera. Another observation is that the validation kappa accuracy decreased more rapidly at the species level than at the genus level by adding a class each time. In particular, the validation accuracy decreased from 78.3% to 72.6% when adding red oak for 6 species classification, and then the validation accuracy declined again to 68.2% after adding sugar maple as the 7<sup>th</sup> class. We derived the largest decrease of validation kappa (3.3%) at genus level when adding *Ulmus* for 3 genera classification (kappa accuracy=75.6%, overall accuracy 84.8%).

We generated two confusion matrices from validation datasets containing our whole sample of nineteen tree species (**Table 3.4**) and 10 genera (**Table 3.5**). The producers' accuracy values for each class ranged from 1.0% to 84.1% for nineteen tree species classification (**Table 3.4**). We found the highest producers' accuracy for Ginkgo (GINKG, Producers' accuracy=84.1%, User's accuracy= 76.9%, validation sample size=170); for the more abundant species red oak (QURU) had a relatively lower producers' accuracy (Producers' accuracy=50.0%, User's accuracy=63.6%, validation sample size=266). The spectral confusion for red oaks mainly occurred with pin oaks (QUPA2) and willow oaks (QUPH). For the ten genera classification, (**Table 3.5**), we observed a relatively lower producers' accuracy for the more abundant genus *Ulmus* (Producers' accuracy=62.8%, User's accuracy= 79.5%, validation sample size=549) compared with the producers' accuracy for less abundant genus *Zelkova* (Producers' accuracy

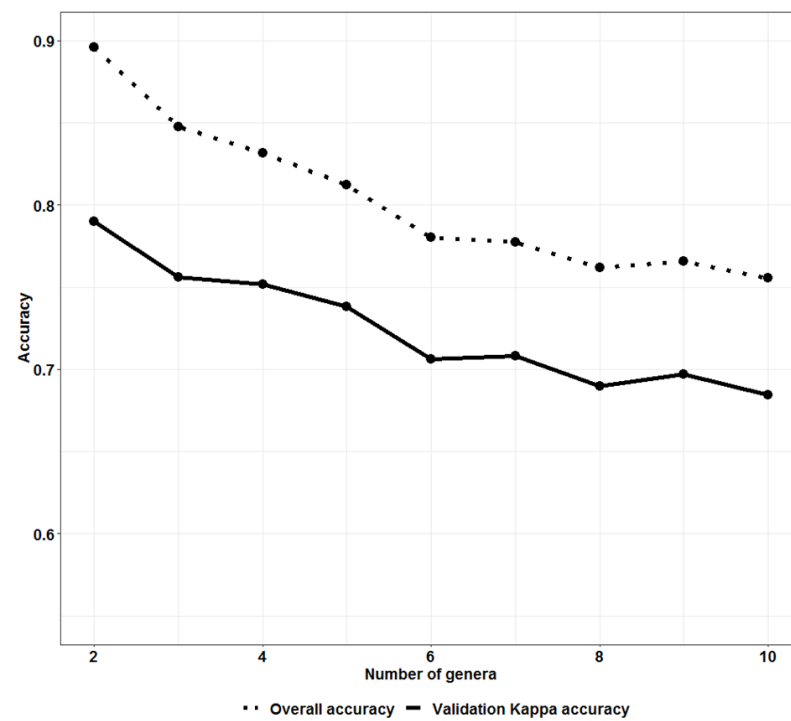
=75.5 %, User's accuracy= 88.2%, validation sample size=306). The lower producers' accuracy for elms was associated with confusion with *Acer* and *Quercus*.

### **3.3.2 Variable importance for 19 species and 10 genera classification**

We selected the most influential ten predictors for nineteen species and ten genera classification based on variable importance provided by RandomForest classifier (**Fig 3.3a & b**). While the variable importance rank was different between each classification, we found a large agreement on the top ten important variables component at both taxonomic levels. We also observed that all these ten top variables, which consist of seven types of vegetation indices (GLI, GR, NDVI1, RE NDVI2, GNDVI1, EVI1, GY), belong to vegetation indices rather than texture variables calculated from spring leaf emergence or fall senescence images. In addition, vegetation indices calculated using the visible band spectrum, such as GLI (green and red), had the highest rank, whereas vegetation indices calculated by infrared bands (i.e. RE NDVI) and texture variables were less important for classification.



a.



b.

**Figure 3.2** Validation Kappa Accuracy from RF classification at species level (a) (from 2 to 19 most common species) and genus level (b) (from 2 to 10 most common genera).



**Table 3.4** Confusion matrix from the validation of the 19 species classification.

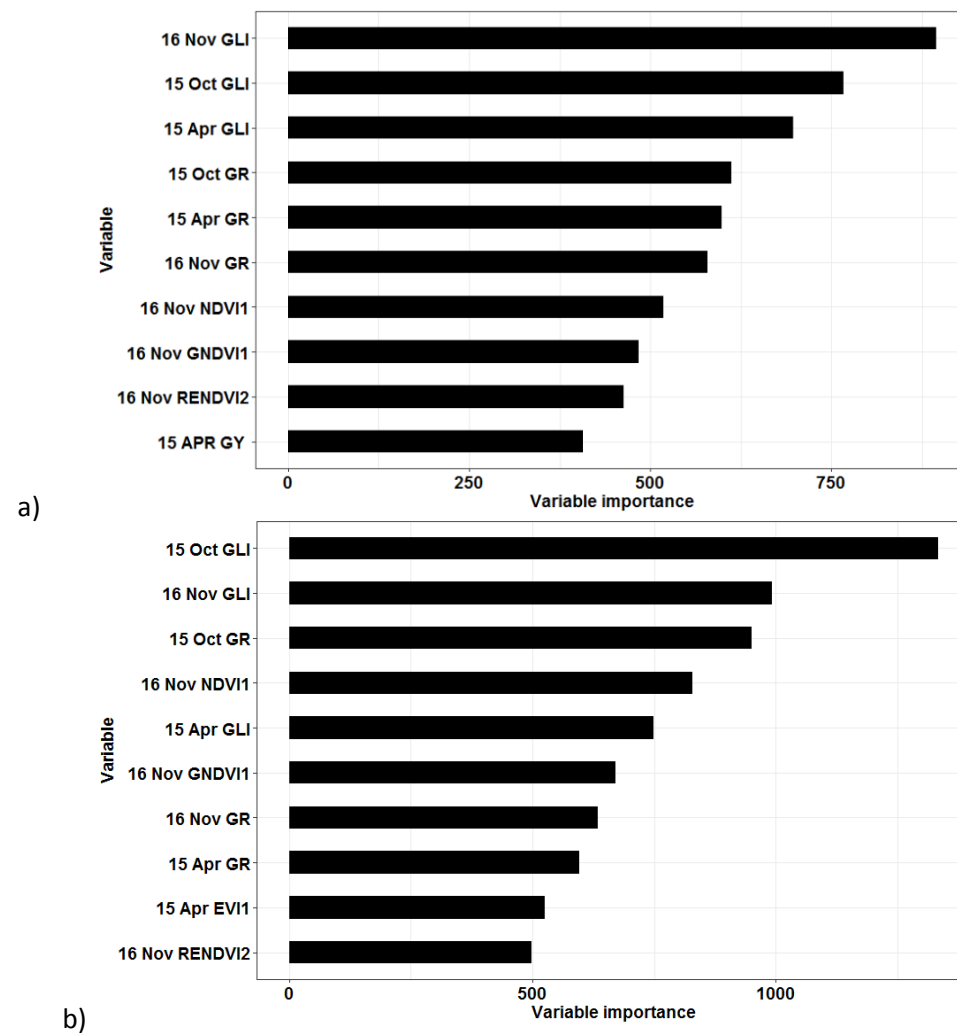
	Reference																			Total	UA (%)
	ULAM	TIAM	PLOC	PYCA80	PRAV	ULCH	GINKG	ZESE80	TICO2	PLHI	ACPL	QUPA2	ACRU	QURU	QUAC80	QUCO2	ACSA	LIST2	QUPH		
ULAM	326	6	0	2	2	10	3	10	5	3	7	5	42	12	0	2	7	0	14	456	71.5
TIAM	1	23	0	0	0	0	0	0	10	0	2	0	0	0	1	0	0	0	0	37	62.2
PLOC	0	0	7	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	11	63.6
PYCA80	0	0	0	21	0	0	0	0	1	1	0	1	2	1	0	1	0	0	4	32	65.6
PRAV	1	0	0	0	14	0	3	2	0	2	1	3	8	1	0	0	3	2	4	44	31.8
ULCH	8	1	0	1	2	40	0	2	2	0	0	3	1	0	1	2	0	0	5	68	58.8
GINKG	8	1	0	0	1	2	143	1	5	1	5	2	8	3	2	0	0	0	4	186	76.9
ZESE80	4	2	0	7	3	7	0	240	3	1	2	5	8	6	1	2	10	3	1	305	78.7
TICO2	10	25	1	3	1	3	3	2	105	1	12	1	10	0	2	0	1	3	2	185	56.8
PLHI	8	2	31	1	11	12	5	2	5	156	6	11	15	4	3	7	3	2	11	295	52.9
ACPL	3	4	1	1	0	1	0	0	9	0	67	2	8	1	0	0	2	0	3	102	65.7
QUPA2	4	0	2	3	2	5	0	12	1	2	3	294	18	62	4	49	7	2	23	493	59.6
ACRU	33	5	8	9	17	15	7	18	20	15	22	33	543	18	8	11	79	4	48	913	59.5
QURU	3	1	0	2	0	4	0	6	0	2	2	27	6	133	0	12	4	0	7	209	63.6
QUAC80	0	0	2	0	0	1	0	0	0	0	0	1	2	2	42	1	0	0	2	53	79.2
QUCO2	0	0	0	0	0	0	0	0	0	0	0	4	1	0	0	1	0	0	0	6	16.7
ACSA	3	2	0	0	0	1	1	7	1	2	7	1	24	3	0	0	102	0	4	158	64.6
LIST2	2	0	1	1	1	1	1	0	0	5	1	3	1	0	3	0	1	29	4	54	53.7
QUPH	18	2	3	1	4	15	4	4	8	9	11	30	17	20	6	11	4	10	341	518	65.8
Total	432	74	56	52	58	117	170	306	175	204	148	426	714	266	73	99	223	55	477		
PA (%)	75.5	31.1	12.5	40.4	24.1	34.2	84.1	78.4	60.0	76.5	45.3	69.0	76.1	50.0	57.5	1.0	45.7	52.7	71.5		
OA (%) = 63.7 Kappa (%) = 59.8																					

PA: Producer's accuracy; UA: User's accuracy; OA: Overall accuracy

**Table 3.5** Confusion matrix for 10 genera classification

	Reference										Total	UA (%)
	<i>Acer</i>	<i>Ginkgo</i>	<i>Liquidambar</i>	<i>Platanus</i>	<i>Prunus</i>	<i>Pyrus</i>	<i>Quercus</i>	<i>Tilia</i>	<i>Ulmus</i>	<i>Zelkova</i>		
<i>Acer</i>	901	13	8	41	34	14	133	64	91	44	1343	67.1
<i>Ginkgo</i>	4	140	0	0	0	0	7	2	2	0	155	90.3
<i>Liquidambar</i>	2	0	21	5	1	1	3	0	4	0	37	56.8
<i>Platanus</i>	15	2	1	178	5	0	19	3	9	3	235	75.7
<i>Prunus</i>	3	0	2	1	7	0	0	1	0	0	14	50.0
<i>Pyrus</i>	1	0	0	2	0	14	4	1	1	0	23	60.9
<i>Quercus</i>	95	9	21	29	9	15	1127	12	83	25	1425	79.1
<i>Tilia</i>	26	3	1	0	0	0	3	152	10	1	196	77.6
<i>Ulmus</i>	30	3	1	4	1	2	34	12	345	2	434	79.5
<i>Zelkova</i>	7	0	0	0	1	6	11	2	4	231	262	88.2
<b>Total</b>	1084	170	55	260	58	52	1341	249	549	306	3116	
<b>PA (%)</b>	83.1	82.4	38.2	68.5	12.1	26.9	84.0	61.0	62.8	75.5		
<b>OA (%) =75.6 Kappa (%) = 68.5</b>												

PA: Producer's accuracy; UA: User's accuracy; OA: Overall accuracy



**Figure 3.3** Top ten important variables at a) 19 species b) 10 genera level. Variable importance is defined by the RF algorithm.

