Application of Landsat time series to the spatially explicit estimation of the forest disturbance history and growing stock volume in Georgia, United States

by

Shingo Obata

(Under the direction of Pete Bettinger and Chris Cieszewski)

Abstract

This dissertation is composed of three studies that evaluated the use of Landsat time series data for the detection of forest disturbance and estimation of the growing stock volume of the forest stands in the state of Georgia, United States. We explored the use of Landsat time series dataset to estimate the disturbance year of the forest and estimated the growing stock volume of forest stands using both Landsat time series and the disturbance year data created using the Landsat time series.

The first study is the preliminary research that is dedicated to creating a model that detects the year of forest disturbance. Landsat time series data and an algorithm were used to estimate the forest disturbance. We selected 7 counties along coastal Georgia as our study area. The main objective of this work was to address the current age-class forest structure in the study area where the pine plantation forest is intensively managed. As an outcome of the study, a disturbance detection map was created. We performed an accuracy assessment for the disturbance map and acquired 52% of overall accuracy, which is acceptable precision.

The second study enlarged preliminary research performed in the first study. The two primary goals of this study are (i) to establish and examine a reliable framework for Georgia's

state-wide monitoring of forest disturbances; (ii) to consider and discuss the use and effect of information on forest disturbance maps. From the first study, we changed our area of interest to the entire state and used all available Landsat time series datasets. The overall accuracy of the data for the year of disturbance was more than 8%. This result is 35 percent higher than the overall accuracy of the first study.

In the third study, We have established a model of random forest regression that estimates the increasing volume of forest growing stock volume. The goal of this work was to to test whether (i) using all available Landsat imagery and (ii) application of the bias correction approach would improve the accuracy of the estimation of the forest stand growing stock volume. We used the Forest Inventory and Analysis dataset that is maintained by the US Forest Service as a field plot data. We used Random Forest as an estimation method. We obtained 65% of relative RMSE in the best model.

These studies demonstrate the importance in the application of Landsat time series dataset to spatially explicit forest inventory. The outcome created in these studies is expected to provide fundamental information of the forest resources in Georgia to allow better decisions for regional scale forest management and conservation.

INDEX WORDS: Landsat time series, forest inventory, remote sensing, random forest, forest disturbance detection, growing stock volume

Application of Landsat time series to the spatially explicit estimation of the forest disturbance history and growing stock volume in Georgia, United States

by

Shingo Obata

M.Sc., the University of Tokyo, 2016

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2020

Application of Landsat time series to the spatially explicit estimation of the forest disturbance history and growing stock volume in Georgia, United States

by

Shingo Obata

Major Professors: Pete Bettinger & Chris Cieszewski

Jacek Siry
Marguerite Madden
Roger Charles Lowe III

Electronic Version Approved:

Ron Walcott
Interim Dean of thse Graduate School
The University of Georgia
May 2020

Acknowledgments

First, I would like to express my gratitude to my co-supervisor Dr. Pete Bettinger and Dr. Chris Cieszewski to give me the chance to obtain my Ph.D. and to give me precious support in my research and in my professional growth. Dr. Bettinger inspired me a lot with his incredible productivity and passion for the research. Your word, “Science is hard”, has been my motto through my studies. Dr. Cieszewski gave me the foundation as a scientist through his scientific attitude toward research and thorough guidance. From conversations with you, I obtained so many research ideas.

I would like to thank my committee members, Dr. Jacek Siry, Dr. Margueritte Madden, and Dr. Roger Charles Lowe III not only for their helpful comments on my research but also for their advice which extends my future research perspective. In addition, I would like to express my special thanks of gratitude to Dr. Sergio Bernardes, who provided me the advanced knowledge and technique of geospatial and remote sensing data analysis. This thesis would not have been possible unless you gave me precious advice. My work was greatly supported by my research team members, Dr. Zennure Ucar, Ramazan Akbulut, Dr. Siyu Zhang. Without the academic insight and encouraging conversations you gave me in Room 4-321, my dissertation won’t be accomplished.

I would also like to extend my gratitude to all members of Warnell School of Forestry and Natural Resources, the University of Georgia (UGA). I deeply appreciate the extraordinary learning and research experience provided by the graduate school.

My doctoral studies would not have been possible without the financial support of from

Japan Student Services Organization through the Student Exchange Support Program. Generous and continuous support I hope that one day I can work for my country to reward the help I have taken throughout my studies.

Nobody has been more important to me in the pursuit of my Ph.D. than the members of my family. I wish to thank my parents who always encourage me when I am discouraged by difficulty and failure. Lastly, but not least, I owe my deepest gratitude to my older brother for caring about my life. Although I am far away from our home, I always feel your love and support. We are always together.

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

Both the density and volume of forests are key quantitative metrics used to guide forest management practices. Timber volume is directly linked to economic value, and, in many cases, related to environmental assessments. In addition, timber volume is a primary environmental monitoring metric for biogeochemical processes and biodiversity loss across the globe (Nguyen et al., 2020). Data collection through ground measurement is the most accurate approach in the process of estimating biomass within forests. Nevertheless, because of their lack of spatial and temporal coverage, especially for broad regions or remote areas, ground measurements are unsuitable for measuring biomass over large areas and across short time frames (Soenen et al., 2010; Wulder et al., 2004). Airborne LiDAR Scanning (ALS) have also been employed to map forest growing stock volumes. However, due to the relatively high cost and intensive effort of operating ALS, the scope of application is often restricted in both temporal and spatial metrics. Accordingly, satellite imagery has remained the most popular data source for large-area surveys of forest volume. A national forest inventory (NFI) can be created by combining satellite imagery and field plot data. The first NFI created from the integration of satellite imagery and field plot data was done for Finland (Tomppo, 1990). Since then, methods for creating an NFI have been expanded upon for nearly three decades. Two trends have emerged in the recent literature concerning the estimation of the volume

for large areas: Firstly, increasingly more satellite imagery is used, stimulated by the free availability of major satellite-based earth observation datasets, such as the Landsat and Sentinel missions; Secondly, various land-cover and land cover change datasets developed in the last ten years have become used as a part of the input data of the volume estimation model. It is noteworthy that the latter trend is also stimulated by the former, as the free availability of satellite imagery has opened the possibility for creating new land cover databases. The series of the research described in greater detail below attempts to create a forest inventory that incorporates the two aforementioned trends for the state of Georgia, United States. This chapter, therefore, reviews the literature concerning the two recent trends of the volume estimation method. The literature review also addresses methods for estimating forest growing stock volume across large areas. Finally, the format of this dissertation is described.

1.1 Literature review of Landsat missions

The series of Landsat 1–8 satellites has been one of the pinnacles of satellite remote sensing programs, primarily because the missions have a long history of operation with high and reliable radiometric, spatial, temporal and spectral resolutions relative to other programs. The Landsat program has collected earth observation data for more than 40 years since the launch of Landsat-1 in 1972, and over that time it has become increasingly complex and elaborate (Roy et al., 2014; Schott et al., 2016; Wulder et al., 2019; Wulder et al., 2016; Zhu, 2019). For each Landsat mission, the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the United States Geological Survey (USGS) worked collaboratively (Goward et al., 2017). The sensors mounted on Landsat missions increased gradually over time, while maintaining continuity in the wavelength ranges. Landsat-1 Multispectral Scanner System captured four spectral band reflectances that are located in the visible and near-infrared ranges of energy. In the Landsat 4- and Landsat-5 Thematic Mapper, the number of bands increased to seven and

an 8-bit quantization method was adapted. The Landsat-8 Operational Land Imager further added enhanced blue and cirrus channels to make databases more useful for aquatic sciences.

In the field of digital image processing of satellite image data, preprocessing refers to the operations which are conducted prior to the principal analysis. Preprocessing is mainly comprised of (i) geometric preprocessing to align the location of a dataset with other datasets, and (ii) radiometric preprocessing to remove the effect of unwanted atmospheric features, such as found in a "hazy" atmosphere. Preprocessing of the remote sensing data necessitates the expenditure of time and effort on the part of remote sensing analysts, as it requires tailored work heavily dependent on the type of the project and the dataset (Campbell et al., 2011). The work performed by the USGS has greatly reduced the exigencies of the preprocessing tasks (U.S. Geological Survey, 2018). Geometric alignment is crucial in the process of combining one imagery with other spatial datasets. Historically, geometric alignment has been facilitated using the Global Land Survey ground control point chip sets (Gutman et al., 2013). Currently, the Landsat chip set is configured to be aligned to the Sentinel-2 Global Reference Image (Dechoz et al., 2015). The USGS completed the reprocessing of the satellite image collection to ensure the compatibility between sensors in May 2017 (Dwyer et al., 2018). Collection 1 data was stratified into Tier 1 and Tier 2 datasets based on the size of the geometric error. Tier 1 data is the data with a geometric error of fewer than twelve meters Root Mean Square Error (RMSE). Tier 2 data, conversely, has a geometric error greater than twelve meters RMSE (U.S. Geological Survey, 2020). As a result of the concerted effort for a standardized geometric alignment, current users of Landsat datasets are not required to manually perform geometric alignment as long as they accept the estimated size of the errors. Radiometric preprocessing affects the brightness values of a remotely sensed image to correct errors caused by the sensor malfunction, or in order to tune the values in an effort to compensate for atmospheric degradation (Campbell et al., 2011). The accuracy of radiometric calibration of Landsat missions has been continuously improved from around 10% in Landsat-1 to approximately 3% in Landsat-8 (Markham et al., 2014; Morfitt et al.,

2015). The majority of analyses that use time series analysis, multi-temporal composites, or time series analyses of aquatic ecosystem (Gordon, 1997; Hermosilla et al., 2015; Pahlevan et al., 2018; Zhu et al., 2014) employ a surface reflectance product that goes through the atmospheric correction at the ground level. Surface reflectance products from the Landsat program are provided using two atmospheric correction algorithms: namely the LaSRC for Landsat-8 (Vermote et al., 2016) and LEDAPS for Landsat-4,5 and 7 (Masek et al., 2006).

Taking advantage of the long temporal data span of Landsat sensors, tracking changes in the Earth's surface over time has become a leading application. Since all of the Landsat datasets became freely available in 2008 (Woodcock et al., 2008), drastic changes in time series analyses with Landsat datasets have been made. Historically, a limited number of Landsat images, usually one per year, were acquired for analysis, as it cost $600 per scene in 2008 (Wulder et al., 2012). After the initiation of the open access policy of Landsat datasets, the number of images utilized for analysis of land cover changes drastically increased (Zhu, 2017). This evolution is supported by the aforementioned advances in automated atmospheric correction algorithms and cloud detection algorithms. In addition, a cloud-based geo computation platform (i.e. Google Earth Engine, Gorelick et al. (2017)) now makes it possible to handle the large Landsat datasets without concern for the computational ability of an analyst's computer.

1.2 Literature review of land cover change

Land cover refers to the density of visible features covering the surface of the land, usually from an elevated perspective. The information about the scope of land coverage generated with remotely sensed data is essential for scientific and administrative purposes (Campbell et al., 2011). Historically, land cover data has composed the basic dataset for climatic and hydrologic modeling (Carlson et al., 2000). From an administrative perspective, land cover data informs local and national level environmental, economic, and social decision-

making processes. In the U.S., periodical National Land Cover Datasets have provided the land cover data of contiguous states since 1992 (Yang et al., 2018). A land cover change analysis is part of a process that examines the temporal trajectory of land coverage in an area of interest. Land cover analysis and land cover change analysis have taken different research paths in the past (Townshend, 1992). However, Wulder et al. (2018) argue that land cover change analysis has become an indispensable part of land cover analyses, which they define as Land Cover 2.0. Land Cover 2.0 requires seven elements to complete an analysis: (1) clearly defined information needs, (2) a proper and feasible definition of the land cover legend, (3) the procurement of independent calibration and validation data, (4) an appropriate image processing process, (5) the application of a legitimate image classification algorithm, (6) the post-classification of the classified data to reduce the size of errors, and (7) a product delivery system (Wulder et al., 2018). Land cover change analyses require the same process, but a change detection algorithm is applied in the fifth step instead of a classification algorithm. Zhu (2017) classifies the existing change detection algorithms as being in type either thresholding, differencing, segmentation, trajectory classification, statistical boundary, or regression. Among them, thresholding, segmentation, and statistical boundary algorithms are three major types of change detection algorithms. The thresholding method utilizes a predefined threshold of a computed metric to determine the land cover. Changes are detected when the value of the metric of interest deviates from the threshold. Most research uses one of the vegetation indices that are sensitive to the change of the land cover type. For example, the normalized z-score based Tasseled Cap Transformation (TCT) (Hilker et al., 2009; Kayastha et al., 2012), or Integrated Forest Z-score (C. Huang et al., 2010; C. Huang et al., 2009c; Thomas et al., 2011) have all been used as the target vegetation indices. The segmentation method cuts the time series Landsat imagery into straight line segments. This method employs the Normalized Burn Ratio (NBR) (Hermosilla et al., 2015; Kennedy et al., 2015) or Tasseled Cap Wetness (Frazier et al., 2015; Grogan et al., 2015). The statistical boundary method assumes that the time series is confined to the statistical

boundary and regards values outside the boundary as changes from the previous land cover. The statistical quality control charts (Brooks et al., 2014), econometrics structural change monitoring (DeVries et al., 2015; Hamunyela et al., 2016), and model prediction (Vogelmann et al., 2016; Zhu et al., 2016; Zhu et al., 2014) are applied using this method.

1.3 Literature review of forest growing stock volume estimation

To create a wall-to-wall nationwide forest inventory dataset, multi-source methods (Tomppo et al., 2008) have been developed. A multi-source method refers to the coupling of spatially referenced forest inventory plot data with remote sensing imagery and auxiliary variables to aid in the mapping of forest attributes (Finley et al., 2008). As an imputation method, both parametric models (Lu, 2005; Lu et al., 2012) and non-parametric models, such as the k nearest neighbors algorithm (Eskelson et al., 2009; Franco-Lopez et al., 2001; Lowe et al., 2014; McRoberts et al., 2002; Tomppo et al., 2004), as well as the random forest algorithm (Wilson et al., 2012), have been developed. National forest inventory (NFI) data including the information on wood volume, age, and height, has been used for the various study areas comprises the field plot data of multi-source methods (Chirici et al., 2016; Reese et al., 2005). Particularly in the U.S., the Forest Inventory and Analysis program (FIA, Bechtold et al. (2015)) which is managed by U.S. Forest Service, has provided field plot data. As remotely sensed data, Landsat imagery has been the most popular dataset utilized. Recently, Sentinel data and airborne LiDAR data have been employed to acquire superior results. In terms of the imputation method, non-parametric methods have been shown to outperform parametric methods (Nguyen et al., 2020). The first operative results were developed in 1990 for Finland. The driving force behind the implementation of the multi-source methods was the need to obtain information on forest resources, at a low cost, for broad landscapes. In addition, new satellite images of natural resources provided opportunities to increase inventory productivity

with relatively small additional costs (Nguyen et al., 2020).

1.4 Dissertation format

This dissertation is written in the manuscript format and it presents the results of three research studies. Chapter 1 (this chapter) provides an introduction to this dissertation and a short outline of previous research on land cover change, Landsat time series, and forest growing stock volume estimation. In Chapter 2, "Preliminary Analysis of Forest Stand Disturbances in Coastal Georgia (USA) Using Landsat Time Series Stacked Imagery", we describe an algorithm that attempts to assign an age to forested areas for seven counties along coastal Georgia, United States. The result was evaluated using established procedures of accuracy assessment of land cover data. As a remaining task, additional modifications to the method were needed to adapt to the specific study area conditions. Chapter 3, "Mapping forest disturbances between 1987-2016 using all available time series Landsat TM/ETM imagery: Developing a reliable methodology for Georgia, United States", expands upon the preliminary research conducted in Chapter 2. The two main objectives of this research are (i) to develop and test a robust methodology for the state-wide monitoring of forest disturbances in Georgia; and (ii) to consider and explore the use and impact of forest disturbance map information. Two main methods of disturbance detection, a threshold algorithm, and a statistical boundary algorithm, were integrated to establish a reliable estimation of the recent history of forest disturbances. The developed model was employed by using all available Landsat TM/ETM data to create a forest disturbance record for the years 1984-2016. We also studied the regional differences in the history of disturbances. This analysis suggests that disturbed urban forests were more likely to be converted to other land-uses. In Chapter 4, "Estimation of forest growing stock stand volumes based on Harmonic regression with ensemble method and 1984-2016 Landsat 5 and 7 imagery", we developed an RF regression model that estimates the growing stock volume in the state of Georgia, United States. This

algorithm used all obtainable Landsat TM/ETM imagery from 1984 to 2016, and then a bias correction method was applied to the primary estimate. In Chapter 5 a conclusion to the dissertation is presented in order to highlight the contributions to science thus far, and to discuss the possible future paths of research. Finally in Chapter 6, all the literature cited in this dissertation is listed.

1.5 References

Bechtold, W. A. and P. L. Patterson (2015). *The enhanced Forest Inventory and Analysis program: National sampling design and estimation procedures*. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. DOI: 10.2737/SRS-GTR-80.

Brooks, E. B., R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston (2014). "On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6, pp. 3316–3332. DOI: 10.1109/TGRS.2013.2272545.

Campbell, J. B. and R. H. Wynne (2011). *Introduction to remote sensing*. 5th edition. New York: The Guilford Press. 667 pp.

Carlson, T. N. and S. T. Arthur (2000). "The impact of land use — land cover changes due to urbanization on surface microclimate and hydrology: A satellite perspective". In: *Global and Planetary Change* 25.1, pp. 49–65. DOI: 10.1016/S0921-8181(00)00021-7.

Chirici, G., M. Mura, D. McInerney, N. Py, E. O. Tomppo, L. T. Waser, D. Travaglini, and R. E. McRoberts (2016). "A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data". In: *Remote Sensing of Environment* 176, pp. 282–294. DOI: 10.1016/j.rse.2016.02.001.

Dechoz, C., V. Poulain, S. Massera, F. Languille, D. Greslou, F. d. Lussy, A. Gaudel, C. L'Helguen, C. Picard, and T. Trémas (2015). "Sentinel 2 global reference image". In:

Image and Signal Processing for Remote Sensing XXI. Image and Signal Processing for Remote Sensing XXI. Vol. 9643. International Society for Optics and Photonics, 96430A. DOI: 10.1117/12.2195046.

DeVries, B., M. Decuyper, J. Verbesselt, A. Zeileis, M. Herold, and S. Joseph (2015). "Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series". In: *Remote Sensing of Environment* 169, pp. 320–334. DOI: 10.1016/j.rse.2015.08.020.

Dwyer, J. L., D. P. Roy, B. Sauer, C. B. Jenkerson, H. K. Zhang, and L. Lymburner (2018). "Analysis ready data: Enabling analysis of the Landsat archive". In: *Remote Sensing* 10.9, p. 1363. DOI: 10.3390/rs10091363.

Eskelson, B. N. I., H. Temesgen, V. Lemay, T. M. Barrett, N. L. Crookston, and A. T. Hudak (2009). "The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases". In: *Scandinavian Journal of Forest Research* 24.3, pp. 235–246. DOI: 10.1080/02827580902870490.

Finley, A. O. and R. E. McRoberts (2008). "Efficient k-nearest neighbor searches for multi-source forest attribute mapping". In: *Remote Sensing of Environment* 112.5, pp. 2203–2211. DOI: 10.1016/j.rse.2007.08.024.

Franco-Lopez, H., A. R. Ek, and M. E. Bauer (2001). "Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method". In: *Remote Sensing of Environment* 77.3, pp. 251–274. DOI: 10.1016/S0034-4257(01)00209-7.

Frazier, R. J., N. C. Coops, and M. A. Wulder (2015). "Boreal Shield forest disturbance and recovery trends using Landsat time series". In: *Remote Sensing of Environment* 170, pp. 317–327. DOI: 10.1016/j.rse.2015.09.015.

Gordon, H. R. (1997). "Atmospheric correction of ocean color imagery in the Earth observing system era". In: *Journal of Geophysical Research: Atmospheres* 102 (D14), pp. 17081–17106. DOI: 10.1029/96JD02443.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". In: *Remote Sensing of Environment* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Goward, S. N., T. Arvidson, D. Williams, L. Rocchio, J. R. Irons, S. S. Johnston, and C. Russel (2017). *Landsat's enduring legacy: Pioneering global land observations from space*. Bethesda, MD: American Society for Photogrammetry and Remote Sensing. 586 pp.

Grogan, K., D. Pflugmacher, P. Hostert, R. Kennedy, and R. Fensholt (2015). "Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series". In: *Remote Sensing of Environment* 169, pp. 438–453. DOI: 10.1016/j.rse.2015.03.001.

Gutman, G., C. Huang, G. Chander, P. Noojipady, and J. G. Masek (2013). "Assessment of the NASA–USGS Global Land Survey (GLS) datasets". In: *Remote Sensing of Environment* 134, pp. 249–265. DOI: 10.1016/j.rse.2013.02.026.

Hamunyela, E., J. Verbesselt, and M. Herold (2016). "Using spatial context to improve early detection of deforestation from Landsat time series". In: *Remote Sensing of Environment* 172, pp. 126–138. DOI: 10.1016/j.rse.2015.11.006.

Hermosilla, T., M. A. Wulder, J. C. White, N. C. Coops, and G. W. Hobart (2015). "Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics". In: *Remote Sensing of Environment* 170, pp. 121–132. DOI: 10.1016/j.rse.2015.09.004.

Hilker, T., M. A. Wulder, N. C. Coops, J. Linke, G. McDermid, J. G. Masek, F. Gao, and J. C. White (2009). "A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS". In: *Remote Sensing of Environment* 113.8, pp. 1613–1627. DOI: 10.1016/j.rse.2009.03.007.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann (2010). "An automated approach for reconstructing recent forest disturbance history using dense

Landsat time series stacks". In: *Remote Sensing of Environment* 114.1, pp. 183–198. DOI: 10.1016/j.rse.2009.08.017.

Huang, C., A. Li, H. Shi, G. Sun, Z. Zhu, S. N. Goward, and J. G. Masek (2009c). "Developing a fine resolution forest height map for Mississippi using Landsat time series observations and GLAS Lidar data". In: *AGU Fall Meeting Abstracts*.

Kayastha, N., V. Thomas, J. Galbraith, and A. Banskota (2012). "Monitoring wetland change using inter-annual Landsat time-series data". In: *Wetlands* 32.6, pp. 1149–1162. DOI: 10.1007/s13157-012-0345-1.

Kennedy, R. E., Z. Yang, J. Braaten, C. Copass, N. Antonova, C. Jordan, and P. Nelson (2015). "Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA". In: *Remote Sensing of Environment* 166, pp. 271–285. DOI: 10.1016/j.rse.2015.05.005.

Lowe, R. C. and C. J. Cieszewski (2014). "Multi-source k-nearest neighbor, Mean balanced forest inventory of Georgia". In: *Mathematical and Computational Forestry & Natural-Resource Sciences* 6.2, pp. 65–79.

Lu, D. (2005). "Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon". In: *International Journal of Remote Sensing* 26.12, pp. 2509–2525. DOI: 10.1080/01431160500142145.

Lu, D., Q. Chen, G. Wang, E. Moran, M. Batistella, M. Zhang, G. Vaglio Laurin, and D. Saah (2012). "Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates". In: *International Journal of Forestry Research* 2012, pp. 1–16. DOI: 10.1155/2012/436537.

Markham, B., J. Barsi, G. Kvaran, L. Ong, E. Kaita, S. Biggar, J. Czapla-Myers, N. Mishra, and D. Helder (2014). "Landsat-8 operational land imager radiometric calibration and stability". In: *Remote Sensing* 6.12, pp. 12275–12308. DOI: 10.3390/rs61212275.

Masek, J. G., E. Vermonte, N. Saleous, R. Wolfe, F. Hall, F. Huemmrich, F. Gao, J. Kulter, and T. Lim (2006). "A Landsat surface reflectance data set for North America, 1990-2000". In: *Geoscience and Remote Sensing Letters* 3, pp. 68–72.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt (2002). "Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique". In: *Remote Sensing of Environment* 82.2, pp. 457–468. DOI: 10.1016/S0034-4257(02)00064-0.

Morfitt, R., J. Barsi, R. Levy, B. Markham, E. Micijevic, L. Ong, P. Scaramuzza, and K. Vanderwerff (2015). "Landsat-8 Operational Land Imager (OLI) radiometric performance on-orbit". In: *Remote Sensing* 7.2, pp. 2208–2237. DOI: 10.3390/rs70202208.

Nguyen, T. H., S. D. Jones, M. Soto-Berelov, A. Haywood, and S. Hislop (2020). "Monitoring aboveground forest biomass dynamics over three decades using Landsat time-series and single-date inventory data". In: *International Journal of Applied Earth Observation and Geoinformation* 84, p. 101952. DOI: 10.1016/j.jag.2019.101952.

Pahlevan, N., S. V. Balasubramanian, S. Sarkar, and B. A. Franz (2018). "Toward long-term aquatic science products from heritage Landsat missions". In: *Remote Sensing* 10.9, p. 1337. DOI: 10.3390/rs10091337.

Reese, H., T. Granqvist-Pahlén, M. Egberth, M. Nilsson, and H. Olsson (2005). "Automated estimation of forest parameters for Sweden using Landsat data and the kNN algorithm". In: *Proceedings of the 31st International Symposium on Remote Sensing*. St. Petersburg, Russia, p. 4.

Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.e., R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R. H. Wynne, and Z. Zhu

(2014). “Landsat-8: Science and product vision for terrestrial global change research”. In: *Remote Sensing of Environment* 145, pp. 154–172. DOI: 10.1016/j.rse.2014.02.001.

Schott, J. R., A. Gerace, C. E. Woodcock, S. Wang, Z. Zhu, R. H. Wynne, and C. E. Blinn (2016). “The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments”. In: *Remote Sensing of Environment* 185, pp. 37–45. DOI: 10.1016/j.rse.2016.04.015.

Soenen, S. A., D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall (2010). “Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain”. In: *Remote Sensing of Environment* 114.7, pp. 1325–1337. DOI: 10.1016/j.rse.2009.12.012.

Thomas, N. E., C. Huang, S. N. Goward, S. Powell, K. Rishmawi, K. Schleeweis, and A. Hinds (2011). “Validation of North American forest disturbance dynamics derived from Landsat time series stacks”. In: *Remote Sensing of Environment* 115.1, pp. 19–32. DOI: 10.1016/j.rse.2010.07.009.

Tomppo, E. (1990). “Designing a satellite image-aided national forest survey in Finland”. In: Usability of remote sensing for forest inventory and planning. Umea, Sweden: Swedish University of Agricultural Sciences, pp. 43–47.

Tomppo, E., M. Haakana, M. Katila, and J. Peräsaari (2008). *Multi-source national forest inventory: Methods and applications*. Springer Netherlands. 373 pp. DOI: 10.1007/978-1-4020-8713-4.

Tomppo, E. and M. Halme (2004). “Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach”. In: *Remote Sensing of Environment* 92.1, pp. 1–20. DOI: 10.1016/j.rse.2004.04.003.

Townshend, J. R. G. (1992). “Land cover”. In: *International Journal of Remote Sensing* 13.6, pp. 1319–1328. DOI: 10.1080/01431169208904193.

U.S. Geological Survey (2018). *Landsat collections*. URL: http://pubs.er.usgs.gov/publication/fs20183049 (visited on 03/02/2020).

U.S. Geological Survey (2020). *Landsat levels of processing*. URL: https://www.usgs.gov/land-resources/nli/landsat/landsat-levels-processing (visited on 02/09/2020).

Vermote, E., C. Justice, M. Claverie, and B. Franch (2016). "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product". In: *Remote Sensing of Environment* 185, pp. 46–56. DOI: 10.1016/j.rse.2016.04.008.

Vogelmann, J. E., A. L. Gallant, H. Shi, and Z. Zhu (2016). "Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data". In: *Remote Sensing of Environment* 185, pp. 258–270. DOI: 10.1016/j.rse.2016.02.060.

Wilson, B. T., A. J. Lister, and R. I. Riemann (2012). "A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data". In: *Forest Ecology and Management* 271, pp. 182–198. DOI: 10.1016/j.foreco.2012.02.002.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne (2008). "Free access to Landsat imagery". In: *Science* 320.5879, pp. 1011–1011. DOI: 10.1126/science.320.5879.1011a.

Wulder, M. A., R. S. Skakun, W. A. Kurz, and J. C. White (2004). "Estimating time since forest harvest using segmented Landsat ETM+ imagery". In: *Remote Sensing of Environment* 93.1, pp. 179–187. DOI: 10.1016/j.rse.2004.07.009.

Wulder, M. A., N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla (2018). "Land cover 2.0". In: *International Journal of Remote Sensing* 39.12, pp. 4254–4284. DOI: 10.1080/01431161.2018.1452075.

Wulder, M. A., T. R. Loveland, D. P. Roy, C. J. Crawford, J. G. Masek, C. E. Woodcock, R. G. Allen, M. C. Anderson, A. S. Belward, W. B. Cohen, J. Dwyer, A. Erb, F. Gao, P. Griffiths, D. Helder, T. Hermosilla, J. D. Hipple, P. Hostert, M. J. Hughes, J. Huntington, D. M. Johnson, R. Kennedy, A. Kilic, Z. Li, L. Lymburner, J. McCorkel, N. Pahlevan,

T. A. Scambos, C. Schaaf, J. R. Schott, Y. Sheng, J. Storey, E. Vermote, J. Vogelmann, J. C. White, R. H. Wynne, and Z. Zhu (2019). "Current status of Landsat program, science, and applications". In: *Remote Sensing of Environment* 225, pp. 127–147. DOI: 10.1016/j.rse.2019.02.015.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock (2012). "Opening the archive: How free data has enabled the science and monitoring promise of Landsat". In: *Remote Sensing of Environment* 122, pp. 2–10. DOI: 10.1016/j.rse.2012.01.010.

Wulder, M. A., J. C. White, T. R. Loveland, C. E. Woodcock, A. S. Belward, W. B. Cohen, E. A. Fosnight, J. Shaw, J. G. Masek, and D. P. Roy (2016). "The global Landsat archive: Status, consolidation, and direction". In: *Remote Sensing of Environment* 185, pp. 271–283. DOI: 10.1016/j.rse.2015.11.032.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian (2018). "A new generation of the United States national land cover database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146, pp. 108–123. DOI: 10.1016/j.isprsjprs.2018.09.006.

Zhu, Z. (2019). "Science of Landsat analysis ready data". In: *Remote Sensing* 11.18, p. 2166. DOI: 10.3390/rs11182166.

Zhu, Z., Y. Fu, C. E. Woodcock, P. Olofsson, J. E. Vogelmann, C. Holden, M. Wang, S. Dai, and Y. Yu (2016). "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014)". In: *Remote Sensing of Environment* 185, pp. 243–257. DOI: 10.1016/j.rse.2016.03.036.

Zhu, Z. and C. E. Woodcock (2014). "Continuous change detection and classification of land cover using all available Landsat data". In: *Remote Sensing of Environment* 144, pp. 152–171. DOI: 10.1016/j.rse.2014.01.011.

Zhu, Z. (2017). "Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 130, pp. 370–384. DOI: 10.1016/j.isprsjprs.2017.06.013.

Chapter 2

Preliminary analysis of forest stand disturbances in coastal Georgia (USA) using Landsat time series stacked imagery [1]

[1] Shingo Obata, Chris J. Cieszewski, Pete Bettinger, Roger C. Lowe III, and Sergio Bernardes, Accepted by FORMATH, 2019. Reprinted here with permission of publisher

Abstract

A determination of forest characteristics across broad areas is of great concern to the forest industry in the southern United States, as timber supply decisions can be based on opportunities, or lack of thereof, across all wood procurement areas. This is important in areas such as the southern United States, where the land ownership distribution is highly fragmented and where no common comprehensive source of forest data exists other than the low intensity USDA Forest Service FIA surveys. In an effort to describe forest characteristics along the lower Coastal Plain of the State of Georgia (USA), we utilized a time series of Landsat data and an algorithm that assesses an integrated forest Z score. The methodology was used to create disturbance maps for over 30 years that represent the year of disturbance for specific locations. The overall accuracy was 52% when all years were considered, and approximately 70% from 1991 forward. Preliminary findings showed moderate levels of accuracy when determining ages for current forests, most of which are even-aged nature stands. Further modifications to the process were necessary to adapt to the unique conditions of study region. The modeling process also prompted several areas for future refinement, including improvement of the temporal resolution of the analysis by using all the available Landsat imagery and detection of the regenerations that normally occurs several years after disturbances.

Keywords: forest planning, forest resource management, land cover change analysis, mathematical modeling, remote sensing

2.1 Introduction

Acquisition of information regarding the state of natural resource availability and their characteristics can be one of the most expensive components of forest management (Bettinger et al., 2017). Satellite technology can provide a cost-effective source of spatial information and can facilitate processes that allow one to estimate various natural resource characteristics needed for the planning and management of forests, sustainability analysis (Cieszewski

et al., 2004; S. Liu et al., 2009) and biomass supply assessment (Cieszewski, 2011). Of the many satellites currently in operation, those belonging to the Landsat series are frequently used for terrestrial resource assessments/monitoring due to the long life of the program (first launched in 1972) and the characteristics of the platform, orbit and sensors, including temporal resolution, consistent footprint and temporal resolution, and its moderate spatial resolution (Roy et al., 2014). Along with the growing power and availability of computing and software technology, the evolving body of scientific research is continuously revealing growing potential to estimate more detailed forest information using satellite imagery. This progression in applied science has evolved beyond static analyses of single satellite images to spatiotemporal analyses of series (stacks) of images captured at different points in time. Time series analysis of remotely-sensed data can help identify forest disturbance, which is one critically important process in an assessment of the current state of forest resources across broad landscapes. Dating disturbance events by using remote sensing can facilitate the determination of forest age and of current stage of forest development. In this paper, the term "disturbance" is used to refer to abrupt loss of forest cover in an area. This includes any types of the causes that lead the change. Disturbance is caused by two types of events. The first one is natural disturbance such as forest fire and windthrow. The other is anthropogenic disturbance including major harvesting that clears all the forest stand and minor harvesting that leaves majority of the forest cover.