### 3.4 Discussion

By using a robust machine learning algorithm to leverage large forest inventory data as well as a suite of high spatial resolution images, our study makes three novel contributions to tree species classification research. First, our study reaffirms that focusing on dates with distinctive phenology, which can enhance inter-species spectral variability, is a key asset for tree species mapping (Key et al. 2001, Hill et al. 2010, Li et al. 2015, Madonsela et al. 2017). Second, we confirmed that the performance of tree species classification is affected by the number of classes and by the taxonomic level of those classes. Third, we successfully classified nineteen tree species at crown-level in an exclusively deciduous forest setting. Previous studies investigated tree species mapping in a mixed forest (Xiao, Ustin and McPherson 2004, Holmgren et al. 2008, Immitzer et al. 2012, Ghosh et al. 2014, Karlson et al. 2016, Pu, Landry and Yu 2018), while our study is more challenging since our study site is dominated by deciduous trees. Our approach confirmed that remote sensing images taken from spring leaf emergence and fall senescence phenological period have the most potential to discriminate trees at both species and genus level.

#### 3.4.1 Effects from number of classes and taxonomic levels

The quality of the training samples such as the number of training samples and class imbalance affects the performance of machine learning algorithms including RF (Maxwell, Warner and Fang 2018). Previous research for tree species classification typically varies from 3 to over 19 species (Xiao and McPherson 2005, Holmgren et al. 2008, Immitzer et al. 2012, Pu and Landry 2012, Ghosh et al. 2014, Karlson et al. 2016, Fang et al., 2018). Our study investigated the relationship among the number of classes to be identified and the validation kappa accuracy of RF classification for mapping urban street trees. We found that both the size of training samples and the taxonomic group can affect the classification accuracy. Specifically, we found a negative relationship between the number of classes being identified and the accuracy of the validation data is measured by the kappa coefficients (**Figure 3.2a**). As the training sample size for each added class decreases between different models, our sample

potentially becomes more biased, and less representative of the within-class spectral variability exhibited by the validation sample for the added class. Therefore, the validation kappa and overall accuracy decreased gradually at both species and genus level. Despite effect from smaller sample size for added classes, the similarity between the added class and the existing classes also resulted in the validation kappa to decline. For example, in **Figure 3.2a**, the largest drop in kappa after adding a new class is the 6 species classification declined 5.7% after adding red oak. Errors in this 6 tree species classification were mainly related to the spectral similarity between red oak and other species of the same *Erythrobalanus* sub genus group, such as pin oak. This is associated with the fact that red oaks and pin oaks are closely related taxonomic groups which share a similar phenological timing and thus share similar spectral characteristics in the expression of the leaf emergence and senescence phases. It is challenging to capture these similar spectral features between red oaks and pin oaks, which resulted in low validation accuracy. Moreover, our confusion matrices in **Table 3.5** further suggested that species with different taxonomic level provided more inter-species spectral variability for tree species discrimination. For instance, ginkgo produced the highest producer's accuracy even though ginkgo was not our most common species. Since ginkgo does not share the same genus with other species, ginkgo has its own division and unique crown-level spectral features at both species and genus level classification.

Our assessment of classification at the genus level provided complementary information for the importance of taxonomic levels for tree species classification. There tends to be greater phenological separation between different genera than different species in the same genus. For instance, at the species level, red oak produced low producer's accuracy due to spectral similarity with other types of oaks. In contrast, in our ten genera classification, oaks produced the highest producer's accuracy, which further validates that the similarity within same taxonomic groups can reduce accuracy while the variability among different taxonomic groups can improve classification accuracy. Thus, we do not recommend comparing accuracies directly from different studies since taxonomic levels have a large impact on classification accuracy. For example, Immitzer et al. (2012) and Karlson et al. (2016) both reported

overall accuracy over 80% for ten species and five species classification. The overall accuracy of ten species classification was as high as five species classification since the study site was a mixed forest where most of the species were in different taxonomic groups. A heterogeneous landscape, a mixed forest, or site with tree species in different taxonomic groups would likely result in a higher species classification accuracy. High spatial resolution images capture more inter-genus spectral variability compared with inter-species spectral variability, which makes it more difficult to classify trees at species level.

### **3.4.2 Key pheno-phases for tree species mapping**

Our assessment of variable importance suggests that high spatial resolution images acquired during spring leaf emergence and fall senescence phenology contributed most for the classification accuracy. In addition, vegetation indices, which describe the pigment-induced change between species, contribute more for tree species classification than texture variables, which can indicate structural features for each crown. This is contrast to Xiao and McPherson (2005) who suggest a spring or summer image rather than an autumn image provides better separation of tree species. Our results indicated that spring and autumn images especially late autumn images from October and November were most helpful to characterize the spectral differences due to leaf senescence. Variables calculated from a June or July image with peak vegetation contributed less to tree species classification, since the spectral variability between different species was relatively low. Moreover, this spectral difference among species was highly related to their species-specific autumn color rather than their crown architecture since the top three important variables (GLI, GR) were calculated using the visible bands of red and green. Specifically, the phenological variability in timing and the expression of leaf pigments caused tree species differences to be most visible during the spring leaf emergence and fall leaf senescence. In addition to the importance of indices calculated by yellow band to describe the leaf-pigment induced spectral difference (Kalson et al. 2016), our results demonstrated that indices calculated from other visible bands i.e. red and green band were the most important factors for successful tree species

mapping. Our late-autumn images successfully captured the leaf pigment-induced changes in visible light spectral bands which aligned with these species-specific autumn color features. For example, ginkgo turns bright yellow (Del 2007) and oak turns red, orange, or brown in late autumn. These species-specific phenological patterns and leaf pigment-induced changes have enabled a better separation between ginkgo and red oak. In conclusion, the transition periods during spring leaf emergence and fall senescence provide the most leaf pigment-induced spectral variability and they are optimum phenology periods for successful tree species mapping.

### **3.5 Conclusion**

Overall, by using RandomForest classifier to leverage high spatial resolution images, large forest inventory datasets and spatially accurate crown objects, this study highlights the opportunities for using VHR images to discriminate broadleaf trees at both species and genus level in all environments. First, this study demonstrated the benefits of phenology using 12 WV-3 images for 16,496 deciduous tree species classification in Washington D.C. Our study comprehensively evaluated the performance of different pheno-phases from the entire leaf-on period from April to November for tree species mapping. We conclude that images taken in spring leaf emergence and fall senescence are the optimum phenological period to discriminate individual tree species. The red and green visible bands are significant in their contribution towards the pigment-induced changes in tree species classification. Second, these remote sensing images have the potential to classify trees with an overall accuracy of 63.7% at the species level and 75.6% at the genus level. VHR remote sensing images have more potential to separate species at genus level than species level due to the higher inter-genus spectral variability. With enough sample size for calibration/validation in the forest inventory database, we also evaluated the performance for tree species mapping on a function of the number of classes (2-19 tree species classification and 2-10 genera classification) and their taxonomic levels. We conclude that the decline of classification accuracy is associated with species in similar taxonomic groups. We suggest future studies to use images taken

during spring leaf emergence or fall senescence period to document the leaf pigment-induced changes for tree species mapping in both urban and rural environments.



## 4. Street tree health from space? An evaluation using WorldView-3 data and the Washington D.C. Street Tree Spatial Database\*

Fang Fang<sup>a</sup>, Brenden McNeil<sup>a</sup>, Timothy Warner<sup>a</sup>, Gregory Dahle<sup>b</sup>, Earl Eutsler<sup>c</sup>

<sup>a</sup> Department of Geology and Geography, West Virginia University, Morgantown, WV, 26505, USA

<sup>b</sup> Division of Forestry & Natural Resources, West Virginia University, Morgantown, WV, 26505, USA

<sup>c</sup> Urban Forestry Division, District Department of Transportation, Washington, DC, 20003, USA

\*Paper currently in revision for re-submission to Remote Sensing of Environment.

### **Abstract**

Spatially accurate and timely information on tree health is an essential component of maintaining sustainable urban environments. We evaluate the potential of WorldView-3 (WV-3) satellite images from June 11<sup>th</sup>, July 30<sup>th</sup> and August 30<sup>th</sup>, 2017 to discriminate the field-measured health condition class of 2538 trees within the District of Columbia Department of Transportation's Street Tree Spatial Database. For each street tree in each image, we measured six vegetation indices (VIs), and find that the NDVI1, defined as the normalized ratio of the red and first near infrared bands, on the July image shows the most potential to discriminate among trees measured in good, fair, and poor health condition classes. Yet, the variability in VIs attributable to health condition class is small, especially relative to the large declines in VIs between the June and August image dates. This greendown phenological pattern occurs similarly for trees in all health condition classes, and is thus essential to consider when comparing VIs from different years or months. Based on these findings, we propose two strategies for using high-resolution satellite data in street tree health management: (1) using anniversary date imagery to track the VIs of individual trees from year to year, and (2) using single-date imagery to prioritize field-based street tree health assessments.

## 4.1 Introduction

More than half of the world's population now lives in cities (Lederbogen et al. 2011) and tree health conditions have been widely recognized as central for the sustainability of these expanding urban populations (Huang et al. 2007, Xiao and McPherson 2005). As a case study representative of other large urban centers worldwide (Gregg et al. 2003, McPherson et al. 1994, Nowak and Crane 2002, Wong and Yu 2005), our study focuses on Washington D.C., where nearly 30% of the land surface is covered by urban trees. Each year, these trees provide D.C. with \$3.6 billion in structural values (e.g. providing shade to cool buildings), and \$2.5 million in functional values like pollution removal (Nowak et al. 2006). There are approximately 1,928,000 trees (Nowak et al. 2006) throughout D.C., and extensive efforts with an annual budget around \$15.5 million are directed toward urban tree management (e.g. planting, pruning, removing, maintaining). To improve this urban tree management process, District Department of Transportation (DDOT) has created a Street Tree Spatial Database to track the location and attributes (e.g. DBH, health condition class, and species) of all street trees since 2006, and they have expended considerable effort to regularly update this database.

Poor tree health can be caused by biotic agents like fungi, bacteria or viruses, or abiotic agents such as drought or freezing injury (Jackson 1986). In addition, because of urban environmental disturbances such as air pollution, compacted soil, damage from traffic or pedestrians, and improper planting, urban street trees typically face a more stressful environment compared with trees in native forests (McIntyre 2006). Previously, one of the most common trees in D.C., the American elm (*Ulmus americana*), was almost entirely lost to Dutch elm disease (*Ophiostoma spp.*, Banfield 1968, Nowak et al. 2006). Currently, over 30% of the tree population in D.C. is at risk of being lost to the Asian Longhorned Beetle (*Anoplophora glabripennis*, Nowak et al. 2006). Less severe, more chronic agents, such as Bacterial leaf scorch (Harris and Balci 2015) also affect street trees in D.C. These diseases and pests are often the initiators of diminished tree health, which can directly lead to tree death, ensuing economic losses, and deteriorating quality of life in urban settings (Harris and Balci 2015, Nowak et al. 2006).

Tree health is traditionally monitored via field surveys. However, there is great potential for using remote sensing to monitor tree health. Poor health is often expressed by lowered leaf water or chlorophyll content, which both affect the spectral reflectance of trees (Knipling 1970, Chaves et al. 2002, Xiao and McPherson 2005, Huang et al. 2007, Sankaran et al. 2010, Eitel et al. 2011, Asmaryan et al. 2013). Indeed, promising research from non-urban forests has demonstrated potential for evaluating vegetation health from multispectral and hyperspectral data collected at the scale of a pixel or leaf (Xiao and McPherson 2005, Franke and Menz 2007, Huang et al. 2007, Delalieux et al. 2009, Eitel et al. 2011, Lambert et al. 2013, Michez et al. 2016). For instance, increased presence of Huanglongbing disease (*Candidatus Liberibacter* spp.) was associated with decreased NIR reflectance in citrus trees (Li et al. 2012). Given the sensitivity of the NIR spectral region to tree health, various vegetation indices like the normalized difference vegetation index (NDVI) are often correlated with poor health (Xiao and McPherson 2005, Eitel et al. 2011). Yet, despite the well-established linkages among spectral reflectance and tree health at pixel and leaf scales in rural settings, there remain substantial challenges for researchers and managers wishing to apply satellite data to assess the health of street trees at the biologically-meaningful, and necessary management scale of individual tree crowns.