With temporal trajectory-based image analysis (change detection), changes in land use can be detected with a help of an algorithm that assesses a time series of satellite images to locate changes in the spectral signatures of arbitrary locations within a landscape (Brooks et al., 2014). The Vegetation Change Tracker (VCT), for example, uses an integrated forest score (IFZ) as a metric to distinguish forested areas from others areas at the resolution of an image single pixel (C. Huang et al., 2009a). Application of these types of algorithms enables the creation of a map, which specifies the years of stand disturbances. LandTrendr is another example algorithm for automatic detection of changes in land use (Kennedy et al.,

2010). Other models (Vogelmann et al., 2012; Jin et al., 2013) share common characteristics: stacking annual images captured during a specific season of the year, using one or more indices for detection of changes, and comparing results with adjacent years to verify the changes. The novelty of this kind of research is of great relevance to forest management. For example, Brooks et al. (2014) used statistical process control tools to detect not only major disturbances (final harvests) but also minor disturbances, such as thinnings. And Zhu et al. (2015b) developed an algorithm to remove seasonal effects from the analysis of satellite imagery. Further, the North American Carbon Program created North American Forest Dynamics (NAFD) products which are a spatially explicit disturbance detection maps for conterminous United States between 1986 and 2010. NAFD product is comprised of 25 annual and two-time integrated forest disturbance map that shows the specific year of disturbance with 30 meter spatial resolution (Goward et al., 2016). Concerning the state of Georgia, Turner, 1987 revealed the land use between 1935 and 1982 was quantified for Georgia. Although it does not cover the specific disturbance year, it estimated the net primary production. Georgia GAP Analysis (Natural Resources Spatial Analysis Lab, 2003) has done the adoptation or development of a land cover classification system for the entire state as of 2003, which is similar to National Land Cover Dataset (Yang et al., 2018).

The objective of this research was to describe the current age class distribution of forests in the lower Coastal Plain of the State of Georgia, an area where forests are managed mainly through even-aged methods, and where forests are intermixed with other land uses. In this paper we present the comprehensive explanation of the work that was briefly summarized in Obata (2018). A long (33 year) time series of Landsat imagery was used to determine the date of the most recent major forest disturbance (final harvest). The forested and agricultural settings in this region include a complex interspersion of croplands (e.g., onions), pasture, conifer plantations, deciduous bottomlands, cypress (*Taxodium*) forests, and mixed species stands of trees that are managed by public agencies, private landowners, and the U.S. military. This landscape heterogeneity, along with the variety of land management

objectives that follow, present both challenge and opportunity to those who seek to describe the current state of forests. While previous research has illustrated sophisticated algorithms for estimating disturbance years, most does not utilize Landsat 8 OLI imagery, which started acquiring images in 2013. Landsat 8 OLI shares common characteristics with Landsat 5 TM in most aspects, yet there is a difference in band centers, where wavelengths located at the center of the each band in Landsat 8 are different than those of Landsat 5. This difference may result in inconsistencies of the disturbance/regeneration map derived from the imagery. Therefore, it is a challenge to the current research to develop an up-to-date disturbance map for our Coastal Plain study site by combining multiple years of Landsat 5 TM and Landsat 8 OLI satellite imagery.

2.2 Methods

2.2.1 Study Area

Our study area consists of seven counties along the Atlantic Ocean coastline in the lower Coastal Plain of the State of Georgia, USA (Figure 2.1), representing approximately 900,000 hectares. In this area, loblolly pine (*Pinus taeda*) and slash pine (*Pinus elliottii*) are the dominant coniferous trees species that are planted, and many areas are managed by the forest product industry. We selected these counties for three reasons. The first reason is that this area has been directly influenced by an intensive forest management program that was initiated in the 1980s on lands that were managed by corporations, companies, and investment management organizations. With the financial support of the U.S. Department of Agriculture (USDA) Conservation Reserve Program (CRP), private landowners have also been able to obtain funds for a portion of the forest management costs (Georgia Forestry Commision, 2018), which gave them an incentive to increase the intensity of forest management on their lands, and even shift some of their land uses from agriculture to forestry. In these ways, intensive forest management has been used to reduce the rotation ages of even-

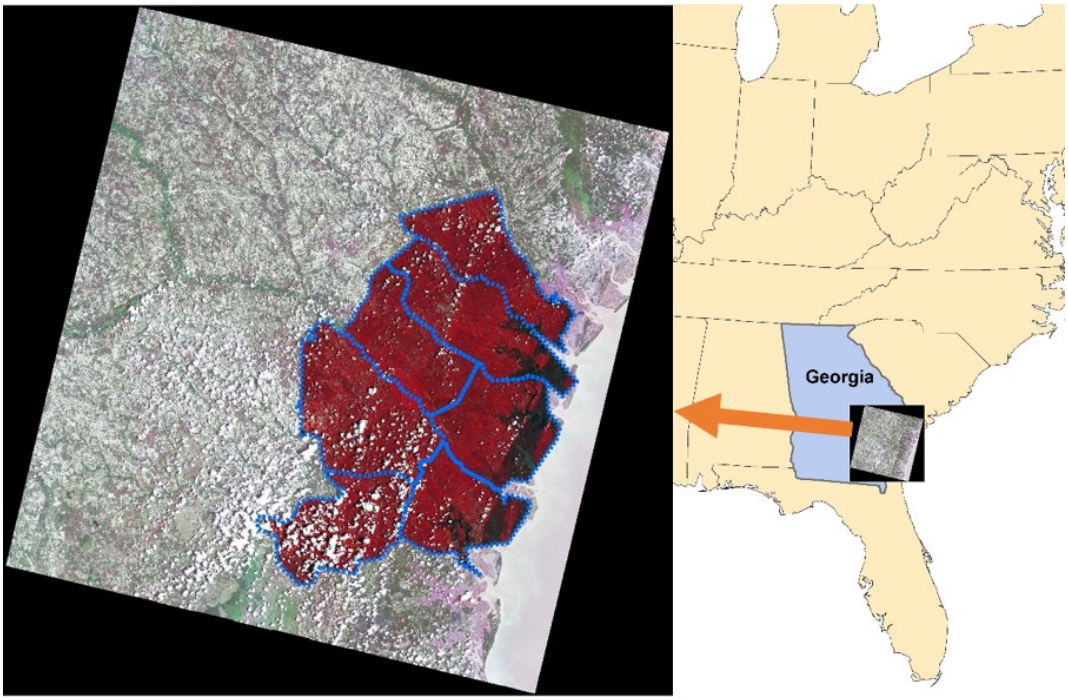


Figure 2.1: Study area.

aged forests in this region. Second, the study area includes mostly rural areas. The largest cities (Savannah, Georgia and Jacksonville, Florida) are on the north and south ends of the study area, respectively. And third, the topography in this area is very consistent, with relatively gentle slopes, compared to the Piedmont or mountainous areas of the southern United States (and further inland in Georgia).

2.2.2 Training Data

We used the Landsat 5 TM and Landsat 8 OLI imagery to create the training data set for the classification process as described below. The creation was subject to two conditions: (i) the imagery needed to be captured during the heart of the growing season, which we assumed to be from the beginning of June to the end of August; and (ii) the imagery needed to be captured between 1984 and 2016. Imagery from 2012 was lacking because neither Landsat 5 nor Landsat 8 was in operation during this year. Landsat 7 data could have been used

for 2012, however, due to the failure of the scan line corrector, Landsat 7 data suffer from significant data gaps (J. Chen et al., 2011).

The images were downloaded via EarthExplorer for each year a single image with the least cloud cover was selected to represent it. We used the C function of the mask (CFMask) band to remove the effect of clouds (presence and shadows). For pixels representing water, we applied criteria to mask a pixel when its normalized difference vegetation index (NDVI) was less than 0.5 and when the surface reflectance value from the near infrared band was less than 0.15. After performing these processes, we visually inspected the quality of the resulting imagery, and decided to exclude the 1993 scene, as the processes failed to correctly eliminate thin, cirrus clouds.

2.2.3 Reference Data

In an effort to verify the accuracy of the created forest age map, we used as reference images from Google Earth, Landsat 5 TM and Landsat 8 OLI imagery. Google Earth was the primary source of reference data for the accuracy assessment for more recent years, as the frequency and quality of high-resolution imagery is greater in the last 15 years than before that. We used Landsat imagery for verification of disturbances for earlier years in our time frame (2000 and before). Annual mosaics of Landsat imagery were created using Google Earth Engine in an effort to acquire cloud-free representations of the landscape. To create annual mosaics, imagery selected to conduct classification were not used so that training data and reference data are not mixed each other. These mosaics were used to verify whether disturbances had occurred during the accuracy assessment process (described below).

2.2.4 Forest Disturbance Detection Method

The algorithm we applied computed the Integrated Forest Z-score (IFZ value) to detect major forest disturbances. The IFZ value is a common index (C. Huang et al., 2010) for this purpose and can be interpreted as the normalized distance between a pixel value of multi-spectral

satellite imagery and the value of a previously identified, reference forest pixel. When a small IFZ value is produced for a given year, this suggests an area of relatively stable mature forest. Large IFZ values, on the other hand, indicate areas where major disturbances have recently occurred. The IFZ value is computed using the mean and standard deviation of identified forest regions within the satellite imagery. An identified forest region is a group of pixels that were manually selected to represent forest areas. Calculation of the IFZ value for each pixel within an image begins with the calculations of stand Z-score (FZ_{jik}) for each pixel j in each band i for each year k.

$$FZ_{jik} = (\frac{b_{ijk} - \bar{b}_{ik}}{SD_{ik}}) \qquad \forall i, j, k \tag{2.1}$$

Where:

b_{ijk} =surface reflectance value of pixel j in band i during year k

$\bar{b}_{ik}$ =the mean surface reflectance of forest regions of training areas in band i during year k

SD_{ik} = the standard deviation of surface reflectance for forest regions of training areas in band i during year k

Each FZ_{jik} is then integrated into a single scalar value for each pixel j during each year k, IFZ_{jk}, using the following formula:

$$IFZ_{jk} = \sqrt{\frac{\sum_{i=1}^{NB}(FZ_{ijk})^2}{NB}} \qquad \forall i, j, k \tag{2.2}$$

Where:

NB = the number of bands used

Three bands (red, shortwave infrared-1, and shortwave infrared-2) were used to determine the IFZ value. The actual band numbers used in this computation were bands 3, 5 and 7 (Landsat 5) and bands 4, 6 and 7 (Landsat 8). In calculating the mean and standard deviation values for forest regions, 15 training areas were selected to represent coniferous

forests and another 15 training areas were selected to represent broadleaf forests. The area of each individual training region was about 100 ha.

2.2.5 Image Classification

The classification algorithm was created to detect the most recent disturbance year of forests in our study area. The algorithm computed the IFZ value for each pixel and each year of the time series. The resulting series of IFZ images was then used to compute the difference in IFZ values between two consecutive images along the series. This allowed a determination of whether a pixel should be categorized as either disturbed or undisturbed. A large difference in IFZ values between consecutive scenes in the series indicated the occurrence of a major disturbance. Non-disturbance related features are represented by minor variations in IFZ values. The chronological order of the application of the algorithm was from the most distant year (1984) to the most current year (2016). For each pixel in year k, when all the following conditions were met, the pixel was updated in year $k+1$ as having been disturbed.

(a) In year k, the IFZ value is lower than x.

(b) The difference in IFZ values between year k and $k+1$ is greater than x.

(c) Through year $k+1$ to $k+y$, the IFZ value is larger than x.

Here, x is set to 3 for two reasons. First, C. Huang et al. (2009b) used $x = 3$ in their analysis. Second, we confirm through the comparison between IFZ value and satellite imagery that pixels of $IFZ < 3$ are forest. Also, y is set to 3 years in the algorithm since IFZ value of the most of the pixels disturbed was greater than 3 for two more years since the year of disturbance. If any of the conditions noted above are not met, the pixel was considered not disturbed in year $k+1$. If these three conditions are met in more than one year in the time horizon (1984 to 2016), the pixel was noted as being last disturbed in the most recent year to 2016. This allowed a determination of the current age of the forest represented by the pixel. As the algorithm requires IFZ value for the year before disturbance, the algorithm

began detecting disturbances in 1985. For years 2015 and 2016, x was assumed to be 2 and 1 years, respectively. Due to the unavailability of the imagery in 1993 and 2012, the algorithm used a slightly different process for years $k-3$ to $k-1$ when $k = 1993$ or 2012. For$k-3$, x was assumed to be to 2 years. For $k-2$, condition (c) was modified:

(c') In year $k+1$ and $k+3$, the IFZ value is larger than 3.

For $k-1$, conditions (b) and (c) were modified:

(b') The difference in IFZ values between year k and $k+2$ is greater than 3.

(c") In year $k+2$ and $k+3$, the IFZ value is larger than 3.

For year k (1993 or 2012), the algorithm does not evaluate each pixel. The algorithm was implemented in the R programming language.

2.2.6 Accuracy Assessment

An accuracy assessment of the classification process was conducted in an attempt to validate the quality of the age class prediction through the forest disturbance maps. One hundred sampling points per year (1984-2016) were randomly selected across the landscape applying a stratified sampling method to generate the points. The total number of sampling points for this purpose was 3,100. For each sampling point, the specific year of the last disturbance was recorded, as evidenced using the reference data noted above. The land use for each point in each year was also visually inspected by using reference data. Two types of classification accuracy were determined, the first using user's and producer's accuracy values reflecting exact (to the year) coincidence of the classified maps and the reference data. The second allows a ±1-year deviation from the reference data, as a ±1-year deviation from the reference data was frequently observed due to cloud coverage and other temporal misalignments of implemented practices and the satellite imagery. Although we examined +/- 2 years deviation from the reference data, we did not compute it as +/- 2 years deviation was not frequently observed in our result.

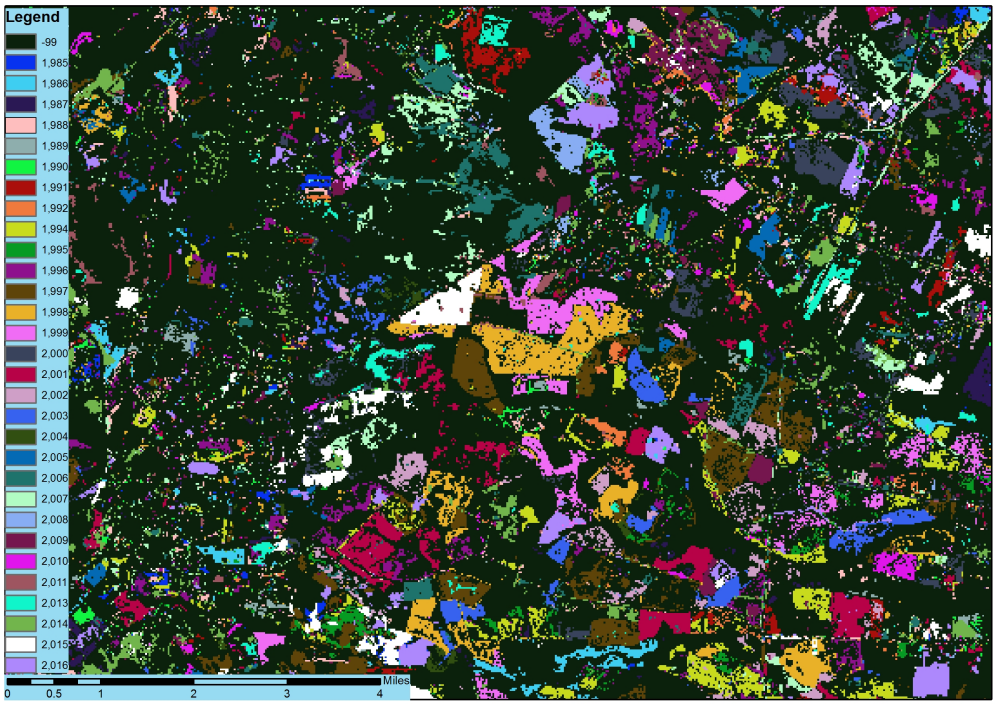
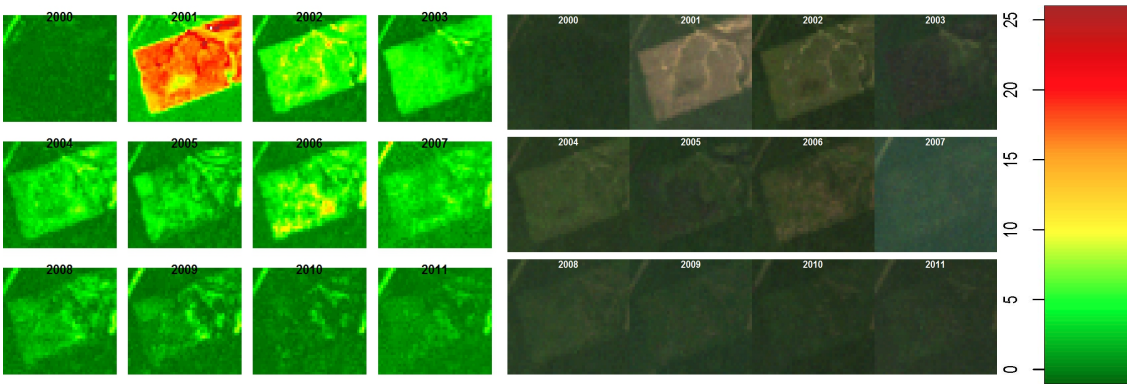


Figure 2.2: Subset of disturbance map. Dates represent the most recent disturbance for the period. The class value of -99 (black color on the map) represents the pixels which are identified by our algorithm as either persistent forest, persistent non-forest between 1984 - 2016.

2.3 Result

The process employed in this study produced a disturbance map (Figure 2.2) that illustrates the years of the most recent major disturbances. Our assumption is that these years correlate directly with the initiations of new stands assuming that forestry will continue to be the designated land use after the occurrence of the major disturbance, such as final harvest and that the landowner or land manager will begin preparation of the new forest soon after the disturbance.

A typical final harvest produces a significant change in the IFZ value, and over time, the IFZ value will return to a range of normal values for these forests. Figure 2.3 illustrates



(a) (b)

Figure 2.3: (a) Time series true color composite of Landsat 5 TM imagery, and (b) time series IFZ value map.

through true color composite (images on the left) and through IFZ value images (to the right) how the landscape would be perceived when a major disturbance occurred in 2001. As time progresses and assuming forestry remains the dominant use of the land area, the imprint of the final harvest gradually diminishes in both true color and IFZ value representations. The change, or difference, that is noted between years 2000 and 2001, if significant enough, would prompt the algorithm to conclude that a major disturbance had occurred. When IFZ values for the land area are plotted over time (Figure 2.4), the change in IFZ value is more dramatic between 2000 and 2001. The minor variation in IFZ value amongst the earlier years is likely a function of local weather conditions and date of acquisition; therefore, the threshold change in IFZ amongst adjacent years (noted earlier) would have prevented the algorithm from concluding a disturbance had occurred in other years when differences among subsequent years were observed.

The result of the accuracy assessment for the 3,100 reference points suggests that the overall accuracy of the classification process was 52% (Table 2.1). If a ± 1 year of error is allowed, the relaxed overall accuracy rises to 71%. Concerning this accuracy assessment, it is worth noting that imagery from Google Earth generally has a higher spatial resolution

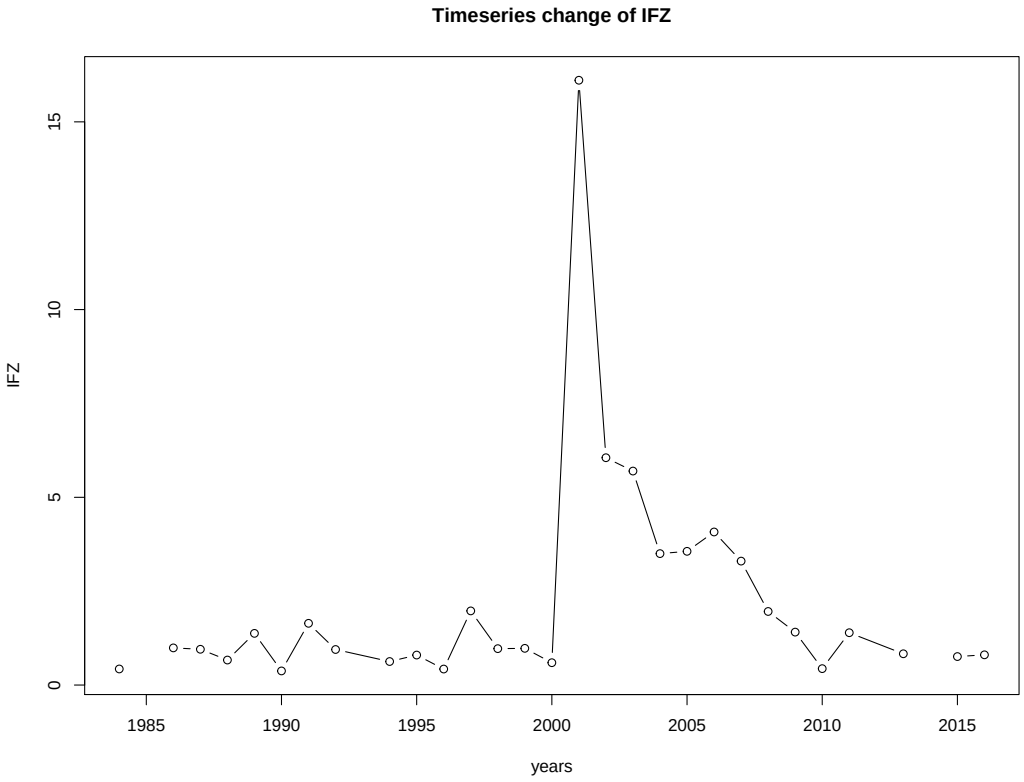


Figure 2.4: Annual IFZ values of a selected pixel on which disturbance occurred in 2001.

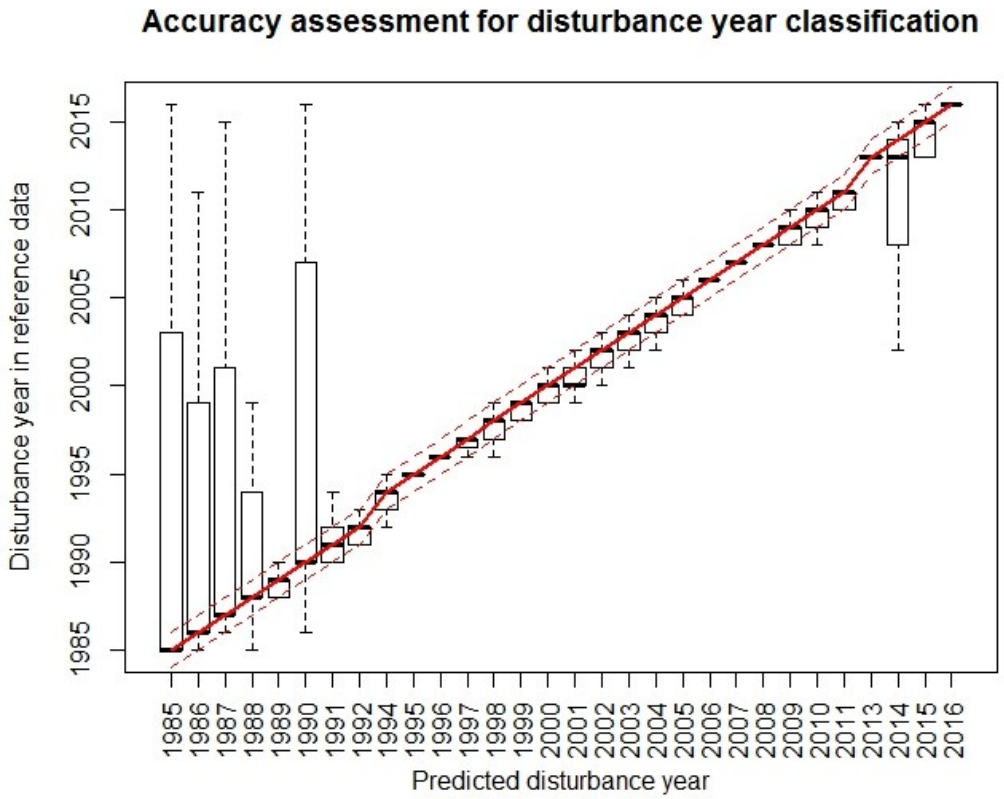


Figure 2.5: Box-and-whisker plot for major disturbances. For each year, the thick horizontal bars represent median, the rectangles are 25 to 75 percentile range and dashed vertical lines cover maximum and minimum values

than Landsat imagery. The historical image tool in Google Earth automatically provides the imagery which has the highest spatial resolution of all the imagery that Google can serve. Imagery captured after 2005 has a higher spatial resolution than that of Landsat 5 TM imagery, which is 27.5 meters. Thus, Google Earth allows one to detect the subtle changes in forest character that may not be detectable with Landsat 5 and 8 imagery.

We created a box-and-whisker plot for major disturbances (Figure 2.5) that illustrates the magnitude of the difference between our predicted year of disturbance and the reference data. This analysis illustrates large differences between subclasses (individual years); the annual user's accuracy ranges between 40% and 90%. User's accuracy in the early 1980s is

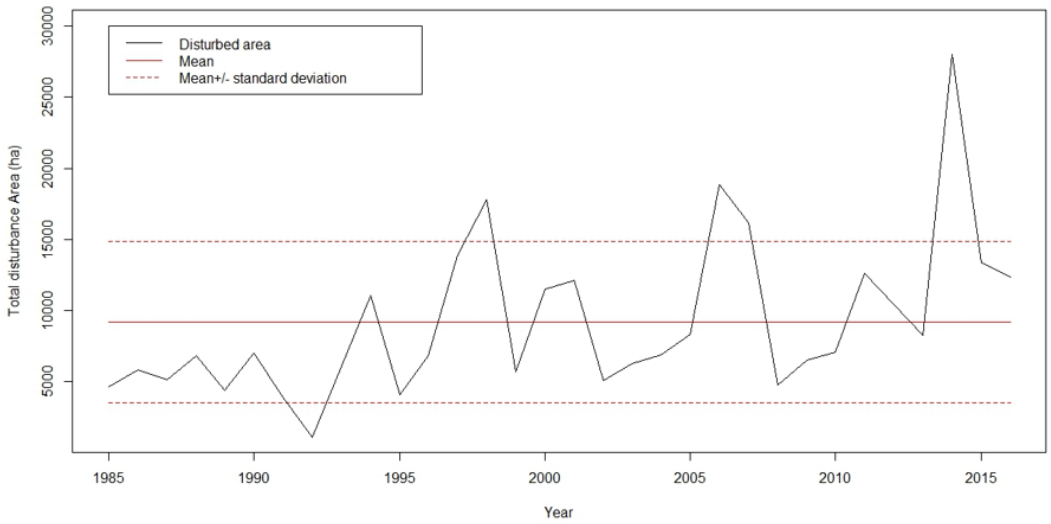


Figure 2.6: Estimated area of major forest disturbance.

low because pixels disturbed after 2005 may have been misclassified to subclasses of 1980s for reasons such as cloud cover (missing the disturbance) or other anomalies. User's accuracy is better than overall accuracy in the 1990s and 2000s. In 2014, the user's accuracy is lower perhaps due to the combined effect of cloud cover and the lack of imagery in 2012. Nonetheless, from 1991 forward the accuracy of the classification process has been approximately 70%. The process to place a forest initiation date on even-aged forests that are 25 years old (a typical even-aged rotation age) or lower in the lower Coastal Plain of the southeastern United States seems plausible.

We also estimated the area of major forest disturbance in each year between 1984 and 2016 (Figure 2.6). The disturbed area for each year ranged from 1118 ha to 27,989 ha. The mean annual disturbance over the period is 10,238 ha. In four of these years, many pixels were detected as having been disturbed. The size of the disturbed area in these years is more than mean over time plus one standard deviation over time. Fluctuations in regional and global markets (and economies) may explain some of this variation. Furthermore, since the study area was a sub-set of the markets for several major wood demand centers, it seems plausible that harvest activity could have shifted into and out of the area from one year to

the next. Interestingly, the area of disturbance has an increasing trend over time, and about 32% of the land area within our area of interest is disturbed at least once in the last 32 years. Previous reports regarding the forest cover status of the southern United States have suggested that commercial, reserved, deferred, and unproductive forest areas accounted for about 45.7% of the landscape in 2012 (Oswalt et al., 2014). Given the long history of forest management in the southern United States (U.S. Forest Service, 1988), activity on 32% of the forests over 32 years does not seem unreasonable.

Table 2.1: Error matrix of accuracy assessment. -99 is assigned to the pixels without any disturbance record.

	Reference																																
Prediction	-99	85	86	87	88	89	90	91	92	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	13	14	15	16	Row Total	User's acc.	User's acc. +/-1 year
-99	68		2	1	1		2	1	1	1		2	1	1		3	3	1				2		1		1	2	2	1	1	2	100	68%	68%
85	25	39	9	4					1		2							1		1	2	1			2		1	8	1	1	2	100	39%	48%
86	18	12	32	9			4			2				2	1					3				2	2	3	2	4	1	1	2	100	32%	53%
87	8		12	50	4	1				1				1						2		1		3		3	2	6	4	2		100	50%	66%
88	21	1		13	42		3							1	1						2		1	1	1	3		6	1	1	2	100	42%	55%
89	12				31	36	1	2		1					2	2		1			1	1			1	2		3	3		1	100	36%	68%
90	28		1	2	4	2	31	5		3		1	1			1	1		1			1		1	3		4	5	2		3	100	31%	38%
91	5						25	44	4	3	1	1			1				1		1	1			1	1		1	4	1	5	100	44%	73%
92	14							27	42	4			1		1	1	1			1								3	2	1	2	100	42%	73%
94	16						1	1	1	70	1				1		3	1							1					1	3	100	70%	72%
95	6		1		1				1	5	75	1				1										1		4		4		100	75%	81%
96	13			1						3	2	71	1	1				1	1					1			3			1	1	100	71%	74%
97	13										2	20	50	2	2	1									1				2	1	6	100	50%	72%
98	23			2		1		2		1		1	18	37	5	1	2	1				1		1	1			1	1		1	100	37%	60%
99	6	1			1				1					27	62													1		1		100	62%	89%
00	15	1					2				1	1			23	43	1	1			1	1					1	4	2	1	2	100	43%	67%
01	18			1											1	45	32	2								1						100	32%	79%
02	16															3	24	48	2	2		2			1	1			1			100	48%	74%
03	18														1	1	3	22	47	1		1	1				1	1	2	1		100	47%	70%
04	13	1																1	41	38	1		1			1	1	2				100	38%	80%
05	12			1																31	47	2	1	1				4		1		100	47%	80%
06	20											1				2	1	1			10	55	2			1	1	4	1		1	100	55%	67%
07	27			1		1		1			1					1					3	4	55	6							100	55%	65%
08	5		2			1			1	1													5	76	6	1			1	1		100	76%	87%
09	6		1	1	1	1				1	2			1		1								24	58	3						100	58%	85%
10	4				1		1			2	1		1											1	18	70	1					100	70%	89%
11	18			1	1	1			2					1						1				2		22	49	2				100	49%	73%
13	1																										10	88	1			100	88%	99%
14	28		1	1				1		2	1		1	1			2	2	1		1	3		2	4	2	1	12	33	1		100	33%	46%
15	18		1	1			1	2			1			4		5	1		2					1				3	10	47	2	99	47%	60%
16	6			2							1	3		1		4					2								1	6	74	100	74%	80%
Column Total	501	55	62	91	87	44	71	86	54	100	91	102	74	80	101	115	74	83	96	80	71	76	66	123	100	116	79	164	74	74	109	3099		
Producer's acc.	14%	71%	52%	55%	48%	82%	44%	51%	78%	70%	82%	70%	68%	46%	61%	37%	43%	58%	49%	48%	66%	72%	83%	62%	58%	60%	62%	54%	45%	64%	68%		52%	71%
Producer's acc. +/-1 year	14%	93%	85%	79%	89%	86%	80%	88%	87%	79%	86%	90%	93%	82%	89%	77%	77%	87%	94%	88%	82%	80%	94%	86%	82%	82%	76%	62%	59%	73%	70%		71%	

2.4 Discussion

As this is a preliminary assessment of a process for determining an age classification of lower Coastal Plain forests in the southeastern United States, some processing challenges were encountered. Many testing areas of land that were actually disturbed around 2010 were not classified to an age class of this vintage. Instead, they are misclassified to the earlier age class (mainly before 1980). Individual visual observation of these age classes using Google Earth and Landsat imagery reveal that our algorithm failed to reliably detect the disturbances around 2010. This might be because we did not use imagery from 2012 and the forest vegetation (as described by the IFZ value) recovered rather quickly. Additionally, given the lack of imagery in 2012, and depending on the relationship between the time of activity and the time of data capture via the satellite, some disturbances from 2012 were likely attributed to 2013. Since the magnitude of the difference in pre- and post-disturbance IFZ values is greatest when the time interval is only one year, we conclude that more frequent (i.e., semi-annual) temporal assessments may be necessary to best ascribe an establishment age to an even-aged forest in the landscape we are studying.

The presence of cloud cover and its differences between adjacent years of imagery can also cause problems in comparing annual IFZ values (and thus determining whether a major disturbance has occurred). Although we selected the imagery with the least cloud coverage for each year, some imagery still contained more than 20% cloud cover. Some of these clouded areas that were masked from the analysis contained forests, and thus it is possible that high cloud cover ratio affects the ratio of forest in our area of interest. Along these lines, another possible factor affecting the accuracy of the classification involves thin, cirrus clouds. Some thin clouds were not recognized and masked by the CFMask process and the resulting IFZ values were likely affected in these areas.