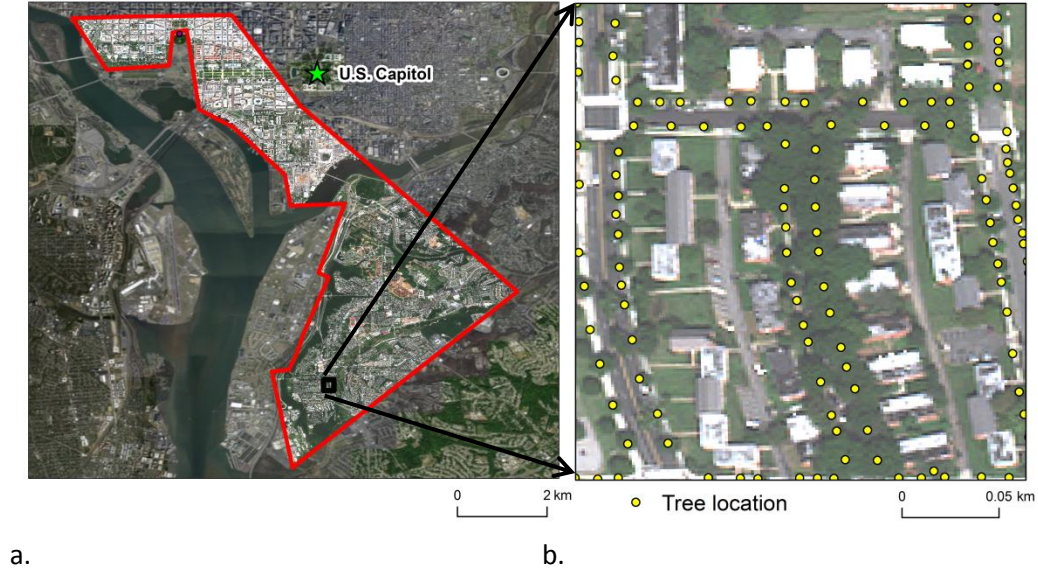
Here, we take advantage of promising new field and image datasets to evaluate the current prospects for meeting these challenges by matching 2308 street trees in the DDOT Street Tree Spatial Database with atmospherically corrected WorldView-3 spectral data collected on June 11<sup>th</sup>, July 30<sup>th</sup>, and August 30<sup>th</sup>, 2017. Recently, an exploratory analysis of trees in various cities around the world found that the high spatial, radiometric, and spectral resolution of WV-3 data was useful for resolving individual urban trees, and for tracking spectral changes consistent with declining tree health (Au 2018). Building from that work, the goal of our study is to systematically evaluate whether WV-3 data can be used to detect a statistical difference among street trees marked by the DDOT as having different field-assessed health condition classes (e.g. good, fair, poor). By robustly evaluating whether a signal of individual street tree health can be detected from satellite imagery, our study can lay a foundation and provide guidelines for urban foresters wishing to incorporate remote sensing imagery into the management of street tree health.

We explicitly examine essential remote sensing data considerations in the areas of spatial resolution, useful spectral regions, and temporal variability. First, to take full advantage of the high spatial resolution of WV-3 data, we present a novel, two-step adaptive process to delineate individual street tree crown-objects based on the field and image data. Second, since tree health may affect the visible, red edge, and NIR portions of the spectrum, we evaluate which spectral vegetation indices (VIs) might best describe the spectral differences in street tree health. Finally, we consider how non-health related spectral variability within the growing season (e.g. related to such factors such as greendown phenology, Elmore et al. 2012, Reaves et al. 2018) may affect prospects for street tree health assessments. Based on our findings, we discuss two promising strategies for using high-resolution satellite data in street tree health management.

## **4.2 Methods**

### **4.2.1 Study area**

Washington D.C. has extensive urban forests covering much of the city (Andrada II et al. 2015). Besides trees under private ownership or various local agencies (e.g. National Capital Planning commission, District of Columbia's Department of Parks and Recreation), the Urban Forestry Administration (UFA) is the District's governmental agency which responsible for planting, pruning, removal and other routine maintenance of the remaining street trees in the public spaces. These street trees comprise approximately 9% of the entire Washington D.C. urban forest (Government of the District of Columbia 2010). In order to maximize the overlap of this street tree database with available cloud-free archived WV-3 imagery, we set our study site within Washington D.C. to a 25 km<sup>2</sup> area (**Figure 4.1**).



**Figure 4.1.** a) Study area (within red boundary) in the District of Columbia. b) Portion of the digitized street tree locations superimposed on a WorldView-3 true color image from August 30, 2017.

#### 4.2.2 Image datasets

We purchased WV-3 cloud-free Atmospheric Compensation (ACOMP) imagery from a DigitalGlobe vendor. The ACOMP products mitigate the effects of haze and atmospheric scattering and provide normalized surface reflectance values (reflectance index) of the surface objects for both panchromatic and multi-spectral images (DigitalGlobe, Inc. 2016). The ACOMP products also include radiometric corrections (e.g. non-response detector filling, calibrations of the relative radiometric response between detectors), as well as sensor correction rectification with respect to the internal detector geometry and scan distortion (Ferreira et al. 2019). To independently validate the ACOMP product, we used 50 pixels from pseudo-invariant objects distributed throughout the study area to confirm that there is no significant difference ( $P=0.88$  in ANOVA test) in the mean NDVI on different dates (Hadjimitsis et al. 2009).

The data contain eight multispectral bands ( $B_{1-8}$ ) at a spatial resolution of 1.2 m and one panchromatic band at a spatial resolution of 0.3m: Panchromatic band (450 – 800 nm), Coastal ( $B_1$ : 400 – 450 nm), Blue ( $B_2$ : 450 – 510 nm), Green ( $B_3$ : 510 – 580 nm), Yellow ( $B_4$ : 585 – 625 nm), Red ( $B_5$ : 630 – 690 nm), Red edge ( $B_6$ : 705 – 745 nm), NIR-1 ( $B_7$ : 770 – 895 nm) and NIR-2 ( $B_8$ : 860 – 1040 nm). In order to evaluate the ability of imagery to assess tree health at different points during the leaf-on period, we acquired three archived WV-3 images from June 11, 2017 (Day of year (DOY) 162), July 30, 2017 (DOY 211), and August 30, 2017 (DOY 242). The off nadir view angles of these images were 22.2°, 14.0° and 14.7° respectively. We used cubic convolution resampling to individually register each image to the DDOT LiDAR Reflective surface (District of Columbia Government 2009) (RMSE < 1 pixel).

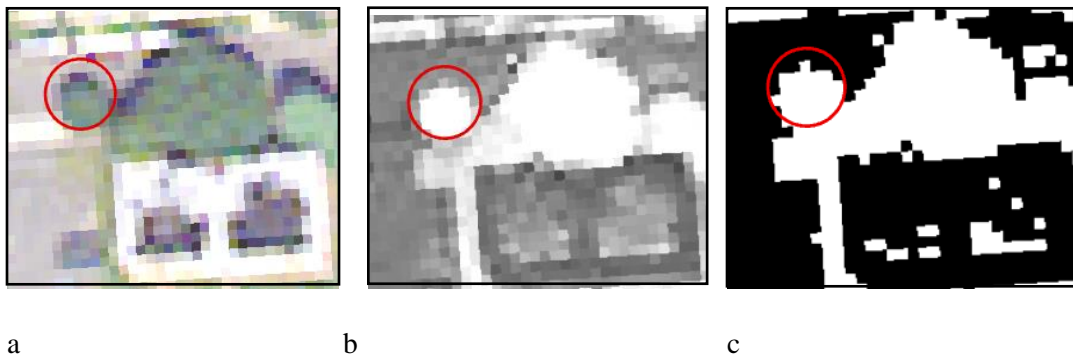
#### 4.2.3 Field dataset

We downloaded the tree inventory data from the DDOT Street Spatial Database on March 8<sup>th</sup> 2018 (District of Columbia Government 2018). This D.C. tree inventory was first created by Casey Trees Foundation in 2002 and has been updated by the Urban Forestry Administration since 2006. The dataset contains attributes of all street trees such as the tree stem coordinates, species common name, tree health condition class, and diameter at breast height (DBH). In total, there are 18,434 tree stems within our 25km<sup>2</sup> study area. Even with the high spatial resolution of WV-3, most trees were too small to provide enough spectral data, so we restricted our study to mature trees with DBH greater than 10.16 cm (4 inches) (Nascimebene et al. 2009, Memiaghe et al.2016). The database describes five classes of tree health condition: excellent, good, fair, poor and dead. A given tree health condition is determined in the field by the quality of a crown and growth vigor (**Appendix B Table s1, s2**). Only 30 trees were marked excellent, so we merged these trees with those marked “good”. Furthermore, we did not include trees classified as “dead”, or trees that did not contain at least 3 pixels that met our NDVI-based tree delineation criteria (see section 2.4). Our final dataset included 1854 trees in the good health condition class, 388 in the fair health condition class, and 66 trees in the poor health condition class, for a total sample size of 2380 trees. Even though there are 95 tree species in this sample, just under half of the

trees are from five tree species: willow oak (*Quercus phellos*), red maple (*Acer rubrum*), pin oak (*Quercus palustris*), American elm (*Ulmus americana*), and red oak (*Quercus rubra*).

#### 4.2.4 Tree crown delineation

We used a two-step process to delineate tree-crown objects, which are objects that contain pixels belonging to the crown of an individual tree. First, we extracted all pixels within a radial buffer from the spatial coordinates of each tree trunk. To avoid introducing uncertainties from allometric equations from the 95 different tree species in our study area, we use the average DBH (34.8cm) and a single allometric equation (Gering and May 1995) to determine an average crown diameter of 9.1m. Accordingly, we applied a buffer radius of 4.55m as the first step in delineating each tree crown-object. Second, in order to avoid the confounding influence of background non-vegetated pixels (e.g. pavement), we masked any pixels with NDVI1 values less than 0.5 (**Figure 4.2**, Li et al. 2015, Karlson et al. 2016). Using a random subset of 100 tree-crown objects, we visually confirmed from pan-sharpened imagery that this two-step tree procedure was effective in delineating individual tree crown-objects. We note that for large trees, the individual tree crown-objects conservatively include only the central portion of the tree crown.



**Figure 4.2** Example of the two-step tree crown-object delineation. Red circle: 4.55m radial tree crown buffer. a) True Color WV-3 image from August 30, 2017. b) NDVI1 calculated from the NIR1 and red bands. c) Application of the 0.5 NDVI1 threshold, showing the delineated tree crown-object as all the white pixels whose majority area falls within the red radial crown buffer.

The tree crown-objects in our study are generally not affected by variable spectral reflectance of the background, because we are only examining public street trees, and not urban trees in park-like settings with backgrounds dominated by other vegetation (e.g. grass). Our field observations and visual examination of available leaf-off imagery revealed that the street trees have two distinct types of backgrounds: 1) a relatively pure background of pavement (often with a metal grate surrounding the trunk), and 2) a mixed background consisting of pavement and a tree box containing soil, bark mulch, and ground vegetation. Yet, even for the 66 trees within the poor health class where thinning leaf area might cause the background to have a consequential effect on the spectral reflectance of a tree crown-object (49 trees with mixed background and 17 trees with pure background of pavement), we found no statistical difference ( $p=0.2, 0.9, 0.2$  for ANOVA test using images acquired from June, July and August respectively) in NDVI of trees sorted by these two background types.

#### **4.2.5 Data analysis**

We focus on VIs because they describe expected changes to multiple parts of the spectral profiles from trees with different tree health conditions. Our assumption is that the shape of the spectrum is a generally more sensitive indicator of tree health than the magnitude of the reflectance in any one band. This was supported by preliminary work, in which we examined other image spectral information, including raw reflectance on each multi-spectral band, and panchromatic (30cm pixel resolution) image texture within the crown objects. This preliminary work indicated that VIs are the most useful and straightforward means to statistically test the difference between tree health classes.