The lower Coastal Plain of the southeastern United States provides the climate and substrate for vegetation to establish and grow very quickly. The dormant season (generally December to March) is short and herbaceous and small woody vegetation can grow quickly

in spring. Therefore, beyond the environmental and technical challenges noted above, more consideration likely needs to be directed to the impact of the form of forest management on the IFZ value that is computed. As an example, assume that the Landsat imagery for an area was captured in June 2009. Later, in July 2009 a final harvest may have been applied to this area. The next Landsat imagery may have been captured in June 2010. During the intervening 11 months, the forest landowner may have the site prepared the area (chopped, burned, piled, or raked the debris and perhaps bedded the ground) or applied a chemical treatment to control the vegetation that would compete for resources with planted trees. Depending on the timing of these activities, the land area may have been given an IFZ value that would prompt the system to conclude a disturbance had occurred. However, had the land area recovered quickly with pioneering vegetation, or had the landowner delayed (or avoided) site preparation and chemical vegetation control treatments, it is plausible that the difference in IFZ value between two subsequent satellite images may not have been dramatic enough for the process to conclude that a major disturbance had occurred. This would also lead us to conclude that more frequent (i.e., semi-annual) temporal assessments may be necessary to best ascribe an establishment age to an even-aged forest in the landscape we are studying.

2.5 Conclusion

The algorithm presented here, which is similar to the Vegetation Change Tracker and other methods for automatic detection of changes in land use, was implemented to help understand the age class distribution of current forests in the lower Coastal Plain of the State of Georgia (USA). The algorithm relied on three bands (visible and infrared regions) within Landsat data and a long time series of growing season images to determine when major forest disturbance events occurred. It was moderately successful in determining the years of disturbances and current age classes of forests in this region, given that many of them

are managed as even-aged stands. The implementation and its testing met some technological challenges. The lack of suitable imagery for two years required modifications to the mathematical assessment of the integrated forest Z score (IFZ value), and the presence of cloud cover, ubiquitous to the area during the growing season, may have led to some classification errors. More specifically, cloud cover, when masked from the analysis, could have required examining non-adjacent years for evidence of disturbance, which may have added to classification error.

2.6 Acknowledgements

This graduate research is funded by Japan Student Service Organization. This work was also supported by the U.S. Department of Agriculture, National Institute of Food and Agriculture, McIntire-Stennis project 1012166, administered by the University of Georgia. We thank our colleagues from the University of Georgia Warnell School of Forestry and Natural Resources who provided valuable advice and expertise that greatly assisted the research.

2.7 References

Bettinger, P., K. Boston, J. P. Siry, and D. L. Grebner (2017). *Forest management and planning*. 2nd edition. Boston, MA: Academic Press. 362 pp.

Brooks, E. B., R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston (2014). "On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6, pp. 3316–3332. DOI: 10.1109/TGRS.2013.2272545.

Chen, J., X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin (2011). "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images". In: *Remote Sensing of Environment* 115.4, pp. 1053–1064. DOI: 10.1016/j.rse.2010.12.010.

Cieszewski, C. J. (2011). "Spatially explicit biomass supply sustainability analysis for bioenergy mill siting in Georgia, USA". In: *The Open Forest Science Journal* 4.1, pp. 2–14. DOI: 10.2174/1874398601104010002.

Cieszewski, C. J., M. Zasada, B. E. Borders, R. C. Lowe, J. Zawadzki, M. L. Clutter, and R. F. Daniels (2004). "Spatially explicit sustainability analysis of long-term fiber supply in Georgia, USA". In: *Forest Ecology and Management* 187.2, pp. 349–359. DOI: 10.1016/j.foreco.2003.08.001.

Georgia Forestry Commision (2018). *Conservation Reserve Program (CRP)*. Cost Share & Incentive Programs. URL: http://www.gfc.state.ga.us/forest-management/private-forest-management/landowner-programs/other-landowner-programs/ (visited on 02/12/2018).

Goward, S. N., C. Huang, F. Zhao, K. Schleeweis, K. Rishmawi, M. Lindsey, J. L. Dungan, and A. R. Michaelis (2016). "NACP NAFD project: Forest disturbance history from Landsat, 1986-2010". In: *ORNL DAAC*. DOI: https://doi.org/10.3334/ORNLDAAC/1290.

Huang, C., S. N. Goward, J. G. Masek, F. Gao, E. F. Vermote, N. Thomas, K. Schleeweis, R. E. Kennedy, Z. Zhu, J. C. Eidenshink, and J. R. Townshend (2009a). "Development of time series stacks of Landsat images for reconstructing forest disturbance history". In: *International Journal of Digital Earth* 2.3, pp. 195–218. DOI: 10.1080/17538940902801614.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann (2010). "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks". In: *Remote Sensing of Environment* 114.1, pp. 183–198. DOI: 10.1016/j.rse.2009.08.017.

Huang, C., S. N. Goward, K. Schleeweis, N. Thomas, J. G. Masek, and Z. Zhu (2009b). "Dynamics of national forests assessed using the Landsat record: Case studies in eastern

United States". In: *Remote Sensing of Environment* 113.7, pp. 1430–1442. DOI: 10.1016/j.rse.2008.06.016.

Jin, S., L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian (2013). "A comprehensive change detection method for updating the national land cover database to circa 2011". In: *Remote Sensing of Environment* 132, pp. 159–175. DOI: 10.1016/j.rse.2013.01.012.

Kennedy, R. E., Z. Yang, and W. B. Cohen (2010). "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms". In: *Remote Sensing of Environment* 114.12, pp. 2897–2910. DOI: 10.1016/j.rse.2010.07.008.

Liu, S. and C. J. Cieszewski (2009). "Impacts of management intensity and harvesting practices on long-term forest resource sustainability in Georgia". In: *Mathematical and Computational Forestry & Natural-Resource Sciences* 1.2, pp. 52–66.

Natural Resources Spatial Analysis Lab (2003). *Georgia GAP Analysis*. University of Georgia College of Engineerings. URL: http://narsal.uga.edu/ (visited on 04/14/2020).

Obata, S. (2018). "Estimation of forest stand disturbance through implementation of vegetation change tracker algorithm using Landsat time series stacked imagery in coastal Georgia". In: *Mathematical and Computational Forestry and Natural-Resource Sciences* 10.1, p. 32.

Oswalt, S. N., W. B. Smith, P. D. Miles, and S. A. Pugh (2014). *Forest resources of the United States, 2012: A technical document supporting the Forest Service 2010 update of the RPA assessment*. DOI: 10.2737/WO-GTR-91.

Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.e., R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R. H. Wynne, and Z. Zhu

(2014). "Landsat-8: Science and product vision for terrestrial global change research". In: *Remote Sensing of Environment* 145, pp. 154–172. DOI: 10.1016/j.rse.2014.02.001.

Turner, M. G. (1987). "Land use changes and net primary production in the Georgia, USA, landscape: 1935–1982". In: *Environmental Management* 11.2, pp. 237–247. DOI: 10.1007/BF01867202.

U.S. Forest Service (1988). *The South's fourth forest : Alternatives for the future.* Vol. 24. Washington, D.C.: U.S. Department of Agriculture, Forest Service. 532 pp.

Vogelmann, J. E., G. Xian, C. Homer, and B. Tolk (2012). "Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems". In: *Remote Sensing of Environment* 122, pp. 92–105. DOI: 10.1016/j.rse.2011.06.027.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian (2018). "A new generation of the United States national land cover database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146, pp. 108–123. DOI: 10.1016/j.isprsjprs.2018.09.006.

Zhu, Z., C. E. Woodcock, C. Holden, and Z. Yang (2015b). "Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time". In: *Remote Sensing of Environment* 162, pp. 67–83. DOI: 10.1016/j.rse.2015.02.009.

Chapter 3

Mapping forest disturbances between 1987-2016 using dense time series Landsat TM/ETM imagery: Implementing a reliable methodology for Georgia, United States [1]

[1] Shingo Obata, Chris J. Cieszewski, Pete Bettinger, and Roger C. Lowe III, Accepted by Forests, 2020. Reprinted here with permission of publisher

Abstract

Forest resources have a high economic value in the State of Georgia (USA) and the landscape is frequently disturbed as a part of forest management activities, such as plantation forest management activities. Thus, tracking the stand-clearing disturbance history in a spatially referenced manner might be pivotal in discussions of forest resource sustainability within the state. The two major objectives of this research are (i) to implement and test a reliable methodology for state-wide tracking of forest disturbances in Georgia, (ii) to consider and discuss the use and implications of the information derived from the forest disturbance map. Two primary disturbance detection methods, a threshold algorithm and a statistical boundary method, were combined to implement a robust estimation of recent forest disturbance history. The implemented model was used to create a forest disturbance record for the years 1984-2016, through the use of dense Landsat TM/ETM data. The final product was a raster database, where each pixel was assigned a value corresponding to the last disturbance year. The overall accuracy of the forest disturbance map was 87%, and it indicated that 4,503,253 ha, equivalent to 29.2% of the total land area in Georgia, experienced disturbances between 1984 and 2016. The estimated disturbed area in each year was highly variable and ranged between 84,651 ha (± 36,354 ha) to 211,780 ha (± 49,504 ha). By combining the use of the disturbance map along with the 2016 database from the National Land Cover Database (NLCD), we also analyzed the regional variation in the disturbance history; This analysis indicated that disturbed forests in urban areas were more likely to be converted to other land-uses. The forest disturbance record created in this research provides the necessary spatial data to address forest resource sustainability in Georgia. Additionally, the methodology used has application in the analysis of other resources, such as the estimation of the above-ground forest biomass.
Keywords: Forest sustainability, forest disturbance, Landsat imagery, Google Earth Engine

3.1 Introduction

Forest disturbances are a primary factor in the dynamics of a forest ecosystem, as the history of disturbances helps keep track of the condition and development of a forest (Pflugmacher et al., 2012). Forest disturbances may affect the forest species composition and stand structure depending on the disturbance agent, magnitude, and other complex factors such as soil type, water and nutrient availability, and ecosystem biodiversity (D. P. Edwards et al., 2014). The spatio-temporal patterns of land disturbances are changing due to both anthropogenic and natural agents of disturbance (Zhu et al., 2020). Hurricanes are a major contributor to deforestation in states bordering the Atlantic and Gulf of Mexico regions of the southern United States (Goldenberg et al., 2001; Zampieri et al., 2019), while in the Western United States, forest fires, which are the second largest cause of disturbances, have increased since the mid-1980s (Westerling et al., 2006; Williams et al., 2016). In Europe, disturbances caused by wind, bark beetles, and wildfires have also increased over the last century (Seidl et al., 2014). Mechanical disturbances caused by human activity are another cause of forest disturbance. In Southeast Asia, for instance, demand for monoculture plantations, such as natural rubber, is the main driver of the disturbance of tropical forests (Grogan et al., 2015). As spatially referenced forest disturbance tracking provides fundamental information in describing the current state of a forest, accurate knowledge of it consequently facilitates the accurate prediction of future conditions and outcomes.

In the Southeastern U.S., harvesting is estimated to be the dominant contributor to disturbances in the region (Cohen et al., 2016); This region encompasses 36% of the timberland, or forested area capable of timber harvest, in the US, (Oswalt et al., 2018). Among the states in the southeastern United States, Georgia has the largest area of timberland. The plantations in Georgia consist predominantly of loblolly pine (*Pinus taeda L.*), which is the dominant commercial tree species in the southeastern United States. Silvicultural prescriptions for intensively managed plantations include site preparation, herbicide control, fertilization, thinning, and the use of genetically improved seedlings (Nilsson et al., 2003).

The continuous efforts to increase plantation volume yields (D'Amato et al., 2017; Fox et al., 2007) have led to the shortening the pine plantation rotations to 20-25 years. Unlike other states having forest practice laws (e.g., Oregon and California) (Hairston-Strang et al., 2008), Georgia employs voluntary best management practices (BMPs) to guide forest management activities (Prud'homme et al., 2002). While the BMP guidelines include adherence to the federal Clean Water Act, forest owners in Georgia do not have to adhere to a specific scheme of forest practices. Thus they have considerable flexibility for optimizing their forest management objectives and economic returns from their plantations, perhaps through intensive forest management and silviculture efforts. Because of the warm and humid climate, the characteristics of the forestland, and common forest management practices, the forest landscape in Georgia evolves rapidly, showing significant changes over relatively short periods of time. The pertinent question regarding the welfare of Georgia forests is whether they are managed on a sustainable basis (Cieszewski et al., 2004; S. Liu et al., 2009). To assess the sustainability of Georgia forests, it is essential to detect stand-level disturbances, such as final harvests, because knowledge of the date of forest regeneration helps predict the growth and yield of plantations. To this end, spatiotemporal data providing information leading to the identification of the disturbances in a spatially referenced way can inform these analyses. It is noteworthy that the Global Forest Change (GFC) dataset provides forest gain and loss with 30-meter resolution globally (Hansen et al., 2013). However, as the GFC dataset provides the disturbance history starting only after the year 2000, it was necessary to implement a model to detect the disturbances occurring from the 1980s onward. Although some research has been dedicated to partially tracking the forest disturbance history in Georgia(C. Huang et al., 2010; Obata et al., 2019; Zhu et al., 2012b), land use concerning net primary product (Turner, 1987), or development of a land cover classification system for the entire state as of 2003 (Natural Resources Spatial Analysis Lab, 2003), with the exception of the GFC, no research has covered the disturbance history of the entire state of Georgia.

The Landsat program is the most common source of data for broad-scale forest distur-

bance analyses because (i) it has a long history starting from 1972, and (ii) its products provide moderate spatial resolution for forest disturbance analyses (Kennedy et al., 2014). Since October 2008, Landsat data managed by the United States Geological Survey (USGS) has been freely available due to a change in the data policy of the Landsat program (Loveland et al., 2012; Woodcock et al., 2008; Wulder et al., 2016). By 2016, more than 3.2 million Landsat images were freely available through the USGS Earth Resources Observation and Science Center archive (Wulder et al., 2016). This policy change has stimulated new research efforts, such as those aimed at the development of algorithms to detect the land-use change over time. One of the achievements of this research field was the assessment of the forest disturbance trends of the conterminous United States between 1985-2005 (Masek et al., 2013). Another example of development within this field is that of the emergence of the LandTrendr algorithm, a univariate approach that segments an imagery time series into a series of trends, including changes in forest conditions (Grogan et al., 2015; Kennedy et al., 2010).

The Vegetation Change Tracker (VCT) algorithm is another process that was developed to reconstruct past forest disturbance history from a series of Landsat imagery. With this algorithm, the change in spectral-temporal properties of vegetation over time is tracked to detect forest disturbances (C. Huang et al., 2010; M. Li et al., 2009a). The index used to detect the change is the Integrated Forest z-score (IFZ), which is computed for each pixel in a Landsat scene. As VCT can be implemented in a reasonable amount of computational time, and since it does not require extensive parameter tuning, application toward various topics has been made. Zhu (2017) categorizes the VCT algorithm as a thresholding method that detects the change of the land-use, where large enough deviations from the assumed threshold are observed. The advantage of a thresholding method is that its implementation process is more straightforward than other change detection methods, and thus it is possible to implement this method using only one satellite image per year. The disadvantage of a thresholding method is that the performance of the model is profoundly affected by the assumed threshold values. Statistical boundary methods are another group of change

detection models that assume time series spectral values fit within a statistical boundary (Brooks et al., 2014; DeVries et al., 2015; Zhu et al., 2016). These models detect a change in forest character when any departure from the assumed statistical boundary is observed. This method has been used to detect changes from dense satellite imagery (Loveland et al., 2012), and seems less affected by the seasonality of the data, as the raw time series satellite data values are subjected to applied statistical procedures, in which the seasonal variation and the noise of the raw data are removed from the transformed values. The difficulty of performing geospatial analysis using dense time series satellite imagery involves basic information technology management and access to high-performance computing. Google Earth Engine (GEE) (Gorelick et al., 2017) has recently made it possible to overcome these two obstacles. GEE is a cloud-based platform that provides the high-performance computing resources necessary to process large-scale satellite imagery without storing and managing the data locally. Recent research has involved using dense time series Landsat imagery with the GEE platform to assess land-uses, agricultural trends, and forest vegetation (H. Huang et al., 2017; Johansen et al., 2015; Shelestov et al., 2017).

We hypothesize that the known adversarial effect of thresholding methods can be reduced when it is combined with statistical boundary methods through the GEE platform. In addition, it is expected that the disturbance detection map created from the forest disturbance detection model helps us to understand the temporal trajectory of each forest stand. Considering these two aspects of our research background, we set forth two major objectives of this research: in the first place, to implement and test a reliable methodology for state-wide tracking of forest disturbances in Georgia; secondly, to consider and discuss the use and implications of the information derived from the forest disturbance map.

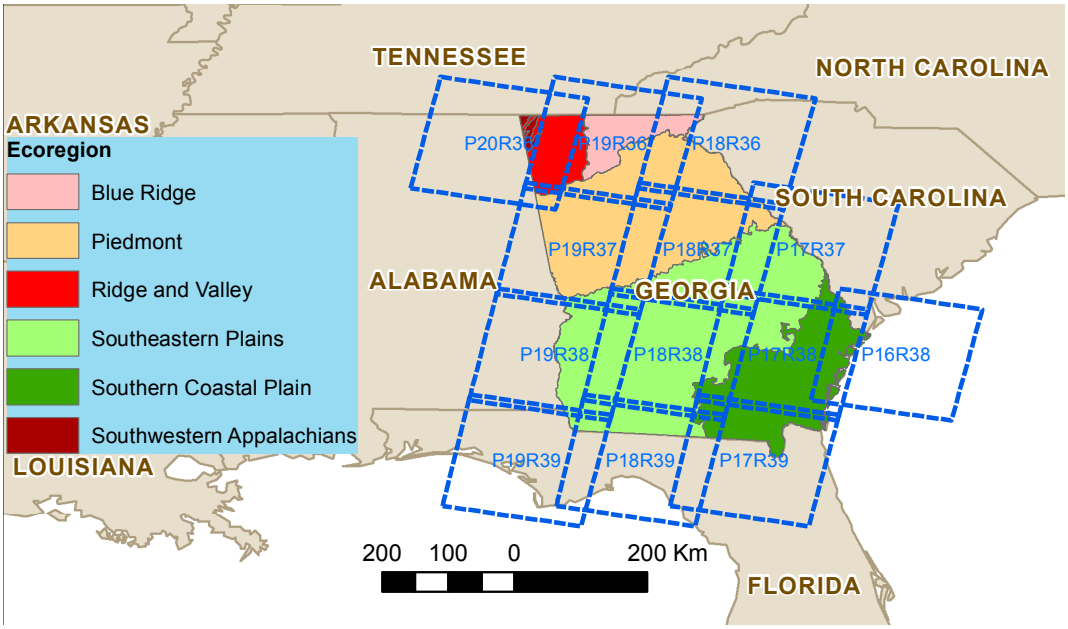


Figure 3.1: Study area. The combination of letter and number in the figure represents the path and row number of WRS-2.

3.2 Study area

The study area encompasses the entire state of Georgia (Figure 3.1). This land area is about 154,000 km^2, and the state contains approximately 100,000 km^2 of forests, covering almost 65% of the total land area (Lambert et al., 2018). Timberland potentially available for commercial use is owned in 89% of cases by private landowners or corporate entities and constitutes 97% of the forested area, which is more than any other state (Oswalt et al., 2018). The state is divided into seven ecoregions using the level III classification of the U.S. Geological Survey (Omernik et al., 2014). The Southern Coastal Plain ecoregion is the southernmost region of the state. It consists of flat plains and low elevations. The primary forest types in this region are loblolly pine and slash pine (*Pinus elliottii Engelm.*), and oak-gum-cypress forests typical of some upland areas and most lowland (wet) areas. The Southeastern Plains ecoregion consists of a mosaic of cropland, pasture, woodland, and forest. Elevations and slopes are greater than those found in the Southern Coastal

Plain. Oak-hickory-pine and southern mixed forests are two of the dominant species in this region, although there are many pine plantations as well. The Piedmont ecoregion is a transitional area between the Appalachian Mountains and the flat Southeastern Plains. The Piedmont includes the Atlanta metropolitan area, where about 57% of the population of Georgia resides (US Census Bureau, 2019); thus, urban areas comprise the most significant portion of the land-use in this ecoregion. The other ecoregions, Blue Ridge, Ridge and Valley, Southwestern Appalachians, and Interior Plateau, contain parts of the Appalachian Mountains (L. Edwards et al., 2013). Among the ecoregions, intensively managed plantation forests are the most densely located in Southern Coastal Plain. The ecoregions correspond closely to four major geologic zones of the state: (a) Coastal Plain, (b) Piedmont, (c) Blue Ridge, and (d) Appalachian Plateau. The boundary between the Coastal Plain and the Piedmont (the fall line), running east-west through the middle of Georgia, represents a distinct change in geologic condition from flatter areas containing relatively sandy soils (the Coastal Plain) to rolling hills with soils derived from the process of metamorphism (the Piedmont, Blue Ridge, and Appalachian Plateau).

3.3 Materials and methods

For the 13 combinations of paths and rows of the Worldwide Reference System-2 shown in Figure 3.1, all Landsat 5 TM and 7 ETM images between April 1984 and December 2016 were queried from the Level-1 Precision and Terrain (L1TP) corrected data set in the Google Earth Engine platform. The L1TP satisfies both radiometric and geometric criteria set by USGS (Roy et al., 2014). We did not use the Landsat 8 OLI data because the initial inspection of using Landsat 5, 7, and 8 altogether produced inconsistencies in the surface reflectance value. We refer our Landsat dataset as dense Landsat imagery since our Landsat time series dataset includes most of the all available Landsat imagery while a portion of the imagery is excluded. For each image in each scene, clouds, the shadows of clouds, water, and snow are masked

out using the C Function of the Mask (CFMask) algorithm (Foga et al., 2017; Zhu et al., 2015a; Zhu et al., 2012a). We filtered out imagery containing a high cloud cover percentage, which is likely to contain cirrus clouds frequently left unmasked by CFMask. After visually inspecting the imagery we determined not to use imagery obfuscated by more than 50% cloud cover. Landsat 7 acquired after May 31, 2003, is provided as SLC-off. We masked out the blank spaces in the imagery provided by GEE. As a result of the preprocessing, 373 to 453 images per scene were available for further analysis, as is summarized in Table 3.1. In total, 5614 images were deployed in the analysis. On average, the mean cloud cover was 14.5%, and the length of the time between images averaged 28.0 days. This indicates that, on average, new imagery was available nearly every month.

To detect forest disturbances, we computed IFZ for each pixel of each scene based on the methodology applied in the VCT algorithm (C. Huang et al., 2010; M. Li et al., 2009b). The IFZ is a vegetation index, which represents the normalized distance between a pixel value of multi-spectral satellite imagery and the value of a previously identified reference forest pixel. A small IFZ value for a given period and pixel implies that the pixel describes a relatively

Table 3.1: Dense Landsat time series data summary by scenes. Four-digit numbers at the header represent the WRS-2 row/path combination of the scene. Top 2 digits represent path number and bottom 2 digits represent row number (e.g. "1638" refers to path 16 and row 38). The number (#) of images represents the total number of available images for each scene captured between April 1984 and December 2016. The Days/Image represents the average length of time from the acquisition of an image to the acquisition of the next image in GEE.

Values/Scene	1638	1737	1738	1739	1836	1837	1838	1839	1936	1937	1938	1939	2036	Mean
Mean Cloud%	15.0	12.6	14.0	17.2	14.3	13.2	14.2	14.9	15.4	13.6	14.7	14.4	14.7	14.5
# of image	446	419	440	453	440	429	442	436	412	429	449	446	373	431.9
Days/Image	27.0	28.8	27.4	26.6	27.4	28.1	27.3	27.6	29.3	28.1	26.8	27.0	32.3	28.0

stable mature forest, while a large IFZ value indicates a non-forested area. An identified forest region is a group of pixels that represent forested areas and are used as reference pixels for the IFZ calculation. Thirty identified forest regions were selected from the entire state. Fifteen of these regions were selected from deciduous forests, while the rest of the regions were selected from coniferous forests. The size of each region encompassed approximately 50 ha. The mean and standard deviation of the area was computed from the single image met the following conditions; (i) was taken between April 1, 2018, and August 31, 2018; and (ii) cloud cover percentage is less than 1%. C. Huang et al. (2010) argue that each Landsat band has a different sensitivity to forest disturbance. The near-infrared band has low sensitivity as most of the NIR electromagnetic energy is reflected by land surface type. The red band, SWIR1 band, and SWIR 2 band of the imagery were left for IFZ calculation after removing the bands insensitive to forest disturbance. For all of the pixels in all the images queried, the IFZ was computed.

To detect a disturbance, we set three conditions related to the time series change in the IFZ. Only if the time series IFZ meets all three conditions is the disturbance detected. The first condition concerns the backward moving average. In the notation below (Eq.3.1), t_i denotes the date in which the ith scene was taken. The t_i is then converted from the standard date format to decimal value. The integer part corresponds to the year, while the fractional part represents the Julian date in the year.

$$\{t_i \in \Re \| t_1, t_2, ...t_n\} \tag{3.1}$$

For the ith imagery, xt_i is the ith imagery captured on t_i. The set of the imagery selected for t_i is denoted as $\mathbf{X_i}$.

$$\mathbf{X_i} : \{x(t_i - 3), x(t_{i+1} - 3), x(t_{i+2} - 3), ..., x(t_i)\} \quad \forall_i \tag{3.2}$$

Let $C(\mathbf{X_i})$ be the number of elements of $\mathbf{X_i}$. Then the backward moving average of IFZ for

jth pixel in the interval of $[t_{i-3}, t_i]$, denoted as $BMA_(t_{ij})$, is formulated as follows.

$$BMA_{t_{ij}} = \frac{\sum_{k \in X_i} F_{kj}}{C(\mathbf{X_i})} \tag{3.3}$$

The first condition is formulated as follows:

$$BMA_{t_{ij}} \leq 3 \tag{3.4}$$

This condition is equivalent to (a) in Figure 3.2, where the backward moving average has to be higher than 3 for a forest disturbance to be detected. The figure shows that the mean of the last three years of the IFZ values is smaller than 3, although some values are higher than 3.

The second condition is related to the size of the change in IFZ value between pre-disturbance and post-disturbance. As is argued in (Kennedy et al., 2014; H. Huang et al., 2017), the raw band value and vegetation index after a forest disturbance can deviate from the model's predicted range. Therefore, we set the condition that detects the presence of the IFZ's deviation from the past trend, which is quantified by the backward moving average. We assumed that the IFZ values do not continuously deviate from a certain range unless a disturbance occurred. The range was set to the backward moving average plus three times the standard deviation in the last three years of the observations. The standard deviation of IFZ for the jth pixel is computed as,

$$SD_{t_{ij}} = \sqrt{\frac{\sum_{k \in \mathbf{X_i}} (Z_{kj} - BMA_(t_{ij}))^2)}{C(\mathbf{X_i})}} \tag{3.5}$$

As the sample of the IFZ values after t(i), the median of the five successive values after t(i)

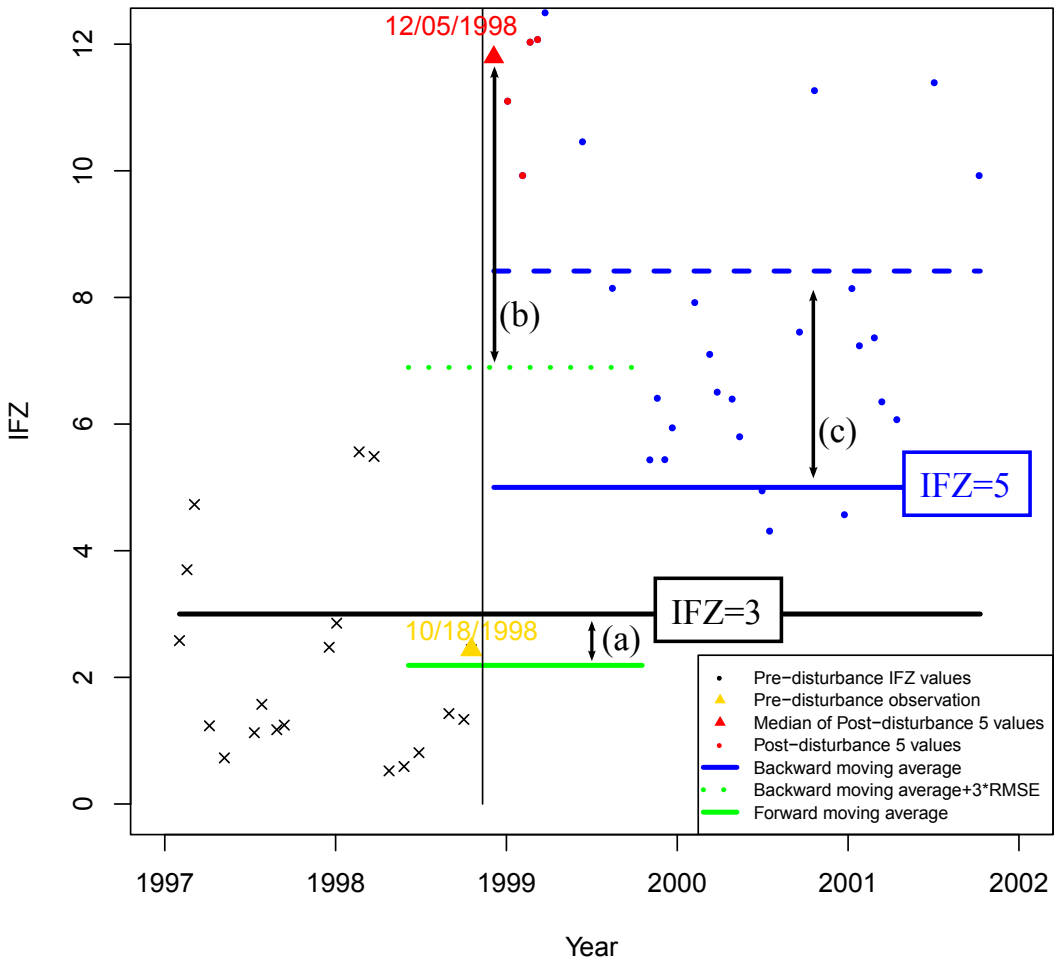


Figure 3.2: Change in time series IFZ values (Lat: 32.097^{deg}, Lon: -84.844^{deg}).

was selected. The second condition is formulated as follows:

$$BMA_{t_{ij}} + 3SD_{t_{ij}} \leq median(\mathbf{Y_i}) \tag{3.6}$$

Where

$$\mathbf{Y_i} : \{x(t_{i+1}), x(t_{i+2}), x(t_{i+3}), x(t_{i+4}), x(t_{i+5}))\} \tag{3.7}$$

This condition is shown as (b) in Figure 3.2, where the backward moving average plus three times the standard deviation has to be greater than the median for a forest disturbance to be

detected. The third condition is related to the moving average of the post-disturbance phase. This includes the same computations as those seen in the computation of the backward moving average. The collection of the imagery taken in $[t_{i+1}, t_{i+3}]$ is denoted as $\mathbf{Z_i}$.

$$\mathbf{Z_i} : \{x(t_{i+1}), x(t_{i+2}), ...x(t_i + 3))\} \tag{3.8}$$

The forward-moving average is computed as follows:

$$FMA_(t_{ij}) = \frac{\sum_{k \in \mathbf{Z_i}} F_{kj}}{C(\mathbf{Z_i}))} \tag{3.9}$$

The third condition requires that

$$FMA_{t_ij} > 5 \tag{3.10}$$

The third condition is presented as (c) in Figure 3.2, where the forward-moving average has to be greater than 5 for a forest disturbance to be detected. As the fourth and final condition of the disturbance, the magnitude of the disturbance was computed. The magnitude is the average distance of the post-disturbance IFZ values from the threshold IFZ value. We used $IFZ = 3$ as the threshold value; it is equivalent to the root mean square difference of post-disturbance IFZ values from the threshold.

All of the conditions were computed for each pixel in each image. The disturbance detection model iteratively applied the four conditions from the oldest image to the most recent image. If the model detects a disturbance for a given pixel in an image, the Julian date was assigned to the disturbance map raster. Otherwise, nothing is assigned to the pixel. Although a specific date when a disturbance was detected was recorded to the disturbance map raster, only the year was retained, since our interest lies only in the annual disturbances. After running the algorithm forward in time toward the most recent image, a single layer raster, of which pixels represent the last disturbance year, was created. The pixels without any disturbance record were assigned non-disturbance value.

3.3.1 Post-processing

We conducted a post-processing effort after creating the raw disturbance map. The first task in the post-processing effort was the reduction of "salt and pepper" in the disturbance map. The "salt and pepper" phenomenon refers to scattered pixels that deviate considerably from their neighbors based on the last disturbance year. We removed small clusters of these that were less than 9 pixels in size (equivalent to 0.81 ha), considering them to be errors. We adopted from the USDA Forest Inventory and Analysis (FIA) program a definition of the minimum size of disturbance to be 0.404 ha (1 acre). Another task completed in the post-processing effort was to cut the pixels at the edge of each image. The edge of the raw disturbance map had a band of areas that recorded disturbances more frequently than the rest of the area of the map. As a result of visual inspection, it was inferred that the individual images within the Landsat time series stack did not have precisely the same geographic extent. Therefore, the edge pixels were frequently located out of the extent of the imagery and subsequently assigned NA value. Next, we calculated the ratio of Clear Pixels (rCP) for each pixel to determine the pixels that were available in smaller quantities (due to clouds, etc.), within the available imagery across the time horizon. For every pixel in a scene, rCP was calculated as the number of the non-NA pixel values throughout the time series divided by the total number of images available within a scene. It was also determined that some of the pixels that were in specific land-uses, such as rivers, ponds, or buildings, had smaller rCP values than the rest of the pixels. Based on visual inspection, we set rCP=0.6 as the threshold value for this post-processing effort. If a pixel had an rCP smaller than the threshold, they were masked out from the forest disturbance map.