Based on knowledge of how tree health might affect spectral reflectance in the NIR, green, red, and yellow portions of the spectrum, we calculated six vegetation indices from each of the three sets of images. NDVI1 using the red and first NIR band, NDVI2 using red and the second NIR band, Red Edge NDVI (REDNVI), Green yellow ratio (GY), and the Green Red ratio (GR) (Boochs et al. 1990, Pontius et al. 2008, Karlson et al. 2016). After we extracted and averaged the pixel VI values within each delineated tree crown-object, we used repeated non-parametric Wilcoxon/Kruskal-Wallis tests for each combination



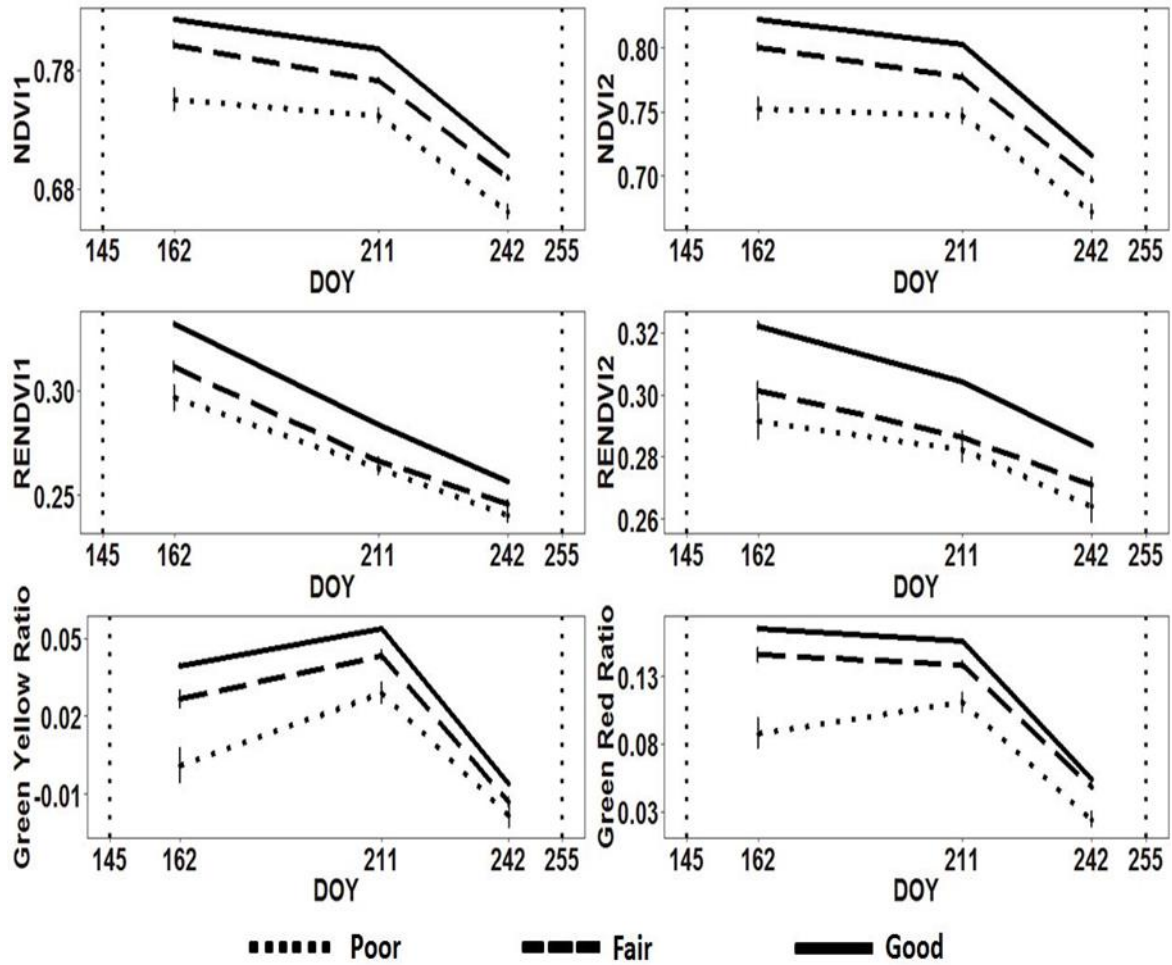
of VI and image date to evaluate the statistical difference among the VI values of tree crown-objects in different health condition classes. We identified the most sensitive VI/image date combination as the one that achieved the highest Chi-square value.

### 4.3 Results

Within each of the three image dates, the VI values of trees in good, fair and poor health condition classes were all highly statistically different (**Table 4.1**). Tree health condition class also had the expected effects on VIs; trees in the poor health condition class had the lowest mean VI values, trees in the fair health condition class had intermediate mean VI values, and trees in the good health condition class had the highest mean VI values (**Figure 4.3**). As measured by the highest Chi-square from the Wilcoxon/Kruskal-Wallis analyses, the most sensitive VI was NDVI on the July 30<sup>th</sup> (DOY 211) image (**Table 4.1**). The VI's that included red edge bands were the next most sensitive, followed by the VIs using only visible light bands.

**Table 4.1** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis tests of statistical difference in VI values among trees in good, fair and poor health condition classes. All tests were significant at  $p < 0.001$ . The bold value denotes the most sensitive image date/VI combination.

<b>Vegetation Index</b>	<b>June</b>	<b>July</b>	<b>August</b>
<b>NDVI1</b>	65.97	<b>102.46</b>	77.80
<b>NDVI2</b>	70.89	94.18	75.57
<b>RENDVI1</b>	73.71	83.78	46.12
<b>RENDVI2</b>	72.17	72.17	41.16
<b>Green Yellow ratio</b>	37.74	48.96	23.61
<b>Green Red ratio</b>	46.43	48.79	25.14



**Figure 4.3** Mean (error bars represent  $\pm 1$  standard error) vegetation index values for trees in three health condition classes. Based on webcam and Landsat phenology data for our study area (from Figure 7 in Elmore et al. 2012), we also show vertical dotted lines to mark the average DOY of full leaf expansion (DOY145), and leaf senescence (DOY255).

## 4.4 Discussion

### 4.4.1 Comparison of VIs for tree health condition class assessment

We find that the most sensitive VI for detecting tree health condition class was NDVI1. Previous studies illustrated the usefulness of red edge NDVI for early observation of trees with poor health conditions (Eitel et al. 2011, Karlson et al. 2016), but we find that the traditional NDVI outperformed red edge NDVI in testing the statistical difference of trees in poor, fair and good condition classes (**Table 4.1**). Within the DDOT field surveys, tree health condition classes are evaluated based on the crown quality and vigor (**Appendix B Table s1, s2**). Our results suggest that NDVI1 best matches with these field-observed health symptoms. Field observers also note changes in leaf color when assigning a tree health condition class, with the increased presence of yellow leaves leading to a fair or poor health assessment. By showing that the Green Yellow ratio (calculated from visible bands) is significantly different between tree health condition classes, our study reconfirms the potential utility of yellow band (Karlson et al. 2016) to describe tree health condition classes. Yet, NDVI1 outperformed Green Yellow ratio, which suggests that the change in NIR band reflectance for trees under poor health class is greater than the change in yellow band. This is likely because field-based tree health evaluators more strongly associate health with a change of leaf area rather than a change in leaf color.

We also focused on the most sensitive index, NDVI1, to evaluate how NDVI1 varied by tree species. Among the five most abundant tree species that had a large enough sample size to statistically analyze, we find that there are significant differences in NDVI1 (**Appendix B Figure s1**). Red maple had the lowest NDVI1, while red oak had the highest. Based on this finding of variability in VI among species, we also tried to repeat our analysis of the most sensitive VI on a species-specific basis. However, for each species, this analysis was hampered by low sample sizes in several health condition classes (**Appendix B Table s3**). Nevertheless, this exploratory species-specific analysis suggests that the most sensitive VI may not be consistent across different tree species (**Appendix B Table s4-8**).

Thus, even while we find NDVI1 to be the best all-purpose VI for assessing tree health condition classes, individual species may exhibit tree health condition symptoms that better align with spectral variability described by other VIs.

#### 4.4.2 Effects of image acquisition date

According to an analysis by Elmore et al. (2012) of the Landsat data record and phenology web camera (phenocam) imagery, the leaf-on period in Washington D.C extends from a full leaf expansion date around DOY 145 to an average date of leaf senescence around DOY 255 (**Figure 4.3**). These Landsat and phenocam data from our study area also both evidence a steady decline in greenness during this leaf-on period (Elmore et al. 2012). Importantly, our analysis from the three images illustrates that the ACOMP WV-3 data also captures this greendown phenology (**Figure 4.3**). As we detail in section 4.3, this decline in greenness, or greendown (Reaves et al. 2018), during the leaf-on period is a pervasive feature in vegetation phenology, and has important implications for selecting high resolution satellite data to assess street tree health.

We found that the July image (DOY 211), midway through the leaf-on period, appears to be the strongest for detecting a signal of statistical difference in NDVI1 based on field-measured tree health condition classes (**Table 1**). We suggest that the June image (DOY 162) is less helpful describing tree health condition symptoms because this date is close to the full leaf expansion date for some tree species. This timing could cause some trees to have the spectral effects of tree health convolved with spectral effects of a late full leaf expansion date. Indeed, in our exploratory species-specific analyses, we found that for red maple, and American elm, the VIs from the June image were actually more sensitive than from the July image (**Appendix B Tables s4-s6**). We suggest that the weaker strength of the June image may be driven by relatively late leaf out dates of other species, including the two most common species, Willow oak and pin oak. In these two oak species, the July image was more sensitive in detecting a statistical difference between tree health classes (**Appendix B Tables s7 & s8**). For all species, the August image (DOY 242) may have been too close to the average leaf senescence date,

which could cause some trees to have the spectral effects of tree health convolved with spectral effects of leaf senescence. Thus, we recommend using image dates that are safely within the middle of the leaf-on period.

A further supplemental analysis revealed that our study greatly benefitted from the availability of field-measured health condition data collected at the same time as the imagery during the 2017 growing season. We used other archived ACOMP WV-3 imagery from 2014, 2015, and 2016 to track the NDVI of all the trees in different 2017 health condition classes. This analysis shows that there was limited statistical difference in NDVI among trees in different health classes in the years preceding the field-based health assessment in 2017 (**Appendix B Figure s2**).

#### **4.4.3 Strategies for using remote sensing data in street tree health research and management**

In our preliminary work with these datasets, we attempted to apply machine learning algorithms to directly discriminate tree health classes from VI's, raw reflectance on each multispectral band, and panchromatic (30cm pixel resolution) image texture metrics extracted from each crown object. However, this analysis revealed that the ability to directly discriminate tree health classes is overwhelmed by other drivers of spectral reflectance in this complex urban environment. Nevertheless, our findings evidencing a robust, remotely detectable signal of statistical difference among the health classes of individual street trees does highlight two realistic and useful approaches for advancing the integration of high-resolution multispectral satellite data into street tree health research and management.

First, researchers and street tree managers could use multi-year, anniversary-date imagery to track VIs of individual trees. Tracking individual trees avoids the large spatial variation in VI likely driven by tree species differences (see e.g. **Appendix B Figure s1**), as well as myriad site-specific effects on VI caused by tree age, past management actions, and tree ecophysiological responses to the complex urban environment. In choosing anniversary-date imagery, we emphasize that researchers and managers

should attempt to minimize VI change related to tree phenology, and especially change associated with leaf expansion, greendown, and leaf senescence. Furthermore, researchers or managers wishing to target health assessments to individual tree species may find it advantageous to match the anniversary date and the selected diagnostic VI to the most sensitive date/VI for the target species (**see e.g. Appendix B Tables s4-s8**). This anniversary date approach will require future research in order to build on our single-year analysis, and confirm that changes in field-observed individual street tree health across years can be linked to inter-annual changes in VIs extracted from imagery.

Second, we suggest that researchers could use a VI threshold on a single-date image during the leaf-on period to identify priority trees for field-based street tree health evaluation. Based on our species-specific exploratory analysis, we suggest that this approach could also be refined by matching the image date and VI for a particular species of interest. For example, rather than using the all-purpose VI of NDVI1 on an image in the middle of the leaf-on period (e.g. our July image), an urban tree manager in Washington D.C. might be best advised to use the NDVI2 VI on an image from the early part of the leaf-on period in order to prioritize fieldwork assessing the health of red maple trees (**Appendix B Table s6**).

We suggest two other productive lines of future research that can overcome remaining challenges to remotely assessing the health of individual trees in urban settings. First, our study found that tree health was most expressed by changes in NDVI1, which is likely driven by reduced reflectance in the NIR as trees lose leaf area in response to declining health. Accordingly, future studies may profit from directly assessing these health-related crown structural changes using high-resolution LiDAR, especially from terrestrial and UAV platforms (Disney 2019). Second, our final dataset for tree health condition analysis included 2308 mature trees, which is only 13% of the entire street tree inventory of Washington D.C.. We suggest other near-surface remote sensing methods leveraging phenocams and high-resolution UAV data (Dash et al. 2018, Fillipa et al. 2018), may be successful in not only overcoming issues of image availability and cloud cover that limited our study area, but can also create opportunities to apply our findings to smaller and younger trees.