3.3.2 Accuracy assessment

To validate the forest disturbance map, we created 30 points for each of the 30 disturbance year class and 2000 points for the undisturbed classes using a stratified sampling method. The ratio of points between each disturbance year class and the undisturbed class was equiv-

alent to the initial result of the disturbance detection map, which classified about 70% of the pixels to the undisturbed class. The last disturbance year was visually inspected for each point. Google Earth Pro Historical Imagery (GEPHI), custom GEE API, and the time series Landsat Surface reflectance value chart were used as the data sources. GEPHI selected the imagery used based on the user's zoom level on display. When a user increases the scale (zooms in), it provides high-resolution imagery, such as QuickBird and NAIP. When the scale is decreased (when the user zooms out), it provides intermediate resolution imagery, such as Landsat. In our inspection, we determined that the combination of Landsat imagery and middle-to-low scale imagery level was not sufficient to verify the occurrence of a disturbance on a sample point. Therefore, when GEPHI provided only Landsat or Sentinel imagery for a given year, we switched to using a custom GEE API. The custom GEE API showed an annual Landsat mosaic imagery during the growing season for each year from 1984-2018. Although the spatial resolution of the API is 30-meters for all of the mosaics, by increasing the scale (zooming in to the points) it turned out that visual verification of disturbance was possible. For the detection of disturbances before 2005, the custom GEE API was used as the primary data source, as GEPHI did not provide annual high-resolution imagery; A time series chart of surface reflectance values was generated using the custom GEE API. Although the occurrence of disturbances was not determined solely from the time series chart, the chart was used to ascertain the time range within which disturbances were likely to occur. Out of the 2900 points created through the stratified sampling method, we first filtered out 187 points that were located on the boundary of two different stands. It is because visually determine the disturbance year for these points is difficult. The last disturbance year was detected for the rest of the points by using the aforementioned tools. We created a confusion matrix based on the sample counts, and a normal confusion matrix in terms of estimated area proportions (Card, 1982; Olofsson et al., 2013). The area-based confusion matrix was then used to develop an unbiased estimator of the total area for each disturbance class, with 95% confidence intervals. In addition, overall and producer's accu-

Table 3.2: Disturbance history classification class.

Class name	NLCD class	Disturbance	Area (ha)	Description
Disturbed forest	Forest	Between 1987-2016	2,347,978	Disturbed at least once, currently forested
Persistent forest	Forest	No disturbance	7,330,405	Persistent forest
Recent disturbance	Non-Forest	2011-2016	370,208	Disturbed after 2011, currently non-forest
Persistent non-forest	Non-Forest	No disturbance	4,944,445	Persistent non-forest
Deforestation	Non-Forest	1987-2010	398,663	Disturbed before 2011, currently non-forest

racies were calculated using the estimated area proportions for each class in the area-based confusion matrix.

Pixels without a disturbance record in our model were assigned to the undisturbed class. To further classify the undisturbed classes into either persistent forest, persistent non-forest, or into a non-forest-to-forest class, the National Landcover Database 2016 (NLCD 2016, Multi-Resolution Land Characteristics Consortium (2019)) was consulted. The goal of the projects involving the NLCD 2016 is to classify all the pixels within an area of interest to forest or non-forest classes from the beginning to the end of the time horizon. To classify the land-use at the end of the time horizon, the NLCD 2016 was used. The NLCD 2016 was developed to create a consistent multitemporal land cover and land cover change map for the conterminous U.S., at the 30-meter spatial resolution, from 2001 to 2016. In the NLCD 2016, deciduous forests, evergreen forests, mixed forests, and woody wetland classes were assigned as the vegetative land-uses. Yang et al. (2018) studied the accuracy of a map created from the model for the Appalachian Mountains in the northern part of Georgia (WRS-2 path 18, row 35), and determined that the overall agreement between map and reference labels was 88%. However, the user's accuracy of evergreen forest, mixed forest, and woody wetland were relatively low (between 32%-62%). The NLCD 2016 raster data was downloaded from the Multi-Resolution Land Characteristics Consortium website (Multi-Resolution Land Characteristics Consortium, 2019). For the sake of our goal, reclassification of the original data was performed. The four aforementioned vegetation classes were aggregated into one

forest class, while the rest of the classes were reclassified into a non-forest class. The temporal trajectory of all the land in the state in terms of the forest class was then examined (Table 3.2). The pixels assigned to the persistent forest class were those with undetected forest disturbances over our time horizon forest disturbance, and were considered as forest in the reclassified NLCD 2016. The persistent non-forest class was assigned to pixels with undetected forest disturbances, and were considered as non-forest in the reclassified NLCD 2016. The disturbed forest class was assigned to pixels with detected forest disturbance within our time horizon, and which were also considered as forest in the reclassified NLCD 2016. Pixels with undetected forest disturbances were assigned to the persistent non-forest class and were considered as non-forest in the reclassified NLCD 2016. The disturbed forest class was assigned to pixels with detected forest disturbance within our time horizon, and which were also considered as forest in the reclassified NLCD 2016. It was observed that many pixels with a recent disturbance record were classified as the disturbed non-forest class. Classifying a newly disturbed pixel as non-forest, as of 2016, was appropriate if the forest represented by the pixel had not yet recovered. However, these pixels are more likely to be regenerated as forest than pixels within the non-forest areas that were disturbed more than five years ago. Therefore, we created the deforested class and recent disturbance class as subclasses of the disturbed non-forest class to examine how much area was left as non-forest for a substantial length of time. In the typical plantation forest management scheme in the southeastern U.S., forest regeneration is initiated soon after a major disturbance, such as a final harvest or fire. Thus, we classified pixels within the disturbed non-forest class detected as being disturbed after 2011 into the recent disturbance subclass, while the rest of the pixels belonging to a disturbed non-forest class were classified as the deforested subclass. As a result, five classes were developed to describe the temporal trajectory of all the land in the state.

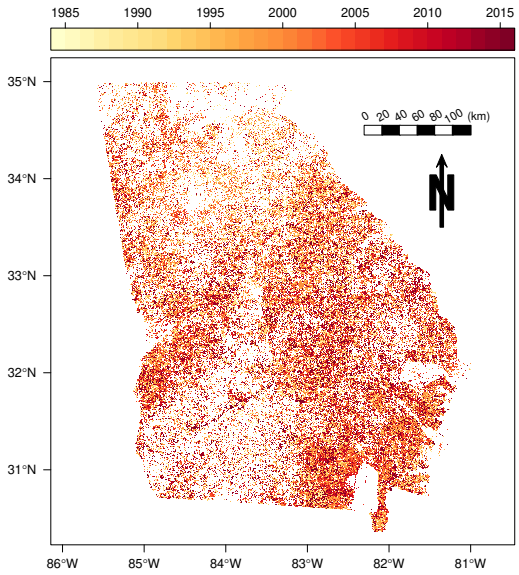


Figure 3.3: Estimated area distribution of the last disturbance year.

3.4 Results

A 30-meter spatial resolution raster database, in which each pixel was assigned a value that represents the last disturbance year of the forest (Figure 3.3), was created as a result of applying the GEE to a time series of Landsat data. As is explained in Section 3.3.2, a count-based confusion matrix was developed to assess gauge the accuracy assessment (Table 3.3). An area-adjusted confusion matrix was also developed to inform the accuracy assessment (Table 3.4). Using the area-adjusted method, the overall accuracy was 87%. As was shown in Table 3.3, when using the area-adjusted method, the commission rate was lower than the omission rate for most of the disturbed classes. Conversely, the commission rate was higher than the omission rate for the undisturbed classes. The high omission rate indicates that the disturbance detection algorithm tends to underestimate the area of disturbance.

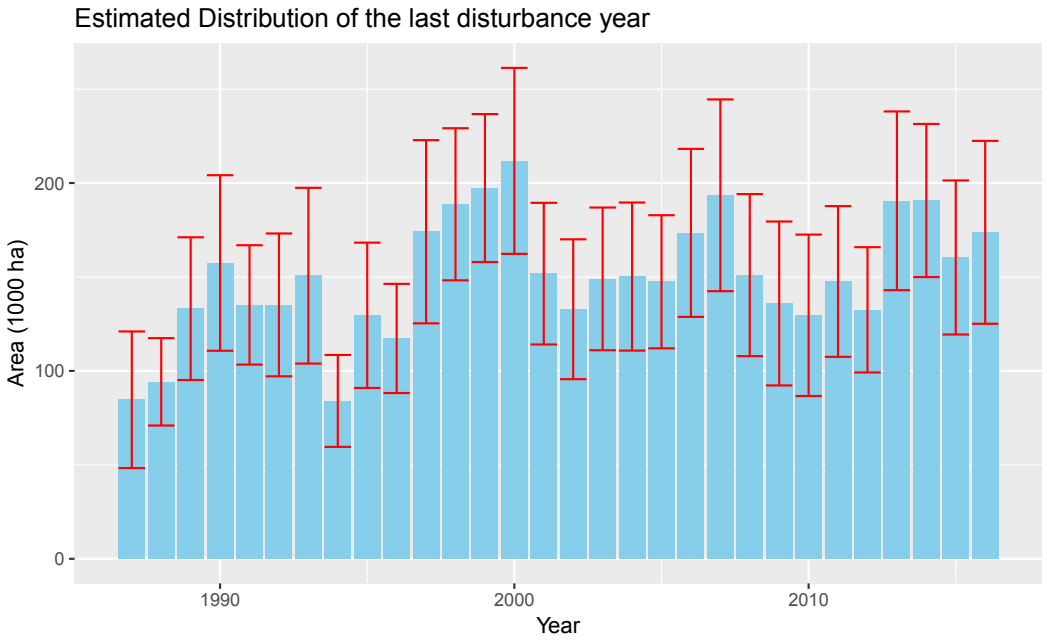


Figure 3.4: The last disturbance year map.

Table 3.3: Count-based confusion matrix. Reference data and prediction class values represent the last disturbance year (e.g., 89 = year 1989). The -1 class represents the undisturbed class. UA = user's accuracy. PA = producer's accuracy.

Prediction	Reference Data																															Total	UA (%)
	-1	87	88	89	90	91	92	93	94	95	96	97	98	99	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
-1	1612	7	2	8	12	5	7	11	3	8	4	13	9	8	13	7	8	8	8	6	9	13	9	10	9	8	5	11	7	8	12	1860	87
87	1	21	4													1																27	78
88	1		27				1																									29	93
89				28																								1				29	97
90	1				27	1																							1			30	90
91	1					27																										28	96
92	2					2	24																				1					29	83
93	3						1	23	1	1			1																			30	77
94	1								23					1																		25	92
95										28														1								29	97
96				1						1	23																	1				26	88
97	3											24									1											28	86
98					1								28																			29	97
99														27								1										28	96
0	1														27						1											29	93
1															1	26					1	1										29	90
2	1																28															29	97
3																		29												1		30	97
4	2																	1	26													29	90
5	2																			26												28	93
6	2																			1	24		1	1								29	83
7	1																					26	1	1								29	90
8	1																					3	23				1					28	82
9	1									1													1	25					1			29	86
10	10																								20							30	67
11	4														1											24						29	83
12																											25					25	100
13	1	1																									2	23				27	85
14	2		1					1																					25			29	86
15	2																		1	1										25		29	86
16	3													1		1														1	22	28	79
Total	1658	29	34	37	40	35	33	35	27	39	27	37	38	37	42	35	36	38	35	34	36	44	35	38	29	32	34	36	34	35	34	2713	
PA (%)	97	72	79	76	68	77	73	66	85	72	85	65	74	73	64	74	78	76	74	76	67	59	66	66	69	75	74	64	74	71	65		

Table 3.4: Area-adjusted confusion matrix. Reference data and prediction class values represent the last disturbance year (e.g., 89 = year 1989). The -1 class represents the undisturbed class. UA = user's accuracy. PA = producer's accuracy.

Prediction	Reference Data																															UA (%)	Area (ha)	ClassArea (acre)
	-1	87	88	89	90	91	92	93	94	95	96	97	98	99	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total			
-1	69.17	0.3	0.09	0.34	0.52	0.22	0.3	0.47	0.13	0.34	0.17	0.56	0.39	0.34	0.56	0.3	0.34	0.34	0.34	0.26	0.39	0.56	0.39	0.43	0.39	0.34	0.22	0.47	0.3	0.34	0.52	79.81	87	12,284,571	30,355,810
87	0.01	0.22	0.04													0.01																0.28	78	42,470	104,946
88	0.02		0.45				0.02																									0.48	93	73,616	181,910
89				0.49																								0.02				0.51	97	78,043	192,848
90	0.02				0.47	0.02																							0.02			0.52	90	80,670	199,339
91	0.02					0.6																										0.62	96	94,926	234,566
92	0.04					0.04	0.53																				0.02					0.64	83	99,305	245,389
93	0.06						0.02	0.46	0.02	0.02			0.02																			0.61	77	93,176	230,243
94	0.02								0.39					0.02																		0.43	92	65,664	162,259
95										0.43														0.02								0.45	97	68,883	170,214
96				0.03						0.03	0.59																	0.03				0.66	88	101,733	251,387
97	0.07											0.57									0.02											0.66	86	101,524	250,870
98					0.03								0.81																			0.84	97	129,322	319,562
99														0.89								0.03										0.92	96	141,228	348,982
0	0.03														0.76						0.03											0.81	93	125,097	309,121
1															0.03	0.64					0.03	0.03										0.72	90	109,877	271,511
2	0.02																0.51															0.53	97	81,738	201,980
3																		0.6												0.02		0.62	97	94,826	234,320
4	0.05																	0.02	0.6													0.67	90	103,049	254,639
5	0.05																			0.64												0.69	93	106,268	262,594
6	0.06																			0.03	0.66		0.03	0.03								0.79	83	122,110	301,741
7	0.02																					0.57	0.02	0.02								0.63	90	96,964	239,604
8	0.02																					0.07	0.52				0.02					0.64	82	97,960	242,065
9	0.02									0.02													0.02	0.38					0.02			0.44	86	68,373	168,954
10	0.23																								0.45							0.68	67	103,912	256,772
11	0.1														0.03											0.61						0.74	83	113,286	279,936
12																											0.53					0.53	100	82,111	202,900
13	0.03	0.03																									0.06	0.72				0.84	85	129,061	318,917
14	0.07		0.04						0.04																				0.9			1.04	86	160,223	395,919
15	0.05																		0.03	0.03										0.64		0.74	86	114,555	283,071
16	0.16													0.05		0.05														0.05	1.16	1.47	79	159,235	393,479
Total	70.33	0.55	0.61	0.86	1.02	0.87	0.87	0.97	0.54	0.84	0.76	1.12	1.22	1.3	1.37	1	0.86	0.96	0.97	0.95	1.12	1.25	0.97	0.88	0.84	0.95	0.86	1.23	1.23	1.06	1.67	100			
PA (%)	98	39	73	57	46	68	61	48	73	52	77	50	67	68	55	64	60	62	62	67	59	45	54	44	54	64	62	58	73	61	69	Overall acc. (%)		87	

Table 3.5: Count-based confusion matrix for Global Forest Change Version 1.4 data set. Reference data and prediction class values represent the last disturbance year (e.g., 01 = year 2001). The -1 class represents the undisturbed class. UA = user's accuracy. PA = producer's accuracy.

Prediction	Reference data																	Total	UA (%)	
	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16			
-1	2058	12	13	10	9	12	13	15	11	12	12	11	8	12	12	10	16	2246		92
1	1	18	2				1	1							2			25		72
2	3	3	20	1				1						1		2		31		65
3				24										2		1		27		89
4	6			1	19	3		1			1		1			1	2	35		54
5	2		1		3	14	2							1		2		25		56
6	2			1	2	1	17							1	1			25		68
7	3					2	2	23	2				1	1		1		35		66
8	5	1					1	2	20	2					1	1	1	34		59
9	6			1	1			1	2	21	3	1				1		37		57
10	10									2	11	3		1				27		41
11	5					2				1	1	14	2					25		56
12	5				1						1	3	21	5				36		58
13	6												1	11				18		61
14	9													1	17	3		30		57
15	12	1													1	10	2	26		38
16	15															3	13	31		42
Total	2148	35	36	38	35	34	36	44	35	38	29	32	34	36	34	35	34	2713		
PA (%)	96	74	78	76	74	76	67	59	66	66	69	75	74	64	74	71	65	Overall acc. (%)		86

The estimated amount of disturbed area by disturbance year is plotted in Figure 3.4. The estimate indicated that 4,503,253 ha of the forest land in Georgia, which occupies approximately 29.2% of the total forest land, was disturbed at least once between 1984 and 2016. The annual amount of disturbance area ranged between 84,651 ha (±36,354 ha) and 211,780 ha (±49,504 ha) for the entire state. It is worth mentioning that the annual last disturbance area tends to be lower in the early period of the time horizon of this research. Conversely, the annual last disturbance area tends to be larger for more recent years. This is because the recent disturbances overwrite the older disturbance record. To examine the regional variations of in forest disturbances, we aggregated the pixel values of the temporal trajectory raster data to the county-scale. The mean last disturbance year was then computed for each county, using pixels that have a disturbance record (Figure 3.5). Subsequently, county-scale data was further aggregated at the ecoregion-scale (Table 3.6) to illustrate the mean last disturbance year for the 159 counties in the state. The mean last disturbance year was approximately 2003.69 for Southern Coastal Plain, while that of Blue Ridge ecoregion was 2000.11. This result indicates that the forest was disturbed more frequently in the areas where intensive management of plantations occupies more land.

To further examine the types of disturbances, the ratio of the recent disturbance subclass to the disturbed non-forest class (Figure 3.6) was mapped. Figure 3.6 indicates that the majority of non-forest pixels with a disturbance record in the Southern Coastal Plain were disturbed after 2011. In these 19 counties located in the Southern Coastal Plain, more than 70% of the non-forest land was disturbed since 2011. On the other hand, about 67% of the non-forest land in Piedmont was disturbed during this time period. This result coincides with previous research, which revealed that the growth of the Atlanta metropolitan area over the last four decades has led to changes in land-use and the fragmentation of forests for urbanization purposes (Miller, 2012).

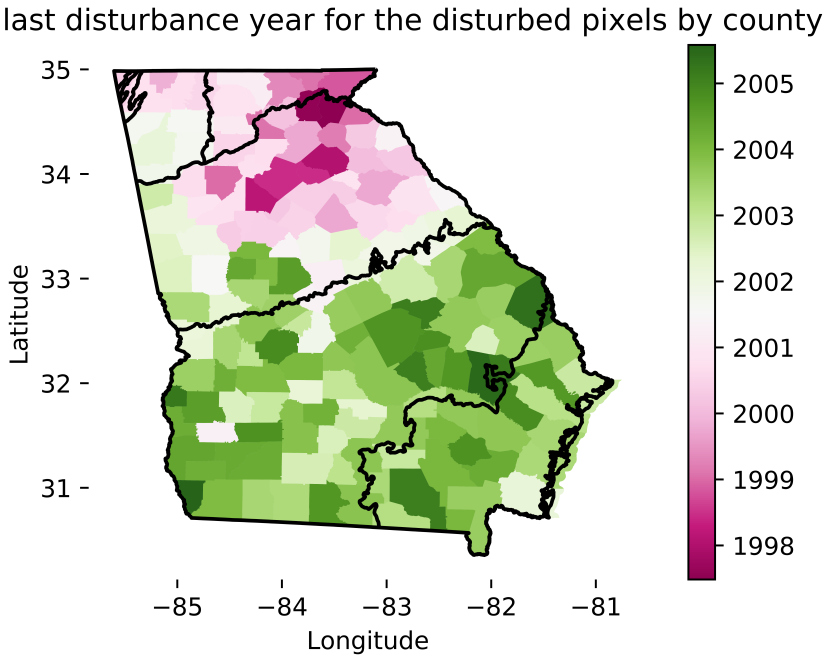


Figure 3.5: Mean last disturbance year for the disturbed pixels.

3.5 Discussion

In this research, we examined the recent forest disturbance history of the state of Georgia. In comparison to the change detection research that employed the VCT algorithm in North Carolina (C. Huang et al., 2015), our results have almost the same overall accuracy (88.6% vs. 87%), although the accuracy assessment methods between two research models are not identical. The most significant difference between the model in C. Huang et al. (2015) and our model is the number of the Landsat time series scenes used to conduct the analysis. As a result, our model less frequently misclassifies the disturbance year, plus or minus one year. This may be because we used dense imagery with less than 50% cloud cover. Due to the improved data availability, it was possible to detect forest disturbances approximately every 30 days, unlike the previous research that used a single set of imagery or mosaic imagery captured only during the growing season.

We attempted to compare our result with the two existing disturbance detection datasets:

the North American Forest Dynamics data set (Masek et al., 2013), and the Global Forest Change dataset (Hansen et al., 2013; Hansen et al., 2010). The North American Forest Dynamics dataset does not match the spatio-temporal range of our study, as it covered only 2 out of 14 WRS-2 scenes used in our research. Additionally, the North American Forest Dynamic dataset was limited to detecting change from 1986-2010. On the other hand, the Global Forest Change dataset includes the entire state of Georgia, although its temporal coverage initiates only after the year 2001. Thus, we selected the Global Forest Change dataset as the basis of the comparison. Version 1.4 of the Global Forest Change dataset was selected among the several available versions, as it covers the years extending between 2000 and 2016. If an accuracy assessment point was assigned a value smaller than that found in 2001, we reclassified it to the undisturbed class. Following this, the pixel values of the band that represents the year of gross forest cover loss event in the Global Forest Change

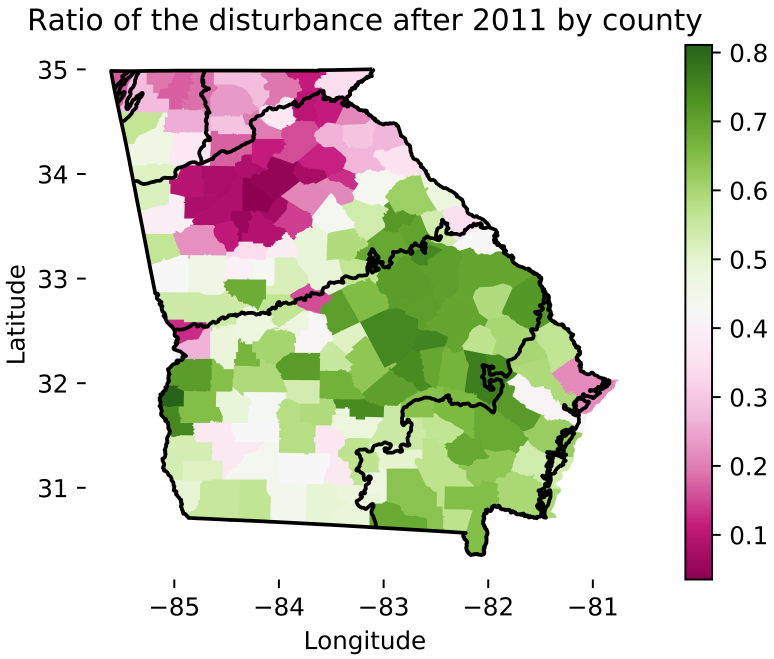


Figure 3.6: Ratio of the disturbance after 2011 to all of the disturbances detected between 1987-2016.

Table 3.6: Land-use pattern classification in Georgia. DistFor: Disturbed forest (%); PerFor: Persistent forest (%); RecDist: Recent disturbance (%); PerNon: Persistent non-forest (%); DeFor: Deforestation (%); AvgYear: Mean last disturbance year.

Ecoregion	DistFor	PerFor	RecDist	PerNon	Defor	AvgYear
Blue Ridge	3.51	76.47	0.36	18.36	1.31	2000.11
Piedmont	11.83	50.23	1.52	33.18	3.23	2000.93
Ridge and Valley	9.07	52.70	1.02	35.30	1.92	2001.12
Southeastern Plains	15.48	42.73	2.80	37.21	1.79	2003.68
Southern Coastal Plain	23.27	45.18	4.55	23.96	3.06	2003.69

dataset was extracted to the accuracy assessment points. Finally, a confusion matrix was constructed (Table 3.5). The overall accuracy of Global Forest Change was 85.9%, which is lower than our result. In most of the user's and producer's, accuracies for disturbance class in Table 3.5 were lower than our result, except for the undisturbed class.

The four conditions described in our new algorithm served different roles in detecting forest disturbances. Among these conditions, two of them, namely the backward and forward-moving average of IFZ, are based on the threshold method, while a third is based on the statistical boundary method. The backward moving average is a condition that quantifies the likelihood of a pixel being forest three years before the target date. The forward-moving average, on the other hand, quantifies the likelihood of a pixel being forest three years after the target date. These two moving averages describe the state of the landscape within a time frame. As is shown in the various research conducted concerning change detection models, it is possible to detect the occurrence of the land-use change by using only the backward and forward moving averages (Kayastha et al., 2012; J. Li et al., 2015; Pickell et al., 2014). However, if only these two conditions are employed, our algorithm suggests that disturbances are detected before they actually occur. This is because the forward-moving average exceeds the threshold when most of the observations come from the post-disturbance period. The final condition, which sets the statistical boundary for the five post-disturbance observations,

helps to detect disturbances at the correct time. Therefore, combining the threshold method and the statistical boundary method alleviates the limitations of each method individually. Although the results of the accuracy assessment suggest that this new method is better than our previous analysis, which attempted to detect the disturbance history of seven counties in Coastal Georgia (Obata et al., 2019), our new method still contained several areas of concern. For instance, on the commission error side of the analysis, we found that some forest thinnings could have been misclassified as a final harvest disturbance. One possible reason for the confusion is that the change in IFZ, when mediated through a high intensity thinning and through a final disturbance, could be similar due to the amount of forest canopy removed, the amount of forest floor exposed, and the size of Landsat pixels. Often, a forest thinning in plantation forests in Georgia removes up to 50% of the basal area, and the residual live tree canopy may require three years or more to close the gaps that are created. Conversely, on the omission error side, omission error rates for the disturbance classes were between 23% to 67% in the area-adjusted confusion matrix. These errors are often physically located at the edges of the disturbed stands. The omission error rate could be reduced by changing the threshold values set for our disturbance detection algorithm. However, a more relaxed set of the threshold values for the disturbance detection process might result in a higher commission error rate, by confusing thinnings with final harvest disturbances. Thus, a delicate balance must be struck if the parameters are to be adjusted in an effort to reduce omission or commission errors. One plausible way of acquiring better overall accuracy may be to develop a system that would optimize the threshold values in the conditions we set in our algorithm.

Future research related to the methods described here might focus on detecting more recent disturbances. To achieve this goal, it is necessary to remove the third and fourth conditions in our model, which observe the IFZ values after the disturbance. A second focus may be to combine the regeneration history data with disturbance year, as the timing of regeneration processes (site preparation, planting, etc.) can range from nearly instantaneous

after harvest, to 2-3 years after harvest. In our regional-scale disturbance analysis we did not directly examine the timing of the forest regeneration. By applying a method for detecting the forest regeneration, it is expected that one may more closely comprehend dynamic changes in forests across broad areas of a landscape.

Finally, the disturbance year map created in this research can be used to conduct an analysis related to the forest inventory in Georgia. Once the disturbance year is combined with a broad-scale forest inventory, the data can be used to estimate the above-ground biomass and perhaps other forest product-related metrics of interest to both society and industry. Previous research has demonstrated the use of these types of methods in estimating biomass in various regions by incorporating environmental data to improve the accuracy of the estimation (Maselli et al., 2005; McRoberts, 2012; Tanaka et al., 2014; Tomppo et al., 2008). Although these research efforts have succeeded in creating databases that estimate above-ground biomass, there is a known issue in which spectral reflectance values of Landsat imagery are insensitive to changes in biomass among dense and multilayered canopy forests, regardless of their age. Consequently, the biomass estimation results in low accuracy for high biomass stands (Lu et al., 2016; Zhao et al., 2016). By incorporating the disturbance year data created in this research, it is expected that the old forest and young forest can be discriminated more easily among dense and multilayered canopy forests.

3.6 Conclusion

In this research we created a raster database, in which each pixel value represented the last disturbance year of forests within the State of Georgia (USA). Then, Google Earth Engine was used to process dense images from the time series of Landsat TM/ETM imagery. Finally, we combined the VCT algorithm, a threshold method, and the statistical boundary method to create our algorithm, which to detects the forest disturbances. The overall accuracy of the disturbance year data was 87%. The data suggest that 29.2% of the land area in Georgia

was disturbed between 1984 and 2016. The estimated annual last disturbance area ranged between 84,651 ± 36,354 ha and 211,780 ± 49,504 ha. As the process records only the final disturbance for each pixel, the annual area of disturbance tends to increase towards the more recent years. The use of dense time series Landsat imagery contributed to the achievement of less frequent misclassification of the disturbance year, plus or minus one year. Combining the threshold method and the statistical boundary method made it possible to detect disturbances with a relatively accurate temporal specification. Furthermore, by using the NLCD 2016 dataset and the disturbance map we created, we were able to further examine regional variations in the disturbance history. The subsequent analysis confirmed common insight on land-use patterns: namely, that disturbed forests near urban areas were more likely to be converted to other land-uses than disturbed forests in rural areas. Meanwhile, in the Southern Coastal Plain, where intensive plantation forest management is widely practiced, forests experienced the greatest intensity of disturbance and regenerated at a higher rate relative to other regions. The remaining task is to increase the size of the dataset used for the disturbance detection by combining sensors or missions that were not used in this analysis. For instance, incorporation of Landsat 8 OLI may improve the accuracy of the estimation since it enables the denser temporal coverage within the time range. In addition, Sentinel 2 could be incorporated to our model if the data fusion between Landsat mission and Sentinel mission is successfully achieved.

3.7 Acknowledgements

We are indebted to the Japan Student Service Organization Student Exchange Support Program (Graduate Scholarship for Degree Seeking Students) for funding this research. This work was also supported by the U.S. Department of Agriculture, National Institute of Food and Agriculture, McIntire-Stennis project (1012166), administered by the Warnell School of Forestry and Natural Resources at the University of Georgia. We would also like to mention

our appreciation for the support and assistance of the Google Earth Engine Development team.

3.8 References

Brooks, E. B., R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston (2014). "On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6, pp. 3316–3332. DOI: 10.1109/TGRS.2013.2272545.

Card, D. H. (1982). "Using known map category marginal frequencies to improve estimates of thematic map accuracy". In: *Photogrammetric Engineering & Remote Sensing* 48.3, pp. 431–439.

Cieszewski, C. J., M. Zasada, B. E. Borders, R. C. Lowe, J. Zawadzki, M. L. Clutter, and R. F. Daniels (2004). "Spatially explicit sustainability analysis of long-term fiber supply in Georgia, USA". In: *Forest Ecology and Management* 187.2, pp. 349–359. DOI: 10.1016/j.foreco.2003.08.001.

Cohen, W. B., Z. Yang, S. V. Stehman, T. A. Schroeder, D. M. Bell, J. G. Masek, C. Huang, and G. W. Meigs (2016). "Forest disturbance across the conterminous United States from 1985–2012: The emerging dominance of forest decline". In: *Forest Ecology and Management* 360, pp. 242–252. DOI: 10.1016/j.foreco.2015.10.042.

D'Amato, A. W., E. J. Jokela, K. L. O'Hara, and J. N. Long (2017). "Silviculture in the United States: An amazing period of change over the past 30 years". In: *Journal of Forestry* 116.1, pp. 55–67. DOI: 10.5849/JOF-2016-035.

DeVries, B., M. Decuyper, J. Verbesselt, A. Zeileis, M. Herold, and S. Joseph (2015). "Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series". In: *Remote Sensing of Environment* 169, pp. 320–334. DOI: 10.1016/j.rse.2015.08.020.

Edwards, D. P., J. A. Tobias, D. Sheil, E. Meijaard, and W. F. Laurance (2014). "Maintaining ecosystem function and services in logged tropical forests". In: *Trends in Ecology & Evolution* 29.9, pp. 511–520. DOI: 10.1016/j.tree.2014.07.003.

Edwards, L., J. Ambrose, and L. K. Kirkman (2013). *The natural communities of Georgia*. Athens, GA: University of Georgia Press. 697 pp.

Foga, S., P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue (2017). "Cloud detection algorithm comparison and validation for operational Landsat data products". In: *Remote Sensing of Environment* 194, pp. 379–390. DOI: 10.1016/j.rse.2017.03.026.

Fox, T. R., E. J. Jokela, and H. L. Allen (2007). "The development of pine plantation silviculture in the Southern United States". In: *Journal of Forestry* 105.7, pp. 337–347. DOI: 10.1093/jof/105.7.337.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray (2001). "The recent increase in Atlantic hurricane activity: Causes and implications". In: *Science* 293.5529, pp. 474–479. DOI: 10.1126/science.1060040.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". In: *Remote Sensing of Environment* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Grogan, K., D. Pflugmacher, P. Hostert, R. Kennedy, and R. Fensholt (2015). "Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series". In: *Remote Sensing of Environment* 169, pp. 438–453. DOI: 10.1016/j.rse.2015.03.001.

Hairston-Strang, A. B., P. W. Adams, and G. G. Ice (2008). "The Oregon Forest Practices Act and forest research". In: *Hydrological and biological responses to forest practices*. Ed. by J. D. Stednick. Vol. 199. New York, NY: Springer, pp. 95–113. DOI: 10.1007/978-0-387-69036-0_6.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). "High-resolution global maps of 21st-Century forest cover change". In: *Science* 342.6160, pp. 850–853. DOI: 10.1126/science.1244693.

Hansen, M. C., S. V. Stehman, and P. V. Potapov (2010). "Quantification of global gross forest cover loss". In: *Proceedings of the National Academy of Sciences* 107.19, pp. 8650–8655. DOI: 10.1073/pnas.0912668107.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann (2010). "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks". In: *Remote Sensing of Environment* 114.1, pp. 183–198. DOI: 10.1016/j.rse.2009.08.017.

Huang, C., P. Ling, and Z. Zhu (2015). "North Carolina's forest disturbance and timber production assessed using time series Landsat observations". In: *International Journal of Digital Earth* 8.12, pp. 947–969. DOI: 10.1080/17538947.2015.1034200.

Huang, H., Y. Chen, N. Clinton, J. Wang, X. Wang, C. Liu, P. Gong, J. Yang, Y. Bai, Y. Zheng, and Z. Zhu (2017). "Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine". In: *Remote Sensing of Environment* 202, pp. 166–176. DOI: 10.1016/j.rse.2017.02.021.

Johansen, K., S. Phinn, and M. Taylor (2015). "Mapping woody vegetation clearing in Queensland, Australia from Landsat imagery using the Google Earth Engine". In: *Remote Sensing Applications: Society and Environment* 1, pp. 36–49. DOI: 10.1016/j.rsase.2015.06.002.

Kayastha, N., V. Thomas, J. Galbraith, and A. Banskota (2012). "Monitoring wetland change using inter-annual Landsat time-series data". In: *Wetlands* 32.6, pp. 1149–1162. DOI: 10.1007/s13157-012-0345-1.

Kennedy, R. E., S. Andréfouët, W. B. Cohen, C. Gómez, P. Griffiths, M. Hais, S. P. Healey, E. H. Helmer, P. Hostert, M. B. Lyons, G. W. Meigs, D. Pflugmacher, S. R. Phinn, S. L. Powell, P. Scarth, S. Sen, T. A. Schroeder, A. Schneider, R. Sonnenschein, J. E. Vogelmann, M. A. Wulder, and Z. Zhu (2014). "Bringing an ecological view of change to Landsat-based remote sensing". In: *Frontiers in Ecology and the Environment* 12.6, pp. 339–346. DOI: 10.1890/130066.

Kennedy, R. E., Z. Yang, and W. B. Cohen (2010). "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms". In: *Remote Sensing of Environment* 114.12, pp. 2897–2910. DOI: 10.1016/j.rse.2010.07.008.

Lambert, S. and T. Brandeis (2018). *Forests of Georgia, 2016. Resource Update FS–176*. Asheville, N.C.: U.S. Department of Agriculture, Forest Service, Southern Research Station. 4 pp.

Li, J., C. E. Zipper, P. F. Donovan, R. H. Wynne, and A. J. Oliphant (2015). "Reconstructing disturbance history for an intensively mined region by time-series analysis of Landsat imagery". In: *Environmental Monitoring and Assessment* 187.9, p. 557. DOI: 10.1007/s10661-015-4766-1.

Li, M., C. Huang, Z. Zhu, H. Shi, H. Lu, and S. Peng (2009a). "Assessing rates of forest change and fragmentation in Alabama, USA, using the vegetation change tracker model". In: *Forest Ecology and Management* 257.6, pp. 1480–1488. DOI: 10.1016/j.foreco.2008.12.023.

Li, M., C. Huang, Z. Zhu, W. Wen, D. Xu, and A. Liu (2009b). "Use of remote sensing coupled with a vegetation change tracker model to assess rates of forest change and fragmentation in Mississippi, USA". In: *International Journal of Remote Sensing* 30.24, pp. 6559–6574. DOI: 10.1080/01431160903241999.

Liu, S. and C. J. Cieszewski (2009). "Impacts of management intensity and harvesting practices on long-term forest resource sustainability in Georgia". In: *Mathematical and Computational Forestry & Natural-Resource Sciences* 1.2, pp. 52–66.

Loveland, T. R. and J. L. Dwyer (2012). "Landsat: Building a strong future". In: *Remote Sensing of Environment* 122, pp. 22–29. DOI: 10.1016/j.rse.2011.09.022.

Lu, D., Q. Chen, G. Wang, L. Liu, G. Li, and E. Moran (2016). "A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems". In: *International Journal of Digital Earth* 9.1, pp. 63–105. DOI: 10.1080/17538947.2014.990526.

Masek, J. G., S. N. Goward, R. E. Kennedy, W. B. Cohen, G. G. Moisen, K. Schleeweis, and C. Huang (2013). "United States forest disturbance trends observed using Landsat time series". In: *Ecosystems* 16.6, pp. 1087–1104. DOI: 10.1007/s10021-013-9669-9.

Maselli, F., G. Chirici, L. Bottai, P. Corona, and M. Marchetti (2005). "Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images". In: *International Journal of Remote Sensing* 26.17, pp. 3781–3796. DOI: 10.1080/01431160500166433.

McRoberts, R. E. (2012). "Estimating forest attribute parameters for small areas using nearest neighbors techniques". In: *Forest Ecology and Management* 272, pp. 3–12. DOI: 10.1016/j.foreco.2011.06.039.

Miller, M. D. (2012). "The impacts of Atlanta's urban sprawl on forest cover and fragmentation". In: *Applied Geography* 34, pp. 171–179. DOI: 10.1016/j.apgeog.2011.11.010.

Multi-Resolution Land Characteristics Consortium (2019). *NLCD 2016 Land Cover*. URL: https://www.mrlc.gov/data/nlcd-2016-land-cover-conus (visited on 10/01/2019).

Natural Resources Spatial Analysis Lab (2003). *Georgia GAP Analysis*. University of Georgia College of Engineerings. URL: http://narsal.uga.edu/ (visited on 04/14/2020).

Nilsson, U. and H. L. Allen (2003). "Short- and long-term effects of site preparation, fertilization and vegetation control on growth and stand development of planted loblolly

pine". In: *Forest Ecology and Management* 175.1, pp. 367–377. DOI: 10.1016/S0378-1127(02)00140-8.

Obata, S., C. J. Cieszewski, P. Bettinger, R. C. Lowe III, and S. Bernardes (2019). "Preliminary analysis of forest stand disturbances in Coastal Georgia (USA) using Landsat time series stacked imagery". In: *FORMATH* 18, pp. 1–11. DOI: 10.15684/formath.001.

Olofsson, P., G. M. Foody, S. V. Stehman, and C. E. Woodcock (2013). "Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation". In: *Remote Sensing of Environment* 129, pp. 122–131. DOI: 10.1016/j.rse.2012.10.031.

Omernik, J. M. and G. E. Griffith (2014). "Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework". In: *Environmental Management* 54.6, pp. 1249–1266. DOI: 10.1007/s00267-014-0364-1.

Oswalt, S. N., W. B. Smith, P. D. Miles, and S. A. Pugh (2018). *Forest resources of the United States, 2017*. WO-GTR-91. Washington, DC: U.S. Department of Agriculture, Forest Service. DOI: 10.2737/WO-GTR-91.

Pflugmacher, D., W. B. Cohen, and R. E. Kennedy (2012). "Using Landsat-derived disturbance history (1972–2010) to predict current forest structure". In: *Remote Sensing of Environment* 122, pp. 146–165. DOI: 10.1016/j.rse.2011.09.025.

Pickell, P. D., T. Hermosilla, N. C. Coops, J. G. Masek, S. Franks, and C. Huang (2014). "Monitoring anthropogenic disturbance trends in an industrialized boreal forest with Landsat time series". In: *Remote Sensing Letters* 5.9, pp. 783–792. DOI: 10.1080/2150704X.2014.967881.

Prud'homme, B. A. and J. G. Greis (2002). "Best management practices in the South". In: *Southern forest resource assessment*. Asheville, NC: Department of Agriculture, Forest Service, Southern Research Station, pp. 519–536.

Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.e., R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y.

Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R. H. Wynne, and Z. Zhu (2014). "Landsat-8: Science and product vision for terrestrial global change research". In: *Remote Sensing of Environment* 145, pp. 154–172. DOI: 10.1016/j.rse.2014.02.001.

Seidl, R., M. Schelhaas, W. Rammer, and P. J. Verkerk (2014). "Increasing forest disturbances in Europe and their impact on carbon storage". In: *Nature Climate Change* 4.9, pp. 806–810. DOI: 10.1038/nclimate2318.

Shelestov, A., M. Lavreniuk, N. Kussul, A. Novikov, and S. Skakun (2017). "Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping". In: *Frontiers in Earth Science* 5. DOI: 10.3389/feart.2017.00017.

Tanaka, S., T. Takahashi, T. Nishizono, F. Kitahara, H. Saito, T. Iehara, E. Kodani, and Y. Awaya (2014). "Stand volume estimation using the k-NN technique combined with forest inventory data, satellite image data and additional feature variables". In: *Remote Sensing* 7.1, pp. 378–394. DOI: 10.3390/rs70100378.

Tomppo, E., M. Haakana, M. Katila, and J. Peräsaari (2008). *Multi-source national forest inventory: Methods and applications*. Springer Netherlands. 373 pp. DOI: 10.1007/978-1-4020-8713-4.

Turner, M. G. (1987). "Land use changes and net primary production in the Georgia, USA, landscape: 1935–1982". In: *Environmental Management* 11.2, pp. 237–247. DOI: 10.1007/BF01867202.

US Census Bureau (2019). *Metropolitan and micropolitan statistical areas totals: 2010-2018*. URL: https://www.census.gov/data/tables/time-series/demo/popest/2010s-total-metro-and-micro-statistical-areas.html (visited on 08/20/2019).

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam (2006). "Warming and earlier spring increase Western U.S. forest wildfire activity". In: *Science* 313.5789, pp. 940–943. DOI: 10.1126/science.1128834.

Williams, C. A., H. Gu, R. MacLean, J. G. Masek, and G. J. Collatz (2016). "Disturbance and the carbon balance of US forests: A quantitative review of impacts from harvests, fires, insects, and droughts". In: *Global and Planetary Change* 143, pp. 66–80. DOI: 10.1016/j.gloplacha.2016.06.002.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne (2008). "Free access to Landsat imagery". In: *Science* 320.5879, pp. 1011–1011. DOI: 10.1126/science.320.5879.1011a.

Wulder, M. A., J. C. White, T. R. Loveland, C. E. Woodcock, A. S. Belward, W. B. Cohen, E. A. Fosnight, J. Shaw, J. G. Masek, and D. P. Roy (2016). "The global Landsat archive: Status, consolidation, and direction". In: *Remote Sensing of Environment* 185, pp. 271–283. DOI: 10.1016/j.rse.2015.11.032.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian (2018). "A new generation of the United States national land cover database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146, pp. 108–123. DOI: 10.1016/j.isprsjprs.2018.09.006.

Zampieri, N. E., S. Pau, and D. K. Okamoto (2019). "The impact of Hurricane Michael on longleaf pine habitats in Florida". In: *bioRxiv*, p. 736629. DOI: 10.1101/736629.

Zhao, P., D. Lu, G. Wang, C. Wu, Y. Huang, and S. Yu (2016). "Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation". In: *Remote Sensing* 8.6, p. 469. DOI: 10.3390/rs8060469.

Zhu, Z., Y. Fu, C. E. Woodcock, P. Olofsson, J. E. Vogelmann, C. Holden, M. Wang, S. Dai, and Y. Yu (2016). "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014)". In: *Remote Sensing of Environment* 185, pp. 243–257. DOI: 10.1016/j.rse.2016.03.036.

Zhu, Z. (2017). "Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 130, pp. 370–384. DOI: 10.1016/j.isprsjprs.2017.06.013.

Zhu, Z., S. Wang, and C. E. Woodcock (2015a). "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images". In: *Remote Sensing of Environment* 159, pp. 269–277. DOI: 10.1016/j.rse.2014.12.014.

Zhu, Z. and C. E. Woodcock (2012a). "Object-based cloud and cloud shadow detection in Landsat imagery". In: *Remote Sensing of Environment* 118, pp. 83–94. DOI: 10.1016/j.rse.2011.10.028.

Zhu, Z., C. E. Woodcock, and P. Olofsson (2012b). "Continuous monitoring of forest disturbance using all available Landsat imagery". In: *Remote Sensing of Environment* 122, pp. 75–91. DOI: 10.1016/j.rse.2011.10.030.

Zhu, Z., J. Zhang, Z. Yang, A. H. Aljaddani, W. B. Cohen, S. Qiu, and C. Zhou (2020). "Continuous monitoring of land disturbance based on Landsat time series". In: *Remote Sensing of Environment* 238, p. 111116. DOI: 10.1016/j.rse.2019.03.009.

Chapter 4

Estimation of forest growing stock stand volumes based on Harmonic regression with ensemble method and 1984-2016 Landsat 5 and 7 imagery [1]

[1] Shingo Obata, Chris J. Cieszewski, Pete Bettinger, and Roger C. Lowe III, To be submitted to Forest Ecology and Management

Abstract

Forest density and volume are some of the most critical inventory estimates used in forest management activity. Since their inception in the 1990s, the development of satellite imagery-based imputation methods has allowed for the estimations of forest inventories on the national-scale. The first inventory created using such a method was developed in Finland in 1990, and since then, the method has been applied in various other regions. In the most recent decade, the random forest (RF) algorithm has gained considerable attention. The RF algorithm can reduce the variance of regression predictors; even though, it does not reduce resulting biases, which can cause systematic underestimation or overestimation of predicted values. We have built an RF regression model that estimates the growing stock volume of the forests in the state of Georgia, United States, using dense Landsat TM/ETM imagery from 1984-2016 and subsequently applied a bias correction method to the primary estimation. The goal of this research was to examine the effectiveness of using dense Landsat TM/ETM imagery and the application of the bias correction method on the accuracy of the growing stock volume estimation. We used dense Landsat satellite imagery from 1984-2016 and computed three Tasseled-cap components were computed. Both short-time (2014-2016) and long-time range (1984-2016) time series were prepared using the Landsat data, and harmonic regressions were subsequently performed on each data entity. Afterward, the coefficients and the fitted values of the harmonic regression were recorded as pixel values in a multi-layer raster. The data from the Forest Inventory and Analysis surveyed in 2016 was used for the field inventory data, while the National Land Cover Database and disturbance history data were used as ancillary databases. The totality of the available data was organized into seven distinct random forest models differing in their variables. The models were subsequently compared in order to ascertain the most satisfactory model. A bias correction method was then applied to all the RF models to examine the effect of the method. Among seven models, the worst model was the model using the coefficients and fitted values of the short-time range Landsat imagery only and the best model was the model using coefficients and fitted values

of both short-term and long-term range Landsat imagery. The best model was 35% better than the worst model, rRMSE, Out-Of-Bag (OOB) score and bias of the top 10% volume class of the best model was better than the worst model by 10%, 35%, and 12%, respectively. The bias correction method reduced the percent bias from 3.79% to -1.47%. The bias correction method also reduced the bottom 10% bias by 18% and the top 10% bias by 7%. The types of the feature variables were varied depending on the species used for individual RF model.
Keyword; Forest sustainability, forest disturbance, Landsat imagery, Google Earth Engine

4.1 Introduction

4.1.1 Importance of the growing stock volume

Forest density and volume are the primary measures of inventory needed in forest management planning. Timber volume is directly related to the economic benefits of the forest product industry (Bettinger et al., 2017). Forest density is also important for various ecosystem functions and wildlife habitats. For instance, the maximum basal area suitable for the habitat of the red-cockaded woodpecker (*Picoides borealis*) is suggested to be 18.4 m^2/ha (U.S. Department of the Interior, Fish and Wildlife Service, 2003). Moreover, timber volume and density are directly related to the carbon stock, which is a critical issue in sustainability considerations (Miksys et al., 2007).

In the Southeastern USA, both private and public organizations managed the forest inventories and their assessments. The U.S. Forest Service's Forest Inventory and Analysis (FIA) Unit provides data for large-scale forest inventory as a publicly available forest survey based on annual measurements. The main objective of the FIA is to determine the extent, condition, volume, and growth of forest land (Brandeis et al., 2012). The inventory splits the conterminous United States into 28,000 constituent hexagons, each with their centers

approximately 17 miles apart. The center of each constituent hexagon subsequently serves as the field survey point. Each established FIA plot represents about 6000 acres of land area. In addition to the central point, three other sample features are located around the central point of the hexagonal grid, each consisting of a subplot and macroplot (Figure 4.1). All the trees within a plot of whose DBH is greater than 5 inches are the subjects of the field measurement (Bechtold et al., 2015). Although this data provides a good estimation of volume on the state-level, they are not suitable for sub-county level estimations of the stand volumes. Some public and private landowners conduct their separate inventory assessments using ground measurements and forest information systems and associated field data. Although these inventories may provide high-resolution volume estimations, the systems are often spatially restricted to the property boundaries of the landowners and as privately owned information they are generally not available to the public.

4.1.2 Satellite-based volume estimation

Satellite-based forest volume estimation was first developed since the 1990s for the national-scale forest inventories. This approach was using multi-source methods to combine field

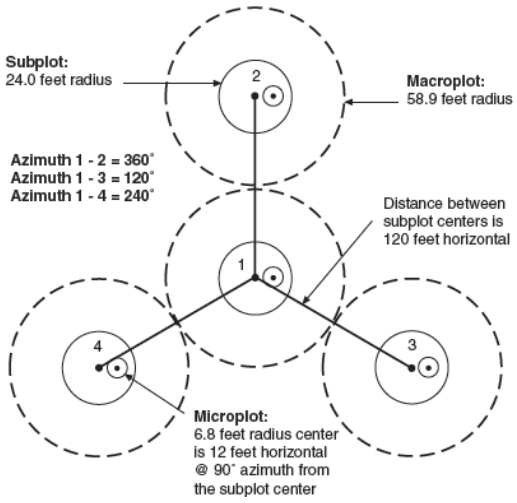


Figure 4.1: Sampling plot design of FIA Bechtold et al., 2015.

inventory data with satellite or airborne sensor measurements contained in the imagery. Statistical models were applied to estimate the volume for each pixel in the raster database represented by each image. Landsat imagery is the most commonly used satellite imagery for this approach, which uses kNN methods for propagating estimation of pixels having spatially corresponding to them ground measurements to other pixels without such additional ground information. With the combination of FIA and Landsat TM imagery, it is possible to develop a distribution of forest resources at a pixel level spatial resolution. The Finish inventory uses 25 m pixels, while in the USA the most common is the use of 30 m pixels). The first operative results based on this kind of approach were developed in the year 1990 for Finland (Tomppo, 1990). The product created through this work was a 30-meter resolution raster database that contained the volume of growing stock. Following the Finnish example, similar inventory data were derived for Sweden (Reese et al., 2002; Reese et al., 2003). After these successful applications of the kNN methods to forest inventories, similar approaches were further applied to many regions and national forest inventories, such as those seen in the USA, Norway, Ireland, and Japan (Franco-Lopez et al., 2001; McRoberts et al., 2002; Maselli et al., 2005; Tanaka et al., 2014; Barrett et al., 2016).

Along with the development of multi-source methods in general, the satellite imagery employed has also evolved. The use of satellite imagery in the applications of the kNN methods has also evolved over time. In the earlier research involving multi-source methods, initially satellite imagery was selected from a single date, leading to potential problems. For instance, To overcome the problems caused by cloud coverage, some of the research analysts created a composite image spanning over a year (Moody et al., 2001; Hird et al., 2009). Over time, the number of satellite images used in research has increased since Landsat imagery became available for free in October 2008 (**roy˙landsat-8:˙2014**; Woodcock et al., 2008). To take advantage of the unlimited and gratis availability of the Landsat archives, some research began to have employed the use of dense Landsat imagery for the estimations of the land use changes, and to detect disturbances (Kennedy et al., 2012; Zhu et al., 2014; Brooks et al.,

2014). In this type of research, multiple images taken within the established spatial and temporal boundaries were employed collectively to construct a time series datasets allowing tracking stand changes over time with the pixel spatial resolution. The time series datasets are usually reduced to trends, seasonal changes, and noise components, prior to the analysis of the land use changes itself. Thus, the raw pixel values of satellite imagery are regarded to be a quasi-systematic sample of the land surface (Brooks et al., 2014). The derivatives of the raw pixel values are then used as inputs of the land use change analysis. Although the considerable computational power necessary to analyze dense satellite imagery has made it more difficult to perform these types of analyses, the rise of cloud-based platforms provides sufficient computation ability for large-scale geospatial analysis. Google Earth Engine (GEE) serves as one of the most common platforms for the implementation of large-scale geospatial analyses (Gorelick et al., 2017).

For instance, using the GEE platform, Wilson et al., 2018 performed a harmonic regression on dense Landsat imagery. The coefficients of the regression were used as the feature variables of the non-parametric modeling approaches. Using all of the available time series Landsat imagery, it was found that the estimations of the continuous variables related to forest stands (i.e., basal area, number of trees and foliage biomass) showed a two- to three-fold increase in the explained variance, relative to the estimations found using per monthly image composites. Although Wilson et al., 2018 thoroughly examined the advantage of using relatively short time range Landsat imagery (2013-2016), the advantage of taking the long-term Landsat imagery was not inspected. Additionally, Wilson et al., 2018 used the coefficients of harmonic regression; the fitted values of the harmonic regression were not examined. Furthermore, the effect of the bias correction method when it is applied to the forest inventory data is not examined.

4.1.3 Volume estimation with random forest methods

Random Forest (RF) is an algorithm that handles large volumes of data within a relatively short computation time (Breiman, 2001). The RF regression initially begins by building many simple classification trees during a training period. Afterwards, the final predictor generated by the model is considered the average of the predictions made by the individual classification trees (Liaw et al., 2002). The bagging and CART-split criterion are the two main components of the RF algorithm. The bagging creates bootstrap samples from the original data set, before using these samples for the individual decision trees. The CART-split criterion is employed in the process of the construction of each tree to determine the best cut of the data set into the branches. For the regression model, the CART-split criterion is based on the prediction squared error. The RF is generally recognized as a suitable method for making predictions using small data sets with high-dimensional feature spaces. This characteristic makes the RF algorithm the preferable method for estimations of the forest stand growing stock volume, as the size of the field inventory data is relatively small and the variety of variables, such as remote sensing data, geographical data, and environmental data, can be high. This supervised learning procedure works based on the “divide and conquer” principle. RF regression is widely used for making data predictions, including forest attribute estimation (Wilson et al., 2018; Gigović et al., 2019; Tompalski et al., 2019). One of the primary advantages of using an RF model is that it can determine the importance of a variable, which indicates the contribution of each feature variable to the model prediction. The mean reduction in prediction accuracy evaluates the importance.

One of the known issues of RF involves bias. It is argued that bagging could diminish the extent of the variance of regression predictors, yet it does not reduce the magnitude of the bias (Breiman, 1996). On the other hand, because extreme observations are estimated using the average of the estimation of each tree, large observations close to the maximum value within the data set are underestimated, and small values of the regression function are overestimated (Zhang et al., 2012). When a data set is imbalanced, estimations using the

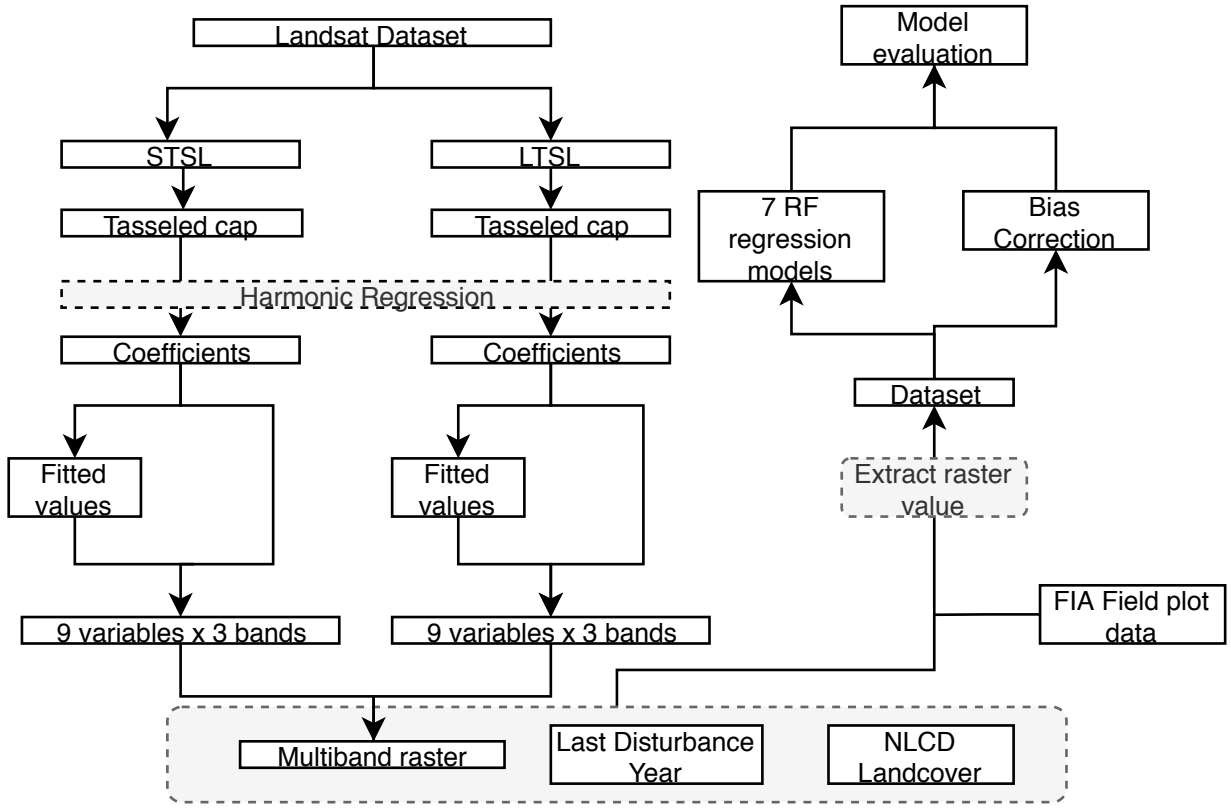


Figure 4.2: Flow chart of the research.

RF algorithm are more susceptible to the risk of bias (C. Chen et al., 2004). Zhang et al., 2012 propose a method to correct the bias in RF.

4.2 Methods

In this research, we built a model that estimates the growing stock volume of the forest in the state of Georgia, United States, using an RF regression. The goal of this research was to determine contributing factors in a model that either improve the model's performance, or reduce the magnitude of the overestimation of the large volume class and magnitude of the underestimation of the small volume class. We specifically focused on the Out-of-bag (OOB) score, the relative RMSE (rRMSE), and the relative bias (rB) to evaluate the model performance. To achieve this goal, we examine the following questions throughout this paper:

1. Does applying different time ranges to the available Landsat imagery data set contribute to a better estimation? If so, which time range options are the most useful?

2. Which feature variables are important? Why are they important?

3. Does using the last disturbance year data as one of the feature variables of the RF regression improve the estimation accuracy?

4. Does the bias correction method for RF regression decrease the size of the bias of the estimation? How is the rRMSE affected by this correction?

To address question 1 and question 2, two types of time series Landsat datasets were prepared, with the Landsat imagery in each dataset originating from a distinct time range. Subsequently, the time series data was transformed into the Fourier series via harmonic regression. From each series, we retrieved the key values that were used as the feature variables of the RF regression and created a multi-layer raster in which the pixel values represent the key values. Publicly available field inventory data was combined with the multi-layer raster data to create the tabular data. Using the tabular data, we performed an RF regression using various combinations of the feature variables to determine the best combination of variables, and to examine the importance of individual features. To address our third question, we added two ancillary data to the combination of the feature variables derived from using the time series Landsat data. The first data is the raster data in which the pixel values represent the last disturbance year of the forest stands since 1987. The other mentioned ancillary data is land cover data. Then, we examined the magnitude of the contribution of these two ancillary databases to the estimation. Next, we applied the bias correction method for RF regression to analyze how much bias is reduced. We also examined the effect of the bias correction model as it was applied to the best RF model among the models we built, in order to answer our fourth question. Finally, we discussed the result of the work and considered the potential reasons for why each of the situations contributed, or did not contribute to the improvement of the estimation.

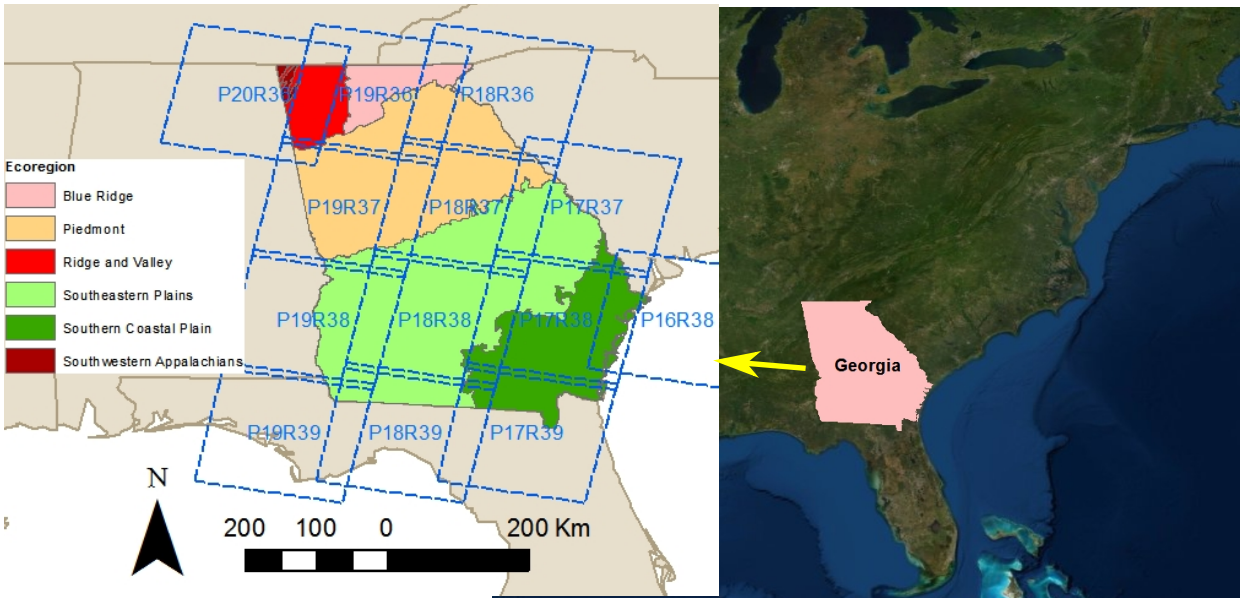


Figure 4.3: Study area and its ecoregion.

4.2.1 Study area

Our study area is the state of Georgia, United States; Georgia is located in the southeastern region of the United States and contains approximately 15,390,000 ha of land area (Figure 4.3). In the Southern Coastal Plain plantation forests are often intensively managed. The main plantation species is loblolly pine (*Pinus taeda*), which has a normal rotation age of 20-25 years (Fox et al., 2007; D'Amato et al., 2017). In the Appalachian mountain area, most of the forests are hardwood or mixed forest and are less frequently disturbed (Obata et al., 2019).

4.2.2 Satellite data

All of the satellite data in this study were queried and processed using the GEE platform. All of the Landsat data was selected from the Level-1 Precision Terrain corrected product (L1TP), for 13 path/row combinations as shown in Figure 4.3. The L1TP satisfies both radiometric and geometric criteria set by the United States Geological Survey (USGS) (U.S. Geological Survey, 2020). From the L1TP collection, we selected Landsat 5 TM and Land-

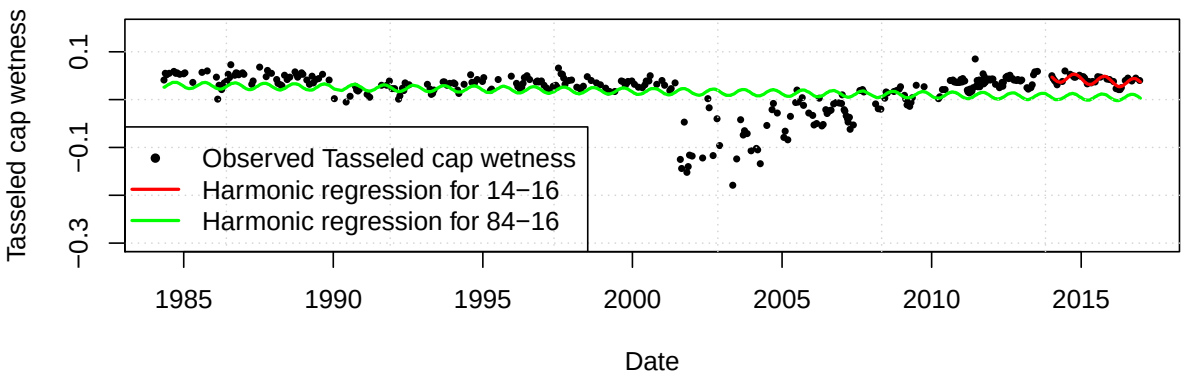


Figure 4.4: Harmonic regression on TCW of the time series Landsat imagery for a decidous forest (Lon. -82.052, Lat. 31.8389).