## 4.5 Conclusion

We evaluate the ability of six WV-3 VIs extracted from delineated tree street tree crown-objects on images from three different dates to find statistical differences among the field-measured health condition class of 2308 street trees in central Washington D.C. We find that spectral data useful for tree health analysis can be effectively extracted from imagery by delineating tree crown-objects using an NDVI-threshold mask within a buffer around a field-measured GPS point of the tree trunk. Extracted VIs from all image dates revealed highly significant differences among tree health condition classes. Moreover, NDVI1 from an image from the middle of the leaf-on period has the most potential to assess the health condition class of street trees in central Washington D.C. Our results thus provide insight into the utility of high-resolution satellite data as a tool for research and management of street tree health. In highlighting avenues for further research in this exciting application, we propose two strategies that account for essential considerations of phenology and tree species.

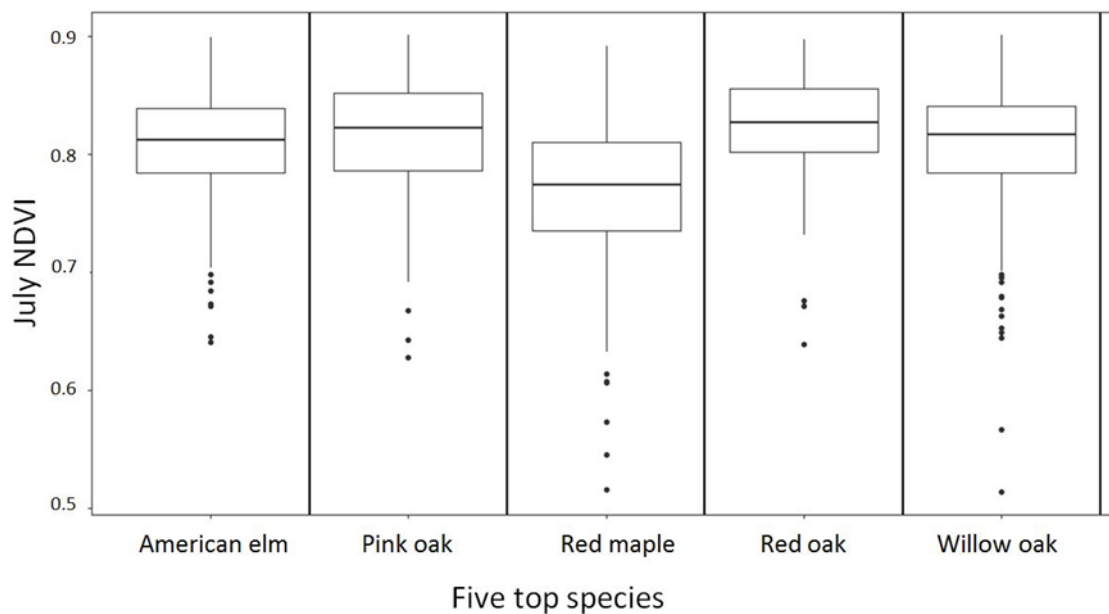
## Appendix B.

**Table s1** Tree health condition assessment parameters in the field

Parameter	Crown Quality (Value)	Vigor (Value)
0-25%	Poor (1)	Poor (1)
25-50%	Fair (2)	Fair (2)
50-75%	Good (3)	Good (3)
75-100%	Excellent (4)	Excellent (4)
Notes	Percentage of crown that is free from dieback, disease, and/or other physiological abnormality	Year-over-year shoot elongation and distribution of growth throughout crown.

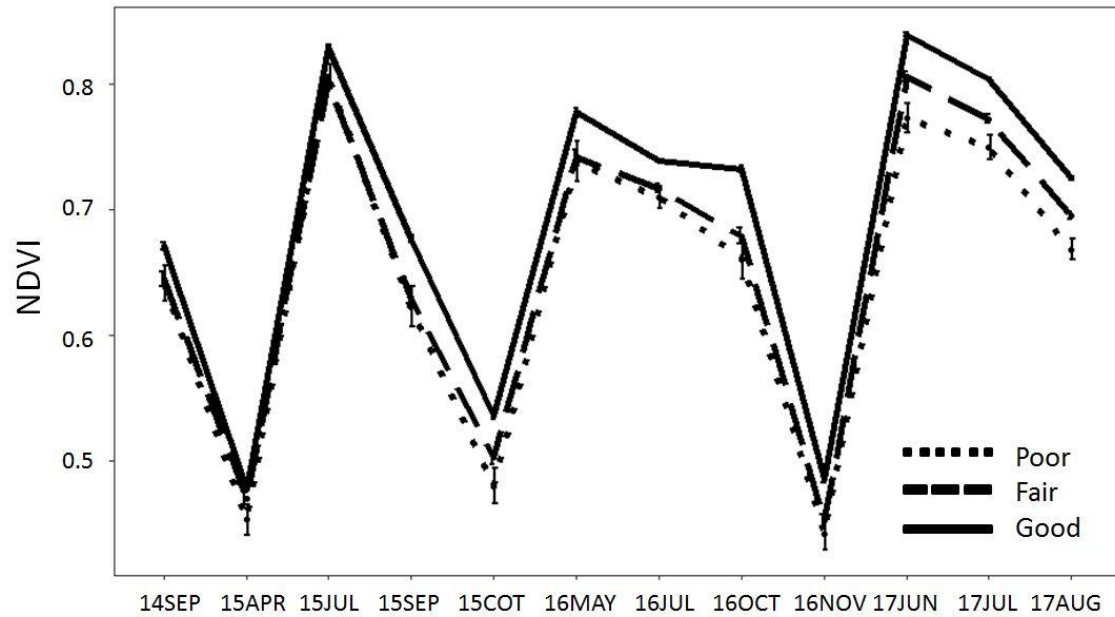
**Table s2** Tree health condition assessment. Score equals crown quality value plus Vigor value.

Condition	Score
Excellent	≥7
Good	5 to 6
Fair	3 to 5
Poor	2 to 3
Dead	0 to 1



**Figure s1.** Significant ( $p < 0.0001$ ) tree species differences in NDVI on the July 30<sup>th</sup>, 2017 image. These significant species differences were generally consistent across other Vis and image dates (data not shown).





**Figure s2** Median NDVI values for each of three health conditions from 2014 to 2017 during leaf-on period.

**Table s3** Number of trees within the three health condition classes for the top five species.

	Poor	Fair	Good	Sum
<b>American elm</b>	1	26	175	202
<b>Red oak</b>	3	18	121	142
<b>Pin oak</b>	9	25	95	129
<b>Red maple</b>	10	86	164	260
<b>Willow oak</b>	6	65	376	447

**Table s4** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis separating VI means among trees in good, fair and poor condition classes for American elm. For this and subsequent tables, the bold value denotes the most sensitive image date/VI combination. Statistically significant ( $p < 0.05$ ) tests are in red.

Vegetation Index	June		July		August	
	Chi-square	P>ChiSq	Chi-square	P>ChiSq	Chi-square	P>ChiSq
<b>NDVI1</b>	4.44	0.11	0.4882	0.7834	3.1973	0.2022
<b>NDVI2</b>	4.0684	0.1308	0.4141	0.813	2.9087	0.2336
<b>RENDVI1</b>	5.6096	0.0605	2.4034	0.3007	1.7963	0.4073
<b>RENDVI2</b>	3.6469	0.1615	2.2661	0.3221	0.7242	0.6962
<b>Green Yellow ratio</b>	5.10	0.078	1.9044	0.3859	2.438	0.2955
<b>Green Red ratio</b>	6.6127	0.0366	0.4615	0.7939	3.2143	0.2005

**Table s5** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis separating VI means among trees in good, fair and poor condition classes for red oak.

Vegetation Index	June		July		August	
	Chi-square	P>ChiSq	Chi-square	P>ChiSq	Chi-square	P>ChiSq
<b>NDVI1</b>	5.2389	0.0728	7.8841	0.0194	9.1013	0.0106
<b>NDVI2</b>	4.4924	0.1058	8.2025	0.0166	6.0391	0.0488
<b>RENDVI1</b>	7.1241	0.0284	5.7159	0.0574	4.8124	0.0902
<b>RENDVI2</b>	3.7627	0.1524	4.633	0.0986	3.6675	0.1598
<b>Green Yellow ratio</b>	6.1451	0.0463	9.4742	0.0088	7.5254	0.0232
<b>Green Red ratio</b>	6.1538	0.0461	9.8965	0.0071	5.574	0.0616

**Table s6** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis separating VI means among trees in good, fair and poor health condition classes for Red maple.

Vegetation Index	June		July		August	
	Chi-square	P>ChiSq	Chi-square	P>ChiSq	Chi-square	P>ChiSq
<b>NDVI1</b>	29.9469	<0.0001	31.3282	<0.0001	16.5029	0.0003
<b>NDVI2</b>	32.7131	<0.0001	32.0169	<0.0001	19.4022	<0.0001
<b>RENDVI1</b>	24.7394	<0.0001	30.424	<0.0001	6.2454	0.044
<b>RENDVI2</b>	30.4282	<0.0001	33.919	<0.0001	6.8311	0.0329
<b>Green Yellow ratio</b>	8.1561	0.0169	11.4977	0.0032	4.2895	0.1171
<b>Green Red ratio</b>	24.9307	<0.0001	11.3198	0.0035	4.0845	0.1297

**Table s7** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis separating VI means among trees in good, fair and poor health condition classes for Pin oak.

Vegetation Index	June		July		August	
	Chi-square	P>ChiSq	Chi-square	P>ChiSq	Chi-square	P>ChiSq
<b>NDVI1</b>	23.279	<0.0001	25.2396	<0.0001	16.7129	0.0002
<b>NDVI2</b>	23.6311	<0.0001	23.9345	<0.0001	17.098	0.0002
<b>RENDVI1</b>	10.3354	0.0057	6.9747	0.0306	0.2662	0.8754
<b>RENDVI2</b>	9.5863	0.0083	5.8788	0.0529	1.1273	0.5691
<b>Green Yellow ratio</b>	17.1886	0.0002	16.1929	0.0003	9.5148	0.0086
<b>Green Red ratio</b>	16.4918	0.0003	18.2798	0.0001	11.9735	0.0025

**Table s8** Chi-square values from non-parametric Wilcoxon/Kruskal-Wallis separating VI means among trees in good, fair and poor health condition classes for Willow oak.

Vegetation Index	June		July		August	
	Chi-square	P>ChiSq	Chi-square	P>ChiSq	Chi-square	P>ChiSq
<b>NDVI1</b>	7.7428	0.0208	8.0214	0.0181	6.9741	0.0306
<b>NDVI2</b>	6.6044	0.0368	6.5345	0.0381	6.2732	0.0434
<b>RENDVI1</b>	15.949	0.0003	19.0574	<0.0001	12.4481	0.002
<b>RENDVI2</b>	17.195	0.0002	15.1112	0.0005	13.1384	0.0014
<b>Green Yellow ratio</b>	3.7075	0.1566	3.1059	0.2116	5.4296	0.0662
<b>Green Red ratio</b>	1.5415	0.4627	2.5702	0.2766	0.087	0.9574

## 5. Conclusion and future work

### 5.1 Dissertation summary

This dissertation expands the sub-field of using remote sensing datasets to discriminate individual tree species and describe tree health classes statistically. For trees in rural forests, I successfully classified four broadleaf tree species using passive multi-spectral remote sensing imagery and active leaf-on LiDAR data in Fernow Experimental Forest, WV. For trees in urban settings, I classified 19 broadleaf tree species or 10 genera using a suit of VHR WorldView-3 images in Washington D.C.. Moreover, I also detected the statistical differences among three health conditions for these urban street trees using VHR WorldView-3 images in 2017. This research not only benefits researchers in remote sensing, but also benefits forest managers and urban foresters by making specific contributions in each of the three manuscripts.

In the first manuscript, the combination of multiple VHR images with leaf-on LiDAR data improves the discrimination of four broadleaf deciduous tree species: red oak, sugar maple, black cherry and tulip poplar in Fernow Experimental Forest. My assessment evaluated the potential of six types of features in 1) spring, (2) summer, and (3) autumn phenology, (4) height and intensity data from leaf-on LiDAR, (5) spectral indices and (6) texture features to discriminate tree species. I concluded that combining high spatial resolution multi-temporal satellite data with LiDAR datasets can enhance the ability to discriminate tree species at the crown level. Summer phenology is the most helpful feature, and crown structural features from LiDAR are the second-best variables for tree species discrimination.