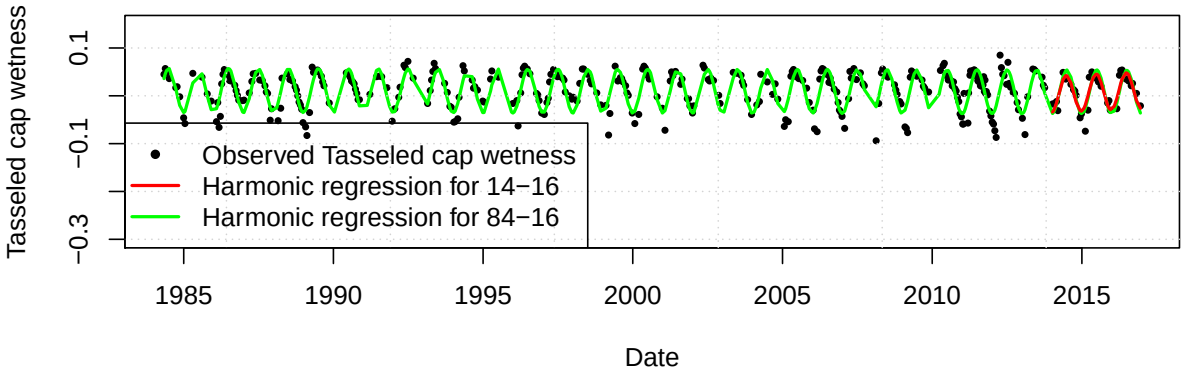


Figure 4.5: Harmonic regression on TCW of the time series Landsat imagery for an evergreen forest (Lon.-81.790, Lat.31.063).

sat 7 ETM+ Surface Reflectance data, which was generated using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm (Masek et al., 2006). Two different time ranges were arranged in order to create distinctive time series Landsat data. The short time period was limited to 3 years and ranged from the beginning of 2014 to the end of 2016. The long-time range was set to 33 years, spanning from 1984 to 2016. All

of the available images were queried for each time range. For each set of images in both time ranges, clouds, cloud shadows, water, and snow interference were masked out using the C Function of Mask (CFMask) algorithm (Zhu et al., 2012a; Zhu et al., 2015a; Foga et al., 2017). For convenience, we refer to the time series Landsat data that encompasses the long time range as the "Long Time Series Landsat" (LTSL) data set. Similarly, we call the time series Landsat data that encompasses the short time range as the "Short Time Series Landsat" (STSL) data set. Additionally, we computed the Tasseled Cap Brightness (TCB), Tasseled Cap Greenness (TCG) and Tasseled Cap Wetness (TCW) using surface reflectance data. The coefficients calculated in Crist, 1985 were applied to compute the TCB, TCG, and TCW. Subsequently, these values were inputted into a multiband time series raster. Additionally, an ordinary least squares, harmonic regression was performed to fit the Fourier series to each Tasseled Cap band for both the short and long data sets. The form of the Fourier curve is derived from (Shumway et al., 2017) and is as follows:

$$\hat{Y}_t = \beta_0 + \beta_1 t + \beta_2 \cos(2\pi\omega t) + \beta_3 \sin(2\pi\omega t) \tag{4.1}$$

where, $\hat{Y}_t$: Fitted value for the imagery taken at t, β_0: Intercept, β_1: Slope, β_2: Cosine term coefficient, β_3: Sine term coefficient.

We fixed $\omega = 1$ so that the Fourier curve has a single-cycle in a year, although there is previous research that fits the Fourier curve to the time series satellite imagery employing the multiple cycles in a year (Zhu et al., 2016). All four coefficients of the equation (1) are taken as the raster values. The amplitude is computed as follows:

$$Amplitude = \sqrt{\beta_2^2 + \beta_3^2} \tag{4.2}$$

The impact of the amplitude is on the height of the wave. Fitted values were computed for all the dates for which the time series Landsat imagery was acquired and for each pixel. Next, we calculated the maximum, minimum, mean and RMSE from both the fitted values

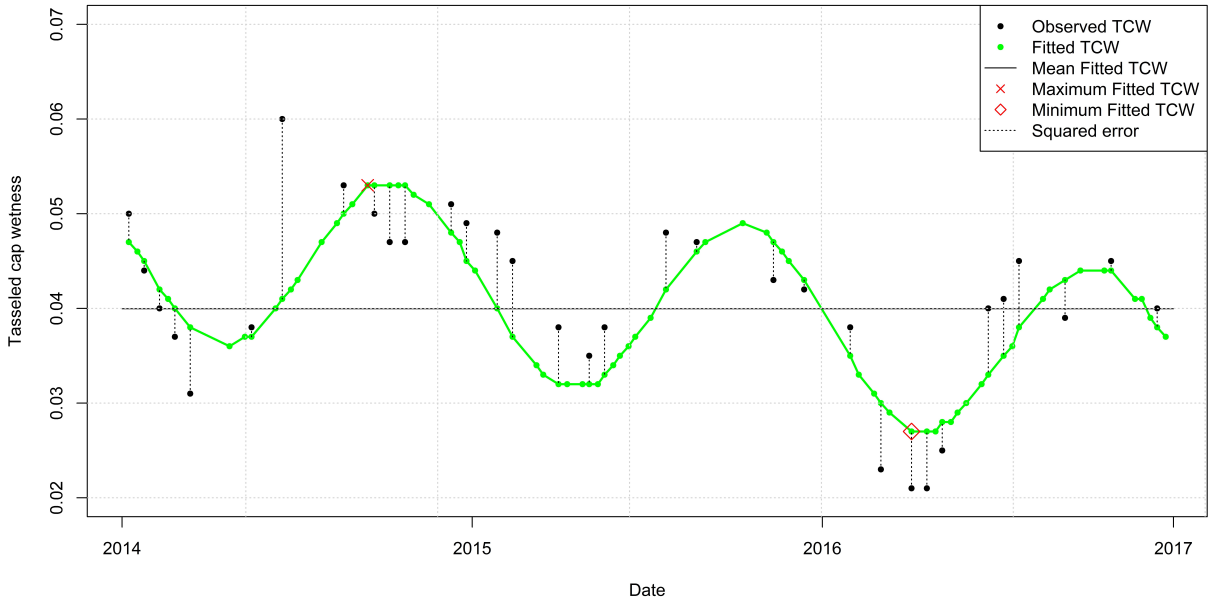


Figure 4.6: Harmonic regression on TCW of the time series Landsat imagery of an evergreen forest (Longitude: -81.789827,31.062906).

and the observed values (Figure 4.6). Consequently, 9-band imagery was created for each band by stacking all of the metrics. Then, each of the bands generated from the STSL and the LTSL were compiled to create the single raster layer used for the subsequent analysis.

4.2.3 Ancillary databases

In addition to the remote sensing data, we selected two ancillary databases, which were expected to help improve the RF prediction. The first of these is the 2016 National Land Cover Database (NLCD) Land Cover products for the conterminous U.S. (Yang et al., 2018). The 2016 NLCD was created by the Multi-Resolution Land Characteristics consortium to provide a consistent multitemporal land cover and land cover change map for the conterminous U.S. at 30-meter spatial resolution. The 2016 NLCD classifies the land into 16 classes. Out of these 16 classes, the land where shrubs or trees cover more than 20% of the area is

Table 4.1: NLCD 2016 Land Cover Class on the FIA field plots.

Land cover class	Class[1]	# of plots	Mean volume (m^3/ha)	# of plots disturbed[2]
Water	0	4	220.61	2
Developed	0	43	330.14	6
Barren land	0	2	146.99	1
Deciduous forest[3,4]	2	191	433.5	21
Evergreen forest[3,4]	1	274	431.61	80
Mixed forest[3,4]	2	75	416.8	10
Shrubland	0	32	138.43	16
Herbaceous	0	28	148.46	19
Planted/Cultivated	0	51	132.47	2
Woody wetlands[3]	2	185	462.28	32

[1] 0:Non-forest. 1: Evergreen Forest. 2: Non-Evergreen. [2] Disturbance record for each plot was retrieved from (**obata˙mapping˙2020**).
[3] Areas where forest or shrubland vegetation accounts for greater than 20% of vegetative cover.
[4] Areas dominated by trees generally greater than 5 meters tall.

classified either as deciduous forest, evergreen forest, mixed forest, or woody wetlands. We note that more than 20% of the FIA field inventory data are positioned on locations where the land cover class is not forested (Table 4.1). We included all of the field inventory data that is classified as non-forest in a later analysis, as the NLCD misclassifies some of the forest pixels as a non-forest class. The second ancillary database adopted in this research was the Last Disturbance Year map of Georgia. This map detects the last disturbance that occurred between 1984-2016 for every land area in the entire state of Georgia at 30-meter spatial resolution (**obata˙mapping˙2020**). Regardless of the current land use, a pixel without any disturbance record between 1984-2016 is classified as undisturbed.

4.2.4 Growing stock estimation

FIA data was selected as the field inventory. Satellite data and ancillary data were stacked into a multilayer raster that contained 56 bands (54 bands from satellite data and 2 bands from ancillary data). To integrate the raster and FIA field inventory data, we requested the USDA Forest Service to extract the pixel values of the raster data onto the field plot

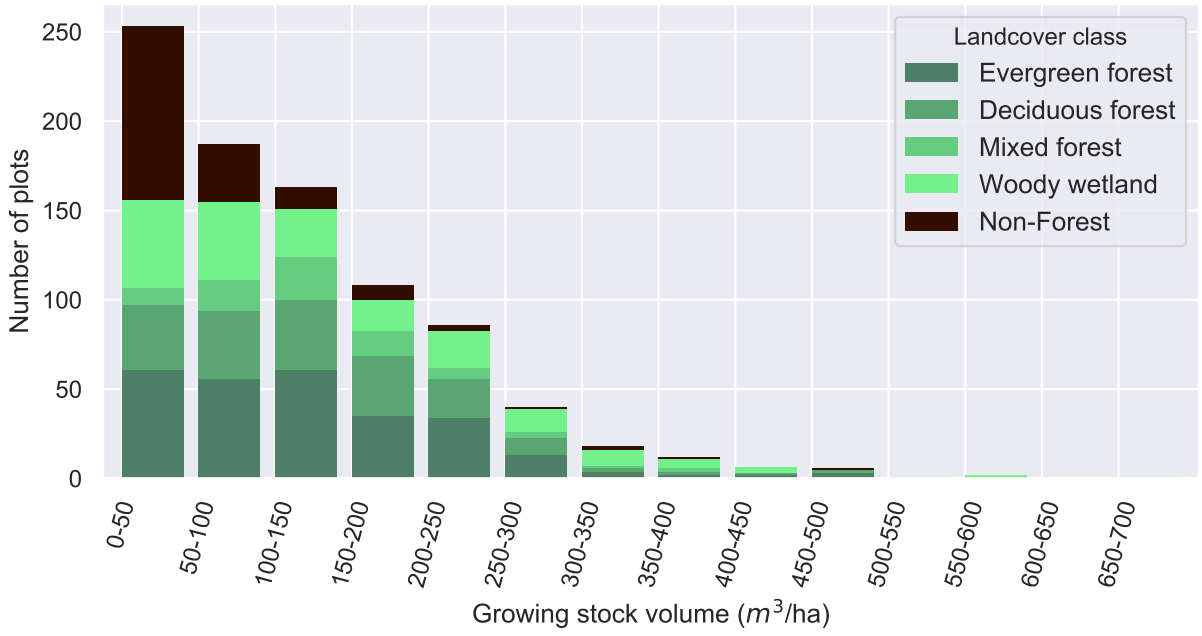


Figure 4.7: Flow chart of the research.

points; They extracted our raster data onto the point data to provide us with tabular data containing plot ID number and the pixel values of our raster data set. We note that all information potentially allowing the data user to detect the exact coordinate of the plot data was removed by USDA Forest Service in compliance with the Privacy Act in 1974 (Smith, 2002); Thus, we do not know the exact locations of the plots. The tabular data was aggregated with the FIA's original database, available from the FIA DataMart (`https://apps.fs.usda.gov/fia/datamart/datamart.html`). For each plot ID, all of the individual tree measurement data were available. Although individual tree measurements are available from the four subplots shown in Figure 4.1, only the data from the central subplot of a plot was used for the volume calculation, in order to avert the problem of spatial correlation among subplot observations (McRoberts et al., 2002). Based on the code found in Burrill et al., 2018, individual tree data were aggregated into the plot-level growing stock volume per acre as follows:

$$V_i = \Big(\sum_{j=1}^{m_i} v_{ij}\Big) \cdot k \tag{4.3}$$

Where,

V_i: Per hectare growing stock volume of plot i.

v_{ij}: Net m^3 volume of jth tree in plot i equivalent to the net volume of wood in the central stem.

m_i: The number of trees in plot i.

k: Expansion factor to convert the total growing stock volume of the plot to per hectare growing stock volume.

For the RF regression, the data set was split into the dependent variable, which is growing stock volume per hectare, and the independent variables, which are all derived from Landsat imagery and ancillary data. We calculated the rRMSE, the relative bias (rB), and OOB score. RMSE has been used as the primary determinant of the model performance (Chirici et al., 2016). As the absolute value of the RMSE is incomparable between research conducted in the different areas of study, the rRMSE, calculated by the following formula, is used in favor of the RMSE.

$$rRMSE = \frac{RMSE}{\hat{y}} * 100 \tag{4.4}$$

The bias is required to know the direction of the error. Subsequently, rB is calculated in the same way as rRMSE, They are formulated as follows.

$$Bias = \frac{\sum_{n_i=1} (y_i - \hat{y}_i)}{n} \tag{4.5}$$

$$rB = \frac{Bias}{\hat{y}} * 100 \tag{4.6}$$

In addition to the evaluation of the entire data set, we focused on the smallest and largest volume group as it is known that the error of the non-parametric estimation of the volume is usually heteroskedastic that the size of the error is not equally distributed (McRoberts, 2009). We grouped the field inventory data into deciles based on the observed growing stock volume.

The bottom 10% ranged between 0.1 m^3/ha to 12.9 m^3/ha, while the top 10% group ranged between 249.1 m^3/ha to 682.1 m^3/ha.We calculated the bias for the bottom 10%, middle 80%, and top 10%, separately. The RF regressor was trained using the training dataset. *Scikit-learn*, a Python module that provides machine learning algorithms for medium-scale supervised and unsupervised problems, was used to perform the RF regression (Pedregosa et al., 2011). The number of decision trees created in the RF algorithm was set to 500. The mean squared error was selected as the function to measure the quality of a split in the individual decision trees. Individual decision trees were trained by the data bootstrapped from the original training data set.

First, we built a base model that used the coefficients of the harmonic regression on the STSL, LTSL and the last disturbance year record as the feature variables (C_{SL} in Table 4.2).

Table 4.2: Summary of the feature variables.

	Vegetation index /Data source	Time Range	# of Variables	Values	RF Models						
					C_{SL}	F_{SL}	CF_S	CF_L	CF_{SL}	E_{SL}	NE_{SL}
Features	Landsat TCB	1984-2016	4	Regression coefficients		✓			✓	✓	✓
		1984-2016	5	Fitted values				✓	✓	✓	✓
		2014-2016	4	Regression coefficients	✓	✓			✓	✓	✓
		2014-2016	5	Fitted values			✓	✓	✓	✓	✓
	Landsat TCG	1984-2016	4	Regression coefficients		✓			✓	✓	✓
		1984-2016	5	Fitted values				✓	✓	✓	✓
		2014-2016	4	Regression coefficients	✓	✓			✓	✓	✓
		2014-2016	5	Fitted values			✓	✓	✓	✓	✓
	Landsat TCW	1984-2016	4	Regression coefficients		✓			✓	✓	✓
		1984-2016	5	Fitted values				✓	✓	✓	✓
		2014-2016	4	Regression coefficients	✓	✓			✓	✓	✓
		2014-2016	5	Fitted values			✓	✓	✓	✓	✓
	NLCD	2016	1	Land use class						✓	✓
	Last disturbance	1984-2016	1	Disturbance year	✓	✓	✓	✓	✓	✓	✓
Response	FIA data	2016	1	Growing stock volume	✓	✓	✓	✓	✓	✓	✓
			57 variables		14	26	17	32	56	57	57

Next, the fitted values of the STSL, the LTSL, and the last disturbance year data were selected as the feature variables of the base model (F_{SL} in Table 4.2). In the third model, both the fitted values and coefficients of the STSL and the last disturbance year data were selected (CF_S in Table 4.2). To make a comparison with the third model, the fourth model takes both the fitted values and coefficients of the LTSL, as well as the last disturbance year data were selected (CF_L in Table 4.2). CF_{SL}, meanwhile, used all of the feature variables from the previous models (CF_{SL} in Table 4.2). After determining the best combination of the variables derived from remote sensing data, we split the data by the forest type, as defined in the 2016 NLCD. The first group contained only the evergreen forest and was denoted as the Evergreen data set. The second group contains the rest of the forest groups listed in Table 4.1 and was denoted as the Non-Evergreen data set. Then, the RF model was trained and evaluated separately (E_{SL} and NE_{SL} in Table 4.2). Following this, predictions for each data were aggregated to compute the OOB score and rRMSE as the eighth model ($E_{SL} + NE_{SL}$ in Table 4.3).

One of the bias correction methods proposed in (Zhang et al., 2012) was applied to each model to reduce the bias observed in the top and bottom 10% of the volume classes. In the model, we conjectured that bias would be attributed to the response variables. To inspect the effect of the bias correction, the data was split into training and test data, respectively. The ratio of the training to test data was then set to a 2:1 ratio. For the training data, the RF model was created, and the residual of the RF regression (e) was computed as follows:

$$e = Y - \hat{f}(\mathbf{X}) - B(Y) + \epsilon \tag{4.7}$$

Where,

Y: The growing stock volume of the observations in the training data.

$f(\mathbf{X})$: Predicted values of the RF regression using feature variables of the training data.

$B(Y)$: Regression bias.

ϵ: The error term. $\epsilon \sim N(0, \sigma^2)$.

$$\hat{B}(Y) = \alpha + \beta_1 Y^2 + \beta_2 Y \tag{4.8}$$

The bias-corrected prediction ($\hat{f}_{bc}$) was calculated by subtracting the estimated bias.

$$\hat{f}_{bc} = \hat{f} - \hat{B}(Y) \tag{4.9}$$

The effect of the correction was evaluated for the test data.

4.3 Results

Eight models described in the previous section are trained and evaluated (Table 4.3). Among all the models using the field inventory data of all species, the best model was CF_{SL} in terms of the OOB score. In comparison to CF_S, in which the OOB score is the worst amongst them, the OOB score was 35.2% better. The OOB score and rRMSE of F_{SL} were better than C_{SL} by 2.2% and 3.0%, respectively. Between the models that used both coefficients and fitted values, CF_L showed a better result than CF_S. The rRMSE of CF_L was improved by 9.8% and the OOB score was improved by 34%. Figure 4.8 shows the feature importance of the top 10 variables in CF_{SL}. The maximum value of the TCW generated from the STSL had the highest feature importance. E_{SL} and NE_{SL} were trained for the smaller sample sizes, as the data set was split based on the forest type. The evergreen forest had a

Table 4.3: Summary of the RF models.

	C_{SL}	F_{SL}	CF_S	CF_L	CF_{SL}	E_{SL}	NE_{SL}	$E_{SL} + NE_{SL}$
Observation Mean	121.21	121.21	121.21	121.21	121.21	113.39	128.16	-
rRMSE	68.93	67.38	71.48	65.09	64.42	59.66	70.50	65.67
rBias	4.19	3.48	3.72	2.66	3.79	3.09	4.82	-
OOB_score	34.8	35.87	23.39	35.63	36.11	46.52	34	39.15
Species	all	all	all	all	all	Evergreen	non-Evergreen	all

better OOB score than CF_{SL} by 9%, whereas NE_{SL} – which takes field inventory data from non-evergreen samples, returned a lower OOB score than CF_{SL}. The OOB predictions of E_{SL} and NE_{SL} were aggregated to compute the rRMSE and OOB score for the entire data ($E_{SL} + NE_{SL}$ in Table 4.3). The rRMSE for the aggregated prediction was similar to that of CF_{SL}, while the OOB score for the aggregated prediction was worse than that of CF_{SL}. Feature importance of E_{SL} and NE_{SL} were presented in Figure 4.8. For evergreen forests, the six most important features were either the TCW or the TCB of the LTSL. For the rest of the species, the maximum fitted value of the harmonic regression derived from the TCW of the STSL. The second important feature was the maximum fitted value of the harmonic regression on the TCW of the LTSL. In comparison with E_{SL} and NE_{SL}, the maximum fitted values are given greater importance in E_{SL} than in NE_{SL}.

The bias correction method was applied to all the models. The relationship between the observed growing stock volume and the estimated bias for CF_{SL} is illustrated in Figure 4.9. For each RF model, we subtracted the estimated bias from predicted volumes to acquire the bias-corrected prediction. Bias-corrected prediction reduced rB, bottom 10% bias, and top 10% bias from the original prediction in all models (Table 4.5).

4.4 Discussion

The multiple models we examined in the previous section showed different results due to the different data sets and the prospective variables applied to each model. In this section, we discuss the possible impact of several factors on the predictions using RF.

The inclusion of the LTSL into the set of feature variables was effective. As is shown in the comparison between CF_S and CF_L, the inclusion of the LTSL into the set of feature variables contributed to improving the OOB score. Table 4.4 shows how many times a variable was selected as the 10 most important variables in terms of feature importance for CF_{SL}, E_{SL}, and NE_{SL}. The number of features created from the LTSL is more than the STSL

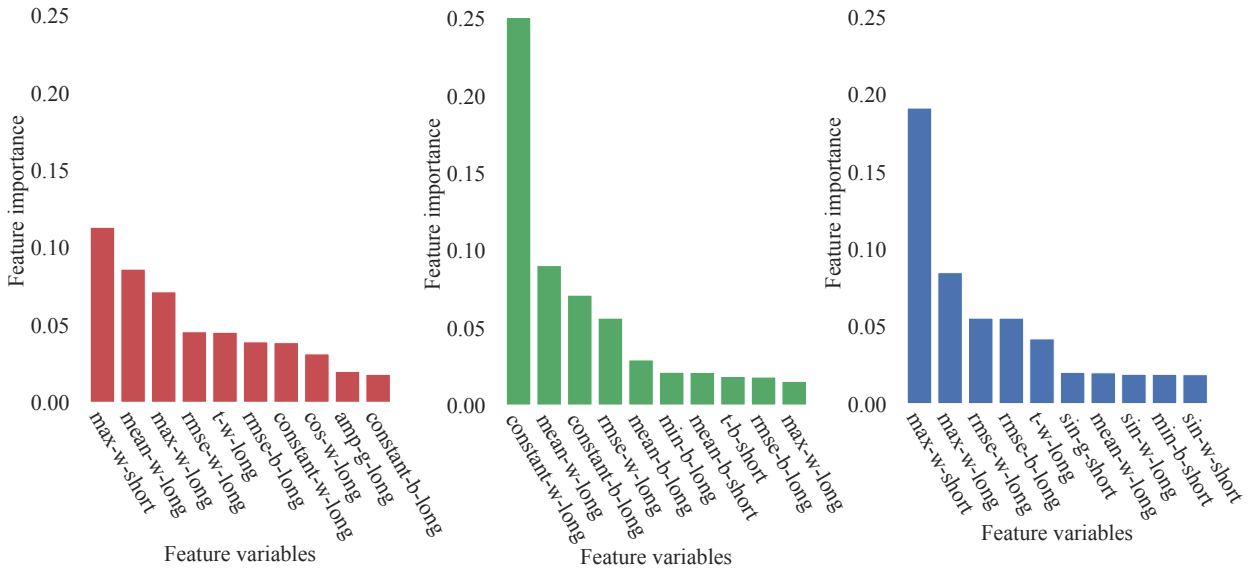


Figure 4.8: Feature importances of three RF models. Left: CF_{SL}, Center: E_{SL}. Right: NE_{SL}. The abbreviated name for the feature variable represents <Type of variable>-<Type of tasseled cap component used>-<Length of the Landsat data used> (i.e. max-w-short represents the maximum value of the fitted value of the harmonic regression on the LTSL of TCW).

in CF_{SL}, E_{SL}, and NE_{SL}. In E_{SL}, features derive from the LTSL were more important than in NE_{SL} model. On the other hand, the STSL maintains a degree of importance for NE_{SL} model. The different effects of the LTSL and STSL on the two models result from the different ratios of the field inventory data with disturbance Table 4.1. While 29% of the field inventory data of evergreen forest has disturbance record, 14% of the field inventory data of the non-Evergreen forest has the disturbance record. Since LTSL convolutes the time series trajectory of Landsat spectral values over long periods of time, features from the LTSL gained importance in E_{SL}, of which field inventory data was taken from the relatively dynamic and young forest. As the STSL captures recent trends more precisely than the LTSL, the STSL gained a degree of importance for NE_{SL}, of which field plot data relate to the relatively stable and mature forest.

Features extracted from fitted values of the regression generally showed more importance than the coefficients of the harmonic regression (Table 4.4) in CF_{SL} model. The importance

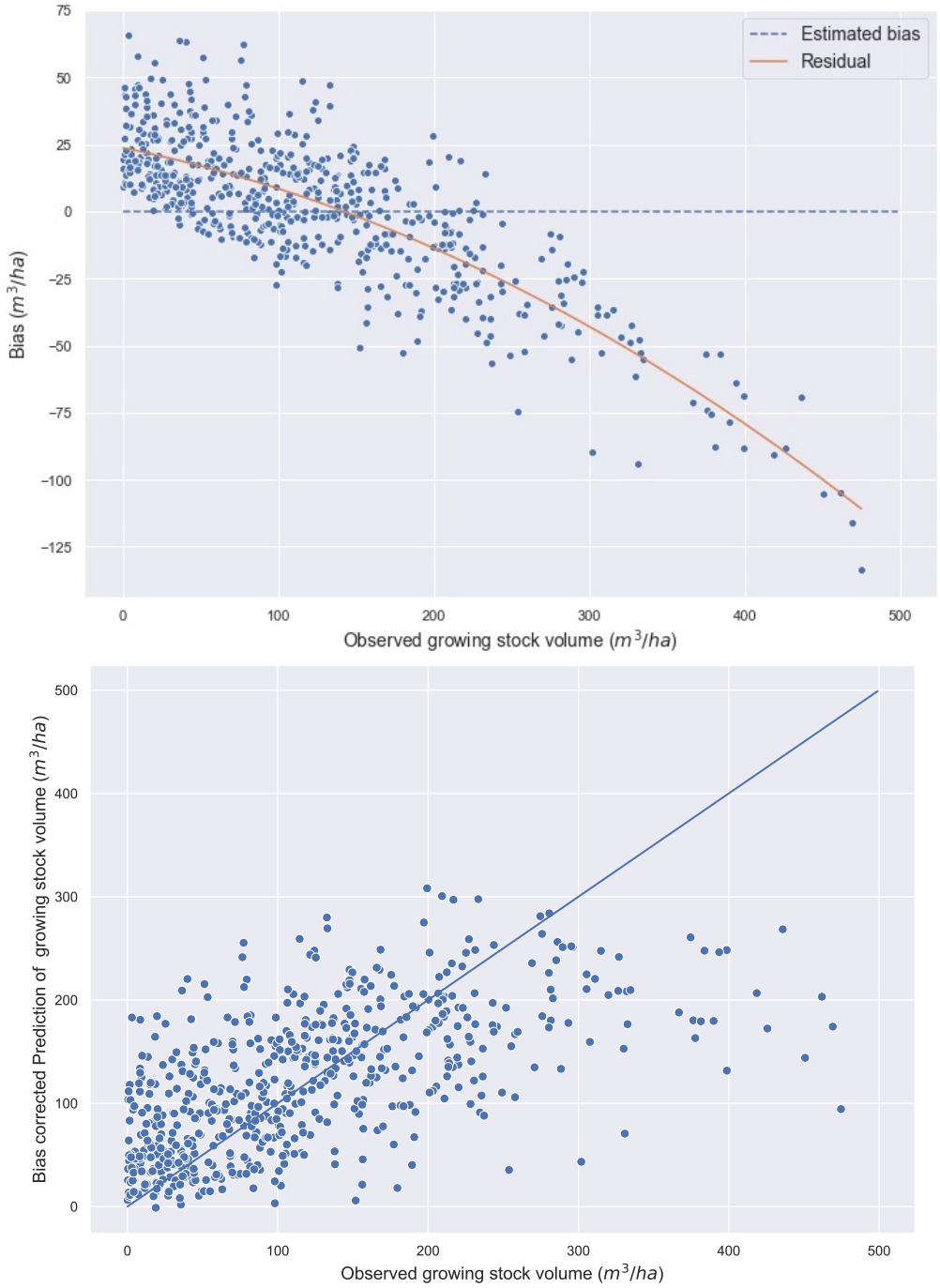


Figure 4.9: Top: Bias correction for CF_{SL}, Bottom: Observed volume vs bias corrected prediction in CF_{SL}.

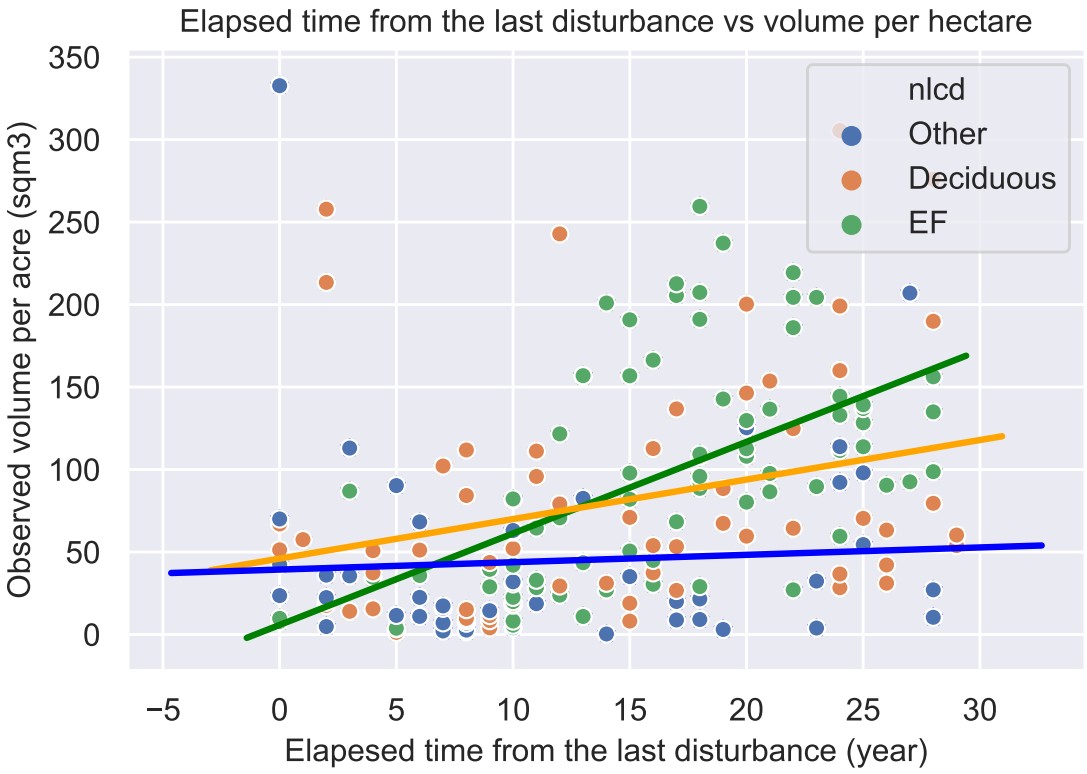


Figure 4.10: Elapsed time from the last disturbance and the observed volume per hectare.

of the features was different between species as is shown in Figure 4.8. The mean of the fitted values of the LTSL had a lower feature importance value for NE_{SL} than for E_{SL}. The mean value combines the fitted value of the leaf-off season and the leaf-on season. As the fitted value of the leaf-off season does not have a clear difference between the large-volume class and the small volume class, the mean of all the fitted values distinguishes between the large and small volume ambiguity. On the other hand, the maximum value of the LTSL of the TCW had lower feature importance value for E_{SL} than for NE_{SL}. This is because the maximum value can capture the highest value in a time range. Thus, we infer that the maximum value captured the value in a leaf-on season.

The feature importance of the last disturbance year data was less than 0.02 in all the models we built. Additionally, the difference between the OOB score of the model without the disturbance year record and the model with the disturbance year record was less than 0.01.

The first plausible reason for the relative unimportance of the last disturbance year data is that 80% of our field inventory data does not have any disturbance record. In line with most disturbance detection models, our last disturbance year data is limited to tracking disturbances that occurred after 1984, when the Landsat 5 was launched (Markham et al., 2004). Thus, disturbances that occurred before the launch time are not tracked, and these pixels are assigned to an "undisturbed" class as the value of the last disturbance year. Accordingly, about 70% of the pixels in our study area were assigned to the undisturbed class. The second plausible reason for the relative unimportance of the last disturbance year data is that the land use after a disturbance is not considered in our model. Figure 4.10 shows the correlation between the observed growing stock volume per hectare and the last disturbance year. We split the original data into three groups: (1) evergreen forest group (EF); (2) wet woodland, deciduous forest and mixed forest group (Deciduous); and (3) rest of the land-use group (Other). Three linear regression models were built for the three different groups. From these, it was observed that the plots with a short lapse of time from the last disturbance resulted in a minuscule degree of variability. On the other hand, the observed volume of the plots with long elapsed time showed a large degree of variability.

This distinction demonstrates that the importance of the elapsed time from the last disturbance decreases as time passes and that consequently, other factors become more important in the determination of the volume of the stand. Additionally, it was noticed that the slope of the regression line for the Other group is nearly equal to zero. Therefore, we infer that even though the FIA collects the relevant data, the other group remains unforested after the disturbance. Conversely, the EF group was found to have the largest slope among the three groups; This is likely because a large amount of the evergreen forests in this area are plantations, and as such they are subject to various prescribed silvicultural management protocols.

The rB ranged between 2.66 and 4.19 in the models using all the field inventory data. CF_L returned the smallest rB and the bias of the top 10% field inventory data. A comparison

Table 4.4: Number of top-10 feature variables for CF_{SL}, E_{SL}, and NE_{SL}.

Length	Model	max	mean	min	rmse	sin	slope	intercept	Total
STSL	CF_SL	1	-	-	2	-	-	-	3
	E_SL	-	1	-	-	-	1	-	2
	NE_SL	1	-	1	-	2	-	-	4
LTSL	CF_SL	1	1	-	2	1	1	1	7
	E_SL	1	2	1	2	-	-	2	8
	NE_SL	1	1	-	2	1	1	-	6

between CF_S and CF_L showed the difference in the bias of the top 10% data. The bias of the top 10% data for CF_S was 13.2% larger than that for CF_L. This difference might be caused by the characteristics of feature variables derived from LTSL that are less likely to spectrally saturate. The saturation of the spectral reflectance value of satellite imagery refers to the situation whereby spectral reflectance values mimic the values normally seen in forest vegetation with dense canopy cover. This phenomenon is considered to be the decisive factor in the low estimation accuracy of the forest above-ground biomass (AGB) and volume estimation, especially when the volume or AGB is high (Foody et al., 2003; Lu et al., 2016). Figure 4.11 contains plots of the observed growing stock volume on the x-axis and the values of the selected feature variables on the y-axis. Some of the feature variables such as the intercept of the LTSL and the mean of the LTSL can differentiate the small volume group and large volume group. The bias correction method reduced the absolute value of

Table 4.5: Summary of the bias metrics for each volume group.

	C_{SL}	F_{SL}	CF_S	CF_L	CF_{SL}	E_{SL}	NE_{SL}
rB	4.19	3.48	3.72	2.66	3.79	3.09	4.82
rB_corr	-2.73	-1.98	-1.48	-0.69	-1.47	-1.26	-2.86
mid80Bias	-16.63	-15.42	-16.19	-11.94	-14.1	-15.17	-15.857
mid80Bias_corr	-14.41	-13.08	-13.6	-9.01	-10.91	-12.61	-13.23
Bottom10Bias	-69.16	-65.42	-71.37	-71.51	-69.89	-53.39	-71.03
Bottom10Bias_corr	-56.43	-52.74	-58.14	-60.13	-57.43	-42.03	-56.31
Top10Bias	151.93	146.08	159.81	138.88	140.93	141.87	142.24
Top10Bias_corr	139.55	132.90	152.97	127.93	131.01	128.26	128.88

the relative bias for all the models. rB changed from 3.79% to -1.47% in CF_{SL}, which is the best model among all the models. The size of the overestimation in the bottom 10% data was reduced by 18% in CF_{SL} model. In addition, the underestimation in the top 10% data is reduced by 7%. These results coincide with the reported findings in (Zhang et al., 2012).

4.5 Conclusion

The primary sources of the presented estimation error come from the algorithmic characteristic of random forest and saturation of the spectral reflectance value of satellite imagery. The developed growing stock volume estimation model for the forest in the state of Georgia, United States, using RF regression, was applied to successfully find the factors that contribute to: (1) reducing the size of the overestimation of the large volume class and the underestimation of the small volume class; and (2) improving the rRMSE and OOB score. As a consequence of the comparison of the different models, we found that using the long-term Landsat imagery reduces the bias of the top 10% volume class by 12%. Additionally, using the bias correction method mitigates the under/overestimation error, while the middle volume rRMSE increases. Furthermore, feature importance performed on each model revealed that the critical features are affected by the various forest types. It should also be noted that relative to the model, which uses the coefficients of the short time range Landsat imagery only, the OOB score of the model, which uses the coefficients and fitted values of both the short-term and long-term Landsat imagery, was improved by 35%. Finally, since the raster data created by the presented model can estimate the growing stock volume of the forest stand with 30-meter spatial resolution, it is expected that the data can be used for sub-county areas volume estimations, which is an important functionality for the forest product industry and land owners in the state of Georgia. Remained task of this research is to incorporate Landsat 8 to increase the number of Landsat imagery.

4.6 Acknowledgements

We are indebted to the Japan Student Service Organization Student Exchange Support Program (Graduate Scholarship for Degree Seeking Students) for funding this research. This work was also supported by the U.S. Department of Agriculture, National Institute of Food and Agriculture, McIntire-Stennis project (1012166), administered by the Warnell School of Forestry and Natural Resources at the University of Georgia. We would also like to mention our appreciation for the support and assistance of the Google Earth Engine Development team.

4.7 References

Barrett, F., R. E. McRoberts, E. Tomppo, E. Cienciala, and L. T. Waser (2016). "A questionnaire-based review of the operational use of remotely sensed data by national forest inventories". In: *Remote Sensing of Environment* 174, pp. 279–289. DOI: 10.1016/j.rse.2015.08.029.

Bechtold, W. A. and P. L. Patterson (2015). *The enhanced Forest Inventory and Analysis program: National sampling design and estimation procedures*. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. DOI: 10.2737/SRS-GTR-80.

Bettinger, P., K. Boston, J. P. Siry, and D. L. Grebner (2017). *Forest management and planning*. 2nd edition. Boston, MA: Academic Press. 362 pp.

Brandeis, T. J., A. J. Hartsell, J. W. Bentley, and C. Brandeis (2012). *Economic dynamics of forests and forest industries in the Southern United States*. SRS-152. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station, p. 86.

Breiman, L. (1996). "Bagging predictors". In: *Machine Learning* 24.2, pp. 123–140. DOI: 10.1007/BF00058655.

Breiman, L. (2001). "Random forests". In: *Machine Learning* 45.1, pp. 5–32. DOI: 10.1023/A:1010933404324.

Brooks, E. B., R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston (2014). "On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6, pp. 3316–3332. DOI: 10.1109/TGRS.2013.2272545.

Burrill, E. A., A. M. Wilson, J. A. Turner, S. A. Pugh, J. Menlove, G. Christensen, B. L. Conkling, and W. David (2018). *The Forest Inventory and Analysis Database: Database description and user guide for phase 2 (version 7.2)*. Washington, DC: U.S. Forest Service.

Chen, C., A. Liaw, and L. Breiman (2004). *Using random forest to learn imbalanced data*. 666. Berkley, CA: Department of Statistics, University of California Berkeley, pp. 1–12.

Chirici, G., M. Mura, D. McInerney, N. Py, E. O. Tomppo, L. T. Waser, D. Travaglini, and R. E. McRoberts (2016). "A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data". In: *Remote Sensing of Environment* 176, pp. 282–294. DOI: 10.1016/j.rse.2016.02.001.

Crist, E. P. (1985). "A TM Tasseled Cap equivalent transformation for reflectance factor data". In: *Remote Sensing of Environment* 17.3, pp. 301–306. DOI: 10.1016/0034-4257(85)90102-6.

D'Amato, A. W., E. J. Jokela, K. L. O'Hara, and J. N. Long (2017). "Silviculture in the United States: An amazing period of change over the past 30 years". In: *Journal of Forestry* 116.1, pp. 55–67. DOI: 10.5849/JOF-2016-035.

Foga, S., P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue (2017). "Cloud detection algorithm comparison and validation for operational Landsat data products". In: *Remote Sensing of Environment* 194, pp. 379–390. DOI: 10.1016/j.rse.2017.03.026.

Foody, G. M., D. S. Boyd, and M. E. J. Cutler (2003). "Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions". In: *Remote Sensing of Environment* 85.4, pp. 463–474. DOI: 10.1016/S0034-4257(03)00039-7.

Fox, T. R., E. J. Jokela, and H. L. Allen (2007). "The development of pine plantation silviculture in the Southern United States". In: *Journal of Forestry* 105.7, pp. 337–347. DOI: 10.1093/jof/105.7.337.

Franco-Lopez, H., A. R. Ek, and M. E. Bauer (2001). "Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method". In: *Remote Sensing of Environment* 77.3, pp. 251–274. DOI: 10.1016/S0034-4257(01)00209-7.

Gigović, L., H. R. Pourghasemi, S. Drobnjak, and S. Bai (2019). "Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park". In: *Forests* 10.5, p. 408. DOI: 10.3390/f10050408.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". In: *Remote Sensing of Environment* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Hird, J. N. and G. J. McDermid (2009). "Noise reduction of NDVI time series: An empirical comparison of selected techniques". In: *Remote Sensing of Environment* 113.1, pp. 248–258. DOI: 10.1016/j.rse.2008.09.003.

Kennedy, R. E., Z. Yang, W. B. Cohen, E. Pfaff, J. Braaten, and P. Nelson (2012). "Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan". In: *Remote Sensing of Environment* 122, pp. 117–133. DOI: 10.1016/j.rse.2011.09.024.

Liaw, A. and M. Wiener (2002). "Classication and regression by randomForest". In: *R News* 2, pp. 18–22.

Lu, D., Q. Chen, G. Wang, L. Liu, G. Li, and E. Moran (2016). "A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems". In: *International Journal of Digital Earth* 9.1, pp. 63–105. DOI: 10.1080/17538947.2014.990526.

Markham, B., J. Storey, D. Williams, and J. Irons (2004). "Landsat sensor performance: History and current status". In: *IEEE Transactions on Geoscience and Remote Sensing* 42.12, pp. 2691–2694. DOI: 10.1109/TGRS.2004.840720.

Masek, J. G., E. Vermonte, N. Saleous, R. Wolfe, F. Hall, F. Huemmrich, F. Gao, J. Kulter, and T. Lim (2006). "A Landsat surface reflectance data set for North America, 1990-2000". In: *Geoscience and Remote Sensing Letters* 3, pp. 68–72.

Maselli, F., G. Chirici, L. Bottai, P. Corona, and M. Marchetti (2005). "Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images". In: *International Journal of Remote Sensing* 26.17, pp. 3781–3796. DOI: 10.1080/01431160500166433.

McRoberts, R. E. (2009). "Diagnostic tools for nearest neighbors techniques when used with satellite imagery". In: *Remote Sensing of Environment. 113: 489-499.* 113.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt (2002). "Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique". In: *Remote Sensing of Environment* 82.2, pp. 457–468. DOI: 10.1016/S0034-4257(02)00064-0.

Miksys, V., I. Varnagiryte-kabasinskiene, I. Stupak, K. Armolaitis, M. Kukkola, and J. Wojcik (2007). "Above-ground biomass functions for Scots pine in Lithuania". In: *Biomass and Bioenergy* 31.10, pp. 685–692. DOI: 10.1016/j.biombioe.2007.06.013.

Moody, A. and D. M. Johnson (2001). "Land-surface phenologies from AVHRR using the discrete fourier transform". In: *Remote Sensing of Environment* 75.3, pp. 305–323. DOI: 10.1016/S0034-4257(00)00175-9.

Obata, S., C. J. Cieszewski, P. Bettinger, R. C. Lowe III, and S. Bernardes (2019). "Preliminary analysis of forest stand disturbances in Coastal Georgia (USA) using Landsat time series stacked imagery". In: *FORMATH* 18, pp. 1–11. DOI: 10.15684/formath.001.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay (2011). "Scikit-learn: Machine learning in Python". In: *Journal of Machine Learning Research* 12, 2825 − 2830.

Reese, H., M. Nilsson, T. G. Pahén, O. Hagner, S. Joyce, U. Tingelöf, M. Egberth, and H. Olsson (2003). "Countrywide estimates of forest variables using satellite data and field data from the national forest inventory". In: *Ambio* 32.8, pp. 542–548. DOI: 10.1579/0044-7447-32.8.542.

Reese, H., M. Nilsson, P. Sandström, and H. Olsson (2002). "Applications using estimates of forest parameters derived from satellite and forest inventory data". In: 37.

Shumway, R. H. and D. S. Stoffer (2017). "Spectral analysis and filtering". In: *Time series analysis and its applications: With r examples*. 4th. New York, NY: Springer Science+Business Media, pp. 165–172.

Smith, W. (2002). "Forest inventory and analysis: A national inventory and monitoring program". In: *Environmental Pollution* 116, pp. 233–242. DOI: 10.1016/S0269-7491(01)00255-X.

Tanaka, S., T. Takahashi, T. Nishizono, F. Kitahara, H. Saito, T. Iehara, E. Kodani, and Y. Awaya (2014). "Stand volume estimation using the k-NN technique combined with forest inventory data, satellite image data and additional feature variables". In: *Remote Sensing* 7.1, pp. 378–394. DOI: 10.3390/rs70100378.

Tompalski, P., J. C. White, N. C. Coops, and M. A. Wulder (2019). "Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data". In: *Remote Sensing of Environment* 227, pp. 110–124. DOI: 10.1016/j.rse.2019.04.006.

Tomppo, E. (1990). "Designing a satellite image-aided national forest survey in Finland". In: Usability of remote sensing for forest inventory and planning. Umea, Sweden: Swedish University of Agricultural Sciences, pp. 43–47.

U.S. Department of the Interior, Fish and Wildlife Service (2003). *Recover plan for the Red-Cockaded Woodpecker (Picoides borealis)*. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service, p. 316.

U.S. Geological Survey (2020). *Landsat levels of processing*. URL: https://www.usgs.gov/land-resources/nli/landsat/landsat-levels-processing (visited on 02/09/2020).

Wilson, B. T., J. F. Knight, and R. E. McRoberts (2018). "Harmonic regression of Landsat time series for modeling attributes from national forest inventory data". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 137, pp. 29–46. DOI: 10.1016/j.isprsjprs.2018.01.006.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne (2008). "Free access to Landsat imagery". In: *Science* 320.5879, pp. 1011–1011. DOI: 10.1126/science.320.5879.1011a.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian (2018). "A new generation of the United States national land cover database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146, pp. 108–123. DOI: 10.1016/j.isprsjprs.2018.09.006.

Zhang, G. and Y. Lu (2012). "Bias-corrected random forests in regression". In: *Journal of Applied Statistics* 39.1, pp. 151–160. DOI: 10.1080/02664763.2011.578621.

Zhu, Z., Y. Fu, C. E. Woodcock, P. Olofsson, J. E. Vogelmann, C. Holden, M. Wang, S. Dai, and Y. Yu (2016). "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014)".

In: *Remote Sensing of Environment* 185, pp. 243–257. DOI: 10.1016/j.rse.2016.03.036.

Zhu, Z. and C. E. Woodcock (2014). "Continuous change detection and classification of land cover using all available Landsat data". In: *Remote Sensing of Environment* 144, pp. 152–171. DOI: 10.1016/j.rse.2014.01.011.

Zhu, Z., S. Wang, and C. E. Woodcock (2015a). "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images". In: *Remote Sensing of Environment* 159, pp. 269–277. DOI: 10.1016/j.rse.2014.12.014.

Zhu, Z. and C. E. Woodcock (2012a). "Object-based cloud and cloud shadow detection in Landsat imagery". In: *Remote Sensing of Environment* 118, pp. 83–94. DOI: 10.1016/j.rse.2011.10.028.

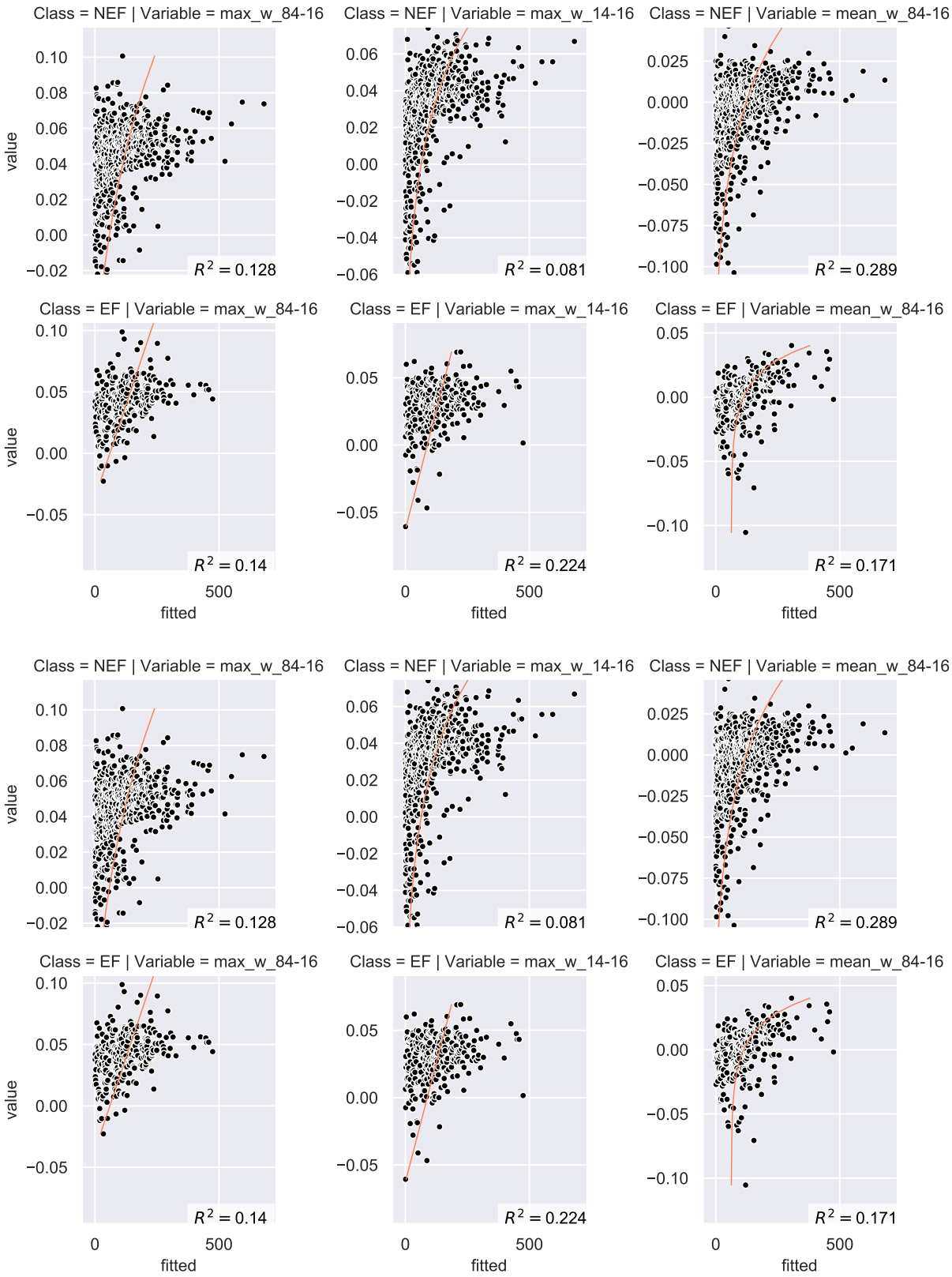


Figure 4.11: Scatter plot of the Feature variables and observed growing stock volume. EF: Evergreen forest, NEF: non-Evergreen forest111

Chapter 5

Conclusion

Estimating the growing stock volume of forest stands is a crucial matter for forest management and conservation. These estimates provide economically, ecologically, and environmentally important information on forest management. In this dissertation, a series of research studies that aimed at estimating the growing stock volume of the forest stands in Georgia with a 30-meter spatial resolution was conducted.

The study described in Chapter 2, "Preliminary analysis of forest stand disturbances in coastal Georgia (USA) using Landsat time series stacked imagery", revealed the result of preliminary research on forest disturbance detection. The main goal of this research was to explain the current age-class structure of forests in the Lower Coastal Plain of the State of Georgia, an area where forests are primarily maintained using even-aged methods and where forests are mixed with other land-uses. The algorithm, which is similar to the Vegetation Change Tracker, was developed to help recognize the age-class distribution of existing forests in the Lower Coastal Plain forests. The algorithm employed three Landsat bands and used both visible and infrared spectrums. The satellite imagery was queried across a long time horizon to determine major forest disturbances since the year 1987. The overall accuracy was 52% when all years were considered, and approximately 70% from 1991 forward. The result of this preliminary research was acceptable for this region. The first issue of this

research was that no suitable imagery was available for 1993 and 2012, meaning that we could not detect the disturbed areas for these years. In addition to this, since we used one image per year, cloud coverage limited the model's capability to detect the disturbance. My contribution to science was, therefore, the development of a new method to determine the age of forests across broad areas of landscapes.

In Chapter 3, "Mapping forest disturbances between 1987-2016 using all available time series Landsat TM/ETM imagery: Developing a reliable methodology for Georgia, the United States", we expanded our area of investigation in two distinct ways from the first research effort. Firstly, we changed our area of interest from Coastal Georgia to the entire state of Georgia, as our ultimate interest was to analyze the forest disturbance history of the entire state. Secondly, we utilized all the available Landsat imagery to improve the accuracy of the disturbance detection. In this research, a raster database of which each pixel represents the last disturbance year of the forests within the State of Georgia was created. We set two primary objectives of the research. The first objective was to develop and test a robust methodology for state-wide detection of forest disturbances in Georgia. The second objective was to examine and discuss the implication of the information obtained from the forest disturbance map created in this research. We used two major disturbance detection methods, namely the threshold algorithm and a statistical boundary method, to establish a sound disturbance detection model. To process over 5000 Landsat images we utilized the Google Earth Engine, a cloud-based platform that provides a high-performance computing environment to process large-scale satellite imagery without concern for computational and data storage issues. The overall accuracy of the disturbance year data in this research was 87%. This figure is better than the overall accuracy of the preliminary research by 35%. This improvement was the result of the two modifications we made from our previous model. The first modification was the use of all the available Landsat imagery. With this change the cloud cover issue was by-passed, as the interval between one imagery and the next imagery was reduced to 28.0 days on average. Thus, even if a large portion of an image was obscured

by cloud coverage, a cloud-free image was likely available a relatively short period of time later over the period of analysis. The second modification was the algorithm; the model developed in this research combined two disturbance detection methods while the preliminary. These modifications to the process of estimating forest ages across broad areas of landscapes were the main contribution to the science of our research. In addition, our result revealed that 29.2% of the land area in Georgia was disturbed at least once between 1984 and 2016. Although equivalent research having the same spatial and temporal coverage as ours does not exist, our estimation of the disturbed land area seems reasonable. The estimated annual last disturbance area ranged between 84,651+/- 36,354 ha and 211,780 +/- 49,504 ha. We also conducted an analysis of land-use patterns. This analysis discovered that disturbed forests close to urban zones were more likely to be converted to other land-uses than disturbed forests in rural areas. In the Southern Coastal Plain, where plantation forests that are managed intensively occupy a large amount of the land area, forests encountered the greatest likelihood of disturbance and recovered at a higher rate compared to different regions.

In Chapter 4, "Estimation of forest growing stock stand volumes based on Harmonic regression with ensemble method and 1984-2016 Landsat 5 and 7 imagery", we estimated the growing stock volume of forest stands in Georgia at a 30-meter resolution. The goal of this analysis was to examine the efficacy of using all available Landsat TM / ETM imagery and to apply a bias correction method on the precision of the estimation of the growing stock volume. To attain this goal, a model that estimates the growing stock volume was created. As remotely sensed data, time series Landsat imagery and the disturbance map that was created in our previous two research were used. As a field plot data, we utilized the Forest Inventory and Analysis dataset managed by US Forest Service. The estimation method employed was Random Forest. For the best model developed, we obtained 65% of relative RMSE. As a result of the comparison of the seven models developed by using the aforementioned datasets, it was revealed that using long-term Landsat imagery provides

better estimates than a more basic model. In addition, the bias correction method reduces the under/overestimation error, yet the median volume rRMSE increased.

To conclude this dissertation, I discuss some of the remaining extensions of this research. Firstly, publicly available land cover change data could be incorporated into our growing stock volume estimation model as ancillary data. As is discussed in Chapter 4, the disturbance year map had low feature importance compared to the feature variables derive from remotely sensed data. Pflugmacher et al. (2012) and Pflugmacher et al. (2014) argue that incorporating the disturbance and regeneration metrics obtained from Landsat time series dataset results in a better above ground biomass estimation accuracy compared to the estimation using single-date images. Other research has shown that forest age data computed from Landsat time series can improve the estimates of above-ground biomass (Helmer et al., 2009; L. Liu et al., 2014). Also, Gómez et al. (2014) argue that identifiers of time series patterns in spectral trajectories are useful to improve the accuracy of the historical above-ground biomass estimations. This research utilizes the publicly available change detection dataset or algorithm. Therefore, it is worth trying to incorporate the publicly available change detection dataset/algorithm, in order to improve the accuracy of the growing stock volume estimation. Another remaining work is to examine the efficacy of another remote sensing dataset. Concerning the satellite imagery, Sentinel 2 (Dechoz et al., 2015) could be used for the replacement of the Landsat imagery. As both the temporal resolution and spatial resolution of Sentinel 2 imagery is better than Landsat missions, it is expected that the accuracy of estimation might be improved by using Sentinel 2 imagery. As airborne imagery, LiDAR data is also widely used in previous research (Deo et al., 2017b; Deo et al., 2017a; Matasci et al., 2018). Additionally, as it is suggested from the previous research that combining LiDAR data and Landsat data may improve the accuracy of the estimation of above-ground biomass, it is highly expected that using LiDAR data would improve our estimation. To successfully incorporate the LiDAR data with field plot data, point cloud data covering the entire study area and field plot data systematically sampled, are required

(Tinkham et al., 2018). However, As of early 2020, publicly available point cloud data covers only a portion of Georgia (U.S. Geological Survey, 2020).

The final remaining task of this research is to create the growing stock volume estimation map of Georgia based on the model we built. To make the map useful for the various stakeholders in Georgia, it is required to mask out the non-forest pixels, as the model itself makes an estimation to every single pixel regardless of its actual land-use. National Land Cover Dataset can be used as the land-cover dataset. However, the NLCD is not perfect, as it has been observed to assign non-forest classes, such as barren land or pasture, to recently regenerated forests. Thus, applying the NLCD as the mask raster to determine the currently forested region could result in the omission of the young forest from the volume map. Another candidate for the land cover dataset is a custom land cover dataset. For example, Wilson et al. (2018) performed a land-use classification using the same data set used to estimate the growing stock volume. These tasks may be pursued in the future if time and resources are available.

5.1 References

Dechoz, C., V. Poulain, S. Massera, F. Languille, D. Greslou, F. d. Lussy, A. Gaudel, C. L'Helguen, C. Picard, and T. Trémas (2015). "Sentinel 2 global reference image". In: *Image and Signal Processing for Remote Sensing XXI*. Image and Signal Processing for Remote Sensing XXI. Vol. 9643. International Society for Optics and Photonics, 96430A. DOI: 10.1117/12.2195046.

Deo, R. K., M. B. Russell, G. M. Domke, H.-E. Andersen, W. B. Cohen, and C. W. Woodall (2017a). "Evaluating site-specific and generic spatial models of aboveground forest biomass based on Landsat time-series and LiDAR strip samples in the Eastern USA". In: *Remote Sensing* 9.6, p. 598. DOI: 10.3390/rs9060598.

Deo, R. K., M. B. Russell, G. M. Domke, C. W. Woodall, M. J. Falkowski, and W. B. Cohen (2017b). "Using Landsat time-series and LiDAR to inform aboveground forest biomass baselines in Northern Minnesota, USA". In: *Canadian Journal of Remote Sensing* 43.1, pp. 28–47. DOI: 10.1080/07038992.2017.1259556.

Gómez, C., J. C. White, M. A. Wulder, and P. Alejandro (2014). "Historical forest biomass dynamics modelled with Landsat spectral trajectories". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 93, pp. 14–28. DOI: 10.1016/j.isprsjprs.2014.03.008.

Helmer, E. H., M. A. Lefsky, and D. A. Roberts (2009). "Biomass accumulation rates of Amazonian secondary forest and biomass of old-growth forests from Landsat time series and the geoscience laser altimeter system". In: *Journal of Applied Remote Sensing* 3, p. 033505. DOI: 10.1117/1.3082116.

Liu, L., D. Peng, Z. Wang, and Y. Hu (2014). "Improving artificial forest biomass estimates using afforestation age information from time series Landsat stacks". In: *Environmental Monitoring and Assessment* 186.11, pp. 7293–7306. DOI: 10.1007/s10661-014-3927-y.

Matasci, G., T. Hermosilla, M. A. Wulder, J. C. White, N. C. Coops, G. W. Hobart, D. K. Bolton, P. Tompalski, and C. W. Bater (2018). "Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots". In: *Remote Sensing of Environment* 216, pp. 697–714. DOI: 10.1016/j.rse.2018.07.024.

Pflugmacher, D., W. B. Cohen, and R. E. Kennedy (2012). "Using Landsat-derived disturbance history (1972–2010) to predict current forest structure". In: *Remote Sensing of Environment* 122, pp. 146–165. DOI: 10.1016/j.rse.2011.09.025.

Pflugmacher, D., W. B. Cohen, R. E. Kennedy, and Z. Yang (2014). "Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics". In: *Remote Sensing of Environment* 151, pp. 124–137. DOI: 10.1016/j.rse.2013.05.033.

Tinkham, W. T., P. R. Mahoney, A. T. Hudak, G. M. Domke, M. J. Falkowski, C. W. Woodall, and A. M. Smith (2018). "Applications of the United States Forest Inventory

and Analysis dataset: A review and future directions". In: *Canadian Journal of Forest Research* 48.11, pp. 1251–1268. DOI: 10.1139/cjfr-2018-0196.

U.S. Geological Survey (2020). *Landsat levels of processing*. URL: https://www.usgs.gov/land-resources/nli/landsat/landsat-levels-processing (visited on 02/09/2020).

Wilson, B. T., J. F. Knight, and R. E. McRoberts (2018). "Harmonic regression of Landsat time series for modeling attributes from national forest inventory data". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 137, pp. 29–46. DOI: 10.1016/j.isprsjprs.2018.01.006.

Reference

Barrett, F., R. E. McRoberts, E. Tomppo, E. Cienciala, and L. T. Waser (2016). "A questionnaire-based review of the operational use of remotely sensed data by national forest inventories". In: *Remote Sensing of Environment* 174, pp. 279–289. DOI: 10.1016/j.rse.2015.08.029.

Bechtold, W. A. and P. L. Patterson (2015). *The enhanced Forest Inventory and Analysis program: National sampling design and estimation procedures*. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. DOI: 10.2737/SRS-GTR-80.

Bettinger, P., K. Boston, J. P. Siry, and D. L. Grebner (2017). *Forest management and planning*. 2nd edition. Boston, MA: Academic Press. 362 pp.

Brandeis, T. J., A. J. Hartsell, J. W. Bentley, and C. Brandeis (2012). *Economic dynamics of forests and forest industries in the Southern United States*. SRS-152. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station, p. 86.

Breiman, L. (1996). "Bagging predictors". In: *Machine Learning* 24.2, pp. 123–140. DOI: 10.1007/BF00058655.

Breiman, L. (2001). "Random forests". In: *Machine Learning* 45.1, pp. 5–32. DOI: 10.1023/A:1010933404324.

Brooks, E. B., R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston (2014). "On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6, pp. 3316–3332. DOI: 10.1109/TGRS.2013.2272545.

Burrill, E. A., A. M. Wilson, J. A. Turner, S. A. Pugh, J. Menlove, G. Christensen, B. L. Conkling, and W. David (2018). *The Forest Inventory and Analysis Database: Database description and user guide for phase 2 (version 7.2)*. Washington, DC: U.S. Forest Service.

Campbell, J. B. and R. H. Wynne (2011). *Introduction to remote sensing*. 5th edition. New York: The Guilford Press. 667 pp.

Card, D. H. (1982). "Using known map category marginal frequencies to improve estimates of thematic map accuracy". In: *Photogrammetric Engineering & Remote Sensing* 48.3, pp. 431–439.

Carlson, T. N. and S. T. Arthur (2000). "The impact of land use — land cover changes due to urbanization on surface microclimate and hydrology: A satellite perspective". In: *Global and Planetary Change* 25.1, pp. 49–65. DOI: 10.1016/S0921-8181(00)00021-7.

Chen, C., A. Liaw, and L. Breiman (2004). *Using random forest to learn imbalanced data*. 666. Berkley, CA: Department of Statistics, University of California Berkeley, pp. 1–12.

Chen, J., X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin (2011). "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images". In: *Remote Sensing of Environment* 115.4, pp. 1053–1064. DOI: 10.1016/j.rse.2010.12.010.

Chirici, G., M. Mura, D. McInerney, N. Py, E. O. Tomppo, L. T. Waser, D. Travaglini, and R. E. McRoberts (2016). "A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data". In: *Remote Sensing of Environment* 176, pp. 282–294. DOI: 10.1016/j.rse.2016.02.001.

Cieszewski, C. J. (2011). "Spatially explicit biomass supply sustainability analysis for bioenergy mill siting in Georgia, USA". In: *The Open Forest Science Journal* 4.1, pp. 2–14. DOI: 10.2174/1874398601104010002.

Cieszewski, C. J., M. Zasada, B. E. Borders, R. C. Lowe, J. Zawadzki, M. L. Clutter, and R. F. Daniels (2004). "Spatially explicit sustainability analysis of long-term fiber supply in Georgia, USA". In: *Forest Ecology and Management* 187.2, pp. 349–359. DOI: 10.1016/j.foreco.2003.08.001.

Cohen, W. B., Z. Yang, S. V. Stehman, T. A. Schroeder, D. M. Bell, J. G. Masek, C. Huang, and G. W. Meigs (2016). "Forest disturbance across the conterminous United States from 1985–2012: The emerging dominance of forest decline". In: *Forest Ecology and Management* 360, pp. 242–252. DOI: 10.1016/j.foreco.2015.10.042.

Crist, E. P. (1985). "A TM Tasseled Cap equivalent transformation for reflectance factor data". In: *Remote Sensing of Environment* 17.3, pp. 301–306. DOI: 10.1016/0034-4257(85)90102-6.

D'Amato, A. W., E. J. Jokela, K. L. O'Hara, and J. N. Long (2017). "Silviculture in the United States: An amazing period of change over the past 30 years". In: *Journal of Forestry* 116.1, pp. 55–67. DOI: 10.5849/JOF-2016-035.

Dechoz, C., V. Poulain, S. Massera, F. Languille, D. Greslou, F. d. Lussy, A. Gaudel, C. L'Helguen, C. Picard, and T. Trémas (2015). "Sentinel 2 global reference image". In: *Image and Signal Processing for Remote Sensing XXI*. Image and Signal Processing for Remote Sensing XXI. Vol. 9643. International Society for Optics and Photonics, 96430A. DOI: 10.1117/12.2195046.