In the second manuscript, by using RandomForest classifier to leverage these three advances 1) 12 high spatial resolution images covering each pheno-phase of the growing season from April to November, 2) large forest inventory datasets and 3) spatially accurate crown objects, I concluded that images acquired from spring leaf emergence and fall senescence are the optimum pheno-phase to discriminate individual tree species in Washington D.C.. I also found that remote sensing images

successfully captured the pigment-induced changes between different species especially using red and green visible bands. Moreover, VHR images have the potential to classify trees with an overall accuracy of 63.7 % at the species level and 75.6 % at the genus level. The confusion for tree species classification is associated with the decreased sample size of smaller class, and species in similar taxonomic groups.

Finally, in the third manuscript, my results confirmed the utility of VHR satellite data as a tool for research and management of street tree health. The six types of vegetation indices calculated from three WorldView-3 images (June 11<sup>th</sup> July 30<sup>th</sup> and August 30<sup>th</sup>, 2017) were statistically different between three categories of tree health conditions. I concluded that VHR remote sensing imagery have the potential to detect differences between tree health conditions statistically, and NDVI from an image from the middle of the leaf-on period performed the best. Since these VHR multi-temporal remote sensing image also have the potential to capture green-down phenology, I proposed two strategies for using VHR satellite data in street tree health management: (1) using anniversary date imagery from similar pheno-phase to track the VIs of individual trees from year to year in order to minimize VI change related to tree phenology, and (2) using single-date imagery to prioritize field-based street tree health assessments.

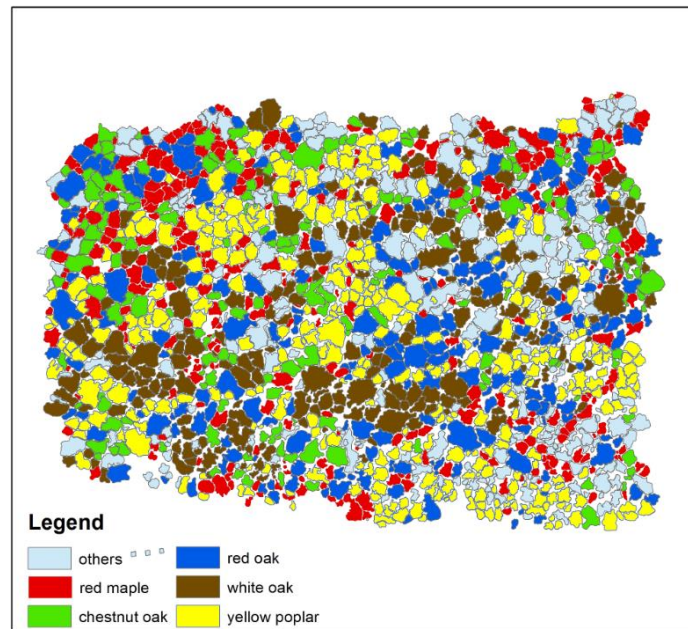
## **5.2 Remaining challenges and future research**

While providing insights and promising results for tree species and health status detection using remote sensing in both rural and urban forests, this dissertation also elucidated other topics worthy of additional investigations. First, one of the limitations for the first chapter is the complexity of terrain conditions in Fernow Experimental Forest. It is more challenging to precisely co-register multiple images from leaf-on period and locate individual trees based on their inaccurate field-surveyed coordinates due to steep slopes. Exploring exactly how slope and complex terrain conditions influence the tree location identification from the imagery would be a major leap forward for tree species classification in rural areas. Second, my dissertation only implemented two machine learning classifiers (RandomForest and

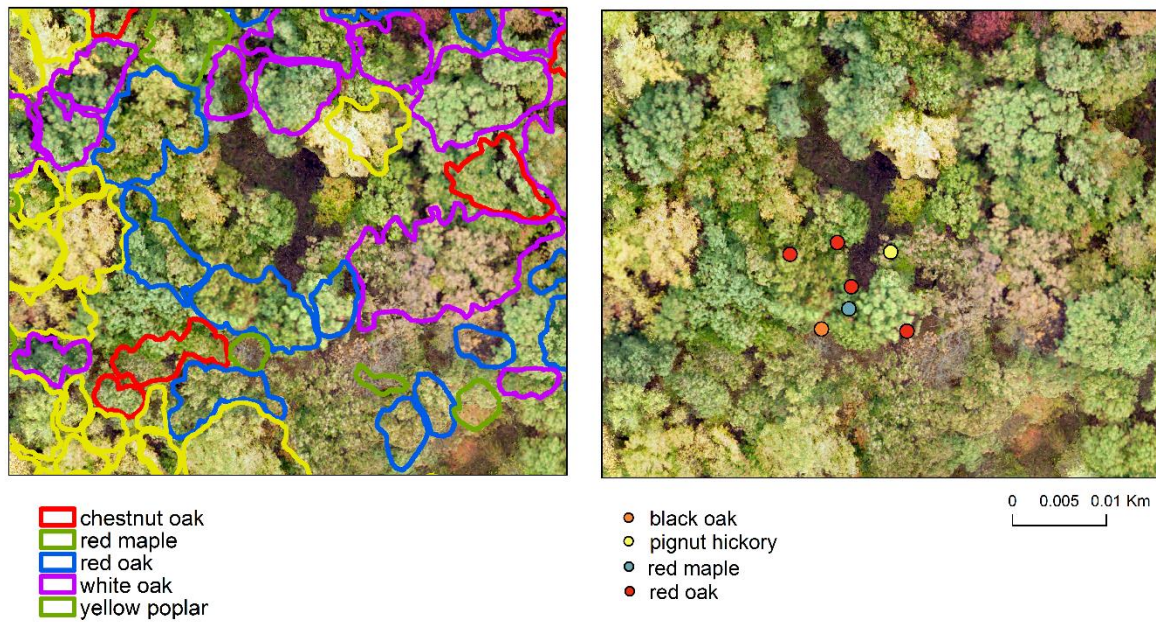
Decision trees) for tree species classification. It is necessary to scientifically determine the optimum classifier by comparing more machine learning classifiers for tree species classification.

Towards these two additional investigations mentioned above, I have used a multi-seasonal dataset of three WorldView2 images (image acquisition date: 5/14/2015, 8/16/2014, 9/23/2010) and the large field inventory (19 species with 1528 individual trees) to discriminate tree species in West Virginia Research Forest (WV Research Forest). For the field inventory, Dr. Landenberger and his colleagues manually generated non-overlapping 1528 tree crowns with 19 species (**Figure 5.1**) by field interpretation of high spatial resolution aerial photos in 2000. The sample size and accuracy of this dataset is considerable for testing machine learning classifiers, especially relative to other published studies which typically have sample sizes much less than 1500 crowns. Crown boundaries were drawn carefully in the field with gaps identified to minimize the effects from background or terrain in 2000. The major concern or drawback of these 1528 crowns is they were created more than 15 years ago. Even though it is impossible for trees to change location or species type, trees may have fallen down and new trees might grow or trees become damaged during these years. In order to increase the reliability and decrease the error due to the 19 year gap, I updated the reference dataset using high spatial resolution (about 5cm resolution) UAV images collected for Dr. McNeil and Dr. Michael Strager in spring, summer, and fall 2016. I identified new gaps and remove crowns if they do no longer exist. However, there are still large discrepancies of tree locations between the updated field inventory map and three WorldView-2 images. Next, I also tried to adopt another inventory from Henry Lieberman, who had collected 1991 crowns, 26 species with location, height and DBH in WVU Research Forest in summer 2016. I tried to match tree locations using spatial join tool for the crown polygons from Dr. Landenberger in 2000 and tree stems surveyed in 2016. However, due to the complexity of terrain conditions (steep slopes) and the overlap of tree crowns in WV Research Forest, it is still challenging to match tree locations based on the measurements from 2000 and 2016 (**Figure 5.2**). There were still discrepancies for tree stem location and species information. In order to ultimately test different machine learning classifiers, accurate forest

inventory reference datasets are necessary. Bringing in high quality LiDAR data with both canopy height model and digital terrain model in WV Research Forest as the next step will help to minimize the influences of slope and complex terrain conditions for tree location identification. With the updated accurate forest inventory as a reference map, I can proceed to test different machine learning classifiers including SVM, RF, Decision Tree, k-means Nearest Neighborhood and Neural Network for tree species classification.



**Figure 5.1** Species distribution for WV Research Forest in 2000



**Figure 5.2** Inconsistency of tree species reference map in WV Research Forest. Crown polygons were delineated by Dr. Landenberger in 2000 (Left). Tree stem locations were provided by Henry Liebermann in 2016 (Right). Dr. Michael Strager provided the UAV image in 2016 as base map

Additionally, for tree health condition detection in Washington D.C., I only considered vegetation indices, texture and phenological patterns provided by different spectral bands between three categories of health conditions. Trees under poor health condition may also result in changes like the loss of branches and leaves, which will change their crown-level structure and architecture compared with healthy trees. High-quality LiDAR provides the opportunity to describe tree health conditions based on their crown-level architecture. Examining the potential of LiDAR using texture and structural features is my third future work direction and it would be another significant insight into tree health condition detection in urban settings. Moreover, by calculating biomass and tree height from LiDAR data, I can also assess the forest loss due to trees under poor health condition in urban settings.



## 6. Reference

- Ahokas, E., Kaasalainen, S., Hyypä, J., & Suomalainen, J. (2006). Calibration of the optech ALTM 3100 laser scanner intensity data using brightness targets. In *Revue Francaise de Photogrammetrie et de Teledetection* (pp. 10–16).
- Alonzo, M., Bookhagen, B., & Roberts, D. A. (2014). Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sensing of Environment*, 148, 70–83.  
<https://doi.org/10.1016/j.rse.2014.03.018>
- Andrada II, R., Deng, J., & Gazal, L. (2015). Exploring peoples preferences on specific attributes of urban forests in Washington DC: A conjoint approach. *Journal of Horticulture and Forestry*, 7(10), 200–209.
- Asmaryan, S., Warner, T. A., Muradyan, V., & Nersisyan, G. (2013). Mapping tree stress associated with urban pollution using the WorldView-2 Red Edge band. *Remote Sensing Letters*, 4(2), 200–209.  
<https://doi.org/10.1080/2150704X.2012.715771>
- Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, 64(3), 234–253. [https://doi.org/10.1016/S0034-4257\(98\)00014-5](https://doi.org/10.1016/S0034-4257(98)00014-5)
- Au. (2018). An integrated approach to tree stress monitoring. *Arborist news*.  
[http://geocarto.com.hk/readings/aug2018\\_issue/#page/1](http://geocarto.com.hk/readings/aug2018_issue/#page/1)
- Banfield, W. M. (1968). Dutch Elm Disease Recurrence and Recovery in American Elm. *Journal of Phytopathology*, 62(1), 21–60. <https://doi.org/10.1111/j.1439-0434.1968.tb02345.x>
- Boochs, F., Kupfer, G., Dockter, K., & Kuhbaüch, W. (1990). Shape of the red edge as vitality indicator for plants. *International Journal of Remote Sensing*, 11(10), 1741–1753.  
<https://doi.org/10.1080/01431169008955127>
- Brandtberg, T. (2007). Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(5), 325–340.  
<https://doi.org/10.1016/j.isprsjprs.2006.10.006>
- Brandtberg, T., Warner, T. A., Landenberger, R. E., & McGraw, J. B. (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, 85(3), 290–303.  
[https://doi.org/10.1016/S0034-4257\(03\)00008-7](https://doi.org/10.1016/S0034-4257(03)00008-7)
- Budei, B. C., St-Onge, B., Hopkinson, C., & Audet, F. A. (2018). Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. *Remote Sensing of Environment*.  
<https://doi.org/10.1016/j.rse.2017.09.037>
- Burnham, M. B., Cumming, J. R., Adams, M. B., & Peterjohn, W. T. (2017). Soluble soil aluminum alters the relative uptake of mineral nitrogen forms by six mature temperate broadleaf tree species: possible implications for watershed nitrate retention. *Oecologia*, 185(3), 327–337.  
<https://doi.org/10.1007/s00442-017-3955-8>
- Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P. P., Osório, M. L., ... Pinheiro, C. (2002). How plants cope with water stress in the field. Photosynthesis and growth. *Annals of Botany*, 89(SPEC. ISS.), 907–916. <https://doi.org/10.1093/aob/mcf105>