Deo, R. K., M. B. Russell, G. M. Domke, H.-E. Andersen, W. B. Cohen, and C. W. Woodall (2017a). "Evaluating site-specific and generic spatial models of aboveground forest biomass based on Landsat time-series and LiDAR strip samples in the Eastern USA". In: *Remote Sensing* 9.6, p. 598. DOI: 10.3390/rs9060598.

Deo, R. K., M. B. Russell, G. M. Domke, C. W. Woodall, M. J. Falkowski, and W. B. Cohen (2017b). "Using Landsat time-series and LiDAR to inform aboveground forest biomass baselines in Northern Minnesota, USA". In: *Canadian Journal of Remote Sensing* 43.1, pp. 28–47. DOI: 10.1080/07038992.2017.1259556.

DeVries, B., M. Decuyper, J. Verbesselt, A. Zeileis, M. Herold, and S. Joseph (2015). "Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series". In: *Remote Sensing of Environment* 169, pp. 320–334. DOI: 10.1016/j.rse.2015.08.020.

Dwyer, J. L., D. P. Roy, B. Sauer, C. B. Jenkerson, H. K. Zhang, and L. Lymburner (2018). "Analysis ready data: Enabling analysis of the Landsat archive". In: *Remote Sensing* 10.9, p. 1363. DOI: 10.3390/rs10091363.

Edwards, D. P., J. A. Tobias, D. Sheil, E. Meijaard, and W. F. Laurance (2014). "Maintaining ecosystem function and services in logged tropical forests". In: *Trends in Ecology & Evolution* 29.9, pp. 511–520. DOI: 10.1016/j.tree.2014.07.003.

Edwards, L., J. Ambrose, and L. K. Kirkman (2013). *The natural communities of Georgia*. Athens, GA: University of Georgia Press. 697 pp.

Eskelson, B. N. I., H. Temesgen, V. Lemay, T. M. Barrett, N. L. Crookston, and A. T. Hudak (2009). "The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases". In: *Scandinavian Journal of Forest Research* 24.3, pp. 235–246. DOI: 10.1080/02827580902870490.

Finley, A. O. and R. E. McRoberts (2008). "Efficient k-nearest neighbor searches for multi-source forest attribute mapping". In: *Remote Sensing of Environment* 112.5, pp. 2203–2211. DOI: 10.1016/j.rse.2007.08.024.

Foga, S., P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue (2017). "Cloud detection algorithm comparison and validation for operational Landsat data products". In: *Remote Sensing of Environment* 194, pp. 379–390. DOI: 10.1016/j.rse.2017.03.026.

Foody, G. M., D. S. Boyd, and M. E. J. Cutler (2003). "Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions". In: *Remote Sensing of Environment* 85.4, pp. 463–474. DOI: 10.1016/S0034-4257(03)00039-7.

Fox, T. R., E. J. Jokela, and H. L. Allen (2007). "The development of pine plantation silviculture in the Southern United States". In: *Journal of Forestry* 105.7, pp. 337–347. DOI: 10.1093/jof/105.7.337.

Franco-Lopez, H., A. R. Ek, and M. E. Bauer (2001). "Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method". In: *Remote Sensing of Environment* 77.3, pp. 251–274. DOI: 10.1016/S0034-4257(01)00209-7.

Frazier, R. J., N. C. Coops, and M. A. Wulder (2015). "Boreal Shield forest disturbance and recovery trends using Landsat time series". In: *Remote Sensing of Environment* 170, pp. 317–327. DOI: 10.1016/j.rse.2015.09.015.

Georgia Forestry Commision (2018). *Conservation Reserve Program (CRP)*. Cost Share & Incentive Programs. URL: http://www.gfc.state.ga.us/forest-management/private-forest-management/landowner-programs/other-landowner-programs/ (visited on 02/12/2018).

Gigović, L., H. R. Pourghasemi, S. Drobnjak, and S. Bai (2019). "Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park". In: *Forests* 10.5, p. 408. DOI: 10.3390/f10050408.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray (2001). "The recent increase in Atlantic hurricane activity: Causes and implications". In: *Science* 293.5529, pp. 474–479. DOI: 10.1126/science.1060040.

Gómez, C., J. C. White, M. A. Wulder, and P. Alejandro (2014). "Historical forest biomass dynamics modelled with Landsat spectral trajectories". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 93, pp. 14–28. DOI: 10.1016/j.isprsjprs.2014.03.008.

Gordon, H. R. (1997). "Atmospheric correction of ocean color imagery in the Earth observing system era". In: *Journal of Geophysical Research: Atmospheres* 102 (D14), pp. 17081–17106. DOI: 10.1029/96JD02443.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". In: *Remote Sensing of Environment* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Goward, S. N., T. Arvidson, D. Williams, L. Rocchio, J. R. Irons, S. S. Johnston, and C. Russel (2017). *Landsat's enduring legacy: Pioneering global land observations from space*. Bethesda, MD: American Society for Photogrammetry and Remote Sensing. 586 pp.

Goward, S. N., C. Huang, F. Zhao, K. Schleeweis, K. Rishmawi, M. Lindsey, J. L. Dungan, and A. R. Michaelis (2016). "NACP NAFD project: Forest disturbance history from Landsat, 1986-2010". In: *ORNL DAAC*. DOI: https://doi.org/10.3334/ORNLDAAC/1290.

Grogan, K., D. Pflugmacher, P. Hostert, R. Kennedy, and R. Fensholt (2015). "Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series". In: *Remote Sensing of Environment* 169, pp. 438–453. DOI: 10.1016/j.rse.2015.03.001.

Gutman, G., C. Huang, G. Chander, P. Noojipady, and J. G. Masek (2013). "Assessment of the NASA–USGS Global Land Survey (GLS) datasets". In: *Remote Sensing of Environment* 134, pp. 249–265. DOI: 10.1016/j.rse.2013.02.026.

Hairston-Strang, A. B., P. W. Adams, and G. G. Ice (2008). "The Oregon Forest Practices Act and forest research". In: *Hydrological and biological responses to forest practices*. Ed. by J. D. Stednick. Vol. 199. New York, NY: Springer, pp. 95–113. DOI: 10.1007/978-0-387-69036-0_6.

Hamunyela, E., J. Verbesselt, and M. Herold (2016). "Using spatial context to improve early detection of deforestation from Landsat time series". In: *Remote Sensing of Environment* 172, pp. 126–138. DOI: 10.1016/j.rse.2015.11.006.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). "High-resolution global maps of 21st-Century forest cover change". In: *Science* 342.6160, pp. 850–853. DOI: 10.1126/science.1244693.

Hansen, M. C., S. V. Stehman, and P. V. Potapov (2010). "Quantification of global gross forest cover loss". In: *Proceedings of the National Academy of Sciences* 107.19, pp. 8650–8655. DOI: 10.1073/pnas.0912668107.

Helmer, E. H., M. A. Lefsky, and D. A. Roberts (2009). "Biomass accumulation rates of Amazonian secondary forest and biomass of old-growth forests from Landsat time series and the geoscience laser altimeter system". In: *Journal of Applied Remote Sensing* 3, p. 033505. DOI: 10.1117/1.3082116.

Hermosilla, T., M. A. Wulder, J. C. White, N. C. Coops, and G. W. Hobart (2015). "Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics". In: *Remote Sensing of Environment* 170, pp. 121–132. DOI: 10.1016/j.rse.2015.09.004.

Hilker, T., M. A. Wulder, N. C. Coops, J. Linke, G. McDermid, J. G. Masek, F. Gao, and J. C. White (2009). "A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS". In: *Remote Sensing of Environment* 113.8, pp. 1613–1627. DOI: 10.1016/j.rse.2009.03.007.

Hird, J. N. and G. J. McDermid (2009). "Noise reduction of NDVI time series: An empirical comparison of selected techniques". In: *Remote Sensing of Environment* 113.1, pp. 248–258. DOI: 10.1016/j.rse.2008.09.003.

Huang, C., S. N. Goward, J. G. Masek, F. Gao, E. F. Vermote, N. Thomas, K. Schleeweis, R. E. Kennedy, Z. Zhu, J. C. Eidenshink, and J. R. Townshend (2009a). "Development of time series stacks of Landsat images for reconstructing forest disturbance history". In: *International Journal of Digital Earth* 2.3, pp. 195–218. DOI: 10.1080/17538940902801614.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann (2010). "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks". In: *Remote Sensing of Environment* 114.1, pp. 183–198. DOI: 10.1016/j.rse.2009.08.017.

Huang, C., S. N. Goward, K. Schleeweis, N. Thomas, J. G. Masek, and Z. Zhu (2009b). "Dynamics of national forests assessed using the Landsat record: Case studies in eastern

United States". In: *Remote Sensing of Environment* 113.7, pp. 1430–1442. DOI: 10.1016/j.rse.2008.06.016.

Huang, C., A. Li, H. Shi, G. Sun, Z. Zhu, S. N. Goward, and J. G. Masek (2009c). "Developing a fine resolution forest height map for Mississippi using Landsat time series observations and GLAS Lidar data". In: *AGU Fall Meeting Abstracts*.

Huang, C., P. Ling, and Z. Zhu (2015). "North Carolina's forest disturbance and timber production assessed using time series Landsat observations". In: *International Journal of Digital Earth* 8.12, pp. 947–969. DOI: 10.1080/17538947.2015.1034200.

Huang, H., Y. Chen, N. Clinton, J. Wang, X. Wang, C. Liu, P. Gong, J. Yang, Y. Bai, Y. Zheng, and Z. Zhu (2017). "Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine". In: *Remote Sensing of Environment* 202, pp. 166–176. DOI: 10.1016/j.rse.2017.02.021.

Jin, S., L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian (2013). "A comprehensive change detection method for updating the national land cover database to circa 2011". In: *Remote Sensing of Environment* 132, pp. 159–175. DOI: 10.1016/j.rse.2013.01.012.

Johansen, K., S. Phinn, and M. Taylor (2015). "Mapping woody vegetation clearing in Queensland, Australia from Landsat imagery using the Google Earth Engine". In: *Remote Sensing Applications: Society and Environment* 1, pp. 36–49. DOI: 10.1016/j.rsase.2015.06.002.

Kayastha, N., V. Thomas, J. Galbraith, and A. Banskota (2012). "Monitoring wetland change using inter-annual Landsat time-series data". In: *Wetlands* 32.6, pp. 1149–1162. DOI: 10.1007/s13157-012-0345-1.

Kennedy, R. E., S. Andréfouët, W. B. Cohen, C. Gómez, P. Griffiths, M. Hais, S. P. Healey, E. H. Helmer, P. Hostert, M. B. Lyons, G. W. Meigs, D. Pflugmacher, S. R. Phinn, S. L. Powell, P. Scarth, S. Sen, T. A. Schroeder, A. Schneider, R. Sonnenschein, J. E. Vogelmann, M. A. Wulder, and Z. Zhu (2014). "Bringing an ecological view of change

to Landsat-based remote sensing". In: *Frontiers in Ecology and the Environment* 12.6, pp. 339–346. DOI: 10.1890/130066.

Kennedy, R. E., Z. Yang, J. Braaten, C. Copass, N. Antonova, C. Jordan, and P. Nelson (2015). "Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA". In: *Remote Sensing of Environment* 166, pp. 271–285. DOI: 10.1016/j.rse.2015.05.005.

Kennedy, R. E., Z. Yang, and W. B. Cohen (2010). "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms". In: *Remote Sensing of Environment* 114.12, pp. 2897–2910. DOI: 10.1016/j.rse.2010.07.008.

Kennedy, R. E., Z. Yang, W. B. Cohen, E. Pfaff, J. Braaten, and P. Nelson (2012). "Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan". In: *Remote Sensing of Environment* 122, pp. 117–133. DOI: 10.1016/j.rse.2011.09.024.

Lambert, S. and T. Brandeis (2018). *Forests of Georgia, 2016. Resource Update FS–176.* Asheville, N.C.: U.S. Department of Agriculture, Forest Service, Southern Research Station. 4 pp.

Li, J., C. E. Zipper, P. F. Donovan, R. H. Wynne, and A. J. Oliphant (2015). "Reconstructing disturbance history for an intensively mined region by time-series analysis of Landsat imagery". In: *Environmental Monitoring and Assessment* 187.9, p. 557. DOI: 10.1007/s10661-015-4766-1.

Li, M., C. Huang, Z. Zhu, H. Shi, H. Lu, and S. Peng (2009a). "Assessing rates of forest change and fragmentation in Alabama, USA, using the vegetation change tracker model". In: *Forest Ecology and Management* 257.6, pp. 1480–1488. DOI: 10.1016/j.foreco.2008.12.023.

Li, M., C. Huang, Z. Zhu, W. Wen, D. Xu, and A. Liu (2009b). "Use of remote sensing coupled with a vegetation change tracker model to assess rates of forest change and

fragmentation in Mississippi, USA". In: *International Journal of Remote Sensing* 30.24, pp. 6559–6574. DOI: 10.1080/01431160903241999.

Liaw, A. and M. Wiener (2002). "Classication and regression by randomForest". In: *R News* 2, pp. 18–22.

Liu, L., D. Peng, Z. Wang, and Y. Hu (2014). "Improving artificial forest biomass estimates using afforestation age information from time series Landsat stacks". In: *Environmental Monitoring and Assessment* 186.11, pp. 7293–7306. DOI: 10.1007/s10661-014-3927-y.

Liu, S. and C. J. Cieszewski (2009). "Impacts of management intensity and harvesting practices on long-term forest resource sustainability in Georgia". In: *Mathematical and Computational Forestry & Natural-Resource Sciences* 1.2, pp. 52–66.

Loveland, T. R. and J. L. Dwyer (2012). "Landsat: Building a strong future". In: *Remote Sensing of Environment* 122, pp. 22–29. DOI: 10.1016/j.rse.2011.09.022.

Lowe, R. C. and C. J. Cieszewski (2014). "Multi-source k-nearest neighbor, Mean balanced forest inventory of Georgia". In: *Mathematical and Computational Forestry & Natural-Resource Sciences* 6.2, pp. 65–79.

Lu, D. (2005). "Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon". In: *International Journal of Remote Sensing* 26.12, pp. 2509–2525. DOI: 10.1080/01431160500142145.

Lu, D., Q. Chen, G. Wang, L. Liu, G. Li, and E. Moran (2016). "A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems". In: *International Journal of Digital Earth* 9.1, pp. 63–105. DOI: 10.1080/17538947.2014.990526.

Lu, D., Q. Chen, G. Wang, E. Moran, M. Batistella, M. Zhang, G. Vaglio Laurin, and D. Saah (2012). "Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates". In: *International Journal of Forestry Research* 2012, pp. 1–16. DOI: 10.1155/2012/436537.

Markham, B., J. Storey, D. Williams, and J. Irons (2004). "Landsat sensor performance: History and current status". In: *IEEE Transactions on Geoscience and Remote Sensing* 42.12, pp. 2691–2694. DOI: 10.1109/TGRS.2004.840720.

Markham, B., J. Barsi, G. Kvaran, L. Ong, E. Kaita, S. Biggar, J. Czapla-Myers, N. Mishra, and D. Helder (2014). "Landsat-8 operational land imager radiometric calibration and stability". In: *Remote Sensing* 6.12, pp. 12275–12308. DOI: 10.3390/rs61212275.

Masek, J. G., S. N. Goward, R. E. Kennedy, W. B. Cohen, G. G. Moisen, K. Schleeweis, and C. Huang (2013). "United States forest disturbance trends observed using Landsat time series". In: *Ecosystems* 16.6, pp. 1087–1104. DOI: 10.1007/s10021-013-9669-9.

Masek, J. G., E. Vermonte, N. Saleous, R. Wolfe, F. Hall, F. Huemmrich, F. Gao, J. Kulter, and T. Lim (2006). "A Landsat surface reflectance data set for North America, 1990-2000". In: *Geoscience and Remote Sensing Letters* 3, pp. 68–72.

Maselli, F., G. Chirici, L. Bottai, P. Corona, and M. Marchetti (2005). "Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images". In: *International Journal of Remote Sensing* 26.17, pp. 3781–3796. DOI: 10.1080/01431160500166433.

Matasci, G., T. Hermosilla, M. A. Wulder, J. C. White, N. C. Coops, G. W. Hobart, D. K. Bolton, P. Tompalski, and C. W. Bater (2018). "Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots". In: *Remote Sensing of Environment* 216, pp. 697–714. DOI: 10.1016/j.rse.2018.07.024.

McRoberts, R. E. (2009). "Diagnostic tools for nearest neighbors techniques when used with satellite imagery". In: *Remote Sensing of Environment. 113: 489-499.* 113.

McRoberts, R. E. (2012). "Estimating forest attribute parameters for small areas using nearest neighbors techniques". In: *Forest Ecology and Management* 272, pp. 3–12. DOI: 10.1016/j.foreco.2011.06.039.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt (2002). "Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique".

In: *Remote Sensing of Environment* 82.2, pp. 457–468. DOI: 10.1016/S0034-4257(02)00064-0.

Miksys, V., I. Varnagiryte-kabasinskiene, I. Stupak, K. Armolaitis, M. Kukkola, and J. Wojcik (2007). "Above-ground biomass functions for Scots pine in Lithuania". In: *Biomass and Bioenergy* 31.10, pp. 685–692. DOI: 10.1016/j.biombioe.2007.06.013.

Miller, M. D. (2012). "The impacts of Atlanta's urban sprawl on forest cover and fragmentation". In: *Applied Geography* 34, pp. 171–179. DOI: 10.1016/j.apgeog.2011.11.010.

Moody, A. and D. M. Johnson (2001). "Land-surface phenologies from AVHRR using the discrete fourier transform". In: *Remote Sensing of Environment* 75.3, pp. 305–323. DOI: 10.1016/S0034-4257(00)00175-9.

Morfitt, R., J. Barsi, R. Levy, B. Markham, E. Micijevic, L. Ong, P. Scaramuzza, and K. Vanderwerff (2015). "Landsat-8 Operational Land Imager (OLI) radiometric performance on-orbit". In: *Remote Sensing* 7.2, pp. 2208–2237. DOI: 10.3390/rs70202208.

Multi-Resolution Land Characteristics Consortium (2019). *NLCD 2016 Land Cover*. URL: https://www.mrlc.gov/data/nlcd-2016-land-cover-conus (visited on 10/01/2019).

Natural Resources Spatial Analysis Lab (2003). *Georgia GAP Analysis*. University of Georgia College of Engineerings. URL: http://narsal.uga.edu/ (visited on 04/14/2020).

Nguyen, T. H., S. D. Jones, M. Soto-Berelov, A. Haywood, and S. Hislop (2020). "Monitoring aboveground forest biomass dynamics over three decades using Landsat time-series and single-date inventory data". In: *International Journal of Applied Earth Observation and Geoinformation* 84, p. 101952. DOI: 10.1016/j.jag.2019.101952.

Nilsson, U. and H. L. Allen (2003). "Short- and long-term effects of site preparation, fertilization and vegetation control on growth and stand development of planted loblolly pine". In: *Forest Ecology and Management* 175.1, pp. 367–377. DOI: 10.1016/S0378-1127(02)00140-8.

Obata, S. (2018). "Estimation of forest stand disturbance through implementation of vegetation change tracker algorithm using Landsat time series stacked imagery in coastal

Georgia". In: *Mathematical and Computational Forestry and Natural-Resource Sciences* 10.1, p. 32.

Obata, S., C. J. Cieszewski, P. Bettinger, R. C. Lowe III, and S. Bernardes (2019). "Preliminary analysis of forest stand disturbances in Coastal Georgia (USA) using Landsat time series stacked imagery". In: *FORMATH* 18, pp. 1–11. DOI: 10.15684/formath.001.

Olofsson, P., G. M. Foody, S. V. Stehman, and C. E. Woodcock (2013). "Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation". In: *Remote Sensing of Environment* 129, pp. 122–131. DOI: 10.1016/j.rse.2012.10.031.

Omernik, J. M. and G. E. Griffith (2014). "Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework". In: *Environmental Management* 54.6, pp. 1249–1266. DOI: 10.1007/s00267-014-0364-1.

Oswalt, S. N., W. B. Smith, P. D. Miles, and S. A. Pugh (2014). *Forest resources of the United States, 2012: A technical document supporting the Forest Service 2010 update of the RPA assessment*. DOI: 10.2737/WO-GTR-91.

Oswalt, S. N., W. B. Smith, P. D. Miles, and S. A. Pugh (2018). *Forest resources of the United States, 2017*. WO-GTR-91. Washington, DC: U.S. Department of Agriculture, Forest Service. DOI: 10.2737/WO-GTR-91.

Pahlevan, N., S. V. Balasubramanian, S. Sarkar, and B. A. Franz (2018). "Toward long-term aquatic science products from heritage Landsat missions". In: *Remote Sensing* 10.9, p. 1337. DOI: 10.3390/rs10091337.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay (2011). "Scikit-learn: Machine learning in Python". In: *Journal of Machine Learning Research* 12, 2825 − 2830.

Pflugmacher, D., W. B. Cohen, and R. E. Kennedy (2012). "Using Landsat-derived disturbance history (1972–2010) to predict current forest structure". In: *Remote Sensing of Environment* 122, pp. 146–165. DOI: 10.1016/j.rse.2011.09.025.

Pflugmacher, D., W. B. Cohen, R. E. Kennedy, and Z. Yang (2014). "Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics". In: *Remote Sensing of Environment* 151, pp. 124–137. DOI: 10.1016/j.rse.2013.05.033.

Pickell, P. D., T. Hermosilla, N. C. Coops, J. G. Masek, S. Franks, and C. Huang (2014). "Monitoring anthropogenic disturbance trends in an industrialized boreal forest with Landsat time series". In: *Remote Sensing Letters* 5.9, pp. 783–792. DOI: 10.1080/2150704X.2014.967881.

Prud'homme, B. A. and J. G. Greis (2002). "Best management practices in the South". In: *Southern forest resource assessment*. Asheville, NC: Department of Agriculture, Forest Service, Southern Research Station, pp. 519–536.

Reese, H., T. Granqvist-Pahlén, M. Egberth, M. Nilsson, and H. Olsson (2005). "Automated estimation of forest parameters for Sweden using Landsat data and the kNN algorithm". In: *Proceedings of the 31st International Symposium on Remote Sensing*. St. Petersburg, Russia, p. 4.

Reese, H., M. Nilsson, T. G. Pahén, O. Hagner, S. Joyce, U. Tingelöf, M. Egberth, and H. Olsson (2003). "Countrywide estimates of forest variables using satellite data and field data from the national forest inventory". In: *Ambio* 32.8, pp. 542–548. DOI: 10.1579/0044-7447-32.8.542.

Reese, H., M. Nilsson, P. Sandström, and H. Olsson (2002). "Applications using estimates of forest parameters derived from satellite and forest inventory data". In: 37.

Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.e., R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner,

J. G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R. H. Wynne, and Z. Zhu (2014). "Landsat-8: Science and product vision for terrestrial global change research". In: *Remote Sensing of Environment* 145, pp. 154–172. DOI: 10.1016/j.rse.2014.02.001.

Schott, J. R., A. Gerace, C. E. Woodcock, S. Wang, Z. Zhu, R. H. Wynne, and C. E. Blinn (2016). "The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments". In: *Remote Sensing of Environment* 185, pp. 37–45. DOI: 10.1016/j.rse.2016.04.015.

Seidl, R., M. Schelhaas, W. Rammer, and P. J. Verkerk (2014). "Increasing forest disturbances in Europe and their impact on carbon storage". In: *Nature Climate Change* 4.9, pp. 806–810. DOI: 10.1038/nclimate2318.

Shelestov, A., M. Lavreniuk, N. Kussul, A. Novikov, and S. Skakun (2017). "Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping". In: *Frontiers in Earth Science* 5. DOI: 10.3389/feart.2017.00017.

Shumway, R. H. and D. S. Stoffer (2017). "Spectral analysis and filtering". In: *Time series analysis and its applications: With r examples*. 4th. New York, NY: Springer Science+Business Media, pp. 165–172.

Smith, W. (2002). "Forest inventory and analysis: A national inventory and monitoring program". In: *Environmental Pollution* 116, pp. 233–242. DOI: 10.1016/S0269-7491(01)00255-X.

Soenen, S. A., D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall (2010). "Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain". In: *Remote Sensing of Environment* 114.7, pp. 1325–1337. DOI: 10.1016/j.rse.2009.12.012.

Tanaka, S., T. Takahashi, T. Nishizono, F. Kitahara, H. Saito, T. Iehara, E. Kodani, and Y. Awaya (2014). "Stand volume estimation using the k-NN technique combined with

forest inventory data, satellite image data and additional feature variables". In: *Remote Sensing* 7.1, pp. 378–394. DOI: 10.3390/rs70100378.

Thomas, N. E., C. Huang, S. N. Goward, S. Powell, K. Rishmawi, K. Schleeweis, and A. Hinds (2011). "Validation of North American forest disturbance dynamics derived from Landsat time series stacks". In: *Remote Sensing of Environment* 115.1, pp. 19–32. DOI: 10.1016/j.rse.2010.07.009.

Tinkham, W. T., P. R. Mahoney, A. T. Hudak, G. M. Domke, M. J. Falkowski, C. W. Woodall, and A. M. Smith (2018). "Applications of the United States Forest Inventory and Analysis dataset: A review and future directions". In: *Canadian Journal of Forest Research* 48.11, pp. 1251–1268. DOI: 10.1139/cjfr-2018-0196.

Tompalski, P., J. C. White, N. C. Coops, and M. A. Wulder (2019). "Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data". In: *Remote Sensing of Environment* 227, pp. 110–124. DOI: 10.1016/j.rse.2019.04.006.

Tomppo, E. (1990). "Designing a satellite image-aided national forest survey in Finland". In: Usability of remote sensing for forest inventory and planning. Umea, Sweden: Swedish University of Agricultural Sciences, pp. 43–47.

Tomppo, E., M. Haakana, M. Katila, and J. Peräsaari (2008). *Multi-source national forest inventory: Methods and applications*. Springer Netherlands. 373 pp. DOI: 10.1007/978-1-4020-8713-4.

Tomppo, E. and M. Halme (2004). "Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach". In: *Remote Sensing of Environment* 92.1, pp. 1–20. DOI: 10.1016/j.rse.2004.04.003.

Townshend, J. R. G. (1992). "Land cover". In: *International Journal of Remote Sensing* 13.6, pp. 1319–1328. DOI: 10.1080/01431169208904193.

Turner, M. G. (1987). "Land use changes and net primary production in the Georgia, USA, landscape: 1935–1982". In: *Environmental Management* 11.2, pp. 237–247. DOI: 10.1007/BF01867202.

U.S. Department of the Interior, Fish and Wildlife Service (2003). *Recover plan for the Red-Cockaded Woodpecker (Picoides borealis)*. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service, p. 316.

U.S. Forest Service (1988). *The South's fourth forest : Alternatives for the future.* Vol. 24. Washington, D.C.: U.S. Department of Agriculture, Forest Service. 532 pp.

U.S. Geological Survey (2018). *Landsat collections*. URL: http://pubs.er.usgs.gov/publication/fs20183049 (visited on 03/02/2020).

U.S. Geological Survey (2020). *Landsat levels of processing*. URL: https://www.usgs.gov/land-resources/nli/landsat/landsat-levels-processing (visited on 02/09/2020).

US Census Bureau (2019). *Metropolitan and micropolitan statistical areas totals: 2010-2018*. URL: https://www.census.gov/data/tables/time-series/demo/popest/2010s-total-metro-and-micro-statistical-areas.html (visited on 08/20/2019).

Vermote, E., C. Justice, M. Claverie, and B. Franch (2016). "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product". In: *Remote Sensing of Environment* 185, pp. 46–56. DOI: 10.1016/j.rse.2016.04.008.

Vogelmann, J. E., A. L. Gallant, H. Shi, and Z. Zhu (2016). "Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data". In: *Remote Sensing of Environment* 185, pp. 258–270. DOI: 10.1016/j.rse.2016.02.060.

Vogelmann, J. E., G. Xian, C. Homer, and B. Tolk (2012). "Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems". In: *Remote Sensing of Environment* 122, pp. 92–105. DOI: 10.1016/j.rse.2011.06.027.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam (2006). "Warming and earlier spring increase Western U.S. forest wildfire activity". In: *Science* 313.5789, pp. 940–943. DOI: 10.1126/science.1128834.

Williams, C. A., H. Gu, R. MacLean, J. G. Masek, and G. J. Collatz (2016). "Disturbance and the carbon balance of US forests: A quantitative review of impacts from harvests, fires, insects, and droughts". In: *Global and Planetary Change* 143, pp. 66–80. DOI: 10.1016/j.gloplacha.2016.06.002.

Wilson, B. T., A. J. Lister, and R. I. Riemann (2012). "A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data". In: *Forest Ecology and Management* 271, pp. 182–198. DOI: 10.1016/j.foreco.2012.02.002.

Wilson, B. T., J. F. Knight, and R. E. McRoberts (2018). "Harmonic regression of Landsat time series for modeling attributes from national forest inventory data". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 137, pp. 29–46. DOI: 10.1016/j.isprsjprs.2018.01.006.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne (2008). "Free access to Landsat imagery". In: *Science* 320.5879, pp. 1011–1011. DOI: 10.1126/science.320.5879.1011a.

Wulder, M. A., R. S. Skakun, W. A. Kurz, and J. C. White (2004). "Estimating time since forest harvest using segmented Landsat ETM+ imagery". In: *Remote Sensing of Environment* 93.1, pp. 179–187. DOI: 10.1016/j.rse.2004.07.009.

Wulder, M. A., N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla (2018). "Land cover 2.0". In: *International Journal of Remote Sensing* 39.12, pp. 4254–4284. DOI: 10.1080/01431161.2018.1452075.

Wulder, M. A., T. R. Loveland, D. P. Roy, C. J. Crawford, J. G. Masek, C. E. Woodcock, R. G. Allen, M. C. Anderson, A. S. Belward, W. B. Cohen, J. Dwyer, A. Erb, F. Gao, P. Griffiths, D. Helder, T. Hermosilla, J. D. Hipple, P. Hostert, M. J. Hughes, J. Huntington, D. M. Johnson, R. Kennedy, A. Kilic, Z. Li, L. Lymburner, J. McCorkel, N. Pahlevan, T. A. Scambos, C. Schaaf, J. R. Schott, Y. Sheng, J. Storey, E. Vermote, J. Vogelmann, J. C. White, R. H. Wynne, and Z. Zhu (2019). "Current status of Landsat program, science, and applications". In: *Remote Sensing of Environment* 225, pp. 127–147. DOI: 10.1016/j.rse.2019.02.015.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock (2012). "Opening the archive: How free data has enabled the science and monitoring promise of Landsat". In: *Remote Sensing of Environment* 122, pp. 2–10. DOI: 10.1016/j.rse.2012.01.010.

Wulder, M. A., J. C. White, T. R. Loveland, C. E. Woodcock, A. S. Belward, W. B. Cohen, E. A. Fosnight, J. Shaw, J. G. Masek, and D. P. Roy (2016). "The global Landsat archive: Status, consolidation, and direction". In: *Remote Sensing of Environment* 185, pp. 271–283. DOI: 10.1016/j.rse.2015.11.032.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian (2018). "A new generation of the United States national land cover database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146, pp. 108–123. DOI: 10.1016/j.isprsjprs.2018.09.006.

Zampieri, N. E., S. Pau, and D. K. Okamoto (2019). "The impact of Hurricane Michael on longleaf pine habitats in Florida". In: *bioRxiv*, p. 736629. DOI: 10.1101/736629.

Zhang, G. and Y. Lu (2012). "Bias-corrected random forests in regression". In: *Journal of Applied Statistics* 39.1, pp. 151–160. DOI: 10.1080/02664763.2011.578621.

Zhao, P., D. Lu, G. Wang, C. Wu, Y. Huang, and S. Yu (2016). "Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve for-

est aboveground biomass estimation". In: *Remote Sensing* 8.6, p. 469. DOI: 10.3390/rs8060469.

Zhu, Z. (2019). "Science of Landsat analysis ready data". In: *Remote Sensing* 11.18, p. 2166. DOI: 10.3390/rs11182166.

Zhu, Z., Y. Fu, C. E. Woodcock, P. Olofsson, J. E. Vogelmann, C. Holden, M. Wang, S. Dai, and Y. Yu (2016). "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014)". In: *Remote Sensing of Environment* 185, pp. 243–257. DOI: 10.1016/j.rse.2016.03.036.

Zhu, Z. and C. E. Woodcock (2014). "Continuous change detection and classification of land cover using all available Landsat data". In: *Remote Sensing of Environment* 144, pp. 152–171. DOI: 10.1016/j.rse.2014.01.011.

Zhu, Z. (2017). "Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 130, pp. 370–384. DOI: 10.1016/j.isprsjprs.2017.06.013.

Zhu, Z., S. Wang, and C. E. Woodcock (2015a). "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images". In: *Remote Sensing of Environment* 159, pp. 269–277. DOI: 10.1016/j.rse.2014.12.014.

Zhu, Z. and C. E. Woodcock (2012a). "Object-based cloud and cloud shadow detection in Landsat imagery". In: *Remote Sensing of Environment* 118, pp. 83–94. DOI: 10.1016/j.rse.2011.10.028.

Zhu, Z., C. E. Woodcock, C. Holden, and Z. Yang (2015b). "Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time". In: *Remote Sensing of Environment* 162, pp. 67–83. DOI: 10.1016/j.rse.2015.02.009.

Zhu, Z., C. E. Woodcock, and P. Olofsson (2012b). "Continuous monitoring of forest disturbance using all available Landsat imagery". In: *Remote Sensing of Environment* 122, pp. 75–91. DOI: 10.1016/j.rse.2011.10.030.

Zhu, Z., J. Zhang, Z. Yang, A. H. Aljaddani, W. B. Cohen, S. Qiu, and C. Zhou (2020). "Continuous monitoring of land disturbance based on Landsat time series". In: *Remote Sensing of Environment* 238, p. 111116. DOI: 10.1016/j.rse.2019.03.009.