- Chavez Jr, P. S., Chavez, P. S., & Chavez Jr., P. S. (1996). Image-Based Atmospheric Corrections - Revisited and Improved. *Photogrammetric Engineering & Remote Sensing*, 62(9), 1025–1036.
- Cho, M. A., Malahlela, O., & Ramoelo, A. (2015). Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. *International Journal of Applied Earth Observation and Geoinformation*, 38, 349–357. <https://doi.org/10.1016/j.jag.2015.01.015>
- Cochrane, M. A. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. *International Journal of Remote Sensing*, 21(10), 2075–2087. <https://doi.org/10.1080/01431160050021303>
- Colgan, M. S., Baldeck, C. A., Féret, J. baptiste, & Asner, G. P. (2012). Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. *Remote Sensing*, 4(11), 3462–3480. <https://doi.org/10.3390/rs4113462>
- Dalponte, M., Bruzzone, L., & Gianelle, D. (2008). Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5), 1416–1427. <https://doi.org/10.1109/TGRS.2008.916480>
- Dalponte, M., Ørka, H. O., Gobakken, T., Gianelle, D., & Næsset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2632–2645. <https://doi.org/10.1109/TGRS.2012.2216272>
- Dash, J. P., Pearce, G. D., & Watt, M. S. (2018). UAV multispectral imagery can complement satellite data for monitoring forest health. *Remote Sensing*, 10(8), 1–22. <https://doi.org/10.3390/rs10081216>
- Delalieux, S., Somers, B., Verstraeten, W. W., van Aardt, J. A. N., Keulemans, W., & Coppin, P. (2009). Hyperspectral indices to diagnose leaf biotic stress of apple plants, considering leaf phenology. *International Journal of Remote Sensing*, 30(8), 1887–1912. <https://doi.org/10.1080/01431160802541556>
- Del Tredici, P. (2007). The Phenology of Sexual Reproduction in Ginkgo biloba: Ecological and Evolutionary Implications. *The Botanical Review*, 73(4), 267–278. [https://doi.org/10.1663/0006-8101\(2007\)73\[267:tposri\]2.0.co;2](https://doi.org/10.1663/0006-8101(2007)73[267:tposri]2.0.co;2)
- DigitalGlobe Inc. (2016). *Digitalglobe Atmospheric Compensation*. Retrieved from <https://explore.digitalglobe.com/AComp.html>
- Disney, M. (2019). Terrestrial LiDAR: a three-dimensional revolution in how we look at trees. *New Phytologist*, 222(4), 1736–1741. <https://doi.org/10.1111/nph.15517>
- District of Columbia Government. (2009). *LiDAR - Reflective Surface*. Retrieved from <https://dcgis.maps.arcgis.com/home/item.html?id=1057f1898171442487816b7f891dc2fc>
- District of Columbia Government. (2018). *Urban Forestry Street Trees*. Retrieved from [http://opendata.dc.gov/datasets/f6c3c04113944f23a7993f2e603abaf2\\_23](http://opendata.dc.gov/datasets/f6c3c04113944f23a7993f2e603abaf2_23)
- Donoghue, D. N. M., Watt, P. J., Cox, N. J., & Wilson, J. (2007). Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, 110(4), 509–522. <https://doi.org/10.1016/j.rse.2007.02.032>

- Dwyer, J. F., Nowak, D. J., & Noble, M. H. (2003). Sustaining urban forests. *Journal of Arboriculture*, 29(1), 49–55.
- Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., ... Vierling, K. T. (2016). Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2016.08.018>
- Eitel, J. U. H., Vierling, L. A., Litvak, M. E., Long, D. S., Schulthess, U., Ager, A. A., ... Stoscheck, L. (2011). Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sensing of Environment*, 115(12), 3640–3646. <https://doi.org/10.1016/j.rse.2011.09.002>
- Elmore, A. J., Guinn, S. M., Minsley, B. J., & Richardson, A. D. (2012). Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests. *Global Change Biology*, 18(2), 656–674. <https://doi.org/10.1111/j.1365-2486.2011.02521.x>
- Eamus, D., A. Huete, and Q. Yu. (2016). “Vegetation Dynamics: a Synthesis of Plant Ecophysiology.” Remote Sensing and Modelling. *New York: Cambridge University Press*.
- Evans, J. S., and M. A. Murphy. (2014) Rfutilities. *R Package Version 1.0-0*. <http://CRAN.R-project.org/package=rfUtilities>
- Fang, F., Im, J., Lee, J., & Kim, K. (2016). An improved tree crown delineation method based on live crown ratios from airborne LiDAR data. *GIScience and Remote Sensing*, 53(3), 402–419. <https://doi.org/10.1080/15481603.2016.1158774>
- Fang, F., McNeil, B. E., Warner, T. A., & Maxwell, A. E. (2018). Combining high spatial resolution multi-temporal satellite data with leaf-on LiDAR to enhance tree species discrimination at the crown level. *International Journal of Remote Sensing*, 39(23), 9054–9072. <https://doi.org/10.1080/01431161.2018.1504343>
- Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., ... Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2016.08.013>
- Ferreira, M. P., Wagner, F. H., Aragão, L. E. O. C., Shimabukuro, Y. E., & de Souza Filho, C. R. (2019). Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 149, 119–131. <https://doi.org/10.1016/j.isprsjprs.2019.01.019>
- Filippa, G., Cremonese, E., Migliavacca, M., Galvagno, M., Sonnentag, O., Humphreys, E., ... Richardson, A. D. (2018). NDVI derived from near-infrared-enabled digital cameras: Applicability across different plant functional types. *Agricultural and Forest Meteorology*, 249(April 2017), 275–285. <https://doi.org/10.1016/j.agrformet.2017.11.003>
- Franke, J., & Menz, G. (2007). Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture*, 8(3), 161–172. <https://doi.org/10.1007/s11119-007-9036-y>
- Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, 114(8), 1805–1816. <https://doi.org/10.1016/j.rse.2010.04.005>
- Gates, D. M., Keegan, H. J., Schleter, J. C., & Weidner, V. R. (1965). Spectral Properties of Plants. *Applied*

*Optics*, 4(1), 11. <https://doi.org/10.1364/ao.4.000011>

- Gering, L. R., & May, D. M. (1995). The relationship of diameter at breast height and crown diameter for four species groups in Hardin County, Tennessee. *Southern Journal of Applied Forestry*, 19(4), 177–181.
- Ghosh, A., Fassnacht, F. E., Joshi, P. K., & Kochb, B. (2014). A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation*, 26(1), 49–63. <https://doi.org/10.1016/j.jag.2013.05.017>
- Government of the District of Columbia/ District Department of Transportation. (2010). *District of Columbia Assessment of Urban Forest Resources and Strategy*. Washington D.C
- Gregg, J. W., Jones, C. G., & Dawson, T. E. (2003). Urbanization effects on tree growth in the vicinity of New York City. *Nature*, 424(6945), 183–187. <https://doi.org/10.1038/nature01728>
- Hadjimitsis, D. G., Clayton, C. R. I., & Retalis, A. (2009). The use of selected pseudo-invariant targets for the application of atmospheric correction in multi-temporal studies using satellite remotely sensed imagery. *International Journal of Applied Earth Observation and Geoinformation*, 11(3), 192–200. <https://doi.org/10.1016/j.jag.2009.01.005>
- Haralick, R. M., Dinstein, I., & Shanmugam, K. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Harris, J. L., & Balci, Y. (2015). Population structure of the bacterial pathogen *Xylella fastidiosa* among street trees in Washington D.C. *PLoS ONE*, 10(3). <https://doi.org/10.1371/journal.pone.0121297>
- He, Y., Bo, Y., de Jong, R., Li, A., Zhu, Y., & Cheng, J. (2015). Comparison of vegetation phenological metrics extracted from GIMMS NDVIg and MERIS MTCI data sets over China. *International Journal of Remote Sensing*, 36(1), 300–317. <https://doi.org/10.1080/01431161.2014.994719>
- Heinzel, J., & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 13(1), 152–160. <https://doi.org/10.1016/j.jag.2010.09.010>
- Hill, R. A., Wilson, A. K., George, M., & Hinsley, S. A. (2010). Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science*, 13(1), 86–99. <https://doi.org/10.1111/j.1654-109X.2009.01053.x>
- Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90(4), 415–423. [https://doi.org/10.1016/S0034-4257\(03\)00140-8](https://doi.org/10.1016/S0034-4257(03)00140-8)
- Holmgren, J., Persson, Å., & Söderman, U. (2008). Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. In *International Journal of Remote Sensing* (Vol. 29, pp. 1537–1552). <https://doi.org/10.1080/01431160701736471>
- Huang, W., Lamb, D. W., Niu, Z., Zhang, Y., Liu, L., & Wang, J. (2007). Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture*, 8(4–5), 187–197. <https://doi.org/10.1007/s11119-007-9038-9>
- Hubert, F., Aragão, L. E. O. C., Roberto, C., & Filho, D. S. (2019). Tree species classification in tropical

forests using visible to shortwave infrared WorldView-3 images and texture analysis.

- Hughes, G. F. (1968). On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Transactions on Information Theory*, 14(1), 55–63. <https://doi.org/10.1109/TIT.1968.1054102>
- Im, J., Quackenbush, L. J., Li, M., & Fang, F. (2014). Optimum Scale in Object-Based Image Analysis. In *Scale Issues in Remote Sensing* (Vol. 9781118305, pp. 197–214). <https://doi.org/10.1002/9781118801628.ch10>
- Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree species classification with Random forest using very high spatial resolution 8-band worldView-2 satellite data. *Remote Sensing*, 4(9), 2661–2693. <https://doi.org/10.3390/rs4092661>
- Jackson, R. D. (1986). Remote Sensing of Biotic and Abiotic Plant Stress. *Annual Review of Phytopathology*, 24(1), 265–287. <https://doi.org/10.1146/annurev.py.24.090186.001405>
- Jensen, R. R., Hardin, P. J., & Hardin, A. J. (2012). Classification of urban tree species using hyperspectral imagery. *Geocarto International*, 27(5), 443–458. <https://doi.org/10.1080/10106049.2011.638989>
- Jones, T. G., Coops, N. C., & Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment*, 114(12), 2841–2852. <https://doi.org/10.1016/j.rse.2010.07.002>
- Karlson, M., Ostwald, M., Reese, H., Bazié, H. R., & Tankoano, B. (2016). Assessing the potential of multi-seasonal WorldView-2 imagery for mapping West African agroforestry tree species. *International Journal of Applied Earth Observation and Geoinformation*, 50, 80–88. <https://doi.org/10.1016/j.jag.2016.03.004>
- Key, T., Warner, T. A., McGraw, J. B., & Fajvan, M. A. (2001). A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sensing of Environment*, 75(1), 100–112. [https://doi.org/10.1016/S0034-4257\(00\)00159-0](https://doi.org/10.1016/S0034-4257(00)00159-0)
- Kim, S (2007) Individual tree species identification using LIDAR- derived crown structures and intensity data. Doctoral dissertation, University of Washington.
- Kim, S., McGaughey, R. J., Andersen, H. E., & Schreuder, G. (2009). Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2009.03.017>
- Klosterman, S. T., Hufkens, K., Gray, J. M., Melaas, E., Sonnentag, O., Lavine, I., ... Richardson, A. D. (2014). Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. *Biogeosciences*, 11(16), 4305–4320. <https://doi.org/10.5194/bg-11-4305-2014>
- Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*, 1(3), 155–159. [https://doi.org/10.1016/S0034-4257\(70\)80021-9](https://doi.org/10.1016/S0034-4257(70)80021-9)
- Knyazikhin, Y., Schull, M. A., Stenberg, P., Mörtus, M., Rautiainen, M., Yang, Y., ... Myneni, R. B. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3), E185–92. <https://doi.org/10.1073/pnas.1210196109>

- Korpela, I., Ole Ørka, H., Maltamo, M., Tokola, T., & Hyypä, J. (2010). Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44(2), 319–339.
- Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, et al. (2016). *Caret: Classification and Regression Training. R package version 6.0-73*. <https://cran.rproject.org/web/packages/caret/index.html>.
- L.Sayn-Wittgenstein. (1978). Recognition of tree species on aerial photographs. *Photogrammetria*. [https://doi.org/10.1016/0031-8663\(80\)90020-4](https://doi.org/10.1016/0031-8663(80)90020-4)
- Lambert, J., Drenou, C., Denux, J. P., Balent, G., & Cheret, V. (2013). Monitoring forest decline through remote sensing time series analysis. *GIScience and Remote Sensing*, 50(4), 437–457. <https://doi.org/10.1080/15481603.2013.820070>
- Lamson, N. I. (1987). D.b.h./Crown Diameter Relationships in Mixed Appalachian Hardwood stands. *US Department of Agriculture. Forest Service Research Paper, NE-610*, 1–3. Retrieved from [http://www.fs.fed.us/ne/newtown\\_square/publications/research\\_papers/pdfs/scanned/OCR/ne\\_r p610.pdf](http://www.fs.fed.us/ne/newtown_square/publications/research_papers/pdfs/scanned/OCR/ne_r p610.pdf)
- Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., ... Meyer-Lindenberg, A. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature*, 474(7352), 498–501. <https://doi.org/10.1038/nature10190>
- Li, D., Ke, Y., Gong, H., & Li, X. (2015). Object-based urban tree species classification using bi-temporal worldview-2 and worldview-3 images. *Remote Sensing*, 7(12), 16917–16937. <https://doi.org/10.3390/rs71215861>
- Li, X., Lee, W. S., Li, M., Ehsani, R., Mishra, A. R., Yang, C., & Mangan, R. L. (2012). Spectral difference analysis and airborne imaging classification for citrus greening infected trees. *Computers and Electronics in Agriculture*, 83, 32–46. <https://doi.org/10.1016/j.compag.2012.01.010>
- Liu, D., & Xia, F. (2010). Remote Sensing Letters Assessing object-based classification: advantages and limitations Assessing object-based classification: advantages and limitations. *Remote Sensing Letters*, 1(4), 187–194. <https://doi.org/10.1080/01431161003743173>
- Liu, T., Im, J., & Quackenbush, L. J. (2015). A novel transferable individual tree crown delineation model based on Fishing Net Dragging and boundary classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110, 34–47. <https://doi.org/10.1016/j.isprsjprs.2015.10.002>
- Madonsela, S., Cho, M. A., Mathieu, R., Mutanga, O., Ramoelo, A., Kaszta, Ž., ... Wolff, E. (2017). Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species. *International Journal of Applied Earth Observation and Geoinformation*, 58, 65–73. <https://doi.org/10.1016/j.jag.2017.01.018>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*. <https://doi.org/10.1080/01431161.2018.1433343>
- Maxwell, A. E., Warner, T. A., Vanderbilt, B. C., & Ramezan, C. A. (2017). Land Cover Classification and Feature Extraction from National Agriculture Imagery Program (NAIP) Orthoimagery: A Review. *Photogrammetric Engineering & Remote Sensing*, 83(11), 737–747. <https://doi.org/10.14358/pers.83.10.737>

- McIntyre, N. E. (2006). Ecology of Urban Arthropods: A Review and a Call to Action. *Annals of the Entomological Society of America*, 93(4), 825–835. [https://doi.org/10.1603/0013-8746\(2000\)093\[0825:eouaar\]2.0.co;2](https://doi.org/10.1603/0013-8746(2000)093[0825:eouaar]2.0.co;2)
- McPherson, G., Nowak, D., & Rowntree, R. (1994). Chicago's urban forest ecosystem: results of the Chicago Urban Forest Climate Project: GTR-NE-186. In *Chicago's urban forest ecosystem: results of the Chicago Urban Forest Climate Project. Gen. Tech. Rep. NE-186*.
- Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1778–1790. <https://doi.org/10.1109/TGRS.2004.831865>
- Memiaghe, H. R., Lutz, J. A., Korte, L., Alonso, A., & Kenfack, D. (2016). Ecological Importance of Small-Diameter Trees to the Structure, Diversity and Biomass of a Tropical Evergreen Forest at Rabi, Gabon. *PLoS ONE*, 11(5). <https://doi.org/10.1371/journal.pone.0154988>
- Michez, A., Piégay, H., Lisein, J., Claessens, H., & Lejeune, P. (2016). Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environmental Monitoring and Assessment*, 188(3), 1–19. <https://doi.org/10.1007/s10661-015-4996-2>
- Miller, R. W. (1996). *Urban forestry: planning and managing urban greenspaces. Urban forestry: planning and managing urban greenspaces*.
- Mohammed, G. H., Noland, T. L., Irving, P. H., Sampson, P. H., Zarco-Tejada, P. J., & Miller, J. R. (2000). *Natural and stress-induced effects on leaf spectral reflectance in Ontario species. Forest Research Report* (Vol. 156).
- Nascimbene, J., Marini, L., Motta, R., & Nimis, P. L. (2009). Influence of tree age, tree size and crown structure on lichen communities in mature Alpine spruce forests. *Biodiversity and Conservation*, 18(6), 1509–1522. <https://doi.org/10.1007/s10531-008-9537-7>
- Nowak, D. J., & Crane, D. E. (2002). Carbon storage and sequestration by urban trees in the USA. In *Environmental Pollution* (Vol. 116, pp. 381–389). [https://doi.org/10.1016/S0269-7491\(01\)00214-7](https://doi.org/10.1016/S0269-7491(01)00214-7)
- Nowak, D. J., Hoehn III, R. E., Crane, D. E., Stevens, J. C., & Walton, J. T. (2006). Assessing urban forest effects and values: Washington, D.C.'s urban forest. *Resource Bulletin - Northern Research Station, USDA Forest Service*, (No.NRS-1), 24 pp.
- Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*. <https://doi.org/10.1111/j.1469-8137.2010.03536.x>
- Omer, G., Mutanga, O., Abdel-Rahman, E. M., & Adam, E. (2016). Empirical prediction of leaf area index (LAI) of endangered tree species in intact and fragmented indigenous forests ecosystems using WorldView-2 data and two robust machine learning algorithms. *Remote Sensing*, 8(4). <https://doi.org/10.3390/rs8040324>
- Pontius, J., Martin, M., Plourde, L., & Hallett, R. (2008). Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies. *Remote Sensing of Environment*, 112(5), 2665–2676. <https://doi.org/10.1016/j.rse.2007.12.011>
- Pu, R., & Landry, S. (2012). A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment*, 124, 516–533.

<https://doi.org/10.1016/j.rse.2012.06.011>

- Pu, R., Landry, S., & Yu, Q. (2018). Assessing the potential of multi-seasonal high resolution Pléiades satellite imagery for mapping urban tree species. *International Journal of Applied Earth Observation and Geoinformation*, 71, 144–158. <https://doi.org/10.1016/j.jag.2018.05.005>
- Rautiainen, M., Stenberg, P., Nilson, T., & Kuusk, A. (2004). The effect of crown shape on the reflectance of coniferous stands. *Remote Sensing of Environment*, 89(1), 41–52. <https://doi.org/10.1016/j.rse.2003.10.001>
- Reaves, V. C., Elmore, A. J., Nelson, D. M., & McNeil, B. E. (2018). Drivers of spatial variability in greendown within an oak-hickory forest landscape. *Remote Sensing of Environment*, 210, 422–433. <https://doi.org/10.1016/j.rse.2018.03.027>
- Reed, B. C., Schwartz, M. D., & Xiao, X. (2009). Remote Sensing Phenology. In *Phenology of Ecosystem Processes* (pp. 231–246). [https://doi.org/10.1007/978-1-4419-0026-5\\_10](https://doi.org/10.1007/978-1-4419-0026-5_10)
- Reed, Bradley C, Brown, J. F., Vanderzee, D., Loveland, T. R., James W, & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5, 703–714. <https://doi.org/10.2307/3235884>
- Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, 96(3–4), 366–374. <https://doi.org/10.1016/j.rse.2005.03.008>
- Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2010.02.007>
- Sooyoung Kim. (2007). Individual Tree Species Identification using LIDAR- derived Crown Structures and Intensity Data, 1–121.
- Therneau, T., B. Atkinson, and B. Ripley. (2017). Rpart: Recursive Partitioning and Regression Trees. *R package version 4.1-11*. <https://cran.rproject.org/web/packages/rpart/index.html>.
- Toutin, T. (2004). Geometric processing of remote sensing images: Models, algorithms and methods. *International Journal of Remote Sensing*, 25(10), 1893–1924. <https://doi.org/10.1080/0143116031000101611>
- Trimble. (2011). *Ecognition Developer 8.64. 1 User Guide*. Munich: Trimble.
- Vain, A., Kaasalainen, S., Hyypä, J., & Ahokas, E. (2009). Calibration of laser scanning intensity data using brightness targets. The method developed by the Finnish Geodetic Institute. *Geodesy and Cartography*, 35(3), 77–81. <https://doi.org/10.3846/1392-1541.2009.35.77-81>
- Van Ewijk, K. Y., Randin, C. F., Treitz, P. M., & Scott, N. A. (2014). Predicting fine-scale tree species abundance patterns using biotic variables derived from LiDAR and high spatial resolution imagery. *Remote Sensing of Environment*, 150, 120–131. <https://doi.org/10.1016/j.rse.2014.04.026>
- Vaughn, N. R., Moskal, L. M., & Turnblom, E. C. (2012). Tree species detection accuracies using discrete point lidar and airborne waveform lidar. *Remote Sensing*, 4(2), 377–403. <https://doi.org/10.3390/rs4020377>
- Vauhkonen, J., Ørka, H. O., Holmgren, J., Dalponte, M., Heinzl, J., & Koch, B. (2014). Tree Species



- Recognition Based on Airborne Laser Scanning and Complementary Data Sources (pp. 135–156). [https://doi.org/10.1007/978-94-017-8663-8\\_7](https://doi.org/10.1007/978-94-017-8663-8_7)
- Verlic, A., Duric, N., Kokalj, Z., Marsetic, A., Simoncic, P., & Ostir, K. (2014). TREE SPECIES CLASSIFICATION USING WORLDVIEW-2 SATELLITE IMAGES AND LASER SCANNING DATA IN A NATURAL URBAN FOREST. *SUMARSKI LIST*, 138(9–10), 477–488.
- Voss, M., & Sugumaran, R. (2008). Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. *Sensors*, 8(5), 3020–3036. <https://doi.org/10.3390/s8053020>
- Wang, T., Zhang, H., Lin, H., & Fang, C. (2016). Textural-spectral feature-based species classification of mangroves in Mai Po nature reserve from worldview-3 imagery. *Remote Sensing*, 8(1). <https://doi.org/10.3390/rs8010024>
- Warner, T. (2011). Kernel-Based Texture in Remote Sensing Image Classification. *Geography Compass*, 5(10), 781–798. <https://doi.org/10.1111/j.1749-8198.2011.00451.x>
- Warner, T. A., Lee, J. Y., & McGraw, J. B. (1999). Delineation and identification of individual trees in the eastern deciduous forest. In *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry* (pp. 81–91).
- Waser, L. T., Küchler, M., Jütte, K., & Stampfer, T. (2014). Evaluating the potential of worldview-2 data to classify tree species and different levels of ash mortality. *Remote Sensing*, 6(5), 4515–4545. <https://doi.org/10.3390/rs6054515>
- Wolter, P. T., Mladenoff, D. J., Host, G. E., & Crow, T. R. (1995). Improved forest classification in the northern Lake States using multi-temporal landsat imagery. *Photogrammetric Engineering and Remote Sensing*, 61(9), 1129–1143.
- Wolter, P. T., & Townsend, P. A. (2011). Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota. *Remote Sensing of Environment*, 115(2), 671–691. <https://doi.org/10.1016/j.rse.2010.10.010>
- Wong, N. H., & Yu, C. (2005). Study of green areas and urban heat island in a tropical city. *Habitat International*, 29(3), 547–558. <https://doi.org/10.1016/j.habitatint.2004.04.008>
- Xiao, Q., & McPherson, E. G. (2005). Tree health mapping with multispectral remote sensing data at UC Davis, California. *Urban Ecosystems*, 8(3–4), 349–361. <https://doi.org/10.1007/s11252-005-4867-7>
- Xiao, Q., Ustin, S. L., & McPherson, E. G. (2004). Using AVIRIS data and multiple-masking techniques to map urban forest tree species. *International Journal of Remote Sensing*, 25(24), 5637–5654. <https://doi.org/10.1080/01431160412331291224>
- Yan, J., Zhou, W., Han, L., & Qian, Y. (2018). Mapping vegetation functional types in urban areas with WorldView-2 imagery: Integrating object-based classification with phenology. *Urban Forestry and Urban Greening*, 31, 230–240. <https://doi.org/10.1016/j.ufug.2018.01.021>
- Zhang, K., & Hu, B. (2012). Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. *Remote Sensing*, 4(6), 1741–1757. <https://doi.org/10.3390/rs4061741>